National Institute for Health and Care Excellence

Final

Urinary tract infection in under 16s

[A] Evidence review for UTI diagnosis in infants and children under 3 months and 3 months to 3 years

NICE guideline NG224

Evidence reviews

[September 2017]

Final

These evidence reviews were developed by NICE's Guideline Updates Team

Update information

July 2021: The recommendations section of this document was updated following a review of the evidence on symptoms and signs of urinary tract infections. See the recommendations section for more details.

ISBN: 978-1-4731-4648-8

Table of Contents

UTI diagnosis in infants and children under 3 months and 3 months to 3 years

Review question

In infants and children up to 3 years old with suspected urinary tract infection (UTI), what is the diagnostic accuracy of urine tests for detecting UTI?

Introduction

Urinary tract infections (UTIs) most commonly occur when bacteria from the gut enter the urinary tract through the urethra.

The recommendations on dipstick testing in the 2007 NICE guideline on urinary tract infection in under 16s were organised by age-group as follows: under 3 months, 3 months or older but younger than 3 years and over 3 years. This reflected the evidence base that dipstick testing was not accurate in children up to 3 years of age. This topic was reviewed in 2016 by the NICE surveillance team and new evidence (5 studies on the diagnostic accuracy of urine dipstick testing) were identified in the younger age group. This evidence suggested that the guideline should be updated to reflect new evidence in this area. This evidence review focuses on the diagnostic accuracy of dipstick tests alone or in combination with other tests in infants under 3 months and 3 months or older but younger than 3 years.

PICO table

Population	Those in whom there is a clinical suspicion of UTI and are: • less than 3 months old • 3 months or older but younger than 3 years

Index test	• Dipstick test
	  ○ Leukocyte esterase
	  ○ Nitrites
	  ○ Protein
	  ○ Blood
	• Dipstick testing with other tests including:
	  ○ microscopy alone (automated or manual)
	  ○ urine culture alone (can include clean catch, bladder catheterisation and suprapubic aspirate samples)
	  ○ microscopy and culture.

Reference test	Clinical diagnosis of UTI. This may include consideration of a urine culture alone or a combination of tests.
Outcomes	• Sensitivity • Specificity • Likelihood ratios

Methods and process

This evidence review was developed using the methods and process described in 'Developing NICE guidelines: the manual'. Methods specific to this review question are described in the review protocol in Appendix A.

Declarations of interest were recorded according to NICE's 2014 conflicts of interest policy.

A systematic review literature search was carried out to identify randomised controlled trials, cohort studies, cross-sectional studies and systematic reviews of diagnostic accuracy studies (see Appendix B for literature search strategy). The search identified 7,158 articles, which were screened on title and abstract. Of these, 61 potentially relevant articles were ordered for full text review. Additionally, 10 articles were identified as potentially relevant from the included studies in the original UTI in children guideline and in references of systematic reviews [Whiting 2006 (included in economics evidence review), Hay 2016 (included), Deville 2004 (excluded studies table). In total, 71 articles were assessed in full. Of these, 13 were included based on their relevance to the review protocol (Appendix A:) and the presentation of data which was in a useful format for analysis. The clinical evidence study selection is available in Appendix C:.

Evidence was stratified into two age groups: under 3 months and 3 months or older but younger than 3 years. Evidence for the age group of 3 months or older but younger than 3 years was separated by reference test of either culture alone or culture with microscopy, while all evidence included for infants under 3 months used culture alone as the reference test.

Sensitivity, specificity, positive and negative likelihood ratios were calculated for each outcome. Where applicable, meta-analysis of diagnostic test accuracy was conducted. Where sufficient data were available (4 or more studies), a bivariate analysis was run in R (version 3.3.1), which accounts for the correlations between positive and negative likelihood ratios, and between sensitivities and specificities. Where fewer than 4 studies were available, separate pooling was conducted for sensitivity, specificity, positive and negative likelihood ratios using Microsoft Excel. This somewhat conservative approach is likely to underestimate test accuracy because it fails to account for the correlation and trade-off between sensitivity and specificity. Where there were sufficient studies, a bivariate meta-analyses was run to test the validity of the univariate method, and no meaningful differences were observed between the bivariate and univariate approaches. Therefore, the univariate analysis is presented in this evidence review. Random-effects models (der Simonian and Laird) were fitted for all syntheses. See Appendix E: for diagnostic meta-analysis forest plots.

The quality of the diagnostic accuracy outcomes were assessed in modified GRADE tables (Appendix F:). The initial quality ratings for outcomes were set as high for prospective cohort or cross-sectional studies, and moderate for retrospective cohort or cross-sectional studies. This is because the risk of bias from patient selection is considered higher in retrospective studies as there is a potential that urine cultures were undertaken dependent on dipstick test result. If 50% or more of the weight in a pooled meta-analysis came from retrospective studies, the quality of the evidence began at moderate and was then downgraded accordingly.

Four domains are taken into account when downgrading evidence from this initial point: risk of bias, inconsistency, indirectness and imprecision:

- Risk of bias was assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool which takes into account patient selection, index and reference test and flow and timing. The assessment for individual studies is included in the clinical evidence tables (Appendix D).
- Inconsistency (heterogeneity occurring when there is unexplained variability in the treatment effect across studies) was assessed using the I^2 statistic and was considered serious, and the outcome downgraded one level, if the statistic was greater than or equal to 50%.
- Indirectness was used as a reason to downgrade the quality of evidence if a single study, or more than a third of the studies in a meta-analysis, were indirect compared to the review protocol. No indirect index or reference tests were included, and therefore indirectness was downgraded for based only on population age. For example, where a study is included for the age group of 3 months or older but younger than 3 years, but

includes all children below 1 or 2 years, this was downgraded for including an indirect population.

- Imprecision was assessed using the 95% confidence intervals (CIs) of likelihood ratios. Minimal important differences (MIDs) of 0.5 and 2 were defined. A positive likelihood ratio which spans 2 was downgraded for serious imprecision as the data was deemed to be consistent with a meaningful increase in risk and no meaningful predictive value. Similarly, a negative likelihood ratio which spans 0.5 led to downgrading for serious imprecision. Any likelihood ratios that spanned both 0.5 and 2 were downgraded twice for very serious imprecision.

The following schema, adapted from the suggestions of Jaeschke et al. (1994), was used to interpret the likelihood ratio findings:

Value of likelihood ratio	Interpretation
LR ≤ 0.1	**Very large** decrease in probability of disease
0.1 < LR ≤ 0.2	**Large** decrease in probability of disease
0.2 < LR ≤ 0.5	**Moderate** decrease in probability of disease
0.5 < LR ≤ 1.0	**Slight** decrease in probability of disease
1.0 < LR < 2.0	**Slight** increase in probability of disease
2.0 ≤ LR < 5.0	**Moderate** increase in probability of disease
5.0 ≤ LR < 10.0	**Large** increase in probability of disease
LR ≥ 10.0	**Very large** increase in probability of disease

Likelihood ratios are statistically significant if the 95% CIs do not cross 1, as a value of 1 means a test is equivalent to random chance.

A sensitivity analysis was conducted on method of urine sampling by including non-invasive methods (clean catch, sterile bag) only. See: Sensitivity analysis: urine collection method.

Clinical evidence

Included studies

Thirteen studies which met the inclusion criteria were included in the clinical evidence review. Of these, diagnostic accuracy data could not be calculated in one study (DUTY study, Hay 2016) and the authors were contacted to provide relevant information. The information obtained from the authors is presented in Appendix L: and included in this guideline update. In the remaining studies, 1 (McGillivray 2005) reported relevant information for both age groups, 3 reported information for under 3 months only and 8 reported relevant information for the 3 months or older but younger than 3 years age group.

Excluded studies

The excluded studies table is available in Appendix K:.

Summary of clinical studies included in the evidence review

A summary of the included studies is provided in Table 1 to Table 3.

Table 1: Included studies for infants under 3 months

Study ID	Primary publication	Study population	Index test	Reference test
Infants under 3 months				
Dayan 2002	Dayan, P.S., Bennett, J., Best, R.	N = 193 Age: < 60 days	Dipstick assessed using Super UA	Positive culture defined as ≥ 10^4

Study ID	Primary publication	Study population	Index test	Reference test
Infants under 3 months				
	et al (2002). Test characteristics of the urine Gram stain in infants≤ 60 days of age with fever. Pediatric emergency care, 18(1), pp.12-14.	Setting: secondary (emergency care) Country = USA Symptoms: not reported, inclusion based on reported or recorded rectal temperature ≥ 38°C Urine sampling method: urethral catheterisation, SPA	automated urine analyser. • Any nitrite alone • Any LE alone • Nitrite and LE • Nitrite or LE	cfu/ml of a single pathogen from a catheterised sample or 10^3 from SPA sample.
Glissmeyer 2014	Glissmeyer, E.W., Korgenski, E.K., Wilkes, J. et al (2014). Dipstick screening for urinary tract infection in febrile infants. Pediatrics, 133(5), pp.e1121-7.	N = 6394 Age: < 90 days Country: USA Setting: various secondary care centres Symptoms: specific symptoms unclear, states very few were asymptomatic Urine sampling method: urethral catheterisation.	Dipstick (using analyser), dipstick and microscopy. Dipstick positive: nitrite or LE (≥trace) positive. Microscopy positive: > 10 WBCs/hpf or any bacteria.	Culture: ≥1 urine pathogens, each with a quantity of ≥50 000 cfu/ml
Velasco 2015	Velasco, R., Benito, H., Mozun, R. et al (2015). Using a urine dipstick to identify a positive urine culture in young febrile infants is as effective as in older patients. Acta Paediatrica, and International Journal of Paediatrics, 104(1), pp.e39-e44.	N= 3401 Age: < 90 days Country: Spain Setting: paediatric emergency department Symptoms: all symptomatic – fever without source, mean 38.4°C Urine sampling method: urethral catheterisation, SPA	Dipstick, visual reading. • LE (if > 1+) • Nitrite • LE or nitrite • LE and nitrite	Culture: ≥50 000cfu/mL of a single pathogen in a urine sample

Abbreviations: colony forming units per millilitre (cfu/ml), high-power microscopic field (HPF), leukocyte esterase (LE), suprapubic aspiration (SPA), white blood cells (WBCs)

Table 2: Included studies for infants and children aged 3 months or older but younger than 3 years

Study ID	Primary publication	Study population	Index test	Reference test
Infants and children 3 months or older but younger than 3 years				

Study ID	Primary publication	Study population	Index test	Reference test
Doley and Nelligan 2003	Doley A, and Nelligan M. (2003). Is a negative dipstick urinalysis good enough to exclude urinary tract infection in paediatric emergency department patients? Emergency Medicine, 15(1), pp.77-80.	N= 160 Age: 0 – 2 years Country: Australia Setting: emergency medical department (single centre) Symptoms: not reported Urine sampling method: bag specimen or clean catch (4 cases via SPA)	Dipstick, using analyser. • Nitrite or LE or blood or protein positive.	Urine culture: > 100,000 cfu/ml
Kanegaye 2014	Kanegaye, J.T., Jacob, J.M. and Malicki, D., (2014). Automated urinalysis and urine dipstick in the emergency evaluation of young febrile children. Pediatrics, 134(3), pp.523-9.	N= 342 Age: median 8.1 months (IQR: 3.6-14.3 months) Country: USA Setting: paediatric emergency department of tertiary hospital Symptoms: all patients were febrile, mean maximum temperature: 38.8°C Urine sampling method: urethral catheterisation	Dipstick, interpreted visually. • Nitrite positive • LE (≥ trace) • LE (≥ trace) or nitrite	Culture, ≥ 50,000 cfu/ml
Kazi 2013	Kazi, B.A., Buffone, G.J., Revell, P.A. et al (2013). Performance characteristics of urinalyses for the diagnosis of pediatric urinary tract infection. The American Journal of Emergency Medicine, 31(9), pp.1405-7.	N= 1639 Age: 6 to 23 months Country: USA Setting: tertiary hospital paediatric emergency department (single centre) Symptoms: Urine sampling method: urethral catheterisation and void. SPA in 0.02%	Dipstick (POCT): LE positive threshold not defined. • LE or nitrite	Culture: 50,000 cfu/ml of a single organism for specimens collected by voiding / catheter, or grew at least 1000 cfu/ml for SPA specimens.
Lejeune 1991	Lejeune, B., Baron, R., Guillois, B. et al (1991). Evaluation of a screening test for detecting urinary tract infection in newborns and infants. Journal of Clinical Pathology, 44(12), pp.1029-30.	N= 243 Age: under 18 months Country: France Setting: secondary care (single centre) Symptoms: not reported	Dipstick read by analyser. LE threshold not reported. • LE • Nitrite • LE and Nitrite • Protein • LE and protein	Culture and microscopy: • Culture: 100,000 cfu/ml • Microscopy: WBC > 25 x 109/L for boys or 50 x 109/L for girls.

Study ID	Primary publication	Study population	Index test	Reference test
		Urine sampling method: not reported	• LE and protein and nitrite	
Reardon 2009	Reardon, J.M., Carstairs, K.L., Rudinsky, S.L. et al (2009). Urinalysis is not reliable to detect a urinary tract infection in febrile infants presenting to the ED. The American Journal of Emergency Medicine, 27(8), pp.930-2.	N= 435 Age: mean 12.6 months (median 12 months) Country: USA Setting: tertiary care hospital emergency department (single centre) Symptoms: symptomatic (mean temperature: 38.9°C) Urine sampling method: urethral catheterisation.	Dipstick (method of assessment not reported) and microscopy: • Dipstick: LE or nitrite and microscopy positive: ≥5 wbc/hpf	Culture: positive if at least 10,000 cfu/ml.
Sharief 1998	Sharief, N., Hameed, M. and Petts, D., (1998). Use of rapid dipstick tests to exclude urinary tract infection in children. British Journal of Biomedical Science, 55(4), pp.242-6.	N= 124 Age: < 1 year Country: UK Setting: secondary care (single centre) Symptoms: fever (not defined) Urine sampling method: clean catch or sterile paediatric collection bag	Dipstick, assessed using analyser. LE read as either positive or negative. • Nitrite • LE or nitrite • LE and nitrite	Culture, ≥ 100,000 cfu/ml
Shaw 1991	Shaw, K.N., Hexter, D., McGowan, K.L. et a (1991). Clinical evaluation of a rapid screening test for urinary tract infections in children. The Journal of Pediatrics, 118(5), pp.733-736.	N= 145 Age: < 2 years Country: USA Setting: paediatric hospital emergency department (single centre) Symptoms: 144/145 had samples as part of fever or sepsis evaluation Urine sampling method: 128 (88%) by urethral catheter; remainder unspecified (study allowed urine bag / midstream specimen / clean catch)	Dipstick, visual reading. • ≥ trace LE or nitrite • ≥ small LE (1+) or nitrite	Culture, catheter: 1000 cfu/ml, clean catch: 100,000 cfu/ml

Study ID	Primary publication	Study population	Index test	Reference test
Shaw 1998	Shaw, K.N., McGowan, K.L., Gorelick, M.H. et al (1998). Screening for urinary tract infection in infants in the emergency department: which test is best?. Pediatrics, 101(6), pp.E1-E1.	N = 3394 Age: mean 9.2 months (SD 5.7) Country: USA Setting: emergency department of one urban children's hospital Symptom: mean temperature: 39.2°C (SD 2.3 Urine sampling method: urethral catheter (99%); midstream urine in sterile container (1%)	Dipstick, read visually • ≥ trace LE or nitrite	Culture, 10000 cfu/ml

Abbreviations: colony forming units per millilitre (cfu/ml), high-power microscopic field (HPF), interquartile range (IQR), leukocyte esterase (LE), point of care testing (POCT), suprapubic aspiration (SPA), white blood cells (WBCs)

Table 3: Included studies for both infants under 3 months and infants and children aged 3 months or older but younger than 3 years

Study ID	Primary publication	Study population	Index test	Reference test
Both infants under 3 months and 3 months or older but younger than 3 years				
McGillivray 2005	McGillivray, D., Mok, E., Mulrooney, E. et al (2005). A head-to-head comparison: "clean-void" bag versus catheter urinalysis in the diagnosis of urinary tract infection in young children. The Journal of Pediatrics, 147(4), pp.451-6.	N = 303 Age: < 90 days and 3 months or older but younger than 3 years Country: Canada Setting: paediatric emergency department (single centre) Symptom: both symptomatic (rectal equivalent temperature of 39.5°C in 53/297 and asymptomatic. Urine sampling method: urethral catheterisation	Dipstick, read by analyser • LE (> trace) or nitrite	Culture, > 10000 cfu/ml NOTE: catheter samples were obtained only from children with specific clinical indications (following bag sample collection and urinalysis). Therefore, this was a high prevalence population.
Hay 2016	Hay A, Birnie K, Busby J, et al. 2016. The Diagnosis of Urinary Tract infection in Young children (DUTY): a diagnostic	N = 2884 infants and children aged under 3 years Country: UK	Dipstick test	Pure (single) or predominant growth of a uropathogen at 100,000 cfu/ml.

Study ID	Primary publication	Study population	Index test	Reference test
	prospective observational study to derive and validate a clinical algorithm for the diagnosis of urinary tract infection in children presenting to primary care with an acute illness. Health Technol Assess 2016;20(51)	Setting: primary care Symptom: n = 104 with temperature ≥ 39 °C Urine sampling method: clean catch (preferred) or nappy pad		

Abbreviations: colony forming units per millilitre (cfu/ml), leukocyte esterase (LE),

See Appendix D: for full evidence tables.

Quality assessment of clinical studies included in the evidence review

All studies included were of prospective or retrospective cohort or cross sectional study design, of which grading of the quality of the evidence started at high for prospective studies and moderate for retrospective studies. All studies included were downgraded due to risk of bias, mainly due to lack of or unclear blinding between index and reference test. Therefore, the quality of the evidence ranged from moderate to very low quality. Other areas of downgrading included serious imprecision, serious inconsistency (heterogeneity in meta-analysis) or indirect age groups of populations included.

See Appendix F: for full GRADE tables.

Economic evidence

Included studies

A systematic literature search was conducted to identify relevant economic analyses, the details of which are shown in appendix B. The search identified a total of 558 articles, of which 4 studies were identified for full text screening, and 2 were included in the final economic evidence review. The economic evidence study selection process is documented in appendix H, and the details of the 2 excluded studies are shown in Appendix K.

Excluded studies

Two studies were excluded during full text screening. Details of exclusion criteria are listed in Appendix K.

Summary of studies included in the economic evidence review

Whiting et al, 2006 is a health technology assessment of the clinical and cost effectiveness of tests for the diagnosis of UTI in children which included an economic analysis. Key economic results are summarised in Table 4. This analysis used a model-based approach to estimate lifetime costs and QALYs associated with various strategies of testing for UTI and subsequent imaging for reflux. Patient subgroups were stratified by gender and by age (<1 year, 1-2 years, 2-3 years, and >3 years). Due to the large number of possible strategies, results were simply reported as the strategy with the highest expected net benefit at a variety of cost effectiveness thresholds, rather than reporting costs, QALYs and ICERs for each strategy.

At a threshold of £20,000 per QALY, for girls <1 year the optimal strategy was dipstick testing with the presence of nitrite *or* LE interpreted as a positive result, followed by confirmatory culture of positive results, followed by MCUG as the imaging test for reflux. For girls 1-2

years and 2-3 years the optimal strategy was dipstick testing with the presence of nitrite *and* LE interpreted as a positive result, followed by MCUG. For girls > 3 years, the optimal strategy was treating all patients with suspected UTI. For boys <1 year and 1-2 to years, the optimal strategy was dipstick testing with the presence of nitrite *and* LE interpreted as being positive followed by MCUG, and for boys 2-3 years and boys >3 years the optimal strategy was treating all patients with suspected UTI.

It should be noted that in this analysis, due to the large number of strategies, the probability of any one particular strategy being the most cost effective is always relatively low. However, results indicate that, in cases where the optimal strategy is not to simply treat everyone, strategies involving dipstick are generally expected to be the most cost effective. A key limitation of this study is the assumption that the accuracy of diagnostic tests is the same across patients of all ages. This is an important shortcoming, considering the objective of the review question is to determine whether the diagnostic accuracy of UTI tests varies with age and, by extension, whether this affects cost effectiveness.

Hay et al, 2016 is also a health technology assessment which included a secondary economic analysis which examined the cost-effectiveness of a dipstick testing strategy in children under 5 years at low risk of UTI (defined by a GP answering yes to the question: 'if this child was NOT in the DUTY study would you have requested a urine sample?').

This analysis used a modelling approach to compare 3 strategies: dipstick testing (all children tested with dipstick and a urine sample sent for laboratory testing, with the dipstick result used to direct antibiotic treatment while awaiting laboratory results), laboratory testing (urine sample sent for laboratory testing, and antibiotic treatment started on receipt of a positive test result), and presumptive treatment (antibiotics prescribed for all children, and a urine sample sent for laboratory testing). Results showed that both the dipstick testing and presumptive treatment strategies were associated with higher costs and a higher number of QALYs than the laboratory testing strategy. However, the laboratory testing strategy was the most cost-effective at threshold of £20,000 per QALY, due to the relatively small incremental QALY benefit produced by the dipstick testing and presumptive treatment strategies.

It should be noted that, in this analysis, the testing and treatment algorithm for dipstick testing differs fundamentally from the algorithm recommended in the 2007 NICE guidance for UTI testing in children over 3. The Hay analysis examines the cost-effectiveness of a strategy of dipstick testing followed by laboratory testing regardless of the result, while 2007 NICE guidance for children over 3 recommends that laboratory testing is only carried out for children with a positive dipstick test for either nitrite or LE.

Table 4: Economic evidence profile

<table>
<tr><th>Study</th><th>Applicability</th><th>Limitations</th><th>Other comments</th><th>Results</th><th>Uncertainty</th></tr>
<tr><td>Whiting et al (2006)

Dipstick testing (with and without culture) versus microscopy and/or culture

UK</td><td>Directly applicable

Study is UK based, and modelling is from the perspective of the NHS.</td><td>Very serious limitations

Study makes the assumption that accuracy of tests does not vary by age.</td><td>Model-based analysis with a lifetime time horizon.

Does not report costs and QALYs for interventions, only the intervention with the highest expected net monetary benefit at a variety of thresholds.</td><td>Strategy with the highest probability of being cost-effective at a £20,000/QALY threshold:

Girl <1 year: Dipstick (positive for nitrite or LE), followed by confirmatory laboratory culture, followed by MCUG
Girl 1-2 years: Dipstick (positive for nitrite and LE), followed by MCUG
Girl 2-3 years: Dipstick (positive for nitrate and LE), followed by MCUG
Girl >3 years: Treat all patients with suspected UTI

Boy <1 year: Dipstick (positive for nitrite and LE) followed by MCUG
Boy 1-2 years: Dipstick (positive for nitrite and LE) followed by MCUG
Boy 2-3 years: Treat all patients with suspected UTI
Boy >3 years: Treat all patients with suspected UTI</td><td>Probabilistic sensitivity analysis showed that, due to the large number of possible strategies, the probability of the strategy with the highest expected net monetary benefit being the most cost effective was generally low at any threshold.

A deterministic sensitivity analysis was conducted in which strategies involving glucose testing with dipsticks were included for children >3 (these were excluded from the main analysis due to poor quality of data). Results indicated that glucose testing followed by MCUG becomes the optimal strategy for girls at thresholds ≥£24,000/QALY for girls and ≥£40,000/QALY for boys.</td></tr>
<tr><td>Hay et al (2016)

Dipstick testing (DT) versus laboratory</td><td>Partially applicable

Population is children <5, so is not identical to</td><td>Minor limitations

Analysis uses a short time horizon of</td><td>Dipstick testing strategy differs from the dipstick strategies in Whiting 2006 and in the de novo analysis. Subsequent laboratory testing is provided to all</td><td><table><tr><th>Strategy</th><th>Cost</th><th>QALDs</th><th>NMB</th><th>INMB</th></tr><tr><td>LT</td><td>£1.100</td><td>20.709</td><td>£1090.44</td><td>-</td></tr><tr><td>DT</td><td>£1.183</td><td>20.709</td><td>£1090.38</td><td>-£0.05</td></tr><tr><td>PT</td><td>£1.187</td><td>20.709</td><td>£1090.4</td><td>-£0.04</td></tr></table></td><td>Bootstrapping of results showed that the incremental net monetary benefit of the ‘laboratory testing’ strategy compared to ‘dipstick testing’ produced was significant (95%</td></tr>
</table>

Study	Applicability	Limitations	Other comments	Results	Uncertainty
testing (LT) versus presumptive treatment (PT)	the review question	21 days, but is appropriate due to modelling a single episode of UTI	children in the Hay analysis, rather than only to children with a positive dipstick result		confidence intervals did not cross 0)

Economic model

Introduction

The 2007 NICE guidance on the diagnosis and management of urinary tract infection in under 16s recommends that children over the age of 3 years with a suspected UTI should initially be tested with a urine dipstick. However, due to a lack of evidence regarding the accuracy of dipstick tests in younger children, this recommendation was not previously extended to children under the age of 3 years, for whom urgent microscopy and culture was recommended. The purpose of this economic evaluation is to determine whether dipstick testing prior to microscopy and culture is cost-effective in this younger age group.

The full economic modelling report is displayed in Appendix J.

Patient population

The patient population consisted of children under 3 with suspected UTI, stratified into two age groups:

- Infants younger than 3 months
- Children 3 months or older but younger than 3 years

Interventions

Two intervention strategies were compared:

- **'No dipstick testing':** A scenario reflective of current practice, in which a urine sample is sent for urgent microscopy and culture in all children with suspected UTI. Antibiotic treatment is started immediately for all children, with treatment adjusted or discontinued as appropriate when test results are received.
- **'Dipstick testing':** All children with suspected UTI are dipstick tested. For children with a positive dipstick test a urine sample is sent for urgent microscopy and culture, and antibiotic treatment is started. Children with a negative dipstick test are assumed to not have UTI, and no further testing or treatment is administered unless symptoms persist. This option consists of four sub-strategies, according to interpretation of nitrite and leukocyte esterase (LE) results:
    - Presence of nitrite alone is considered a positive test result
    - Presence of LE alone is considered a positive test result
    - Presence of nitrite or LE is considered a positive test result
    - Presence of nitrite and LE is considered a positive test result

Methods

The economic model consists of two elements:

- A short-term decision tree, which simulates testing and treatment of the initial UTI episode
- A long-term Markov model, which estimates lifetime cost and QALY outcomes, and captures any downstream effects of UTI

As there is considerable uncertainty regarding the possible outcomes of a false negative dipstick test (i.e. the consequences of a delay in treating UTI), the model uses three scenarios in order to explore these consequences:

- **Basic scenario:** A false negative test result for UTI only results in a longer duration of symptoms, after which there are no further adverse consequences
- **Scenario 1:** In addition to the basic scenario assumption, a false negative result also increases the risk of children with UTI developing septicaemia
- **Scenario 2:** In addition to the basic scenario assumption, a false negative result also increases the risk of progressive renal scarring (PRS) in the future, and hence the risk of progressing to end-stage renal disease (ESRD).
- **Scenario 3:** In addition to the basic scenario assumption, a false negative result also increases the risk of septicaemia and PRS.

Base case

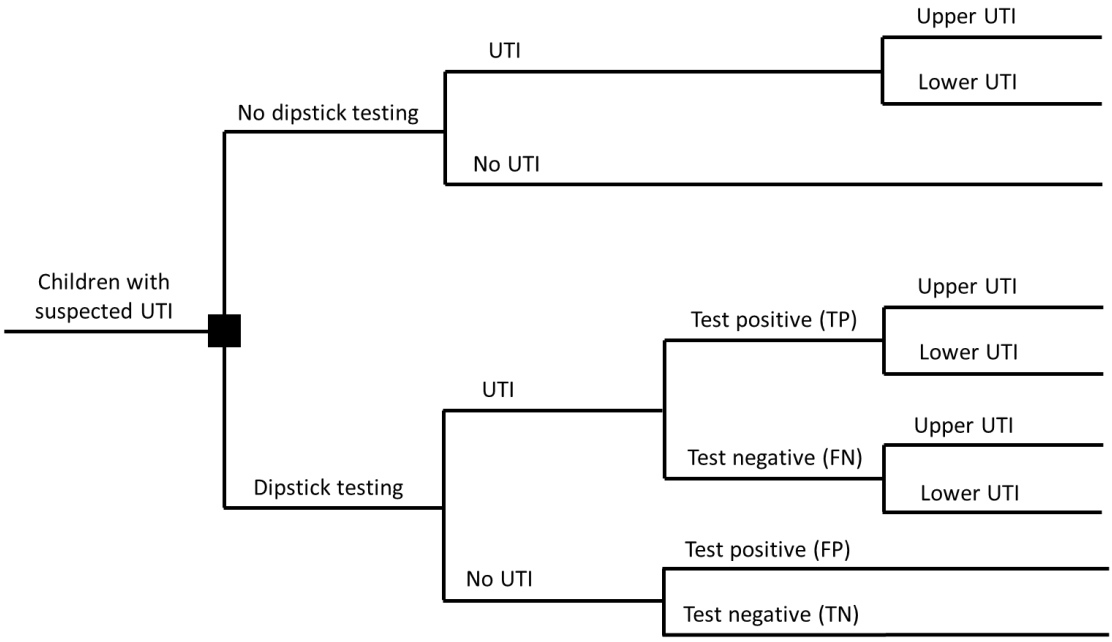
In the base case scenario, a short-term decision tree (shown in Figure 1) is used to model children's UTI status, test results, and treatment of the initial UTI episode.

At the start of the tree, the decision is made between the two strategies: 'no dipstick testing' or 'dipstick testing'. In the 'no dipstick testing' arm, a microscopy and culture test is carried out for all patients with UTI, which is assumed to provide a definite test, and shows their underlying UTI status (either UTI or no UTI). All children are appropriately treated with a course of antibiotics, which is assumed to resolve the infection. UTI may either take the form of upper UTI (pyelonephritis) or lower UTI, which affects the duration of symptoms and cost of treatment.

In the 'dipstick testing' arm, all children are initially tested with dipstick. As with the 'no dipstick testing' arm, all children have an underlying UTI status. For each of these groups a dipstick test can produce a positive or negative test result, with probabilities according to the sensitivity and specificity of the test. Outcomes for each UTI status/test result combination are as follows:

- **True positive:** Child receives antibiotic treatment and a urine sample is sent for microscopy and culture
- **False positive:** Child receives antibiotic treatment and a urine sample is sent for microscopy and culture, which reveals the child's UTI-negative status
- **True negative:** Child appropriately receives no further treatment or testing for UTI
- **False negative:** Child experiences a delay in treatment of 4 days, after which the true UTI-positive status is discovered, antibiotic treatment is administered and a urine sample is sent for microscopy and culture. Treatment is also assumed to be more expensive as a result.

Figure 1: Decision tree structure



Following the short-term section of the model, a Markov model is used to estimate lifetime QALYs. In the base case, following the resolution of UTI, all children are assumed to return to a healthy state, and age-specific baseline mortality rates for the general population are used to estimate survival.

Including risk of progressive renal scarring

For the scenario in which false negative results are associated with an increased risk of PRS, all children with UTI are associated with a baseline risk of developing PRS in the future, with differing probabilities according to whether the infection is upper or lower UTI. Children with a false negative test result have an increased risk of developing PRS. Since this value is unknown, an estimated increased risk of 100% compared to baseline was used, and this value was varied widely in sensitivity analysis in order to capture the level of uncertainty around the parameter.

For children developing PRS, the long-term Markov model simulates progress through various stages of disease. Patients with PRS have a probability of developing ESRD after a number of years, which results in an elevated probability of death and reduced quality of life. From this state, patients are eligible for renal transplantation, which improves quality of life, but is associated with an increased probability of death following surgery, and a chance of transplant failure, which results in returning to the ESRD state.

Including risk of septicaemia

For the scenario in which false negative results are associated with an increased risk of septicaemia, all children with UTI are associated with a baseline risk of developing septicaemia. Children with a false negative test result have an increased risk of septicaemia relative to baseline. However, since this value is unknown, an estimated increased risk of 100% compared to baseline was used, and this value was varied widely in sensitivity analysis in order to capture the level of uncertainty around the parameter.

Children who develop septicaemia also have a probability of death. In order to capture the lifetime QALY loss from septicaemia-related death, these children do not progress to the long-term Markov phase of the model.

Model inputs

Values for all parameters used to populate the model are displayed in Appendix J.

Sensitivity analysis

In order to characterise the uncertainty surrounding model results, extensive deterministic and probabilistic sensitivity analyses were carried out.

One-way sensitivity analyses conducted on the following parameters:

- Prevalence of UTI
- Accuracy of dipstick tests
- Additional duration of untreated UTI
- Quality of life associated with UTI
- Cost of microscopy, culture and antibiotic treatment
- Baseline probability of PRS
- Relative risk of PRS in untreated versus treated UTI
- Baseline probability of septicaemia
- Probability of death from septicaemia
- Relative risk of septicaemia in untreated versus treated UTI

In addition, four deterministic scenarios were included in the one-way sensitivity analysis:

- Cost of dipstick test added to the 'no dipstick testing' strategy: This scenario was included to reflect a pathway in which all children receive a dipstick test, but a urine sample is also sent for microscopy and culture regardless of the result.
- Antibiotic adverse events included: This scenario used pessimistic estimates of the potential consequences of antibiotic treatment. This comprised a 0.05% probability of anaphylactic shock, 0.33% probability of death from anaphylactic shock, and a 1% probability of 'other adverse events', which were assumed to cause a reduction in QoL by 0.5 for 3 days.
- Probability of ESRD set to 0.65%: In order to explore the uncertainty in the progression of UTI to ESRD, a pessimistic upper-bound value of 0.65% for the probability of eventually developing ESRD from a UTI was used in the model, rather than using the probability of developing PRS as an intermediate step.
- Probability of death from septicaemia set to 1.6%: As the base case analysis for scenarios 2 and 3 used a probability of death from bacteraemia which was not necessarily related to UTI, an alternative upper bound value, estimated from the same source as the baseline probability of bacteraemia, was used in the model.

For the probabilistic sensitivity analysis, all model input parameters were assigned probability distributions (rather than being expressed as point estimates) to reflect the uncertainty surrounding the available clinical and cost data. 1,000 iterations of the model were run, each drawing random values from parameter distributions.

Results

Infants under 3 months

For infants under 3 months, of the four possible dipstick interpretations, the 'LE alone' strategy consistently produced a higher number of QALYs than the other three. Therefore, results are presented in this section as the incremental costs and QALYs of the 'no dipstick testing' strategy compared to 'dipstick LE'. For each scenario, base case and one-way sensitivity analysis results are shown here. In-depth threshold analyses, two-way sensitivity analyses and probabilistic sensitivity analysis results are shown in the full economic analysis report in Appendix J.

Basic scenario

Base case and one-way sensitivity analysis results for the basic scenario are shown in Table 5. These results show that, in the base case and in all sensitivity analyses, the 'no dipstick' strategy is not cost-effective at a threshold of £20,000 per QALY for this scenario.

Table 5: Basic scenario one-way sensitivity analysis results for infants under 3 months

Scenario	Δ Costs – 'No dipstick' versus 'dipstick LE'	Δ QALYs – 'No dipstick' versus 'dipstick LE'	ICER – 'No dipstick' versus 'dipstick LE'
Base case	£21.85	0.00003	£776,964
UTI prevalence set to 1%	£24.86	0.00000	£6,365,712
UTI prevalence set to 25%	£13.20	0.00010	£135,159
Additional duration of untreated UTI set to 20 days	£21.85	0.00014	£155,393
Accuracy of dipstick tests set to lower 95% CI	£17.84	0.00003	£519,134
Accuracy of dipstick tests set to upper 95% CI	£23.54	0.00002	£1,369,546
Quality of life of UTI set to 0.1	£21.85	0.00004	£505,916
Cost of microscopy, culture and antibiotic treatment doubled	£45.48	0.00003	£1,617,430
'No dipstick' strategy also associated with the cost of a dipstick test	£21.97	0.00003	£781,228
Antibiotic adverse events included	£21.85	-0.00004	Dipstick dominates no dipstick

Scenario 1: Untreated UTI associated with an increased risk of PRS

Base case and one-way sensitivity analysis results for scenario 1 are shown in Table 6. These results show that, due to the increased risk of PRS in children with an untreated UTI, the ICER of the 'no dipstick' strategy compared to 'dipstick testing' strategies is generally lower than those of the basic scenario. However, in all sensitivity analyses the ICER remains substantially above £20,000 QALY.

Table 6: Scenario 1 one-way sensitivity analysis results for infants under 3 months

Scenario	Δ Costs – 'No dipstick' versus 'dipstick LE'	Δ QALYs – 'No dipstick' versus 'dipstick LE'	ICER – 'No dipstick' versus 'dipstick LE'
Base case	£21.60	0.00006	£334,327
UTI prevalence set to 1%	£24.83	0.00001	£2,766,974
UTI prevalence set to 25%	£12.33	0.00022	£54,965
Baseline probability of PRS halved relative to base case	£21.72	0.00005	£468,566
Baseline probability of PRS doubled relative to base case	£21.35	0.00010	£211,194
Probability of ESRD set to upper bound from Round 2012	£19.62	0.00035	£55,509
Relative risk of PRS for untreated UTI versus treated UTI set to 1.2	£21.80	0.00004	£615,483
Relative risk of PRS for untreated UTI versus treated UTI set to 4	£21.10	0.00014	£153,370
Cost of microscopy, culture and antibiotic treatment doubled	£45.23	0.00006	£700,162
Antibiotic adverse events included	£21.60	-0.00001	Dipstick dominates no dipstick

Scenario 2: Untreated UTI associated with an increased risk of septicaemia

Base case and one-way sensitivity analysis results for scenario 2 are shown in Table 7. These results show that, in the base case, including an increased risk of septicaemia for children with untreated UTI results in an ICER of £11,914 for the 'no dipstick' strategy. However, one-way sensitivity analysis shows that decreasing UTI prevalence, baseline probability of bacteraemia, probability of death from bacteraemia, or the relative risk of bacteraemia in untreated versus treated UTI results in an ICER of £20,000.

Table 7: Scenario 2 one-way sensitivity analysis results for infants under 3 months

Scenario	Δ Costs – 'No dipstick' versus 'dipstick LE'	Δ QALYs – 'No dipstick' versus 'dipstick LE'	ICER – 'No dipstick' versus 'dipstick LE'
Base case	£20.01	0.00168	£11,914
UTI prevalence set to 1%	£24.61	0.00023	£105,471
UTI prevalence set to 25%	£6.82	0.00583	£1,170
Baseline probability of bacteraemia set to 1%	£21.57	0.00028	£76,983
Baseline probability of bacteraemia set to 20%	£16.25	0.00507	£3,205
Probability of death from bacteraemia set to 1%	£20.01	0.00024	£82,667
Probability of death from bacteraemia set to 20%	£20.02	0.00431	£4,647
Probability of death from bacteraemia estimated from Schnadower 2010 (1.6%)	£20.01	0.00037	£54,021

Scenario	Δ Costs – 'No dipstick' versus 'dipstick LE'	Δ QALYs – 'No dipstick' versus 'dipstick LE'	ICER – 'No dipstick' versus 'dipstick LE'
Relative risk of bacteraemia for untreated UTI versus treated UTI set to 1.2	£21.48	0.00036	£59,927
Relative risk of bacteraemia for untreated UTI versus treated UTI set to 4	£16.34	0.00498	£3,279
Cost of microscopy, culture and antibiotic treatment doubled	£43.65	0.00168	£25,983
Cost of septicaemia doubled	£18.18	0.00168	£10,820
Antibiotic adverse events included	£20.01	0.00161	£12,430

Scenario 3: Untreated UTI associated with an increased risk of septicaemia and PRS

Base case and one-way sensitivity analysis results for scenario 2 are shown in Table 8. These results show that, in the base case, including an increased risk of septicaemia for children with untreated UTI results in an ICER of £11,517 for the 'no dipstick' strategy. One-way sensitivity analysis shows that the ICER is relatively sensitive to changes in the prevalence of UTI, and parameters relating to septicaemia, but relatively insensitive to parameters relating to PRS.

Table 8: Scenario 3 one-way sensitivity analysis results for infants under 3 months

Scenario	Δ Costs – 'No dipstick' versus 'dipstick LE'	Δ QALYs – 'No dipstick' versus 'dipstick LE'	ICER – 'No dipstick' versus 'dipstick LE'
Base case	£19.76	0.00172	£11,517
UTI prevalence set to 1%	£24.57	0.00024	£103,095
UTI prevalence set to 25%	£9.84	0.00477	£2,064
Baseline probability of bacteraemia set to 1%	£21.32	0.00032	£67,294
Baseline probability of bacteraemia set to 20%	£16.00	0.00510	£3,135
Probability of death from bacteraemia set to 1%	£19.76	0.00028	£70,905
Probability of death from bacteraemia set to 20%	£19.77	0.00434	£4,552
Probability of death from bacteraemia estimated from Schnadower 2010 (1.6%)	£19.76	0.00041	£48,549
Baseline probability of PRS halved relative to base case	£19.89	0.00170	£11,712
Baseline probability of PRS doubled relative to base case	£19.52	0.00175	£11,138
Relative risk of bacteraemia for untreated UTI versus treated UTI set to 1.2	£21.23	0.00039	£53,765
Relative risk of bacteraemia for untreated UTI versus treated UTI set to 4	£16.09	0.00502	£3,207
Relative risk of PRS for untreated UTI versus treated UTI set to 1.2	£19.96	0.00169	£11,833
Relative risk of PRS for untreated UTI versus treated UTI set to 4	£19.27	0.00179	£10,771
Cost of microscopy, culture and antibiotic treatment doubled	£43.40	0.00172	£25,289
Cost of septicaemia doubled	£17.93	0.00172	£10,446

Scenario	Δ Costs – 'No dipstick' versus 'dipstick LE'	Δ QALYs – 'No dipstick' versus 'dipstick LE'	ICER – 'No dipstick' versus 'dipstick LE'
Antibiotic adverse events included	£19.76	0.00165	£12,005

Children 3 months or older but younger than 3 years

For infants 3 months or older but younger than 3 years, of the four possible dipstick interpretations, the 'nitrite or LE' strategy consistently produced a higher number of QALYs than the other three. Therefore, results are presented in this section as the incremental costs and QALYs of the 'no dipstick testing' strategy compared to 'dipstick nitrite or LE'. For each scenario, base case and one-way sensitivity analysis results are shown here. In-depth threshold analyses, two-way sensitivity analyses and probabilistic sensitivity analysis results are shown in the full economic analysis report in Appendix J.

Basic scenario

Base case and one-way sensitivity analysis results for the basic scenario are shown in Table 9. These results show that, in the base case and in all sensitivity analyses, the 'no dipstick' strategy is not cost-effective at a threshold of £20,000 per QALY for this scenario.

Table 9: Basic scenario one-way sensitivity analysis results for children 3 months or older but younger than 3 years

Scenario	Δ Costs – 'No dipstick' versus 'dipstick nitrite or LE'	Δ QALYs – 'No dipstick' versus 'dipstick nitrite or LE'	ICER – 'No dipstick' versus 'dipstick nitrite or LE'
Base case	£22.89	0.00003	£849,353
UTI prevalence set to 1%	£25.12	0.00000	£5,033,701
UTI prevalence set to 25%	£12.95	0.00012	£103,778
Additional duration of untreated UTI set to 20 days	£22.89	0.00013	£169,871
Accuracy of dipstick tests set to lower 95% CI	£19.68	0.00003	£599,976
Accuracy of dipstick tests set to upper 95% CI	£23.39	0.00001	£1,995,969
Quality of life of UTI set to 0.1	£22.89	0.00005	£416,349
Cost of microscopy, culture and antibiotic treatment doubled	£47.25	0.00003	£1,753,196
'No dipstick' strategy also associated with the cost of a dipstick test	£23.01	0.00003	£853,802
Antibiotic adverse events included	£22.89	-0.00004	Dipstick dominates no dipstick

Scenario 1: Untreated UTI associated with an increased risk of PRS

Base case and one-way sensitivity analysis results for scenario 1 are shown in Table 10. These results show that, due to the increased risk of PRS in children with an untreated UTI, the ICER of the 'no dipstick' strategy compared to 'dipstick testing' strategies is generally lower than those of the basic scenario. However, in all sensitivity analyses the ICER remains substantially above £20,000 QALY.

Table 10: Scenario 1 one-way sensitivity analysis results for children 3 months or older but younger than 3 years

Scenario	Δ Costs – 'No dipstick' versus 'dipstick nitrite or LE'	Δ QALYs – 'No dipstick' versus 'dipstick nitrite or LE'	ICER – 'No dipstick' versus 'dipstick nitrite or LE'
Base case	£22.65	0.00006	£364,766
UTI prevalence set to 1%	£25.08	0.00001	£2,180,882
UTI prevalence set to 25%	£11.83	0.00029	£41,168
Baseline probability of PRS halved relative to base case	£22.77	0.00004	£511,433
Baseline probability of PRS doubled relative to base case	£22.41	0.00010	£230,458
Probability of ESRD set to upper bound from Round 2012	£20.59	0.00036	£56,797
Relative risk of PRS for untreated UTI versus treated UTI set to 1.2	£22.84	0.00003	£672,245
Relative risk of PRS for untreated UTI versus treated UTI set to 4	£22.17	0.00013	£167,460
Cost of microscopy, culture and antibiotic treatment doubled	£47.01	0.00006	£757,058
Antibiotic adverse events included	£22.65	-0.00001	Dipstick dominates no dipstick

Scenario 2: Untreated UTI associated with an increased risk of septicaemia

Base case and one-way sensitivity analysis results for scenario 2 are shown in Table 11. For the base case, these results show that, unlike in the infants under 3 months population, the ICER of 'no dipstick' is considerably higher than £20,000 per QALY (£172,917 per QALY). This is due to the lower baseline probability of bacteraemia in children 3 months or older but younger than 3 years.

Sensitivity analysis shows that a substantial increase in UTI prevalence and baseline probability of bacteraemia results in an ICER below £20,000, but in all other sensitivity analysis scenarios the ICER remains cost-ineffective.

Table 11: Scenario 2 one-way sensitivity analysis results for children 3 months or older but younger than 3 years

Scenario	Δ Costs – 'No dipstick' versus 'dipstick nitrite or LE'	Δ QALYs – 'No dipstick' versus 'dipstick nitrite or LE'	ICER – 'No dipstick' versus 'dipstick nitrite or LE'
Base case	£22.68	0.00013	£172,917
UTI prevalence set to 1%	£25.08	0.00002	£1,032,696
UTI prevalence set to 25%	£11.97	0.00061	£19,720
Baseline probability of bacteraemia set to 0.1%	£22.86	0.00004	£567,977
Baseline probability of bacteraemia set to 20%	£16.25	0.00507	£3,205
Probability of death from bacteraemia set to 1%	£22.68	0.00005	£441,398

Scenario	Δ Costs – 'No dipstick' versus 'dipstick nitrite or LE'	Δ QALYs – 'No dipstick' versus 'dipstick nitrite or LE'	ICER – 'No dipstick' versus 'dipstick nitrite or LE'
Probability of death from bacteraemia set to 20%	£22.68	0.00052	£43,989
Probability of death from bacteraemia estimated from Schnadower 2010 (1.6%)	£22.68	0.00007	£343,422
Relative risk of bacteraemia for untreated UTI versus treated UTI set to 1.2	£22.85	0.00005	£478,072
Relative risk of bacteraemia for untreated UTI versus treated UTI set to 4	£22.26	0.00034	£65,548
Cost of microscopy, culture and antibiotic treatment doubled	£47.04	0.00013	£358,634
Cost of septicaemia doubled	£22.47	0.00013	£171,312
Antibiotic adverse events included	£22.68	0.00006	£382,290

Scenario 3: Untreated UTI associated with an increased risk of septicaemia and PRS

Base case and one-way sensitivity analysis results for scenario 3 are shown in Table 12. For the base case, these results show that, as with scenario 2 for this population, the ICER of 'no dipstick' is considerably higher than £20,000 per QALY (£134,939 per QALY).

Again, sensitivity analysis shows that a substantial increase in UTI prevalence and baseline probability of bacteraemia results in an ICER below £20,000, but in all other sensitivity analysis scenarios the ICER remains cost-ineffective. In general ICERs are substantially more sensitive to changes in parameters relating to the incidence and consequences of septicaemia than to parameters relating to PRS.

Table 12: Scenario 3 one-way sensitivity analysis results for children 3 months or older but younger than 3 years

Scenario	Δ Costs – 'No dipstick' versus 'dipstick nitrite or LE'	Δ QALYs – 'No dipstick' versus 'dipstick nitrite or LE'	ICER – 'No dipstick' versus 'dipstick nitrite or LE'
Base case	£22.44	0.00017	£134,939
UTI prevalence set to 1%	£25.04	0.00003	£813,077
UTI prevalence set to 25%	£10.86	0.00077	£14,107
Baseline probability of bacteraemia set to 0.1%	£22.62	0.00008	£300,011
Baseline probability of bacteraemia set to 20%	£17.28	0.00272	£6,348
Probability of death from bacteraemia set to 1%	£22.44	0.00009	£259,318
Probability of death from bacteraemia set to 20%	£22.44	0.00055	£40,754
Probability of death from bacteraemia estimated from Schnadower 2010 (1.6%)	£22.44	0.00010	£221,759
Baseline probability of PRS halved relative to base case	£22.56	0.00015	£151,679
Baseline probability of PRS doubled relative to base case	£22.20	0.00020	£110,216
Relative risk of bacteraemia for untreated UTI versus treated UTI set to 1.2	£22.61	0.00008	£272,600

Scenario	Δ Costs – 'No dipstick' versus 'dipstick nitrite or LE'	Δ QALYs – 'No dipstick' versus 'dipstick nitrite or LE'	ICER – 'No dipstick' versus 'dipstick nitrite or LE'
Relative risk of bacteraemia for untreated UTI versus treated UTI set to 4	£22.02	0.00037	£58,765
Relative risk of PRS for untreated UTI versus treated UTI set to 1.2	£22.63	0.00014	£163,776
Relative risk of PRS for untreated UTI versus treated UTI set to 4	£21.96	0.00024	£92,825
Cost of microscopy, culture and antibiotic treatment doubled	£46.80	0.00017	£281,421
Cost of septicaemia doubled	£22.23	0.00017	£133,673
Antibiotic adverse events included	£22.44	0.00009	£237,559

Discussion

The results of this analysis show that, in the majority of scenarios and for the majority of sensitivity analyses, a strategy in which all children are treated with antibiotics and a urine sample sent for laboratory testing is not cost-effective compared to a strategy in which initial dipstick testing is used to determine which children receive treatment and further tests. This is for 3 key reasons. First, the prevalence of UTI in children with suspected UTI is relatively low. Second, the accuracy of dipstick testing is relatively high. Third, in the majority of scenarios, the consequences of a UTI are relatively mild. In combination, these factors mean that the proportion of children with a false negative test result and the resulting QALY loss is relatively small.

For infants under 3 months, only in scenarios in which untreated UTI is associated with an increased risk of septicaemia is 'no dipstick testing' potentially a cost-effective strategy. However, this cost-effectiveness relies heavily on a number of parameters, most importantly a high baseline probability of UTI, probability of septicaemia, probability of death from septicaemia, and relative risk of septicaemia in untreated versus treated UTI. Sensitivity analyses show that a relatively small reduction in any of these parameters results in the ICER of the 'no dipstick' strategy exceeding £20,000 per QALY.

For children 3 months or older but younger than 3 years, unlike in the younger cohort, the base case ICER for scenarios in which untreated UTI is associated with an increased risk of septicaemia remains substantially higher than £20,000 per QALY. This is because of the considerably lower baseline risk of bacteraemia in this group of patients. Sensitivity analyses for these scenarios show that the prevalence of UTI, baseline incidence of bacteraemia, probability of death from bacteraemia, or the relative risk of septicaemia would have to be substantially higher for the 'no dipstick' strategy to be cost-effective.

Of the 4 possible interpretations of dipstick results, the 'LE' and 'nitrite or LE' strategies are consistently more cost-effective than 'nitrite' and 'nitrite or LE' strategies. This is because the former 2 interpretations are associated with a substantially higher sensitivity than the latter 2, without a substantially reduced specificity.

A key limitation of this analysis is the considerable uncertainty surrounding the consequences of a false negative test result for UTI. This issue is addressed by

exploring a wide range of possible outcomes of untreated UTI of varying severity. While these scenarios are highly speculative, and may in some cases not be fully reflective of clinical reality, they serve to demonstrate that the consequences of an untreated UTI would need to be relatively severe in order for a 'no dipstick' strategy to be cost-effective.

Another limitation of the analysis is that the full complexity of potentially overlapping symptoms and conditions which may occur in children with possible UTI is not captured. Particularly in infants under 3 months, symptoms are frequently non-specific. This means that, for such children, a single test alone may not be sufficient to determine that a child is in no need of further investigation for other causes, and therefore the model assumption that children without a UTI are otherwise healthy is potentially unrealistic.

In conclusion, this analysis shows that, in the majority of exploratory scenarios, a strategy in which all children with suspected UTI are prescribed antibiotics and a urine sample sent for microscopy and culture is not cost-effective compared to a scenario in which initial dipstick testing is used to determine which children should receive treatment and further testing. Only in scenarios in which a substantial added risk of septicaemia is assumed to result from untreated UTI is a 'no dipstick' strategy potentially cost-effective.

Evidence statements

Under 3 months: reference test culture

The 3 included studies used 4 methods of urine collection: catheterisation, suprapubic aspirate (SPA), clean catch and nappy pad. The reference threshold set for urine culture ranged from 10^3 cfu/ml and 5 x10^4 cfu/ml for SPA and 10^3 cfu/ml for catheterisation and was set to 10^5 cfu/ml for clean catch and nappy pad.

Results which increase the probability of finding a positive urine culture

The following test results **increase** the probability of finding a positive urine culture in an infant under 3 months to a degree that is likely to be **very large**:

- Nitrite positive, assessed visually and using analyser (moderate quality, 95% CIs ranged from moderate to very large)
- Nitrite and LE positive, assessed visually and using analyser (moderate quality)

The following test results **increase** the probability of finding a positive urine culture in an infant under 3 months to a degree that is likely to be **large**:

- Nitrite or LE positive, assessed visually and using analyser (very low quality, 95% CIs ranged from large to very large)
- LE positive, assessed visually and using analyser (low quality, 95% CIs ranged from moderate to very large)
- Nitrite or LE positive and microscopy positive (very low quality)

Results which decrease the probability of finding a positive urine culture

The following test results **decrease** the probability of finding a positive urine culture in an infant under 3 months to a degree that is likely to be **very large:**

- Nitrite or LE negative and microscopy negative, assessed using analyser (low quality)

The following test results **decrease** the probability of finding a positive urine culture in an infant under 3 months to a degree that is likely to be **moderate:**

- LE negative, assessed visually and using analyser (moderate quality, 95% CIs ranged from moderate to large)
- Nitrite or LE negative, assessed visually and using analyser (very low quality, 95% CIs ranged from moderate to large)

The following test results **decrease** the probability of finding a positive urine culture in an infant under 3 months to a degree that is likely to be **slight:**

- Nitrite negative, assessed visually and using analyser (moderate quality)
- Nitrite and LE negative, assessed visually and using analyser (moderate quality)

3 months or older but younger than 3 years: reference test culture

The 9 studies included both invasive methods of urine collection (SPA and catheter) and non-invasive (clean catch, sterile bag and nappy pad). A range of thresholds were used to determine positive culture: 10^3 to $5x10^4$ cfu/ml for catheter samples or 10^5 cfu/ml for clean catch, bag samples and nappy pad. Some studies include all children below 1 year or 2 years and as these may include the age group of below 3 months, these were downgraded for serious indirectness.

Results that increase the probability of finding a positive urine culture

The following test results **increase** the probability of finding a positive urine culture in an infant or child 3 months or older but younger than 3 years to a degree that is likely to be **very large**:

- Nitrite positive, assessed visually and using analyser (low quality, 95% CI ranged from moderate to very large)
- Nitrite or LE positive, assessed visually and using analyser (low quality, 95% CI ranged from moderate to very large)

The following test results **increase** the probability of finding a positive urine culture in an infant or child 3 months or older but younger than 3 years to a degree that is likely to be **large**:

- Nitrite and LE positive, assessed visually and using analyser (very low quality, 95% CI ranged from slight to very large)
- Nitrite or LE positive (method of assessment unclear) and microscopy positive (≥ 5 WBCs/hpf) (very low quality, 95% CI ranged from large to very large)
- LE positive, assessed visually and using analyser (low quality, 95% CI ranged from moderate to very large)

The following test results **increase** the probability of finding a positive urine culture in an infant or child 3 months or older but younger than 3 years to a degree that is likely to be **slight:**

- Nitrite or LE or blood or protein positive, assessed by analyser (very low quality)

Results that decrease the probability of finding a positive urine culture

The following test results **decrease** the probability of finding a positive urine culture in an infant or child 3 months or older but younger than 3 years to a degree that is likely to be **moderate:**

- LE negative, assessed visually and using analyser (very low quality, 95% CI ranged from slight to large)

- Nitrite or LE negative, assessed visually and using analyser (moderate quality)
- Nitrite or LE negative (method of assessment unclear) and microscopy negative (≤ 5 WBCs/hpf) (very low quality, 95% CI ranged from slight to moderate)
- Nitrite or LE or blood or protein negative, assessed by analyser (very low quality, 95% CI ranged from slight to large)

The following test results **decrease** the probability of finding a positive urine culture in an infant or child 3 months or older but younger than 3 years to a degree that is likely to be **slight:**

- Nitrite negative, assessed visually and using analyser (very low quality)

The following results were **not significantly different** from random chance:

- Nitrite and LE negative, assessed visually and using analyser (low quality)

3 months or older but younger than 3 years: reference test culture and microscopy

One study was included which did not report on the method of urine collection. The threshold for culture was 10^5 cfu/ml and for microscopy: WBC > 25 x 10^9/L for boys or 50 x 10^9/L for girls.

Results that increase the probability of finding a positive urine culture

The following test results **increase** the probability of finding a positive urine culture in an infant or child 3 months or older but younger than 3 years to a degree that is likely to be **very large**:

- Nitrite and LE positive, assessed by analyser (moderate quality)
- LE and protein positive, assessed by analyser (moderate quality, 95% CI ranged from large to very large)

The following test results **increase** the probability of finding a positive urine culture in an infant or child 3 months or older but younger than 3 years to a degree that is likely to be **large:**

- Nitrite positive, assessed by analyser (moderate quality, 95% CI ranged from moderate to very large)

The following test results **increase** the probability of finding a positive urine culture in an infant or child 3 months or older but younger than 3 years to a degree that is likely to be **moderate:**

- LE positive, assessed by analyser (moderate quality, 95% CI ranged from moderate to large)
- LE and protein and nitrite positive, assessed by analyser (moderate quality)

The following results were **not significantly different** from random chance:

- Protein positive, assessed by analyser (very low quality)

Results that decrease the probability of finding a positive urine culture

The following test results **decrease** the probability of finding a positive urine culture in an infant or child 3 months or older but younger than 3 years to a degree that is likely to be **large:**

- LE negative, assessed by analyser (moderate quality, 95% CI ranged from moderate to very large)

- Nitrite and LE negative, asses by analyser (moderate quality, 95% CI ranged from moderate to very large)
- LE and protein negative, assessed by analyser (moderate quality, 95% CI ranged from moderate to very large)
- LE and protein and nitrite negative, assessed by analyser (moderate quality, 95% CI ranged from moderate to very large)

The following test results **decrease** the probability of finding a positive urine culture in an infant or child 3 months or older but younger than 3 years to a degree that is likely to be **slight:**

- Nitrite negative, assessed by analyser (moderate quality, 95% CI ranged from moderate to very large)

The following results were **not significantly different** from random chance:

- Protein negative, assessed by analyser (moderate quality)

Sensitivity analysis: urine collection method

The 2007 NICE guideline CG54 states that clean catch is the recommended method for urine collection (recommendation 1.1.3.1). If clean catch is unobtainable, further non-invasive methods are recommended. Invasive methods (catheterisation and suprapubic aspirate) are only recommended where urine collection by non-invasive methods are not possible or practical. For this reason, the urine sampling method was included in the review protocol as a subgroup analysis. To assess the impact of including the recommended non-invasive methods alone in the analysis, a sensitivity analysis was undertaken. Where data from different urine sampling methods were pooled in a meta-analysis, non-invasive methods were removed from the primary analysis and the sensitivity analysis included only non-invasive methods (clean catch or sterile bag). For results of this, please see Table 38.

Health economic evidence statements

One economic analysis (Whiting et al, 2006) found that the testing strategy with the highest probability of being cost effective at a threshold of £20,000/QALY featured dipstick testing as the initial test for girls <3 years and boys <2 years. The optimal strategy for children above this age was to treat all children without testing. This study was directly applicable to the NHS, but was characterised by serious limitations; the assumption was made that accuracy of dipstick testing does not vary according to children's age.

One economic analysis (Hay et al, 2016) found that a strategy in which children under 5 years who are at low risk of UTI are tested with dipstick to guide initial treatment, followed by laboratory testing was not cost-effective compared to a strategy of laboratory testing and treating based on results. This analysis found that dipstick testing is not cost-effective if a urine sample is subsequently sent for laboratory testing in all children, but did not investigate the cost-effectiveness in which laboratory testing is only arranged for children with a positive dipstick result. This evaluation was assessed as being partially applicable, due to the age of the patient population not matching that of the review question, and was categorised as having minor limitations.

The de novo economic analysis for this guideline found that, in the majority of exploratory scenarios, a strategy of initiating antibiotics and sending a urine sample for microscopy and culture in all children with suspected UTI was not cost-effective compared to a strategy of dipstick testing, with treatment and laboratory testing only

arranged in children with a positive dipstick result. Only in children under 3 months, in extreme scenarios in which an increased risk of septicaemia resulting from a delay in UTI treatment was assumed, was the strategy of treating and laboratory testing all children potentially cost-effective. However, this relied on a high additional risk of septicaemia resulting from untreated UTI, and a high probability of death associated with septicaemia. This analysis was assessed as being directly relevant, as it was designed to answer the review question, and was categorised as having potentially serious limitations, due to the exploratory nature of the analysis necessitated by limited data.

Recommendations

1. **Refer all infants under 3 months with a suspected UTI (see table 13) to paediatric specialist care, and**
    - send a urine sample for urgent microscopy and culture
    - manage in line with the NICE guideline on fever in under 5s. [2017]

Table 13: Presenting symptoms and signs in infants and children with UTI

This table was deleted in 2022 as it referred to the symptoms and signs identified in the 2007 guideline. These were updated following a new review in 2022. Please see evidence review B: symptoms and signs for more details and the guideline Table 1 for the symptoms and signs that the recommendation above now refers to.

The recommendations cross referred to by number below refer to the 2017 version of the guideline and have been renumbered in the 2022 version. Please see NICE guideline NG224 for the up-to-date cross references.

2. **Use dipstick testing for infants and children 3 months or older but younger than 3 years with suspected UTI**
    - If both leukocyte esterase and nitrite are negative: do not start antibiotic treatment; do not send a urine sample for microscopy and culture unless at least 1 of the criteria in recommendation 1.1.6.1 apply.
    - If leukocyte esterase or nitrite, or both are positive: start antibiotic treatment; send a urine sample for microscopy and culture. **[2017]**

3. **Urine samples should be sent for culture:**
    - in infants and children who are suspected to have acute pyelonephritis/upper urinary tract infection (see 1.1.8.1)
    - in infants and children with a high to intermediate risk of serious illness
    - in infants under 3 months

- in infants and children with a positive result for leukocyte esterase or nitrite
- in infants and children with recurrent UTI
- in infants and children with an infection that does not respond to treatment within 24–48 hours, if no sample has already been sent
- when clinical symptoms and dipstick tests do not correlate. **[2017]**

Rationale and impact

Why the committee made the recommendations

Evidence showed that a positive urine dipstick test for leukocyte esterase or nitrite in children 3 months or older but younger than 3 years greatly increases the likelihood of finding a positive urine culture. Sending only positive samples for culture offered a better balance of benefits and costs for these children than prescribing antibiotics and urine culture for all children. The committee agreed that there are concerns about sepsis in infants under 3 months with suspected UTI, and usual practice is referral rather than the GP managing symptoms. So the committee recommended that all children under 3 months should be referred to specialist paediatric care and have a urine sample sent for urgent microscopy and culture. In children aged 3 months or older but younger than 3 years, symptoms are easier to identify, and antibiotics should only be started if a dipstick test is positive for either or both leukocyte esterase or nitrite. Children in this age group with a positive dipstick test should also have a urine sample sent for culture.

Impact of the recommendations on practice

Recommending dipstick testing in infants and children aged 3 months or older but younger than 3 years clarifies the role of dipstick testing in this age group and encourages immediate diagnosis and treatment in primary care. The committee believe the new recommendations will provide concise and clear guidance for health care professionals, more efficient diagnosis for infants and children, and cost savings and a reduced burden on laboratories by reducing the amount of urine samples sent for culture.

Interpreting the evidence

The outcomes that matter most

The committee agreed that sensitivity, specificity and likelihood ratios would be considered in the guideline update, and did not prioritise other diagnostic accuracy measures.

The quality of the evidence

Overall, the quality of evidence ranged from very low quality to moderate quality. The main reasons for downgrading the evidence were unclear blinding between index and reference tests, and heterogeneity in meta-analyses. The committee noted that some included studies were published more than 20 years ago and queried whether dipstick testing had changed significantly in recent years. It was agreed that dipstick testing has not changed meaningfully in recent years. The included studies used either visual interpretation of dipstick tests or automated analysers, and the

committee queried whether the accuracy of dipstick tests would differ between these methods. It was noted that visual interpretation is commonly used in primary care, while secondary and tertiary care may be more reliant on analysers. The committee agreed that both methods are prone to errors, but did not believe there to be a substantial difference between them.

The committee noted that different thresholds for positive urine culture were included in the studies, and this is reflective of UK practice as there is no guideline which sets a standard threshold for positive urine culture in infants and children younger than 3 years. Thresholds of predominant growth of a single bacteria can be 10^2 or 10^3 colony forming units per millilitre (cfu/ml) for sterile samples obtained from suprapubic aspirate or urethral catheterisation or up to 10^5 cfu/ml for non-invasive cultures such as clean catch samples. The committee noted that these thresholds are based on the Kass criteria for diagnosing UTI infection, yet there are various limitations in using these criteria. One limitation is that the Kass criteria was based on an adult study population and considered pyelonephritis (upper UTI), whereas bacterial counts may be lower in cystitis (lower UTI). The committee noted similar concerns for the use of microscopy, as there is no standard threshold for white blood cell (WBC) count in infants and children. Additionally, it was noted that there may be differences in procedures in different laboratories and the studies included had different definitions of a uropathogen. For example, Hay 2016 defined a uropathogen as a member of the *Enterobacteriaceae* group, while other studies considered growth of any single pathogen as a uropathogen.

The committee queried the quality of the Lejeune 1991 study, which uses a relatively high threshold for WBC count in microscopy. The committee noted that high WBC does not always indicate the presence of infection, but can be indicative of inflammation. Therefore, the committee did not make use of this study when making recommendations. Additionally, it was noted that the Sharief 1998 study used both clean catch and bag urine sampling techniques, but bag urine sampling is rarely performed in clinical practice. As the proportions of infants and children under 1 year with bag sample was not reported, the committee took this study into consideration when formulating recommendations.

The committee discussed the potential of the accuracy of dipstick testing being influenced by the urine sampling method and baseline prevalence of UTI in the study populations. The committee agreed that these factors should not substantially influence the accuracy of dipstick testing and therefore, downgrading studies for indirectness on the basis of urine sampling method or baseline prevalence was not deemed necessary.

Benefits and harms

The committee noted the importance of a clinical diagnosis of UTI, taking into account symptoms, physical examination and urine testing. It agreed that diagnosing UTI early and effectively is important in preventing recurrence and further complications. Febrile infection can potentially lead to renal scarring and recurrent renal scarring may lead to hypertension. Cystitis can potentially resolve without treatment, however treatment reduces symptoms and eradicates infection quicker.

In primary care, the main concern amongst general practitioners is febrile infants and children, as fever can be indicative of a wide range of conditions. In secondary and tertiary care, the main concern amongst health care professionals is the risk of septicaemia and other complications. It was noted that it can be difficult to distinguish UTI from septicaemia in infants, particularly in neonates. In these settings, urine samples may be sent for laboratory tests before clinical assessment and this could

lead to false positives due to contaminated samples or over treatment with antibiotics.

The committee reflected that the clinical evidence shows that a dipstick test can be useful in increasing or decreasing the likelihood of finding a positive urine culture, and this is particularly true for nitrite alone, LE alone or both nitrite and LE. The committee noted that the specificity of a positive LE on dipstick can vary in practice, as some infants or children with a viral infection can have elevated LE in their urine. In contrast, a positive nitrite on dipstick testing can have a high specificity and low sensitivity due to the presence of nitrites in the urine which indicates the presence of gram negative bacteria. In infants and children UTIs are usually caused by this group of bacteria although not all bacteria convert nitrates to nitrites. Additionally, younger infants, especially those under 3 months, may not retain urine in their bladder long enough to break down nitrates into nitrites. The committee noted that in practice infants and children are often encouraged to drink more to be able to produce a urine sample, but this can lead to diluted urine samples which could affect the performance of diagnostic tests. The committee discussed the risks of missing a UTI caused by non-nitrite forming bacteria and felt that the importance attributed to the presence of LE in the urine safety netted these children. The committee agreed that there was a benefit to considering both nitrite and LE when interpreting dipstick results. Additionally, the committee agreed that the final bullet of recommendation 3, which indicates that a urine sample should be sent for culture when clinical symptoms and dipstick tests do not correlate, covered concerns in patients where symptoms of UTI are present but dipstick test is negative.

The 2007 NICE guideline recommends that non-invasive methods of urine sampling, in particular clean catch urine sampling, is the preferred method of urine sampling. In accordance with this, a sensitivity analysis stratified the results of clinical evidence by non-invasive urine sampling. The committee noted that in the sensitivity analysis for infants under 3 months, likelihood ratios of nitrite, LE and nitrite or LE decrease in significance. However, it was noted that this finding was from 1 study with 144 participants and that the results may be different with a larger sample size. The committee discussed the possibility that the dipstick test itself is less sensitive or specific when non-invasive methods of urine sampling are used and if this could explain the differences observed in the sensitivity analysis. Samples from non-invasive methods could be contaminated and potentially decrease the accuracy of dipstick testing. It was noted that obtaining a good quality clean catch sample is dependent on the ability of the person collecting the sample and that this technique is easier with older children. It was also noted that in the committee's experience invasive sampling methods are never considered in primary care and are rarely considered in secondary care. This is because there is a risk of adverse events including infection and kidney or bladder damage. In the rare cases where urine collection is via a suprapubic aspirate, for example where an infant or child is very unwell, ultrasound guidance is used to determine if there is urine in the bladder and this requires training and experience.

When considering infants under 3 months who are febrile and in whom UTI is a possible diagnosis, the committee noted that these cases are usually not seen in primary care. If they do present to primary care, the infant will be immediately referred to secondary care and will usually be admitted for intravenous antibiotics and further investigation. For these cases there are various additional clinical concerns in these infants alongside the suspicion of a UTI, including: an immature immune system; a risk of meningitis, in which case the infant may have a lumbar puncture; the risk of sepsis; differential responses to antibiotics and concern of other causes of the symptoms such as congenital abnormalities. For this group, the committee discussed the option of conducting a dipstick test first and only sending

samples which tested positive for either or both nitrite or LE for culture, as the evidence presented showed that a positive nitrite or LE or both greatly increases the likelihood of finding a positive urine culture. However, it was agreed that this would not inform management as antibiotics would be given immediately in all cases where UTI is suspected as the risk of false negatives would pose a high risk of complications in this age group. The committee agreed that any infant under 3 months with a suspected UTI, even if they are not febrile, should always be referred to a centre offering specialist paediatric care and treated under the fever in under 5s guideline (CG160).

Having considered the evidence, the committee agreed not to change the 2007 recommendations which state that infants under 3 months with suspected UTI should be referred to paediatric specialist care and urgent microscopy and culture (recommendation 1). Removal of urgent microscopy from the recommendation was discussed, as microscopy alone as a reference test was not considered in the evidence review and no evidence from studies including microscopy and culture as a reference test was found for this age group. However, the committee agreed the use of both microscopy and culture is clinically important in this age group as urgent microscopy can provide information on the state of infection and inflammation prior to culture results and, in combination with culture, can inform management. Additionally, the committee agreed retaining the recommendation will position the need for urgent microscopy. The committee also agreed to amend recommendation 3, which provides indication for urine culture, from all infants and children younger than 3 years to infants under 3 months. This is because for infants and children aged 3 months or older but younger than 3 years, the committee recommended that if either or both of nitrite or LE are positive, antibiotic treatment should be started and a urine sample should be sent for culture.

In infants and children 3 months or older but younger than 3 years, the committee noted that the 2007 recommendations do not reflect the updated evidence. As the evidence presented demonstrated that positive dipstick testing for either or both nitrite or LE greatly increases the likelihood of finding a positive urine culture. The committee agreed that a dipstick strategy should be considered for this age group. It was noted that the sensitivity analysis based on the use of non-invasive urine sampling methods showed that the likelihood ratios of both nitrite and LE is not significant. This finding was based on 1 study with 124 participants (Sharief 1998) and the committee agreed that in clinical practice, a positive result for both nitrite and LE could be a clear sign of UTI. Additionally, Sharief 1998 used a mixture of clean catch and bag urine sampling techniques, of which bag samples are uncommon in current clinical practice. Therefore, the committee recommended in recommendation 2 that if one or both of nitrite or LE is positive, a urine test should be sent for culture and antibiotics should be started. The committee discussed the addition of microscopy to culture in this age group. One study was included (Lejeune 1991) which had culture and microscopy as a reference test, and the committee did not take this into account due to various biases associated with the study. Additionally, the committee noted that in current practice, flow cytometry is usually preferred over microscopy as it reduces the number of plates required. The committee agreed that urgent microscopy, in addition to culture, would not further inform clinical management and therefore did not recommend the use of urgent microscopy in this age group. The committee recommended that if both LE and nitrite are negative, antibiotic treatments should not be started and a culture should not be carried out, unless the criteria specified in recommendation 3 apply. The fourth bullet of recommendation 3, which specified that a single positive result for LE or nitrite is an indication for culture, was amended to remove the word ‘single’ to align with the

recommendations for infants and children 3 months or older but younger than 3 years.

The committee noted that the 2007 guidelines for diagnosis of UTI in infants and children 3 months or older but younger than 3 years are stratified by low, intermediate or high risk of UTI. However, it was noted that this stratification originated from the fever guideline (CG47 feverish illness in children (2007) which was replaced by CG160 fever in under 5s (2013)), and was consensus based. Additionally the new evidence presented to the committee did not consider stratification based on risk. Therefore the committee agreed that recommendation 3, which outlines indications for culture, is sufficient in recommending that urine samples should be sent for culture in those with intermediate to high risk of infection. The committee agreed that the new recommendations will provide a more concise and clearer guidance for health care professionals, more efficient diagnosis for infants and children and less burden on laboratories by reducing the amount of urine samples sent for culture.

Cost effectiveness and resource use

The committee considered evidence from 2 economic evaluations in the literature. Whiting et al. (2006) used a modelling approach to assess the cost-effectiveness of a number of testing strategies for UTI in children aged under 5 years, stratified by age and gender. The analysis found that, for girls under 3 years and for boys under 2 years, the most cost-effective strategy involved dipstick testing. However, this evaluation was judged to suffer from a number of limitations. Most importantly, the analysis did not stratify input data on the accuracy of tests by patients' age group – a factor which is key to answering the review question.

Hay et al (2016) used a modelling approach to determine the cost-effectiveness of dipstick testing in children at low risk of UTI under the age of 5 years compared to a strategy of 'laboratory testing' (laboratory testing in all children, with treatment delayed until results are returned) and 'presumptive treatment' (antibiotics prescribed for all children and a urine sample sent for laboratory testing). Results showed that the dipstick testing strategy was not cost-effective compared to the 'laboratory testing' strategy. However, the committee noted that the dipstick testing algorithm in the Hay analysis differed fundamentally from the algorithm recommended in NICE guidance for children over 3 years: in the model, all children receive a dipstick test and have a urine sample sent for laboratory testing, with dipstick results used to guide antibiotic treatment while awaiting laboratory results.

The committee considered the results of the de novo economic modelling, which compared the following strategies:

- Treat all children with suspected UTI with antibiotics and send a urine sample for microscopy and culture ('no dipstick testing')
- Dipstick test all children with suspected UTI. Only initiate antibiotic treatment and send a sample for microscopy and culture in children with a positive dipstick result ('dipstick testing')

Due to uncertainty regarding the outcomes of a delay in treatment of UTI, the committee was presented with results from a number of scenarios, in which a range of potential consequences of a false negative dipstick result of varying severity were explored. The committee noted that, in the majority of scenarios, a 'no dipstick testing' strategy was not cost-effective compared to a 'dipstick' strategy, due to the relatively high accuracy of dipstick tests, and relative infrequency of serious adverse events.

Only in scenarios in which an untreated UTI was assumed to result in an increased risk of septicaemia was the 'no dipstick testing' strategy potentially cost-effective compared to 'dipstick testing' strategies. However, the committee felt that this assumption did not accurately reflect current clinical practice as, while UTI and septicaemia may be co-incident at presentation, there is little evidence that UTI can develop into septicaemia if left untreated. Moreover, the committee felt that the baseline probability of septicaemia in children with UTI and probability of death in septicaemia used in the model base case were unrealistically high, and the ICER of 'no dipstick' was therefore likely to be higher than £20,000 for this scenario.

The committee also considered evidence from the economic model which compared different interpretations of dipstick results. The committee noted that the economic modelling showed 'LE alone' and 'nitrite or LE' strategies were consistently more cost-effective than 'nitrite alone' or 'nitrite and LE' strategies. This result was consistent with the clinical evidence, which showed that the former two strategies are associated with a considerably higher sensitivity than the latter two, at the expense of a modestly small decrease in specificity.

Based on the clinical evidence, and the economic evidence that the current 'no dipstick' testing strategy is not cost-effective in the large majority of modelled scenarios, the committee decided to recommend dipstick testing in children aged 3 months or older but younger than 3 years with symptoms suggestive of UTI.

It was noted that data on the baseline prevalence of UTI were taken from a meta-analysis which was predominantly based on studies conducted in an emergency department setting. Since children in such a setting are likely to exhibit more severe symptoms than in a general practice setting, it is possible that this source provides an overestimate of the true prevalence of UTI in the population of interest. However, using a lower prevalence of UTI in the analysis would be very unlikely to result in a different decision, as this would only serve to increase the ICER of laboratory testing compared to dipstick testing in children aged 3 months or older but younger than 3 years.

The committee discussed the appropriateness of the economic evidence for children under 3 months, and noted that that the symptoms, risks, and treatment pathway for children in this age groups differed substantially for those of children aged over 3 months. Specifically, the committee noted that symptoms of UTI are generally less specific for young infants, and that there is no plausible group of children of this age in whom only UTI, and no other potential causes, could be suspected. Children under 3 months are also routinely referred to secondary care immediately, rather than being managed in the first instance by a GP. For these reasons, the committee felt that the implicit modelling assumption that children who do not have a UTI are otherwise healthy was not appropriate for children in the younger age group; a dipstick test alone would not be sufficient to determine that a child is healthy and in need of no further investigation. Therefore, the committee determined that children of under 3 months should be referred to a paediatric specialist and a urine sample should be sent for laboratory testing. The committee also discussed whether to recommend that dipstick testing is also provided for this age group (in addition to microscopy and culture in all children), but concluded that this would provide little additional diagnostic benefit.

The committee discussed the potential resource impact of the recommendations, and concluded that they are likely to result in a substantial cost saving in children aged 3 months or older but younger than 3 years, due to the relatively low cost of dipstick testing compared to microscopy and culture.

Other factors the committee took into account

The committee raised that the first bullet of recommendation 3, which specifies that diagnosis of acute pyelonephritis/upper UTI is an indication for culture, is unclear in relation to recommendation 1.1.8.1 which outlines the clinical differentiation between upper and lower UTI. To clarify the recommendation, the first bullet of recommendation 3 was amended to in infants and children who are considered to have acute pyelonephritis/upper urinary tract infection.

The committee considered any possible equalities issues and noted that the evidence and recommendations from this guideline update are not generalisable to children over 3 years of age. It was also noted that the prevalence of congenital abnormalities is higher in boys under 3 months and by retaining recommendation 2, it is ensured that these boys will be referred to secondary care where congenital abnormalities would be taken into consideration. A research recommendation from the old guideline, for research to investigate nitrite or LE dipstick testing and stratify this by age in under 3 years, was deleted as new studies have been published addressing this question and were included in this guideline update.

Glossary

Dipstick	A diagnostic test consisting of a chemically sensitive strip which when dipped into a sample can be used to detect the presence of leucocyte esterase, nitrites, glucose or protein.
Leukocyte esterase (LE)	An enzyme present in white blood cells which can be detected in the urine during infection.
Negative likelihood ratio (LR−)	The negative likelihood ratio describes the probability of having a negative test result in the diseased population compared with that of a non-diseased population and corresponds to the ratio of the false negative rate divided by the true negative rate (1 − sensitivity/specificity).
Nitrite	Nitrite is a chemical compound produced by bacterial metabolism. Its presence in urine is used as a marker of the presence of bacteria. Not all bacteria are able to produce nitrite.
Positive likelihood ratio (LR+)	The positive likelihood ratio describes the probability of having a positive test result in the diseased population compared with that of a non-diseased population and corresponds to the ratio of the true positive rate divided by the false positive rate (sensitivity/(1−specificity)).
Sensitivity	In diagnostic testing, sensitivity refers to the chance of having a positive test result given that you have the disease. 100% sensitivity means that all those with the disease will test positive, but this is not the same the other way around. A patient could have a positive test result but not have the disease – this is called a 'false positive'. The sensitivity of a test is also related to its negative predictive value (true negatives) – a test with a sensitivity of 100% means that all those (or almost all those in very large studies) who get a negative test result do not have the disease. To fully judge the accuracy of a test, its specificity must also be considered.
Specificity	In diagnostic testing, specificity refers to the chance of having a negative test result given that you do not have the disease. 100% specificity means that all those without the disease will test negative, but this is not the same the other way around. A patient could have a negative test result yet

	still have the disease – this is called a 'false negative'. The specificity of a test is also related to its positive predictive value (true positives) – a test with a specificity of 100% means that all those (or almost all those in very large studies) who get a positive test result definitely have the disease. To fully judge the accuracy of a test, its sensitivity must also be considered.

Appendices

Appendix A: Review protocol

A.1 Review protocol for UTI diagnosis in under 3 months and 3 months or older but younger than 3 years

Review Protocol		
Components	**Details**	**Comments during development**
Review question	In infants and children up to 3 years old with suspected UTI, what is the diagnostic accuracy of urine tests for detecting UTI?	
Background/objectives	The recommendations on dipstick testing in the NICE guideline on Urinary tract infection in under 16s were organised by age-group as follows 0 to 3 months, >3 months to 3 years and older than 3 years. This reflected the evidence base that dipstick testing was not accurate in children up to 3 years of age. This topic was reviewed in 2016 by the surveillance team and new evidence (5 studies on diagnostic accuracy of urine dipstick testing) were identified in the younger age group; this evidence suggested that the guideline should therefore be updated to reflect new evidence in this area. This evidence review will focus on the diagnostic accuracy of dipstick tests alone or with other tests in children up to 3 years of age	16/02: edited typo of 4 years to 3 years.
Population	Those in whom there is a clinical suspicion of UTI and are: < 3 months old 3 months to 3 years old	
Index test	Dipstick test: • Leukocyte esterase • Nitrites • Protein • Blood Dipstick testing with other tests including: • microscopy alone (automated or manual) • urine culture alone (can include clean catch, bladder catheterisation and suprapubic aspirate samples) • microscopy and culture.	06/02: no studies included used culture as index test. Dipstick alone or with microscopy included.
Reference test	Clinical diagnosis of UTI. This may include consideration of a urine culture alone or a combination of tests.	01/02: reference tests included include urine culture alone and urine culture with microscopy. Different thresholds for culture positive were used in studies (usually dependant on urine sampling method).
Outcomes	Sensitivity Specificity	06/02: Agreed to include sensitivity,

Review Protocol		
	Likelihood ratios Area under the curve Negative predictive values	specificity and likelihood ratios in analysis.
Type of review question	Diagnostic	
Types of study to be included	RCT (if any available), Cohort studies, Cross-sectional studies	
Language	English	
Status	Published and studies that will be published by the time the guideline update is published (June 2017).	
Any other information or criteria for inclusion/exclusion	The committee will be sent the list of included and excluded studies prior to the committee meeting. The committee will be requested to check whether any studies have been excluded inappropriately, and whether there are any relevant studies they know of which haven't been picked up by the searches or have been wrongly sifted out.	
Analysis of subgroups or subsets	Data will be analysed separately by age (less than 3 months and 3 months to 3 years) Method of urine collection: clean catch, bladder catheterisation or suprapubic aspirate samples.	
Data extraction and quality assessment	Sifting Relevant studies will be identified through sifting the abstracts and excluding studies clearly not relevant to the review question (measured against protocol). In the case of relevant or potentially relevant studies, the full paper will be ordered and reviewed, whereupon studies considered being not relevant to the topic will be excluded. i) Selection based on titles and abstracts A full double-sifting of titles and abstracts will not be conducted due to narrow question with clearly defined straightforward inclusion and exclusion criteria. However in cases of uncertainty the following mechanisms will be in place: • technical analyst will discuss with a support technical analyst • comparison with included studies of other current (within 5 years) systematic reviews • recourse to members of the committee ii) Selection based on full papers A full double-selecting of full papers for inclusion/exclusion will not be conducted (see above). However in cases of uncertainty the same mechanisms stated in i) above will be followed. Data extraction	

Review Protocol		
	Relevant information from included studies will be extracted into standardised evidence tables [adapted to suit this particular question]. Critical appraisal The risk of bias of each included study will be assessed using standardised checklists available in the NICE manual appropriate for the design of each included study. Quality assessment GRADE methodology will be used to assess the quality of evidence on an outcome basis: • Risk of bias will be assessed using critical appraisal checklists • Inconsistency will be assessed using I^2 • Indirectness will be assessed after considering the population, intervention and outcomes of included studies, relative to the target population as specified in the review protocol • Imprecision will be assessed using whether the confidence intervals around point estimates cross the MIDs for each outcome. COMET and published literature including related NICE guidelines will be checked for appropriate minimal important differences (MID) for each outcome. If none are available, the topic experts will be consulted on the appropriateness of using default MIDs as suggested by the GRADE working group. Quality Assurance: A full double-scoring quality assessment will not be conducted due to the nature of the review question (see above). Other quality assurance mechanisms will be in place as follows: Internal QA (10%) by CGUT technical adviser on the risk of bias and quality assessment that is being conducted. Any disagreement will be resolved through discussion. The Committee will be sent the evidence synthesis prior to the committee meeting and will be requested to comment on the quality assessment, which will serve as another QA function.	
Strategy for data synthesis	If possible a meta-analysis of available study data will be carried out to provide a more complete picture of the evidence body as a whole. A fixed effects model will be used as it is expected that the studies will be homogenous in terms of population and we can assume a similar effect size across studies. A random effects model will be used if this assumption is not correct.	09/02: Random-effects models (der Simonian and Laird) were fitted for all syntheses, as recommended in the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy (Deeks et al. 2010).

Review Protocol		
Searches	**Sources to be searched** • Clinical searches - Medline, Medline in Process, Embase, Cochrane CDSR, CENTRAL, DARE (legacy records), HTA and PubMed. • Economic searches - Medline, Medline in Process, Embase, EconLit, PubMed, NHS EED (legacy records) and HTA, with economic evaluations and quality of life filters applied. **Supplementary search techniques** • None identified **Limits** • Studies reported in English • The McMaster diagnostic filter for best sensitivity/specificity will be used and adapted • An age limit will be applied • Animal studies will be excluded from the search results • Conference abstracts will be excluded from the search results in Embase • A date limit from the original search of July 2005-present will be applied	10/01: Whiting 2006 systematic review was found in search and references assessed for relevant articles. This systematic review was updated in a recent HTA (the DUTY study) which was also searched for relevant literature. 23/01: Studies included in the original guideline ordered. 22/02: added in search strategy summary.
Key papers	Kanegaye JT, Jacob JM, and Malicki D (2014) .Automated urinalysis and urine dipstick in the emergency evaluation of young febrile children. Pediatrics, 134, 3: 523-529. Glissmeyer EW, Korgenski EK, Wilkes J, Schunk JE, et al (2014) Dipstick Screening for Urinary Tract Infection in Febrile Infants. Pediatrics . Free full text. Shah AP, Cobb BT, Lower DR, Shaikh, N, et al (2014) Enhanced versus automated urinalysis for screening of urinary tract infections in children in the emergency department. The Pediatric infectious disease journal, 33, 3: 272-275 Velasco R, Benito H, Mozun R et al. (2015) Using a urine dipstick to identify a positive urine culture in young febrile infants is as effective as in older patients. Acta Paediatrica 104:e39-e44. Velasco R, Benito H, Mozun R et al. (2015) Febrile young infants with altered urinalysis at low risk for invasive bacterial infection. A Spanish Pediatric Emergency Research Network's Study. [Erratum appears in Pediatr Infect Dis J. 2015 Mar;34(3):295 Note: Tiago, San [corrected to Mintegi, Santiago]]. Pediatric Infectious Disease Journal 34:17-21	10/01: Shah 2014 was excluded based on abstract as this study compares manual urinalysis with automated urinalysis.

Appendix B: Literature search strategy

B.1 Clinical search summary

Databases	Date searched	Version/files	No. retrieved	EndNote data
Cochrane Central Register of Controlled Trials (CENTRAL)	28/10/2016	Issue 9 of 12, September 2016	257	168
Cochrane Database of Systematic Reviews (CDSR)	28/10/2016	Issue 10 of 12, October 2016	15	9
Database of Abstracts of Reviews of Effect (DARE)	28/10/2016	Issue 2 of 4, April 2015	8	4
Embase (Ovid)	28/10/2016	Embase 1974 to 2016 Week 43	6171	4469
MEDLINE (Ovid)	28/10/2016	Ovid MEDLINE(R) 1946 to October Week 3 2016	2555	2373
MEDLINE In-Process (Ovid)	28/10/2016	Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations October 27, 2016	253	107
PubMeda	28/10/2016	-	28	27
Health Technology Assessment (HTA Database)	28/10/2016	Issue 3 of 4, July 2016	0	0

B.2 Clinical search terms (Medline)

Database: Medline
Strategy used: 1 exp Urinary Tract/ (408154) 2 ((urin$ or renal$) adj (system$ or tract$)).tw. (61757) 3 exp Urinary Tract Infections/ (42068) 4 ((bladder$ or genitourin$ or kidney$ or pyelo$ or renal$ or ureter$ or ureth$ or urin$ or urolog$ or urogen$) adj5 infect$).tw. (53135) 5 UTI.tw. (5928) 6 ((upper or lower) adj5 urin$).tw. (21907) 7 Cystitis/ (7074) 8 Cystitis, Interstitial/ (1840) 9 cystiti$.tw. (8957) 10 (bladder$ adj5 (inflamm$ or ulcer$ or ulcus)).tw. (1397) 11 or/1-10 (498103) 12 fever/ or "fever of unknown origin"/ (38932) 13 (fever$ or pyrexia$).tw. (131667) 14 (febrile adj2 (response$ or reaction$)).tw. (1866)

Database: Medline
15 ((high$ or elevate$ or increas$ or hot) adj2 temp$).tw. (71188)
16 Proteinuria/ (22316)
17 proteinuri$.tw. (31589)
18 Albuminuria/ (13474)
19 albuminuri$.tw. (7733)
20 ((protein$ or albumin$) adj5 urin$).tw. (23459)
21 bacteriuria$.tw. (5039)
22 ((bacteria$ or microbial$) adj5 (bladder$ or genitourin$ or kidney$ or renal$ or ureter$ or ureth$ or urin$ or urolog$ or urogen$)).tw. (5377)
23 pyuri$.tw. (1479)
24 leukocyte esterase.tw. (392)
25 Nitrites/ (17046)
26 nitrite$.tw. (25315)
27 Vesico-Ureteral Reflux/ (7775)
28 ((vesicorenal$ or vesico?ureteral$ or vesicour$) adj reflux).tw. (5018)
29 VUR.tw. (1685)
30 ((backflow$ or bladder$ or cystoureteral$ or ureter$ or urether$ or urin$) adj5 reflux$).tw. (4295)
31 Pyelonephritis/ (13882)
32 (pyelonephriti$ or pyonephrosi$ or pyelocystiti$ or pyelitis).tw. (12042)
33 or/12-32 (347383)
34 or/11,33 (793653)
35 Diagnostic Techniques, Urological/ (761)
36 (urolog$ adj2 (diagnostic$ or exam$ or technic$ or technique$)).tw. (811)
37 Urinalysis/ (6466)
38 urinalys$.tw. (6361)
39 ((urine or urinary) adj2 (analys$ or collect$ or exam$ or investigation$ or sample$ or specimen$ or test$)).tw. (56126)
40 clean catch.tw. (204)
41 (suprapubic adj2 aspirat$).tw. (210)
42 exp Reagent Kits, Diagnostic/ (18684)
43 "Indicators and Reagents"/ (51191)
44 (reagent$ adj (kit$ or strip$)).tw. (988)
45 (dipstick$ or dip?stick$).tw. (2495)
46 multistix.tw. (109)
47 exp Microscopy/ (507757)
48 microscop$.tw. (500315)
49 (dipslide$ or dip?slide$).tw. (81)
50 (urin$ adj culture$).tw. (4116)
51 Culture Techniques/ (47659)
52 exp Microbiological Techniques/ (268860)
53 ((bacteri$ or culture$ or microbial$ or phage$) adj2 (biotyp$ colon$ or techni$ or typ$)).tw. (24724)
54 Antibody-Coated Bacteria Test, Urinary/ (152)
55 ((urine or urinary) adj3 (antibody$ adj coated)).tw. (56)
56 ((bacteri$ or microbial$) adj3 sensitive$).tw. (2235)
57 Colorimetry/ (19528)
58 (colorimet$ or colourimet$).tw. (20887)
59 Catalase/ (28741)
60 uriscreen$.tw. (13)
61 "Nephelometry and Turbidimetry"/ (7002)

Database: Medline
62 (turbidimetry or nephelometry).tw. (2136) 63 (triphen$ adj tetrazolium).tw. (623) 64 TTC.tw. (2698) 65 filtracheck$.tw. (5) 66 sysmex$.tw. (775) 67 headspace$.tw. (5414) 68 impendence$.tw. (33) 69 or/35-68 (1281643) 70 34 and 69 (84243) 71 "Sensitivity and Specificity"/ (313007) 72 (sensitiv: or predictive value:).mp. or accurac:.tw. (1591122) 73 specificity.tw. (350636) 74 roc curve/ (39925) 75 (roc adj2 (curve$ or analys$)).tw. (23324) 76 receiver operat$ characteristic$.tw. (40095) 77 likelihood functions/ (19389) 78 (likelihood adj (estimate$ or ratio$)).tw. (11450) 79 "Predictive Value of Tests"/ (170069) 80 Mass Screening/ (90363) 81 screen$.tw. (510691) 82 exp Diagnostic Errors/ (104834) 83 (diagnos$ adj2 error$).tw. (3878) 84 misdiagnos$.tw. (21222) 85 (false adj (negative$ or positive$)).tw. (59462) 86 "reproducibility of results"/ (328203) 87 Diagnosis, Differential/ (418679) 88 (differential adj diagnos$).tw. (99790) 89 or/71-88 (2864728) 90 70 and 89 (20672) 91 exp Child/ (1715584) 92 Child, Hospitalized/ (6116) 93 exp Infant/ (1035659) 94 (child$ or infant$ or toddler$ or baby or babies or kid$ or girl$ or boy$ or junior$ or neonate$ or newborn$ or preschool or young$).tw. (2218836) 95 exp Pediatrics/ (51580) 96 (pediatric$ or paediatric$).tw. (239344) 97 (under adj (1* or one$ or 2* or two$ or 3* or three$)).tw. (39088) 98 ((1* or one$ or 2* or two$ or 3* or three$) adj year$ adj2 (old$ or age)).tw. (421296) 99 or/91-98 (3474670) 100 90 and 99 (7073) 101 animals/ not humans/ (4298080) 102 100 not 101 (5746) 103 limit 102 to english language (4860) 104 (200507* or 200508* or 200509* or 20051* or 2006* or 2007* or 2008* or 2009* or 201*).ed. (8809246) 105 103 and 104 (2555)

B.3 Economic search summary

Database	Date searched	Version/files	No. retrieved	EndNote data
MEDLINE (Ovid)	31/10/2016	Ovid MEDLINE(R) 1946 to October Week 3 2016	221	200
MEDLINE in Process (Ovid)	31/10/2016	Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations October 27, 2016	16	15
Embase (Ovid)	31/10/2016	Embase 1974 to 2016 Week 44	475	313
EconLit (Ovid)	31/10/2016	Econlit 1886 to September 2016	0	0
NHS Economic Evaluation Database (NHS EED) (legacy database)	28/10/2016	Issue 2 of 4, April 2015	4	2
Health Technology Assessment (HTA Database)	28/10/2016	Issue 3 of 4, July 2016	0	0
PubMedb	28/10/2016	-	28	28

B.4 Economic search terms (Medline)

Database: Medline
Strategy used: 1 exp Urinary Tract/ (408154) 2 ((urin$ or renal$) adj (system$ or tract$)).tw. (61757) 3 exp Urinary Tract Infections/ (42068) 4 ((bladder$ or genitourin$ or kidney$ or pyelo$ or renal$ or ureter$ or ureth$ or urin$ or urolog$ or urogen$) adj5 infect$).tw. (53135) 5 UTI.tw. (5928) 6 ((upper or lower) adj5 urin$).tw. (21907) 7 Cystitis/ (7074) 8 Cystitis, Interstitial/ (1840) 9 cystiti$.tw. (8957) 10 (bladder$ adj5 (inflamm$ or ulcer$ or ulcus)).tw. (1397) 11 or/1-10 (498103) 12 fever/ or "fever of unknown origin"/ (38932) 13 (fever$ or pyrexia$).tw. (131667) 14 (febrile adj2 (response$ or reaction$)).tw. (1866) 15 ((high$ or elevate$ or increas$ or hot) adj2 temp$).tw. (71188) 16 Proteinuria/ (22316) 17 proteinuri$.tw. (31589) 18 Albuminuria/ (13474) 19 albuminuri$.tw. (7733) 20 ((protein$ or albumin$) adj5 urin$).tw. (23459)

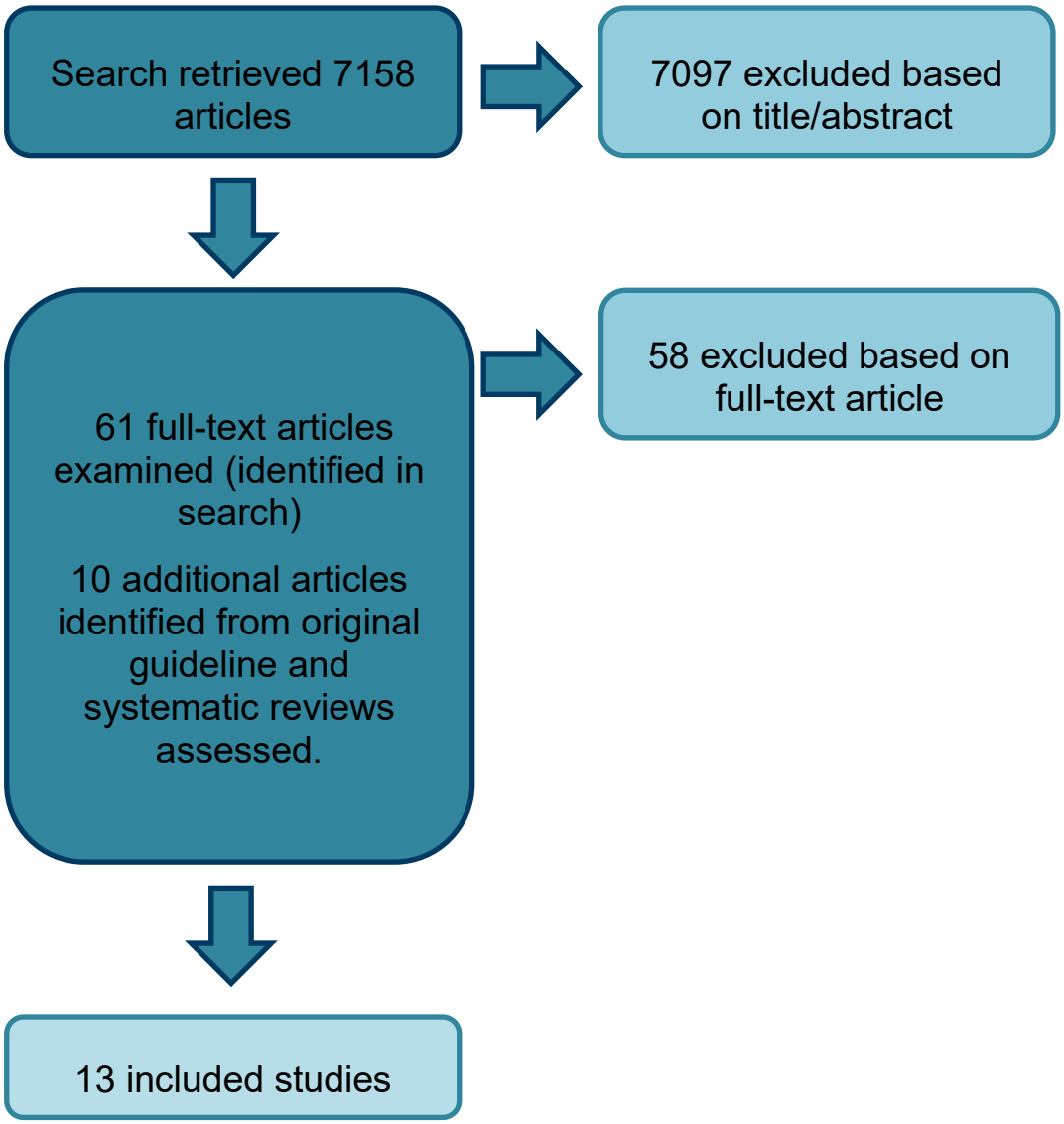
Database: Medline
21 bacteriuria$.tw. (5039)
22 ((bacteria$ or microbial$) adj5 (bladder$ or genitourin$ or kidney$ or renal$ or ureter$ or ureth$ or urin$ or urolog$ or urogen$)).tw. (5377)
23 pyuri$.tw. (1479)
24 leukocyte esterase.tw. (392)
25 Nitrites/ (17046)
26 nitrite$.tw. (25315)
27 Vesico-Ureteral Reflux/ (7775)
28 ((vesicorenal$ or vesico?ureteral$ or vesicour$) adj reflux).tw. (5018)
29 VUR.tw. (1685)
30 ((backflow$ or bladder$ or cystoureteral$ or ureter$ or urether$ or urin$) adj5 reflux$).tw. (4295)
31 Pyelonephritis/ (13882)
32 (pyelonephriti$ or pyonephrosi$ or pyelocystiti$ or pyelitis).tw. (12042)
33 or/12-32 (347383)
34 or/11,33 (793653)
35 Diagnostic Techniques, Urological/ (761)
36 (urolog$ adj2 (diagnostic$ or exam$ or technic$ or technique$)).tw. (811)
37 Urinalysis/ (6466)
38 urinalys$.tw. (6361)
39 ((urine or urinary) adj2 (analys$ or collect$ or exam$ or investigation$ or sample$ or specimen$ or test$)).tw. (56126)
40 clean catch.tw. (204)
41 (suprapubic adj2 aspirat$).tw. (210)
42 exp Reagent Kits, Diagnostic/ (18684)
43 "Indicators and Reagents"/ (51191)
44 (reagent$ adj (kit$ or strip$)).tw. (988)
45 (dipstick$ or dip?stick$).tw. (2495)
46 multistix.tw. (109)
47 exp Microscopy/ (507757)
48 microscop$.tw. (500315)
49 (dipslide$ or dip?slide$).tw. (81)
50 (urin$ adj culture$).tw. (4116)
51 Culture Techniques/ (47659)
52 exp Microbiological Techniques/ (268860)
53 ((bacteri$ or culture$ or microbial$ or phage$) adj2 (biotyp$ colon$ or techni$ or typ$)).tw. (24724)
54 Antibody-Coated Bacteria Test, Urinary/ (152)
55 ((urine or urinary) adj3 (antibody$ adj coated)).tw. (56)
56 ((bacteri$ or microbial$) adj3 sensitive$).tw. (2235)
57 Colorimetry/ (19528)
58 (colorimet$ or colourimet$).tw. (20887)
59 Catalase/ (28741)
60 uriscreen$.tw. (13)
61 "Nephelometry and Turbidimetry"/ (7002)
62 (turbidimetry or nephelometry).tw. (2136)
63 (triphen$ adj tetrazolium).tw. (623)
64 TTC.tw. (2698)
65 filtracheck$.tw. (5)
66 sysmex$.tw. (775)
67 headspace$.tw. (5414)

Database: Medline
68 impendence$.tw. (33)
69 or/35-68 (1281643)
70 34 and 69 (84243)
71 "Sensitivity and Specificity"/ (313007)
72 (sensitiv: or predictive value:).mp. or accurac:.tw. (1591122)
73 specificity.tw. (350636)
74 roc curve/ (39925)
75 (roc adj2 (curve$ or analys$)).tw. (23324)
76 receiver operat$ characteristic$.tw. (40095)
77 likelihood functions/ (19389)
78 (likelihood adj (estimate$ or ratio$)).tw. (11450)
79 "Predictive Value of Tests"/ (170069)
80 Mass Screening/ (90363)
81 screen$.tw. (510691)
82 exp Diagnostic Errors/ (104834)
83 (diagnos$ adj2 error$).tw. (3878)
84 misdiagnos$.tw. (21222)
85 (false adj (negative$ or positive$)).tw. (59462)
86 "reproducibility of results"/ (328203)
87 Diagnosis, Differential/ (418679)
88 (differential adj diagnos$).tw. (99790)
89 or/71-88 (2864728)
90 70 and 89 (20672)
91 exp Child/ (1715584)
92 Child, Hospitalized/ (6116)
93 exp Infant/ (1035659)
94 (child$ or infant$ or toddler$ or baby or babies or kid$ or girl$ or boy$ or junior$ or neonate$ or newborn$ or preschool or young$).tw. (2218836)
95 exp Pediatrics/ (51580)
96 (pediatric$ or paediatric$).tw. (239344)
97 (under adj (1* or one$ or 2* or two$ or 3* or three$)).tw. (39088)
98 ((1* or one$ or 2* or two$ or 3* or three$) adj year$ adj2 (old$ or age)).tw. (421296)
99 or/91-98 (3474670)
100 90 and 99 (7073)
101 animals/ not humans/ (4298080)
102 100 not 101 (5746)
103 limit 102 to english language (4860)
104 (200507* or 200508* or 200509* or 20051* or 2006* or 2007* or 2008* or 2009* or 201*).ed. (8809246)
105 103 and 104 (2555)
106 Economics/ (26804)
107 exp "Costs and Cost Analysis"/ (203488)
108 Economics, Dental/ (1892)
109 exp Economics, Hospital/ (21918)
110 exp Economics, Medical/ (13978)
111 Economics, Nursing/ (3944)
112 Economics, Pharmaceutical/ (2660)
113 Budgets/ (10611)
114 exp Models, Economic/ (12189)
115 Markov Chains/ (11679)

Database: Medline	
116	Monte Carlo Method/ (23376)
117	Decision Trees/ (9758)
118	econom$.tw. (183781)
119	cba.tw. (9238)
120	cea.tw. (18078)
121	cua.tw. (852)
122	markov$.tw. (14003)
123	(monte adj carlo).tw. (24337)
124	(decision adj3 (tree$ or analys$)).tw. (9866)
125	(cost or costs or costing$ or costly or costed).tw. (360333)
126	(price$ or pricing$).tw. (26610)
127	budget$.tw. (19559)
128	expenditure$.tw. (40103)
129	(value adj3 (money or monetary)).tw. (1588)
130	(pharmacoeconomic$ or (pharmaco adj economic$)).tw. (3066)
131	or/106-130 (750758)
132	"Quality of Life"/ (144521)
133	quality of life.tw. (169480)
134	"Value of Life"/ (5528)
135	Quality-Adjusted Life Years/ (8913)
136	quality adjusted life.tw. (7667)
137	(qaly$ or qald$ or qale$ or qtime$).tw. (6254)
138	disability adjusted life.tw. (1674)
139	daly$.tw. (1582)
140	Health Status Indicators/ (21944)
141	(sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).tw. (18133)
142	(sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six).tw. (1100)
143	(sf12 or sf 12 or short form 12 or shortform 12 or sf twelve or sftwelve or shortform twelve or short form twelve).tw. (3459)
144	(sf16 or sf 16 or short form 16 or shortform 16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).tw. (22)
145	(sf20 or sf 20 or short form 20 or shortform 20 or sf twenty or sftwenty or shortform twenty or short form twenty).tw. (350)
146	(euroqol or euro qol or eq5d or eq 5d).tw. (5310)
147	(qol or hql or hqol or hrqol).tw. (30965)
148	(hye or hyes).tw. (54)
149	health$ year$ equivalent$.tw. (38)
150	utilit$.tw. (133084)
151	(hui or hui1 or hui2 or hui3).tw. (1024)
152	disutili$.tw. (270)
153	rosser.tw. (72)
154	quality of wellbeing.tw. (8)
155	quality of well-being.tw. (356)
156	qwb.tw. (189)
157	willingness to pay.tw. (2878)
158	standard gamble$.tw. (705)
159	time trade off.tw. (849)
160	time tradeoff.tw. (220)
161	tto.tw. (697)

Database: Medline	
162	or/132-161 (380248)
163	131 or 162 (1078629)
164	105 and 163 (221)

Appendix C: Clinical evidence study selection



Appendix D: Clinical evidence tables

D.1 Dayan 2002

Bibliographic reference	**Dayan P, Bennett J, Best R, et al 2002. Test characteristics of the urine Gram stain in infants≤ 60 days of age with fever. . Pediatric emergency care, 18(1), pp.12-14.**
Study type	Prospective cohort
Aim	Gram stain has shown favourably low false positive and false-negative rates in young children, but its test characteristics have not been defined specifically for infants. The aim was to evaluate the test characteristics of the Gram stain in infants ≤60 days of age and compare them to the standard UA for nitrites and leukocyte esterase and microscopy (WBCs/hpf).
Patient characteristics	Enrolment: Consecutive sample of infants presenting to a paediatric emergency department during 2 consecutive winter seasons Inclusion criteria: • Reported or recorded rectal temperature ≥ 38°C Exclusion criteria: • Received antibiotics within 48 hours of evaluation • Urine collection attempted but not obtained • If a gram stain was not completed secondary to laboratory unavailability Patient characteristics: N=232 for which Gram stain and culture were analysed Age: 60 (26%) aged 1-30 days old; 172 (74%) aged 31-60 days Gender: 114 (49%) male; 118 (51%) female Number circumcised (if reported): not reported Symptomatic / asymptomatic: not reported Prevalence of UTI: 10% Method of urine collection:

Bibliographic reference	Dayan P, Bennett J, Best R, et al 2002. Test characteristics of the urine Gram stain in infants≤ 60 days of age with fever. . Pediatric emergency care, 18(1), pp.12-14.
	225/232 (97%) by catheterisation 7/232 (3%) by suprapubic aspiration
Number of patients	N = 232 for which Gram stain and culture were analysed (baseline characteristics pertain to this sample) N = 193 for which Gram stain, culture and microscopy were completed and compared (analyses pertain to this sample)
Index test	Dipstick assessed using Super UA automated urine analyser. - Any nitrite alone - Any LE alone - Nitrite(+) and LE(+) - Nitrite(+) or LE(+)
Reference standard (or Gold standard)	Positive culture defined as ≥ 10^4 cfu/ml of a single pathogen from a catheterised sample or 10^3 from suprapubic aspirate sample. Culture prepared by inoculating 0.001 mL of urine from a calibrated loop onto MacConkey agar and Columbia agar with 5% sheep blood, incubated at 35C, and examined at 24 and 48 hours. Urine microscopy was automated using the Yellow IRIS System and measured as WBC/hpf. UA and microscopy were completed using uncentrifuged urine.
Time between testing & treatment	24 and 48 hours Participants were excluded if antibiotics were received within 48 hours of evaluation and urine sample collection.
Length of follow-up	24 and 48 hours
Location	Setting: USA secondary (emergency care) – single centre study
Results	**Any nitrite alone:** True Positive: 7 False Negative: 13 False Positive: 4 True Negative: 169

Bibliographic reference	Dayan P, Bennett J, Best R, et al 2002. Test characteristics of the urine Gram stain in infants≤ 60 days of age with fever. . Pediatric emergency care, 18(1), pp.12-14.
	Sensitivity: 35.0% (14.1–55.9%) Specificity: 97.7% (95.4–99.9%) LR+ = 15.1 LR - = 0.67 **LE alone:** True Positive: 16 False Negative: 4 False Positive: 10 True Negative: 163 Sensitivity: 80.0% (62.5–97.5%) Specificity: 94.2% (90.7–97.7%) **Nitrite(+) and LE(+):** True Positive: 6 False Negative: 14 False Positive: 0 True Negative: 173 Sensitivity: 30.0% (10–50%) Specificity: 100% (98.3–100%) **Nitrite(+) or LE(+):** True Positive: 17 False Negative: 3 False Positive: 14 True Negative: 159 Sensitivity: 85.0% (69.4–100%) Specificity: 91.9% (87.8–96.0%)

Bibliographic reference	Dayan P, Bennett J, Best R, et al 2002. Test characteristics of the urine Gram stain in infants≤ 60 days of age with fever. . Pediatric emergency care, 18(1), pp.12-14.
Source of funding	Not reported.
Comments	QUADAS 2 assessment: www.bristol.ac.uk/social-community-medicine/projects/quadas/quadas-2/ Risk of bias and applicability judgements: **DOMAIN 1: PATIENT SELECTION** A. Risk of Bias • Was a consecutive or random sample of patients enrolled? Yes • Was a case-control design avoided? Yes • Did the study avoid inappropriate exclusions? Yes Could the selection of patients have introduced bias? RISK: LOW B. Is there concern that the included patients do not match the review question? CONCERN: LOW **DOMAIN 2: INDEX TEST(S)** A. Risk of Bias • Were the index test results interpreted without knowledge of the results of the reference standard? Unclear • If a threshold was used, was it pre-specified? not applicable Could the conduct or interpretation of the index test have introduced bias? RISK: UNCLEAR B. Concerns regarding applicability Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: UNCLEAR **DOMAIN 3: REFERENCE TEST** A. Risk of Bias • Is the reference standard likely to correctly classify the target condition? Yes • Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: Unclear B. Concerns regarding applicability Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW

Bibliographic reference	**Dayan P, Bennett J, Best R, et al 2002. Test characteristics of the urine Gram stain in infants≤ 60 days of age with fever. . Pediatric emergency care, 18(1), pp.12-14.**
	DOMAIN 4: FLOW AND TIMING A. Risk of Bias • Was there an appropriate interval between index test(s) and reference standard? Yes • Did all patients receive a reference standard? Yes • Did patients receive the same reference standard? Yes • Were all patients included in the analysis? 11 excluded due to inadequate samples Could the patient flow have introduced bias? RISK: LOW

D.2 Doley and Nelligan 2003

Bibliographic reference	**Doley A, and Nelligan M., 2003. Is a negative dipstick urinalysis good enough to exclude urinary tract infection in paediatric emergency. , 15(1), pp.77-80.**
Study type	Retrospective cross sectional
Aim	To determine if negative dipstick urinalysis is adequate to exclude urinary tract infection in children aged 0-10 years (data extracted only for the 0-2 years sub-sample).
Patient characteristics	Enrolment: Retrospective case note review, conducted between May to December 2000, of paediatric presentation. Notes reviewed at least 3 months after initial presentation Inclusion: - With printed urinalysis record in case notes (n=720 aged 0-10 years) - With full urine culture result (n=375 aged 0-10 years)* - Age 0-2 years (n=160 of above sample) Exclusion: - No urinalysis conducted *Note: likely to have been a high prevalence population Patient characteristics:

Bibliographic reference	Doley A, and Nelligan M., 2003. Is a negative dipstick urinalysis good enough to exclude urinary tract infection in paediatric emergency. , 15(1), pp.77-80.
	Age: 0–2 years (mean age not reported) Gender: not reported Number circumcised (if reported): not reported Symptomatic / asymptomatic: not reported Prevalence of UTI: 10.7% Method of urine collection: bag specimen or clean catch (4 cases via suprapubic aspiration)
Number of patients	N=160 (sub-sample in 0-2 year age group with urinalysis and urine culture results)
Index test	Multistix 10 SG, using analyser: Clinitek 50 urinalysis Negative urinalysis defined as: negative got blood, protein, leukocytes and nitrites Leukocytes graded by machine as either negative, trace, mile, moderate or large
Reference standard (or Gold standard)	Urine culture, > 100,000 cfu/ml Does not specify criteria for performing urine culture but only 52% of patients with printed urinalysis record also had urine culture result)
Time between testing & treatment	Not reported
Length of follow-up	Not reported
Location	Setting: Australia (single centre) – emergency medical department of one hospital
Results	Nitrite or LE or blood or protein positive True Positive: 21 False Negative: 3 False Positive: 82 True Negative: 54 Sensitivity: 87.5% (74.3 – 100) Specificity: 39.7% (31.5 – 47.9) LR+: 1.45 LR: 0.32
Source of funding	Not reported.

Bibliographic reference	Doley A, and Nelligan M., 2003. Is a negative dipstick urinalysis good enough to exclude urinary tract infection in paediatric emergency. , 15(1), pp.77-80.
Comments	- Risk of bias for index test, not fully clear which is included in dipstick… QUADAS 2 assessment: www.bristol.ac.uk/social-community-medicine/projects/quadas/quadas-2/ Risk of bias and applicability judgements: **DOMAIN 1: PATIENT SELECTION** A Risk of Bias • Was a consecutive or random sample of patients enrolled? No • Was a case-control design avoided? Yes • Did the study avoid inappropriate exclusions? Yes Could the selection of patients have introduced bias? RISK: HIGH B Is there concern that the included patients do not match the review question? CONCERN: LOW **DOMAIN 2: INDEX TEST(S)** A. Risk of Bias • Were the index test results interpreted without knowledge of the results of the reference standard? Unclear • If a threshold was used, was it pre-specified? - Threshold of LE (trace, small etc) classified as positive not specified. Could the conduct or interpretation of the index test have introduced bias? RISK: HIGH – LE threshold not specified, clear definition of positive dipstick not provided B. Concerns regarding applicability Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: HIGH **DOMAIN 3: REFERENCE TEST** A. Risk of Bias • Is the reference standard likely to correctly classify the target condition? Yes • Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: UNCLEAR

Bibliographic reference	Doley A, and Nelligan M., 2003. Is a negative dipstick urinalysis good enough to exclude urinary tract infection in paediatric emergency. , 15(1), pp.77-80.
	B. Concerns regarding applicability Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW **DOMAIN 4: FLOW AND TIMING** A. Risk of Bias • Was there an appropriate interval between index test(s) and reference standard? Unclear – retrospective design • Did all patients receive a reference standard? Yes • Did patients receive the same reference standard? Yes • Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: UNCLEAR

D.3 Glissmeyer 2014

Bibliographic reference	Glissmeyer, E., Korgenski, E., Wilkes, J. et al 2014. Dipstick screening for urinary tract infection in febrile infants. Pediatrics, 133(5), pp.e1121-e1127.
Study type	Retrospective cohort
Aim	To compare the performance of urine dipstick alone with urine microscopy and with both tests combined as a screen for urinary tract infection (UTI) in febrile infants aged 1 to 90 days.
Patient characteristics	Enrolment: Retrospectively identified from children's healthcare system database (covers 23 hospitals; provides care for >90% of Utah infants under 1 year). Inclusion: - febrile infants aged 1 to 90 days assessed between 2004 and 2011 - catheterized urine dipstick, microscopic urinalysis, and urine bacterial cultures performed simultaneously* * If multiple urinalysis tests were performed during an encounter, only the first urine specimen was included in analyses.

Bibliographic reference	**Glissmeyer, E., Korgenski, E., Wilkes, J. et al 2014. Dipstick screening for urinary tract infection in febrile infants. Pediatrics, 133(5), pp.e1121-e1127.**
	Exclusion: - urine obtained by a method specified as bag specimen or suprapubic aspirate - equivocal urine cultures (growth of urine pathogens with quantities between 10,000 – 49,999 CFUs per mL) Patient characteristics: Age: 1,745 (27%) aged 1 to 28 days; 4649 (73%) aged 29 to 90 days. Gender: not reported Number circumcised (if reported): not reported Symptomatic / asymptomatic: States that very few, if any, infants were asymptomatic because subjects were identified in the database using a definition for fever and other diagnostic codes Prevalence of UTI: 12% Method of urine collection: urethral catheterisation.
Number of patients	N=6394
Index test	Dipstick (using analyser), dipstick and microscopy Dipstick was considered positive if either leukocyte esterase or nitrite was positive (≥trace). Microscopy was considered positive if under high-power microscopic field (HPF) the technician observed either >10 white blood cells (WBCs) or any bacteria. A positive combined urinalysis was defined as any positive finding for either dipstick or microscopy or both.
Reference standard (or Gold standard)	Positive for UTI was defined as growth of ≥1 urine pathogens, each with a quantity of ≥50 000 colony forming units (CFUs) per mL
Time between testing & treatment	Not stated
Length of follow-up	Not stated
Location	Setting: USA – various secondary care centres around Utah (79% seen at one tertiary referral centre)
Results	Dipstick alone: Nitrite or LE positive:

Bibliographic reference	Glissmeyer, E., Korgenski, E., Wilkes, J. et al 2014. Dipstick screening for urinary tract infection in febrile infants. Pediatrics, 133(5), pp.e1121-e1127.
	True Positive: 699 False Negative: 71 False Positive: 349 True Negative: 5275 Sensitivity: 90.8% (90.4 – 96.2) Specificity: 93.8 (93.5 – 94.1) <u>Dipstick and microscopy:</u> Nitrite or LE positive and > 10 WBCs per HPF True Positive: 729 False Negative: 41 False Positive: 697 True Negative: 4927 Sensitivity: 94.7 (94.4 – 95.0) Specificity: 87.6 (87.2 – 88.0)
Source of funding	Not reported.
Comments	QUADAS 2 assessment: www.bristol.ac.uk/social-community-medicine/projects/quadas/quadas-2/ Risk of bias and applicability judgements: **DOMAIN 1: PATIENT SELECTION** C. Risk of Bias • Was a consecutive or random sample of patients enrolled? Unclear • Was a case-control design avoided? Yes • Did the study avoid inappropriate exclusions? Yes Could the selection of patients have introduced bias? RISK: LOW D. Is there concern that the included patients do not match the review question? CONCERN: LOW

Bibliographic reference	Glissmeyer, E., Korgenski, E., Wilkes, J. et al 2014. Dipstick screening for urinary tract infection in febrile infants. Pediatrics, 133(5), pp.e1121-e1127.
	DOMAIN 2: INDEX TEST(S) A. Risk of Bias • Were the index test results interpreted without knowledge of the results of the reference standard? Unclear • If a threshold was used, was it pre-specified? Yes, LE ≥ trace • Could the conduct or interpretation of the index test have introduced bias? RISK: UNCLEAR B. Concerns regarding applicability Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: UNCLEAR **DOMAIN 3: REFERENCE TEST** A. Risk of Bias • Is the reference standard likely to correctly classify the target condition? Yes • Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: UNCLEAR B. Concerns regarding applicability Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW **DOMAIN 4: FLOW AND TIMING** A. Risk of Bias • Was there an appropriate interval between index test(s) and reference standard? Yes • Did all patients receive a reference standard? Yes • Did patients receive the same reference standard? Yes • Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW

D.4 Hay 2016

Bibliographic reference	**Hay A, Birnie K, Busby J, et al. 2016. The Diagnosis of Urinary Tract infection in Young children (DUTY): a diagnostic prospective observational study to derive and validate a clinical algorithm for the diagnosis of urinary tract infection in children presenting to primary care with an acute illness. Health Technol Assess 2016;20(51)**
Study type	Prospective cohort
Aim	To develop algorithms to accurately identify pre-school children in whom urine should be obtained; assess whether or not dipstick urinalysis provides additional diagnostic information; and model algorithm cost-effectiveness.

Patient characteristics	Enrolment:
	Between April 2010 and April 2012, 516 clinicians from 233 primary care sites enrolled children presenting with an acute illness and/or new urinary symptoms.
	
	Inclusion:
	- aged before their fifth birthday
	- presenting to primary care with a new acute illness episode of ≤ 28 days' duration
	- at least one 'constitutional' symptom or sign identified by NICE as a potential marker for UTI:
	  o fever, vomiting, lethargy/malaise, irritability, poor feeding and failure to thrive, *and/or*
	- at least one urinary symptom identified by NICE as a potential marker of UTI:
	  o abdominal pain, jaundice (children < 3 months only), haematuria, offensive urine, cloudy urine, loin tenderness, frequency, apparent pain on passing urine, changes to continence
	
	Exclusion:
	- Presenting with trauma as a predominant concern
	- Known neurogenic (e.g. spina bifida) or surgically reconstructed bladder or urinary permanent or intermittent catheterisation (for whom different bacterial concentration cut points are used)
	- Taking any antibiotics in the last 7 days
	- Taking immunosuppressant medication (e.g. antirejection drugs, oral or intramuscular steroids or chemotherapy)
	
	Patient characteristics:
	
	Age: up to 3 years: 2884 infants and children

Bibliographic reference	**Hay A, Birnie K, Busby J, et al. 2016. The Diagnosis of Urinary Tract infection in Young children (DUTY): a diagnostic prospective observational study to derive and validate a clinical algorithm for the diagnosis of urinary tract infection in children presenting to primary care with an acute illness. Health Technol Assess 2016;20(51)**
	Gender: 49.9% male Number circumcised (if reported): not reported Symptomatic / asymptomatic: n = 104 with temperature ≥ 39 °Prevalence of UTI (up to 3 years) = 1.7% Method of urine collection: clean catch (preferred) or nappy pad
Number of patients	2884
Index test	Dipstick test
Reference standard (or Gold standard)	Pure (single) or predominant growth of a uropathogen at 100,000 cfu/ml (urine culture carried out in research laboratory).
Time between testing & treatment	Unclear, patient follow-up was conducted 14 days after recruitments and number of children who responded to treatment < 48 hours was noted.
Length of follow-up	Follow-up interview at 14 days, medical review at 3 months.
Location	Setting: UK primary care (multicentre)
Results	Please see Appendix L for the results obtained from the authors and included in the evidence review.
Source of funding	NIHR Health Technology Assessment programme.
Comments	QUADAS 2 assessment: www.bristol.ac.uk/social-community-medicine/projects/quadas/quadas-2/ Risk of bias and applicability judgements: **DOMAIN 1: PATIENT SELECTION** A. Risk of Bias • Was a consecutive or random sample of patients enrolled? Yes • Was a case-control design avoided? Yes • Did the study avoid inappropriate exclusions? Yes Could the selection of patients have introduced bias? RISK: LOW B. Is there concern that the included patients do not match the review question? CONCERN: LOW

Bibliographic reference	Hay A, Birnie K, Busby J, et al. 2016. The Diagnosis of Urinary Tract infection in Young children (DUTY): a diagnostic prospective observational study to derive and validate a clinical algorithm for the diagnosis of urinary tract infection in children presenting to primary care with an acute illness. Health Technol Assess 2016;20(51)
	DOMAIN 2: INDEX TEST(S) A. Risk of Bias • Were the index test results interpreted without knowledge of the results of the reference standard? Yes • If a threshold was used, was it pre-specified? yes Could the conduct or interpretation of the index test have introduced bias? RISK: LOW B. Concerns regarding applicability Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW **DOMAIN 3: REFERENCE TEST** A. Risk of Bias • Is the reference standard likely to correctly classify the target condition? Yes • Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: Unclear B. Concerns regarding applicability Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW **DOMAIN 4: FLOW AND TIMING** A. Risk of Bias • Was there an appropriate interval between index test(s) and reference standard? Unclear • Did all patients receive a reference standard? Yes • Did patients receive the same reference standard? Yes • Were all patients included in the analysis? No Could the patient flow have introduced bias? RISK: LOW

D.5 Kanegaye 2014

Bibliographic reference	**Kanegaye J, Jacob J, and Malicki D. 2014. Automated urinalysis and urine dipstick in the emergency evaluation of young febrile children. Pediatrics, 134(3), pp.523-9.**
Study type	Prospective cohort
Aim	To determine the diagnostic performance of automated cell counts and emergency department point-of-care (POC) dipstick urinalyses in the evaluation of young febrile children.
Patient characteristics	Enrolment: Prospectively identified a convenience sample of febrile paediatric patients attending the emergency department of a tertiary hospital between May 2009 and May 2010. Inclusion: - temperature ≥38°C in the ED or tactile or documented fevers at home within 24 hours - clinical need to evaluate for UTI (no further detail) Exclusion: - incomplete data or urine testing - received systemic antibiotics in the previous 24 hours - immunocompromised or at risk for neutropenia - conditions that predispose to asymptomatic genitourinary bacterial colonization (including neurogenic bladder, chronic or intermittent bladder instrumentation, or surgical diversion of the urinary tract). Patient Characteristics: Age: median 8.1 months (IQR: 3.6-14.3 months) Gender: 142 (41%) male; 202 (59%) Number circumcised: 29% of 129 males Symptomatic / asymptomatic: not reported (all patients were febrile: max mean recorded ED temperature: 38.8°C (SD 1.1)) Prevalence of UTI: 12.3% Method of urine collection: urethral catheterisation
Number of patients	N=342

Bibliographic reference	**Kanegaye J, Jacob J, and Malicki D. 2014. Automated urinalysis and urine dipstick in the emergency evaluation of young febrile children. Pediatrics, 134(3), pp.523-9.**
Index test	Dipstick, using Multistix 10 SG –trained ED nurses visually interpreted reagent strips according to standard colour charts. Urinary nitrite was recorded as positive or negative and leukocyte esterase (LE) as negative, trace, 1+ (small), 2+ (moderate), or 3+ (large). Test strips were then interpreted in laboratories with the Siemens Clinitek 500 Urine Chemistry Analyzer (Bayer Corporation, Elkhart, IN).
Reference standard (or Gold standard)	Culture Positive culture defined as ≥ 50,000 cfu/ml
Time between testing & treatment	Not reported.
Length of follow-up	Not reported.
Location	Setting: USA – paediatric emergency department of tertiary hospital (single centre)
Results	Point of care tests <u>Nitrites +</u> True Positive: 22 False Negative: 20 False Positive: 2 True Negative: 300 Sensitivity: 52% (38 – 67) Specificity: 99% (98 – 99.8) <u>LE ≥ trace</u> True Positive: 38 False Negative: 4 False Positive: 10 True Negative: 290

Bibliographic reference	Kanegaye J, Jacob J, and Malicki D. 2014. Automated urinalysis and urine dipstick in the emergency evaluation of young febrile children. Pediatrics, 134(3), pp.523-9.
	Sensitivity: 91% (78 – 96) Specificity: 97% (94 – 98) <u>LE ≥ trace or nitrite +</u> True Positive: 40 False Negative: 2 False Positive: 11 True Negative: 289 Sensitivity: 95% (84 – 99) Specificity: 96% (94 – 98)
Source of funding	Authors state they had no external funding.
Comments	QUADAS 2 assessment: www.bristol.ac.uk/social-community-medicine/projects/quadas/quadas-2/ Risk of bias and applicability judgements: **DOMAIN 1: PATIENT SELECTION** A. Risk of Bias • Was a consecutive or random sample of patients enrolled? No • Was a case-control design avoided? Yes • Did the study avoid inappropriate exclusions? Yes Could the selection of patients have introduced bias? RISK: High B. Is there concern that the included patients do not match the review question? CONCERN: LOW **DOMAIN 2: INDEX TEST(S)** A. Risk of Bias • Were the index test results interpreted without knowledge of the results of the reference standard? Yes • If a threshold was used, was it pre-specified? yes Could the conduct or interpretation of the index test have introduced bias? RISK: UNCLEAR

Bibliographic reference	Kanegaye J, Jacob J, and Malicki D. 2014. Automated urinalysis and urine dipstick in the emergency evaluation of young febrile children. Pediatrics, 134(3), pp.523-9.
	B. Concerns regarding applicability Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: UNCLEAR **DOMAIN 3: REFERENCE TEST** A. Risk of Bias • Is the reference standard likely to correctly classify the target condition? Yes • Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: UNCLEAR B. Concerns regarding applicability Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW **DOMAIN 4: FLOW AND TIMING** A. Risk of Bias • Was there an appropriate interval between index test(s) and reference standard? Yes • Did all patients receive a reference standard? Yes • Did patients receive the same reference standard? Yes • Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW

D.6 Kazi 2013

Bibliographic reference	Kazi B, Buffone G, Revell Pet al. 2013. Performance characteristics of urinalyses for the diagnosis of pediatric urinary tract infection. The American journal of emergency medicine, 31(9), pp.1405-7.
Study type	Retrospective cross-sectional
Aim	To determine whether point-of-care (POCT) urinalysis is as accurate as laboratory-performed urinalysis in diagnosing urinary tract infections (UTIs) in the paediatric emergency department.
Patient characteristics	Enrolment:

Bibliographic reference	Kazi B, Buffone G, Revell Pet al. 2013. Performance characteristics of urinalyses for the diagnosis of pediatric urinary tract infection. The American journal of emergency medicine, 31(9), pp.1405-7.
	Retrospective review of children evaluated for UTI at one hospital emergency department between July 2008 to December 2012. Inclusion: - both a point of care dipstick urinalysis and urine culture were obtained - aged 6-23 months (reported data for <2 months, 2-5 months and ≥2 years not extracted for this review) Exclusion: - Urine culture specimens collected via a bag, Foley catheter, indwelling stent, or urinary tract fistula Patient characteristics: Age: 6-11 months: n=802; 12-23 months: n=837 (reported data for <2 months and 2-5 months not extracted due to small sample sizes) Gender: not reported Number circumcised (if reported): not reported Symptomatic / asymptomatic: not reported Prevalence of UTI: 7.5% Method of urine collection: urethral catheterisation and void. suprapubic aspirate in 0.02%
Number of patients	N=1,639 (subsample aged 6-23 months)
Index test	Point of care testing: Urisys 1000 Urine Analyzer and Clinitek Status Analyzer Urinalyses were considered positive if leukocyte esterase and/or nitrites were positive (not defined).
Reference standard (or Gold standard)	50,000 cfu/ml of a single organism for specimens collected by voiding / catheter, or grew at least 1000 cfu/ml for suprapubic aspirate specimens. Urine cultures were held for up to 3 days after being plated on blood agar and MacConkey agar plates. Specimens received more than 2 hours (or more than 24 hours refrigerated) after collection were not processed.
Time between testing & treatment	2-24 hours between dipstick testing and culture plating. Does not report proportion of patients taking antibiotics

Bibliographic reference	Kazi B, Buffone G, Revell Pet al. 2013. Performance characteristics of urinalyses for the diagnosis of pediatric urinary tract infection. The American journal of emergency medicine, 31(9), pp.1405-7.
Length of follow-up	Not clear.
Location	Setting: USA (single centre) – one tertiary hospital paediatric emergency department
Results	Point of care tests <u>LE or nitrites</u> 6 – 11 months True Positive: 227 False Negative: 88 False Positive: 19 True Negative: 467 Sensitivity: 72% (67 – 77) Specificity: 96% (94, 98) 12 – 23 months True Positive: 53 False Negative: 11 False Positive: 26 True Negative: 747 Combined 6 to 23 months True Positive: 280 False Negative: 99 False Positive: 45 True Negative: 1214
Source of funding	Not reported.

Bibliographic reference	Kazi B, Buffone G, Revell Pet al. 2013. Performance characteristics of urinalyses for the diagnosis of pediatric urinary tract infection. The American journal of emergency medicine, 31(9), pp.1405-7.
Comments	• This study also reports results for lab performed urinalysis by age. However, positive was defined as LE, nitrite or microscopy positive and was therefore not included in the analysis (positive result could be attributed to microscopy positive alone). • 2 by 2 table was calculated using a code run in R software. This is more effective at calculating the 2 by 2 table in larger samples and therefore data for < 2 months (with 39 participants for point of care testing) was not included. QUADAS 2 assessment: www.bristol.ac.uk/social-community-medicine/projects/quadas/quadas-2/ Risk of bias and applicability judgements: **DOMAIN 1: PATIENT SELECTION** C. Risk of Bias • Was a consecutive or random sample of patients enrolled? No • Was a case-control design avoided? Yes • Did the study avoid inappropriate exclusions? Yes Could the selection of patients have introduced bias? RISK: High D. Is there concern that the included patients do not match the review question? CONCERN: LOW **DOMAIN 2: INDEX TEST(S)** A. Risk of Bias • Were the index test results interpreted without knowledge of the results of the reference standard? Yes • If a threshold was used, was it pre-specified? yes Could the conduct or interpretation of the index test have introduced bias? RISK: UNCLEAR B. Concerns regarding applicability Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: UNCLEAR **DOMAIN 3: REFERENCE TEST** A. Risk of Bias • Is the reference standard likely to correctly classify the target condition? Yes • Were the reference standard results interpreted without knowledge of the results of the index test? Unclear

Bibliographic reference	Kazi B, Buffone G, Revell Pet al. 2013. Performance characteristics of urinalyses for the diagnosis of pediatric urinary tract infection. The American journal of emergency medicine, 31(9), pp.1405-7.
	Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: UNCLEAR B. Concerns regarding applicability Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW **DOMAIN 4: FLOW AND TIMING** A. Risk of Bias • Was there an appropriate interval between index test(s) and reference standard? Yes • Did all patients receive a reference standard? Yes • Did patients receive the same reference standard? Yes • Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW

D.7 Lejeune 1991

Bibliographic reference	Lejeune B, Baron R, Guillois B et al. 1991. Evaluation of a screening test for detecting urinary tract infection in newborns and infants.. Journal of Clinical Pathology, 44(12), pp.1029-30.
Study type	Prospective cohort
Aim	To identify the dipstick test which gave the highest diagnostic accuracy in routine practice for infants under 18 months.
Patient characteristics	Enrolment: Consecutive urine samples of neonates and infants < 18 months. Inclusion: - Not reported. Exclusion: - Not reported. Patient characteristics: Age: under 18 months: 85 (35%) less than 1 month; 81 (33%) aged 1-6 months; 77 (32%) >6 months-18 months Gender: not reported.

Bibliographic reference	**Lejeune B, Baron R, Guillois B et al. 1991. Evaluation of a screening test for detecting urinary tract infection in newborns and infants.. Journal of Clinical Pathology, 44(12), pp.1029-30.**
	Number circumcised (if reported): not reported. Symptomatic / asymptomatic: not reported. Prevalence of UTI: 15.2% Method of urine collection: not reported.
Number of patients	N=243
Index test	Urine reagent strips for nitrate, leucocyte esterase and protein (Multistick 8 SG AMES) read by the Clinitek System photometer (AMES). Does not report criteria for determining a positive test for LE or N
Reference standard (or Gold standard)	Diagnosis of UTI based on a combination of: • Microscopy: WBC > 25 x 10^9/L for boys or 50 x 10^9/L for girls. • Culture: 100,000 (10^5) cfu/ml
Time between testing & treatment	Not reported
Length of follow-up	Not reported
Location	Setting: France (single centre) – secondary care
Results	LE + True Positive: 33 False Negative: 4 False Positive: 45 True Negative: 161 Sensitivity: 89.2% Specificity: 78.2% Nitrite + True Positive: 6 False Negative: 31

Bibliographic reference	Lejeune B, Baron R, Guillois B et al. 1991. Evaluation of a screening test for detecting urinary tract infection in newborns and infants.. Journal of Clinical Pathology, 44(12), pp.1029-30.
	False Positive: 5 True Negative: 201 Sensitivity: 16.2% Specificity: 97.6% LE and Nitrite + True Positive: 33 False Negative: 5 False Positive: 4 True Negative: 201 Sensitivity: 87% (72, 96) Specificity: 97.6% Protein + True Positive: 3 False Negative: 34 False Positive: 10 True Negative: 196 Sensitivity: 8.11% Specificity: 95.1% LE and protein True Positive: 33 False Negative: 4 False Positive: 10 True Negative: 196 Sensitivity: 89.2%

Bibliographic reference	Lejeune B, Baron R, Guillois B et al. 1991. Evaluation of a screening test for detecting urinary tract infection in newborns and infants.. Journal of Clinical Pathology, 44(12), pp.1029-30.
	Specificity: 95.1% LR+: 17.4 LR-: 0.12 LE and protein and nitrite True Positive: 33 False Negative: 4 False Positive: 58 True Negative: 148 Sensitivity: 89.2% Specificity: 95.1% LR+: 3.1 LR-: 0.17
Source of funding	Not reported.
Comments	QUADAS 2 assessment: www.bristol.ac.uk/social-community-medicine/projects/quadas/quadas-2/ Risk of bias and applicability judgements: **DOMAIN 1: PATIENT SELECTION** A. Risk of Bias • Was a consecutive or random sample of patients enrolled? Yes • Was a case-control design avoided? Yes • Did the study avoid inappropriate exclusions? Unclear, not reported Could the selection of patients have introduced bias? RISK: Low B. Is there concern that the included patients do not match the review question? CONCERN: LOW **DOMAIN 2: INDEX TEST(S)** A. Risk of Bias • Were the index test results interpreted without knowledge of the results of the reference standard? No, same investigator • If a threshold was used, was it pre-specified? Threshold for LE not specified

Bibliographic reference	Lejeune B, Baron R, Guillois B et al. 1991. Evaluation of a screening test for detecting urinary tract infection in newborns and infants.. Journal of Clinical Pathology, 44(12), pp.1029-30.
	Could the conduct or interpretation of the index test have introduced bias? RISK: HIGH B. Concerns regarding applicability Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: HIGH – method of urine collection also not reported. Cannot deduce if the method used introduces contamination or a more accurate urine sampling method. **DOMAIN 3: REFERENCE TEST** A. Risk of Bias • Is the reference standard likely to correctly classify the target condition? Yes • Were the reference standard results interpreted without knowledge of the results of the index test? No Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: HIGH B. Concerns regarding applicability Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW **DOMAIN 4: FLOW AND TIMING** A. Risk of Bias • Was there an appropriate interval between index test(s) and reference standard? Unclear • Did all patients receive a reference standard? Yes • Did patients receive the same reference standard? Yes • Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW

D.8 McGillivray 2005

Bibliographic reference	McGillivray D, Mok E, Mulrooney E et al. 2005. A head-to-head comparison: "clean-void" bag versus catheter urinalysis in the diagnosis of urinary tract infection in young children. The Journal of pediatrics, 147(4), pp.451-6.
Study type	Prospective cross-sectional
Aim	To compare the validity of dipstick and microscopic urinalysis on clean-voided bag versus catheter urine specimens from the same child using the catheter culture as the ‘‘gold’’ standard.

Bibliographic reference	**McGillivray D, Mok E, Mulrooney E et al. 2005. A head-to-head comparison: "clean-void" bag versus catheter urinalysis in the diagnosis of urinary tract infection in young children. The Journal of pediatrics, 147(4), pp.451-6.**
Patient characteristics	Enrolment: Prospective enrolment of children attending a hospital emergency department between June 15, 2000 and December 31, 2001. Inclusion: - Non-toilet trained, aged <3 years - At risk of UTI based on following criteria: o fever without source plus male sex <6 months or female sex <12 months; uncircumcised boys of any age; past history of UTI or abnormal renal anatomy; fever >39°C or any fever ≥48 hours duration, *or* o without fever but who were either ill-appearing without identifiable focus of infection or infants age <3 months, exhibited signs or symptoms of UTI (eg, dysuria, foul-smelling urine, change in urine color), or had unexplained abdominal pain Exclusion: - Children needing urgent medical intervention e.g. immediate administration of antibiotics or resuscitation - Children already receiving antibiotics Patient characteristics: Age: 54 (18%) < 90 days old; 249 (82%) aged 3 months-3 years. Gender: 102 (33.6%) male; 201 (66.3%) female Number circumcised: circumcision status in 69 of 102 boys, of whom 14/69 were circumcised Symptomatic / asymptomatic: rectal equivalent temperature obtained in 297 (98%) of children, of whom 17.5% (53/297) had a temperature >39.5°C. Prevalence of UTI: 26% Method of urine collection: urethral catheter (urine also collected first in sterile bags for each child but these samples were not cultured). Note: catheter samples were obtained only from children with specific clinical indications (following bag sample collection and urinalysis). The physician who ordered the catheter specimens was not blinded to the results of the bag urinalysis. This was therefore a high prevalence population.
Number of patients	N=303

Bibliographic reference	**McGillivray D, Mok E, Mulrooney E et al. 2005. A head-to-head comparison: "clean-void" bag versus catheter urinalysis in the diagnosis of urinary tract infection in young children. The Journal of pediatrics, 147(4), pp.451-6.**
Index test	Dipstick: multistix 10 SG using automated machine (Clinitek 100/200) Positive dipstick defined as presence of > trace LE OR nitrite positive Microscopy positive: > 5 WBC/HPF
Reference standard (or Gold standard)	Culture (only samples obtained via catheter sent for culture). Positive is > 10,000 cfu/ml of a single organism.
Time between testing & treatment	Not reported.
Length of follow-up	Not reported.
Location	Setting: Canada – paediatric emergency department (single centre)
Results	<u>LE positive (> trace) or nitrite positive:</u> <u>All age groups, 0 – 3 years (n = 303)</u> True Positive: 58 False Negative: 24 False Positive: 7 True Negative: 214 Sensitivity: 71% (61%- 81%) Specificity: 97% (95% - 99%) <u>≤ 90 days (n = 54)</u> True Positive: 6 False Negative: 7 False Positive: 0 True Negative: 41

Bibliographic reference	**McGillivray D, Mok E, Mulrooney E et al. 2005. A head-to-head comparison: "clean-void" bag versus catheter urinalysis in the diagnosis of urinary tract infection in young children. The Journal of pediatrics, 147(4), pp.451-6.**
	Sensitivity: 46% (19% - 73%) Specificity: 100% (93% - 100%) <u>> 90 days (n = 249)</u> True Positive: 52 False Negative: 17 False Positive: 5 True Negative: 175 Sensitivity: 75% (65% - 86%) Specificity: 97% (94% - 99%)
Source of funding	Supported in part by a grant from the Canadian Association of Emergency Physicians.
Comments	Information in the study was available to calculate 2x2 table for dipstick testing and reference test of urine culture using catheterisation. Information was not available to calculate 2x2 table for dipstick and microscopy combined. The article states “no bag urine specimens were sent for culture” and no information was available to calculate 2x2 table, therefore accuracy using bag samples was not calculable. QUADAS 2 assessment: www.bristol.ac.uk/social-community-medicine/projects/quadas/quadas-2/ Risk of bias and applicability judgements: **DOMAIN 1: PATIENT SELECTION** C. Risk of Bias • Was a consecutive or random sample of patients enrolled? Unclear • Was a case-control design avoided? Yes • Did the study avoid inappropriate exclusions? Yes Could the selection of patients have introduced bias? RISK: UNCLEAR D. Is there concern that the included patients do not match the review question? CONCERN: LOW **DOMAIN 2: INDEX TEST(S)**

Bibliographic reference	McGillivray D, Mok E, Mulrooney E et al. 2005. A head-to-head comparison: "clean-void" bag versus catheter urinalysis in the diagnosis of urinary tract infection in young children. The Journal of pediatrics, 147(4), pp.451-6.
	A. Risk of Bias • Were the index test results interpreted without knowledge of the results of the reference standard? Unclear • If a threshold was used, was it pre-specified? yes Could the conduct or interpretation of the index test have introduced bias? RISK: UNCLEAR B. Concerns regarding applicability Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: UNCLEAR **DOMAIN 3: REFERENCE TEST** A. Risk of Bias • Is the reference standard likely to correctly classify the target condition? Yes • Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW /HIGH/UNCLEAR B. Concerns regarding applicability Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW **DOMAIN 4: FLOW AND TIMING** A. Risk of Bias • Was there an appropriate interval between index test(s) and reference standard? Yes • Did all patients receive a reference standard? Yes • Did patients receive the same reference standard? Yes • Were all patients included in the analysis? Yes • Could the patient flow have introduced bias? RISK: LOW

D.9 Reardon 2009

Bibliographic reference	**Reardon J, Carstairs K, Rudinsky S, et al. 2009. Urinalysis is not reliable to detect a urinary tract infection in febrile infants presenting to the ED. The American journal of emergency medicine, 27(8), pp.930-2.**
Study type	Retrospective cross-sectional (data from high prevalence population)
Aim	To compare urinalysis with urine culture in the emergency department evaluation of febrile infants.
Patient characteristics	Enrolment: A febrile infant registry was instituted at a tertiary care hospital emergency department from December 2002 to December 2003. Treatment records were reviewed for results of urinalysis and urine culture. Inclusion: younger than 3 months with home or ED temperature of at least 100.4°F, or aged 3 to 24 months with home or ED temperature of at least 102.2°F had urinalysis and urine culture test results in medical record Exclusion: Not reported Patient characteristics (data relate to N=985 entered onto febrile registry and not analysis subset with data for both tests): Age: mean 12.6 months (median 12 months) Gender: 542 (55%) male; 443 (45%) female Number circumcised (if reported): not reported Symptomatic / asymptomatic: symptomatic (mean temperature: 102.1°F) Prevalence of UTI: 10.3% Method of urine collection: sterile catheterised UA obtained for all females, all males younger than 6 months and on uncircumcised males < 12 months. Criteria for determining whether urine culture was performed are not stated.
Number of patients	N=435 with both test results
Index test	Urinalysis – combination of tests. Analyser or visual – not reported Dipstick test - LE positive OR nitrite positive AND microscopy (≥5 wbc/hpf)
Reference standard (or Gold standard)	Urine culture, positive if at least 10,000 cfu/ml.

Bibliographic reference	**Reardon J, Carstairs K, Rudinsky S, et al. 2009. Urinalysis is not reliable to detect a urinary tract infection in febrile infants presenting to the ED. The American journal of emergency medicine, 27(8), pp.930-2.**
Time between testing & treatment	Not reported.
Length of follow-up	Not reported.
Location	Setting: USA (single centre) - tertiary care hospital emergency department
Results	Nitrite or LE positive with microscopy positive: True Positive: 29 False Negative: 16 False Positive: 34 True Negative: 356 Sensitivity: 64% (49% – 78%) Specificity: 91% (88% - 94%)
Source of funding	The Chief, Bureau of Medicine and Surgery, Navy Department, Washington, DC, Clinical Investigations Program, sponsored this report #S-05-075 as required by NSHBETHINST 6000.41B
Comments	QUADAS 2 assessment: www.bristol.ac.uk/social-community-medicine/projects/quadas/quadas-2/ Risk of bias and applicability judgements: **DOMAIN 1: PATIENT SELECTION** E. Risk of Bias • Was a consecutive or random sample of patients enrolled? Unclear • Was a case-control design avoided? Yes • Did the study avoid inappropriate exclusions? Unclear, not reported Could the selection of patients have introduced bias? RISK: UNCLEAR F. Is there concern that the included patients do not match the review question? CONCERN: LOW **DOMAIN 2: INDEX TEST(S)** A. Risk of Bias • Were the index test results interpreted without knowledge of the results of the reference standard? Unclear • If a threshold was used, was it pre-specified? yes

Bibliographic reference	**Reardon J, Carstairs K, Rudinsky S, et al. 2009. Urinalysis is not reliable to detect a urinary tract infection in febrile infants presenting to the ED. The American journal of emergency medicine, 27(8), pp.930-2.**
	Could the conduct or interpretation of the index test have introduced bias? RISK: LOW B. Concerns regarding applicability Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW **DOMAIN 3: REFERENCE TEST** A. Risk of Bias • Is the reference standard likely to correctly classify the target condition? Yes • Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW B. Concerns regarding applicability Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW **DOMAIN 4: FLOW AND TIMING** A. Risk of Bias • Was there an appropriate interval between index test(s) and reference standard? Unclear • Did all patients receive a reference standard? No, 435 / 495 who had urinalysis also had culture • Did patients receive the same reference standard? Yes • Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: UNCLEAR

D.10 Sharief 1998

Bibliographic reference	**Sharief N, Hameed M, and Petts D. 1998. Use of rapid dipstick tests to exclude urinary tract infection in children.. British Journal of Biomedical Science, 55(4), pp.242-6.**
Study type	Prospective cohort
Aim	To evaluate the use of rapid dipstick tests in screening paediatric patients (0-16 years) for the absence of UTI, and to examine whether they reduce the workload of the laboratory.

Bibliographic reference	**Sharief N, Hameed M, and Petts D. 1998. Use of rapid dipstick tests to exclude urinary tract infection in children.. British Journal of Biomedical Science, 55(4), pp.242-6.**
	NB - Only data for subsample aged <1 year are extracted
Patient characteristics	Enrolment: Urine was examined in unselected patients admitted to paediatric ward of one district general hospital with fever. Inclusion: - Aged 0-16 years with fever (criteria for defining fever not specified)* Exclusion: - Receiving antibiotics at time of urine sample collection * only data from sub-sample who were aged <1 year were extracted for analysis. Patient characteristics (based on total sample of N=325 patients aged 0–16 years) Age: 0–1 years: 124 (38%); >1-16 years: 201 (62%) – older subsample not included in analyses. Gender: 194/325 (60%) males; 131/325 (40%) females Number circumcised (if reported): not reported. Symptomatic / asymptomatic: symptomatic - fever (not defined) was an inclusion criterion Prevalence of UTI: 4.8% Method of urine collection: either clean catch or sterile paediatric collection bag (proportions not reported for infants <1 year).
Number of patients	N=124 infants <1 year
Index test	Multistix 8 SG, read on analyser: Clinitex 10 Tested immediately for presence of albumin, blood, nitrate and LE. LE read as either positive or negative
Reference standard (or Gold standard)	UTI defined as ≥ 100,000 cfu/ml and pyuria (pyuria defined as ≥20 WBC/mm$^{3)}$ Pure growth of ≥ 100,000 organisms without pyuria was taken as negative. Culture performed on all samples following dipstick test – laboratory staff were blind to results of dipstick test.
Time between testing & treatment	Not reported

Bibliographic reference	**Sharief N, Hameed M, and Petts D. 1998. Use of rapid dipstick tests to exclude urinary tract infection in children.. British Journal of Biomedical Science, 55(4), pp.242-6.**
Length of follow-up	Not reported
Location	Setting: UK (single centre) secondary care
Results	LE+ True Positive: 6 False Negative: 2 False Positive: 30 True Negative: 86 Sensitivity: 75% Specificity: 74% Nitrite+ True Positive: 1 False Negative: 7 False Positive: 2 True Negative: 114 Sensitivity: 12.5% Specificity: 98% LE+ or nitrite+ True Positive: 6 False Negative: 2 False Positive: 31 True Negative: 85 Sensitivity: 75% Specificity: 73%

Bibliographic reference	**Sharief N, Hameed M, and Petts D. 1998. Use of rapid dipstick tests to exclude urinary tract infection in children.. British Journal of Biomedical Science, 55(4), pp.242-6.**
	LE+ and nitrite+ True Positive: 1 False Negative: 7 False Positive: 1 True Negative: 115 Sensitivity: 12.5% Specificity: 99.1%
Source of funding	Not reported
Comments	- Risk of bias for index test, not fully clear which is included in dipstick… QUADAS 2 assessment: www.bristol.ac.uk/social-community-medicine/projects/quadas/quadas-2/ Risk of bias and applicability judgements: **DOMAIN 1: PATIENT SELECTION** G. Risk of Bias • Was a consecutive or random sample of patients enrolled? Yes, consecutive • Was a case-control design avoided? Yes • Did the study avoid inappropriate exclusions? Yes, no exclusion criteria applied Could the selection of patients have introduced bias? RISK: LOW H. Is there concern that the included patients do not match the review question? CONCERN: LOW **DOMAIN 2: INDEX TEST(S)** A. Risk of Bias • Were the index test results interpreted without knowledge of the results of the reference standard? Unclear • If a threshold was used, was it pre-specified? No – LE recorded as either positive or negative Could the conduct or interpretation of the index test have introduced bias? RISK: UNCLEAR B. Concerns regarding applicability

Bibliographic reference	**Sharief N, Hameed M, and Petts D. 1998. Use of rapid dipstick tests to exclude urinary tract infection in children.. British Journal of Biomedical Science, 55(4), pp.242-6.**
	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: UNCLEAR **DOMAIN 3: REFERENCE TEST** A. Risk of Bias • Is the reference standard likely to correctly classify the target condition? Yes • Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: UNCLEAR B. Concerns regarding applicability Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW **DOMAIN 4: FLOW AND TIMING** A. Risk of Bias • Was there an appropriate interval between index test(s) and reference standard? Yes • Did all patients receive a reference standard? Yes • Did patients receive the same reference standard? Yes • Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW

D.11 Shaw 1991

Bibliographic reference	**Shaw K, Hexter D, McGowan K, et al. 1991. Clinical evaluation of a rapid screening test for urinary tract infections in children.". The Journal of pediatrics, 118(5), pp.733-736.**
Study type	Prospective cohort
Aim	To compare the diagnostic performance of the LE-nitrate urine dipstick with microscopy and quantitative urine culture in a paediatric emergency department and provide guidelines for its use.
Patient characteristics	Enrolment: All children examined during an 8-month period in the emergency department of one Children's Hospital who had a urine specimen collection for culture.

Bibliographic reference	**Shaw K, Hexter D, McGowan K, et al. 1991. Clinical evaluation of a rapid screening test for urinary tract infections in children.". The Journal of pediatrics, 118(5), pp.733-736.**
	Inclusion: - Age <2 years (subsample extracted for analyses) - Had a urine specimen collected for culture Exclusion: - None specified Patient characteristics: (data correspond to subsample aged <2 years only, n=145) Age: < 2 years (mean age not reported) Gender: not reported Number circumcised (if reported): not reported Symptomatic / asymptomatic: 144/145 (79%) had urine cultured as part of evaluation of fever or sepsis Prevalence of UTI: 9.6% Method of urine collection: 128 (88%) by urethral catheter; remainder unspecified (study allowed urine bag / MSU / clean catch methods)
Number of patients	N=145 (subsample aged < 2 yrs)
Index test	Dipstick: multistix 10 SG. Visual reading LE measurement read after 2 minutes and recorded as trace, small (+1), moderate (+2) or large (+3) Nitrate measurement read at 60 seconds and recorded as negative or positive.
Reference standard (or Gold standard)	Culture, catheter: 1000 cfu/ml, clean catch: 100,000 cfu/ml Urine received in microbiology laboratory in sterile containers was inoculated onto blood and MacConkey agar plates with 0.01mL calibrated loop, incubated at 35°C and examined daily for growth for 2 days. Unclear if assessor was blind to results of dipstick test
Time between testing & treatment	Not reported
Length of follow-up	2 days

Bibliographic reference	**Shaw K, Hexter D, McGowan K, et al. 1991. Clinical evaluation of a rapid screening test for urinary tract infections in children.". The Journal of pediatrics, 118(5), pp.733-736.**
Location	Setting: USA (single centre) – children's hospital emergency department.
Results	≥ trace LE or nitrite positive True Positive: 10 False Negative: 4 False Positive: 10 True Negative: 121 Sensitivity: 71% Specificity: 92% ≥ small LE (1+) and nitrite positive True Positive: 2 False Negative: 12 False Positive: 3 True Negative: 128 Sensitivity: 14% Specificity: 98%
Source of funding	Not reported
Comments	QUADAS 2 assessment: www.bristol.ac.uk/social-community-medicine/projects/quadas/quadas-2/ Risk of bias and applicability judgements: **DOMAIN 1: PATIENT SELECTION** A. Risk of Bias • Was a consecutive or random sample of patients enrolled? Yes, all meeting criteria during 8 months enrolled • Was a case-control design avoided? Yes • Did the study avoid inappropriate exclusions? Unclear

Bibliographic reference	Shaw K, Hexter D, McGowan K, et al. 1991. Clinical evaluation of a rapid screening test for urinary tract infections in children.". The Journal of pediatrics, 118(5), pp.733-736.
	Could the selection of patients have introduced bias? RISK: LOW B. Is there concern that the included patients do not match the review question? CONCERN: LOW **DOMAIN 2: INDEX TEST(S)** A. Risk of Bias • Were the index test results interpreted without knowledge of the results of the reference standard? Unclear • If a threshold was used, was it pre-specified? Yes LE threshold specified Could the conduct or interpretation of the index test have introduced bias? RISK: Unclear B. Concerns regarding applicability Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: UNCLEAR **DOMAIN 3: REFERENCE TEST** A. Risk of Bias • Is the reference standard likely to correctly classify the target condition? Yes • Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: UNCLEAR B. Concerns regarding applicability Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW **DOMAIN 4: FLOW AND TIMING** A. Risk of Bias • Was there an appropriate interval between index test(s) and reference standard? Yes • Did all patients receive a reference standard? Yes • Did patients receive the same reference standard? Yes • Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW

D.12 Shaw 1998

Bibliographic reference	**Shaw K, McGowan K, Gorelick M et al. 1998. Screening for urinary tract infection in infants in the emergency department: which test is best?. Pediatrics, 101(6), pp.E1-E1.**
Study type	Prospective cross-sectional
Aim	Comparison of rapid tests and screening strategies for detecting urinary tract infection (UTI) in infants.
Patient characteristics	Enrolment: Prospective enrolment of infants attending the emergency department of an urban children's hospital between December 1994 and February 1996. Inclusion: Boys <1 year of age or girls <2 years with fever (≥38.3°C) and no definite cause, or with UTI symptoms (not otherwise defined). Exclusion: Not reported. Patient characteristics n=3873 patients who had urine samples cultured (reference standard) Age: mean age: 9.2 months (SD 5.7) Gender: 1,510 (39%) male; 2363 (61%) female Number circumcised (if reported): not reported Symptomatic / asymptomatic: symptomatic. Mean temperature: 39.2°C (SD 2.3) Prevalence of UTI: 2.7%Method of urine collection: urethral catheter (99%); MSU in sterile container (1%)
Number of patients	N = 3394 with urine culture and dipstick test result
Index test	Dipstick: multistix 10 SG, interpreted visually. Performed immediately on fresh urine by technologists in haematology laboratory. LE measurement read after 2 minutes and recorded as trace, small (+1), moderate (+2) or large (+3) Nitrate measurement read at 60 seconds and recorded as negative or positive. Microscopic UA performed on all dipstick tests with any positive finding.

Bibliographic reference	**Shaw K, McGowan K, Gorelick M et al. 1998. Screening for urinary tract infection in infants in the emergency department: which test is best?. Pediatrics, 101(6), pp.E1-E1.**
Reference standard (or Gold standard)	Culture: 10000 cfu/ml Urine for culture was refrigerated if not plated within 10 minutes of receipt from sterile container was inoculated onto blood and MacConkey agar plates with 0.01mL calibrated loop, incubated at 35°C and examined daily for growth for 2 days. Performed in hospital microbiology lab. Unclear if assessor was blind to results of dipstick test.
Time between testing & treatment	Not reported.
Length of follow-up	2 days
Location	Setting: USA (single centre) – emergency department of one urban children's hospital.
Results	≥ trace LE or nitrite positive True Positive: 75 False Negative: 20 False Positive: 99 True Negative: 3200 Sensitivity: 79% (69 – 86) Specificity: 97% (97 – 98)
Source of funding	Supported by the Maternal and Child Health Bureau (Title V, Social Security Act), Health Resource and Services Administration, Department of Health and Human Services.
Comments	- Microscopy was only performed if dipstick was positive. Therefore, dipstick + microscopy index test not included. QUADAS 2 assessment: www.bristol.ac.uk/social-community-medicine/projects/quadas/quadas-2/ Risk of bias and applicability judgements: **DOMAIN 1: PATIENT SELECTION** A. Risk of Bias

Bibliographic reference	Shaw K, McGowan K, Gorelick M et al. 1998. Screening for urinary tract infection in infants in the emergency department: which test is best?. Pediatrics, 101(6), pp.E1-E1.

- Was a consecutive or random sample of patients enrolled? Unclear
- Was a case-control design avoided? Yes
- Did the study avoid inappropriate exclusions? Unclear

Could the selection of patients have introduced bias? RISK: High

B. Is there concern that the included patients do not match the review question? CONCERN: LOW

DOMAIN 2: INDEX TEST(S)

A. Risk of Bias

- Were the index test results interpreted without knowledge of the results of the reference standard? Unclear
- If a threshold was used, was it pre-specified? Yes for LE

Could the conduct or interpretation of the index test have introduced bias? RISK: UNCLEAR

B. Concerns regarding applicability

Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: UNCLEAR

DOMAIN 3: REFERENCE TEST

A. Risk of Bias

- Is the reference standard likely to correctly classify the target condition? Yes
- Were the reference standard results interpreted without knowledge of the results of the index test? Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: UNCLEAR

B. Concerns regarding applicability

Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW

DOMAIN 4: FLOW AND TIMING

A. Risk of Bias

- Was there an appropriate interval between index test(s) and reference standard? Yes
- Did all patients receive a reference standard? Yes
- Did patients receive the same reference standard? Yes
- Were all patients included in the analysis? Yes

Bibliographic reference	Shaw K, McGowan K, Gorelick M et al. 1998. Screening for urinary tract infection in infants in the emergency department: which test is best?. Pediatrics, 101(6), pp.E1-E1.
	Could the patient flow have introduced bias? RISK: LOW

D.13 Velasco 2015

Bibliographic reference	Velasco R, Benito H, Mozun R et al. 2015. Using a urine dipstick to identify a positive urine culture in young febrile infants is as effective as in older patients. Acta Paediatrica, and International Journal of Paediatrics, 104(1), pp.e39-e44.
Study type	Prospective cohort
Aim	To determine whether urine dipsticks would identify positive urine cultures in febrile infants of less than 90 days of age.
Patient characteristics	Enrolment: Patients admitted via participating hospital paediatric emergency departments between October 2011 and September 2013. Study is a sub-analysis of one designed to determine risk of invasive bacterial infection in febrile infants with altered urinalysis according to their general appearance, age and laboratory tests. Blood and urine samples were obtained from all infants <90 days who had fever without source (FWS). Inclusion: - Age <90 days - FWS defined as axillary or rectal temperature ≥ 38°C (100.4°F) either at home or emergency department, without catarrhal or other respiratory signs/symptoms or diarrhoea Exclusion: - No collection of urine or blood culture by sterile method - No determination of white blood cell count or C-reactive protein values

Bibliographic reference	**Velasco R, Benito H, Mozun R et al. 2015. Using a urine dipstick to identify a positive urine culture in young febrile infants is as effective as in older patients. Acta Paediatrica, and International Journal of Paediatrics, 104(1), pp.e39-e44.**
	- Patients in whom medical history or physical exam suggested the source of the fever Patient Characteristics (N=3,401 patients meeting inclusion/exclusion criteria) Age: mean (days): 46.6 (SD 23.6) Gender: 2,029 (59.7%) male; 1,372 (40.3%) Number circumcised (if reported): Nor reported Symptomatic / asymptomatic: 100% symptomatic, 'fever without source' - maximum home temperature (mean): 38.4oC (SD: 0.49) Prevalence of UTI: 19.1% Method of urine collection: urethral catheter or suprapubic aspiration (does not report proportion by each method)
Number of patients	3401, of which 649 had a positive urine culture
Index test	Dipstick: combur-test strips, visual reading by trained nurses in emergency department. LE positive if > 1+
Reference standard (or Gold standard)	≥50 000cfu/mL of a single pathogen in a urine sample
Time between testing & treatment	Not reported.
Length of follow-up	Not reported.
Location	Setting: Spain (multi-centre) – 19 hospital paediatric emergency departments
Results	LE+ True Positive: 437 False Negative: 96 False Positive: 13 True Negative: 196 Sensitivity: 82.1% (79 – 85) Specificity: 92.4% (91.4 – 93.4)

Bibliographic reference	Velasco R, Benito H, Mozun R et al. 2015. Using a urine dipstick to identify a positive urine culture in young febrile infants is as effective as in older patients. Acta Paediatrica, and International Journal of Paediatrics, 104(1), pp.e39-e44.
	Nitrite positive True Positive: 89 False Negative: 152 False Positive: 0 True Negative: 29 Sensitivity: 37.1% (33.4 – 41) Specificity: 98.9% (98.5– 99.3) LE or Nitrite positive True Positive: 456 False Negative: 88 False Positive: 18 True Negative: 204 Sensitivity: 3.8 (80.8 – 86.6) Specificity: 91.9 (90.9-92.9) LE and Nitrite positive True Positive: 6 False Negative: 10 False Positive: 1 True Negative: 229 Sensitivity: 35.4% (31.8 – 39.3) Specificity: 99.4 (99.1 – 99.7)

Bibliographic reference	**Velasco R, Benito H, Mozun R et al. 2015. Using a urine dipstick to identify a positive urine culture in young febrile infants is as effective as in older patients. Acta Paediatrica, and International Journal of Paediatrics, 104(1), pp.e39-e44.**
Source of funding	Not reported
Comments	QUADAS 2 assessment: www.bristol.ac.uk/social-community-medicine/projects/quadas/quadas-2/ Risk of bias and applicability judgements: **DOMAIN 1: PATIENT SELECTION** C. Risk of Bias • Was a consecutive or random sample of patients enrolled? Yes, consecutive • Was a case-control design avoided? Yes • Did the study avoid inappropriate exclusions? Yes Could the selection of patients have introduced bias? RISK: LOW Is there concern that the included patients do not match the review question? CONCERN: LOW **DOMAIN 2: INDEX TEST(S)** A. Risk of Bias • Were the index test results interpreted without knowledge of the results of the reference standard? Unclear • If a threshold was used, was it pre-specified? Yes Could the conduct or interpretation of the index test have introduced bias? RISK: UNCLEAR B. Concerns regarding applicability Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: UNCLEAR **DOMAIN 3: REFERENCE TEST** A. Risk of Bias • Is the reference standard likely to correctly classify the target condition? Yes • Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: UNCLEAR B. Concerns regarding applicability

Bibliographic reference	Velasco R, Benito H, Mozun R et al. 2015. Using a urine dipstick to identify a positive urine culture in young febrile infants is as effective as in older patients. Acta Paediatrica, and International Journal of Paediatrics, 104(1), pp.e39-e44.
	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW **DOMAIN 4: FLOW AND TIMING** A. Risk of Bias • Was there an appropriate interval between index test(s) and reference standard? Yes • Did all patients receive a reference standard? Yes • Did patients receive the same reference standard? Yes • Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW

Appendix E: Forest plots

E.1 Dipstick versus culture

Figure 2: Nitrite positive

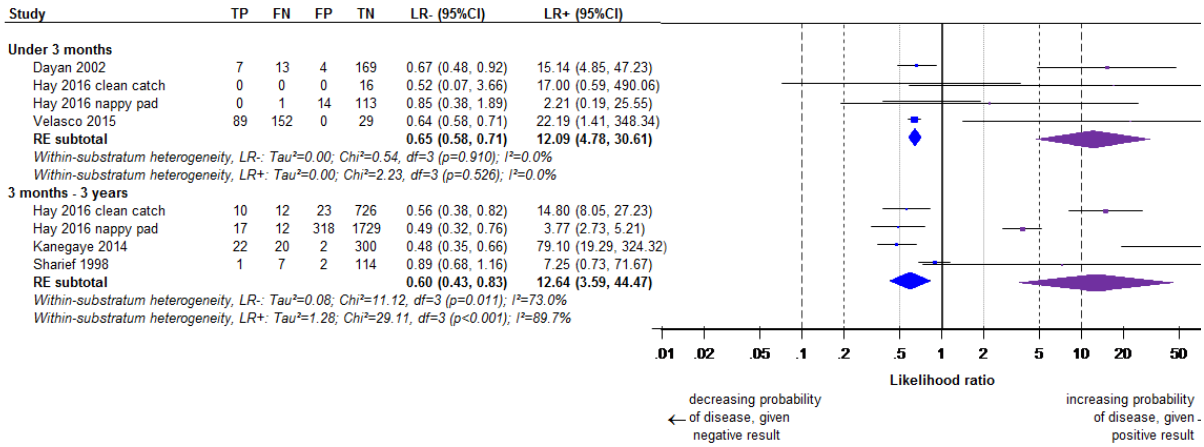


Figure 3: LE positive

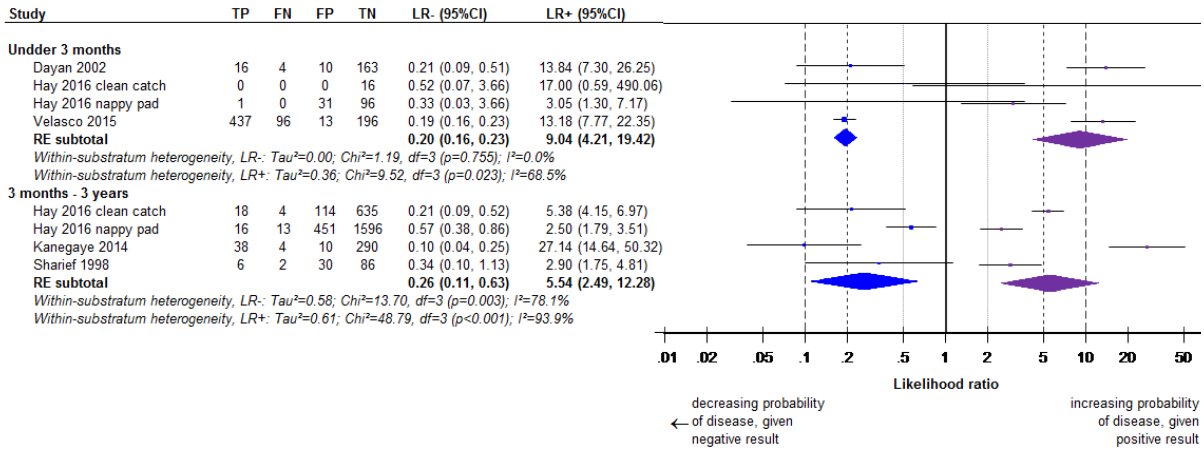


Figure 5: Nitrite or LE positive

Study	TP	FN	FP	TN	LR- (95%CI)	LR+ (95%CI)
Under 3 months						
Dayan 2002	17	3	14	159	0.16 (0.06, 0.46)	10.50 (6.15, 17.93)
Glissmeyer 2014	699	71	349	4926	0.10 (0.08, 0.12)	13.72 (12.37, 15.22)
Hay 2016 clean catch	0	0	3	13	0.63 (0.09, 4.54)	2.43 (0.28, 21.29)
Hay 2016 nappy pad	1	0	37	90	0.35 (0.03, 3.91)	2.56 (1.10, 5.95)
McGillivray 2014	6	7	0	41	0.54 (0.33, 0.88)	39.00 (2.34, 649.17)
Velasco 2015	456	88	18	204	0.18 (0.14, 0.21)	10.34 (6.63, 16.12)
RE subtotal					**0.22 (0.12, 0.39)**	**8.81 (5.42, 14.33)**
Within-substratum heterogeneity, LR-: Tau²=0.34; Chi²=45.13, df=5 (p<0.001); I²=88.9%						
Within-substratum heterogeneity, LR+: Tau²=0.20; Chi²=19.77, df=5 (p=0.001); I²=74.7%						
3 months - 3 years						
Hay 2016 clean catch	19	3	124	625	0.16 (0.06, 0.47)	5.22 (4.14, 6.57)
Hay 2016 nappy pad	23	6	655	1392	0.30 (0.15, 0.62)	2.48 (2.04, 3.02)
Kanegaye 2014	40	2	11	289	0.05 (0.01, 0.19)	25.97 (14.49, 46.57)
Kazi 2013	280	99	45	1214	0.27 (0.23, 0.32)	20.67 (15.42, 27.71)
McGillivray 2014	52	17	5	175	0.25 (0.17, 0.38)	27.13 (11.31, 65.06)
Sharief 1998	6	2	31	85	0.34 (0.10, 1.14)	2.81 (1.70, 4.63)
Shaw 1991	10	4	10	121	0.31 (0.13, 0.71)	9.36 (4.73, 18.50)
Shaw 1998	75	20	99	3200	0.22 (0.15, 0.32)	26.31 (21.11, 32.78)
RE subtotal					**0.25 (0.21, 0.30)**	**10.17 (4.57, 22.64)**
Within-substratum heterogeneity, LR-: Tau²=0.01; Chi²=8.13, df=7 (p=0.321); I²=13.9%						
Within-substratum heterogeneity, LR+: Tau²=1.27; Chi²=342.72, df=7 (p<0.001); I²=98.0%						

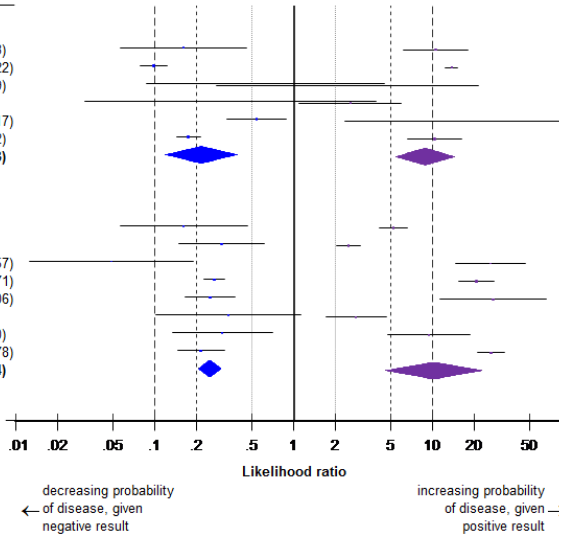
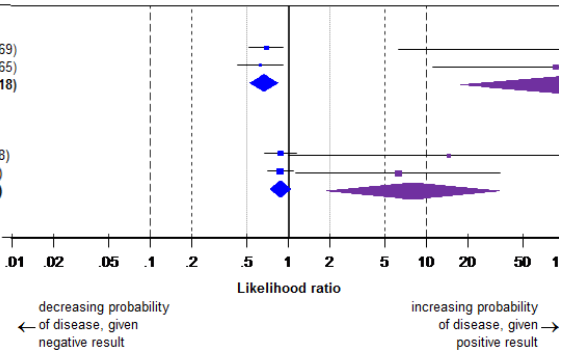


Figure 4: Nitrite and LE positive

Study	TP	FN	FP	TN	LR- (95%CI)	LR+ (95%CI)
Under 3 months						
Dayan 2002	6	14	0	173	0.69 (0.52, 0.92)	107.71 (6.29, 1844.69)
Velasco 2015	6	10	1	229	0.63 (0.43, 0.92)	86.25 (11.04, 673.65)
RE subtotal					**0.67 (0.53, 0.84)**	**93.10 (17.61, 492.18)**
Within-substratum heterogeneity, LR-: Tau²=0.00; Chi²=0.16, df=1 (p=0.686); I²=0.0%						
Within-substratum heterogeneity, LR+: Tau²=0.00; Chi²=0.02, df=1 (p=0.901); I²=0.0%						
3 months - 3 years						
Sharief 1998	1	7	1	115	0.88 (0.68, 1.15)	14.50 (1.00, 210.98)
Shaw 1991	2	12	3	128	0.88 (0.71, 1.09)	6.24 (1.14, 34.22)
RE subtotal					**0.88 (0.74, 1.04)**	**7.95 (1.89, 33.45)**
Within-substratum heterogeneity, LR-: Tau²=0.00; Chi²=0.00, df=1 (p=0.972); I²=0.0%						
Within-substratum heterogeneity, LR+: Tau²=0.00; Chi²=0.27, df=1 (p=0.602); I²=0.0%						



Appendix F:GRADE tables

F.1 UTI diagnosis in infants younger than 3 months: reference test culture

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	LRs	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Nitrite (assessed visually and using analyser) vs culture (10^3, 10^4cfu/ml and $5x10^4$ for SPA and catheter; 10^5 cfu/ml for clean catch and nappy pad)											
3 (Dayan 2002, Hay 2016, Velasco 2015)	Prospective	613	37% (31, 43)	96% (86, 99)	**LR+**	12.09 (4.78, 30.61)	Serious[1]	No serious	No serious	No serious	MODERATE
					LR-	0.65 (0.58, 0.71)	Serious[1]	No serious	No serious	No serious	MODERATE
LE (assessed visually and using analyser) vs culture (10^3, 10^4cfu/ml and $5x10^4$ for SPA and catheter; 10^5 cfu/ml for clean catch and nappy pad)											
3 (Dayan 2002, Hay 2016, Velasco 2015)	Prospective	1,083	82% (78, 85)	91% (77, 97)	**LR+**	9.04 (4.21, 19.42)	Serious[1]	Serious[2]	No serious	No serious	LOW
					LR-	0.20 (0.16, 0.23)	Serious[1]	No serious	No serious	No serious	MODERATE
Nitrite and LE (assessed visually and using analyser) vs culture (10^3, 10^4cfu/ml and $5x10^4$ for SPA and catheter)											
2 (Dayan 2002 and Velasco 2015)	Prospective	439	34% (21, 50)	100% (98, 100)	**LR+**	93.10 (17.61, 492.18)	Serious[1]	No serious	No serious	No serious	MODERATE
					LR-	0.67 (0.53, 0.84)	Serious[1]	No serious	No serious	No serious	MODERATE
Nitrite or LE (assessed visually and analyser) vs culture (10^3, 10^4cfu/ml and $5x10^4$ for SPA and catheter; 10^5 cfu/ml for clean catch and nappy pad)											
5 (Dayan 2002, Glissmeyer 2014, Hay 2016, McGillivray 2005,	Prospective and retrospective	7,208	82% (70, 89)	89% (79, 95)	**LR+**	8.81 (5.42, 14.33)	Serious[1]	Serious[2]	Serious[3]	No serious	VERY LOW
					LR-	0.22 (0.12, 0.39)	Serious[1]	Serious[2]	Serious[3]	No serious	VERY LOW

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	LRs	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Velasco 2015)											
Nitrite or LE (assessed using analyser) and microscopy positive vs culture (10^4 cfu/ml catheter)											
1 (Glissmeyer 2014)	Retrospective	6,394	95% (94.4, 95.0)	88% (87, 88)	**LR+**	7.64 (7.11, 8.21)	Serious[1]	N/A[4]	Serious[5]	No serious	VERY LOW
					LR-	0.06 (0.05, 0.08)	Serious[1]	N/A[4]	No serious	No serious	LOW

1. Evidence downgraded one level for unclear blinding between index and reference test.
2. Evidence downgraded one level due as $I^2 \geq 50\%$.
3. Evidence downgraded one level for serious indirectness as 3/5 studies had infants < 90 days old.
4. Inconsistency not applicable as evidence from a single study.
5. Evidence downgraded one level for serious indirectness as study had infants < 90 days old.

F.2 UTI diagnosis in infants and children 3 months or older but younger than 3 years: reference test culture

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	LRs	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Nitrite (assessed visually and using analyser) vs culture ($5x10^4$ cfu/ml for catheter and 10^5 cfu/ml for bag, clean catch and nappy pad)											
3 (Hay 2016, Kanegaye 2014, Sharief 1998)	Prospective	3,270	50% (37, 62)	97% (88, 99)	**LR+**	12.64 (3.59, 44.47)	Serious[1]	Serious[2]	No serious	No serious	LOW
					LR-	0.60 (0.43, 0.83)	Serious[1]	Serious[2]	No serious	Serious[3]	VERY LOW
LE (assessed visually and using analyser) vs culture ($5x10^4$ cfu/ml for catheter and 10^5 cfu/ml for bag, clean catch and nappy pad)											
3 (Hay 2016, Kanegaye 2014 and Sharief 1998)	Prospective	3,313	77% (55, 90)	85% (77, 91)	**LR+**	5.54 (2.49, 12.28)	Serious[1]	Serious[2]	No serious	No serious	LOW
					LR-	0.26 (0.11, 0.63)	Serious[1]	Serious[2]	No serious	Serious[3]	VERY LOW

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	LRs	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Nitrite and LE (assessed visually and using analyser) vs culture (10^3 cfu/ml for catheter, 10^5 cfu/ml for clean catch)											
2 (Shaw 1991 and Sharief 1998)	Prospective	269	14% (4, 35)	98% (95, 99)	**LR+**	7.95 (1.89, 33.45)	Serious[4]	No Serious	Serious[5]	Serious[3]	VERY LOW
					LR-	0.88 (0.74, 1.04)	Serious[4]	No Serious	Serious[5]	No serious	LOW
Nitrite or LE (assessed visually and analyser) vs culture (10^3 to $5x10^4$ cfu/ml for catheter, 10^5 cfu/ml for clean catch)											
7 (Hay 2016, Kanegaye 2014; Kazi 2013; McGillivray 2014; Sharief 1998; Shaw 1991; Shaw 1998)	Prospective and retrospective	8739	77% (72, 82)	92% (81, 97)	**LR+**	10.17 (4.57, 22.64)	Serious[1]	Serious[2]	No serious	No serious	LOW
					LR-	0.25 (0.21, 0.30)	Serious[1]	No Serious	No serious	No serious	MODERATE
Nitrite or LE (method of assessment unclear) and microscopy positive vs culture (10^3 cfu/ml catheter)											
1 (Reardon 2009)	Retrospective	435	64% (49, 78)	91% (88, 94)	**LR+**	7.39 (5.02, 10.89)	Serious[1]	N/A[6]	Serious[7]	No serious	VERY LOW
					LR-	0.39 (0.26, 0.58)	Serious[1]	N/A[6]	Serious[7]	Serious[3]	VERY LOW
Nitrite or LE or blood or protein (assessed by analyser) vs culture (10^5 cfu/ml for clean bag)											
1 (Doley and Nelligan 2003)	Retrospective	160	88% (68, 97)	39% (31, 49)	**LR+**	1.45 (1.18, 1.78)	Very serious[8]	N/A[6]	Serious[7]	No serious	VERY LOW
					LR-	0.31 (0.11, 0.93)	Very serious[8]	N/A[6]	Serious[7]	Serious[3]	VERY LOW

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	LRs	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality

1. Evidence downgraded one level due to unclear blinding of reference and index test.
2. Evidence downgraded one level due as $I^2 \geq 50\%$.
3. Evidence downgraded one level as 95% confidence interval of likelihood ratio crosses one MID (0.5 or 2).
4. Evidence downgraded two levels due to very serious risk of bias from patient selection and unclear dipstick testing criteria.
5. Evidence downgraded one level due to indirect age group (<2 years and < 1 year).
6. Inconsistency not applicable as evidence from a single study.
7. Evidence downgraded one level due to indirect age group (< 24 months) in one study.

F.3 UTI diagnosis in infants and children 3 months or older but younger than 3 years: reference test culture and microscopy

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	LRs	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Nitrite (assessed by analyser) vs culture (10^5 cfu/ml, method of collection not reported) and microscopy (WBC > 25 x 10^9/L for boys or 50 x 10^9/L for girls)											
1 (Lejeune 1991)	Prospective	243	16% (6, 32)	98% (94, 99)	LR+	6.68 (2.15, 20.77)	Serious[1]	N/A[2]	No serious	No serious	MODERATE
					LR-	0.86 (0.74, 0.99)	Serious[1]	N/A[2]	No serious	No serious	MODERATE
LE (assessed by analyser) vs culture (10^5 cfu/ml, method of collection not reported) and microscopy (WBC > 25 x 10^9/L for boys or 50 x 10^9/L for girls)											
1 (Lejeune 1991)	Prospective	243	89% (75, 97)	78% (72, 84)	LR+	4.08 (3.08, 5.41)	Serious[1]	N/A[2]	No serious	No serious	MODERATE
					LR-	0.14 (0.05, 0.35)	Serious[1]	N/A[2]	No serious	No serious	MODERATE
Nitrite and LE (assessed by analyser) vs culture (10^5 cfu/ml, method of collection not reported) and microscopy (WBC > 25 x 10^9/L for boys or 50 x 10^9/L for girls)											
1 (Lejeune 1991)	Prospective	243	87% (72, 96)	98% (95, 99)	LR+	44.51 (16.73, 118.38)	Serious[1]	N/A[2]	No serious	No serious	MODERATE
					LR-	0.13 (0.06, 0.30)	Serious[1]	N/A[2]	No serious	No serious	MODERATE
Protein (assessed by analyser) vs culture (10^5 cfu/ml, method of collection not reported) and microscopy (WBC > 25 x 10^9/L for boys or 50 x 10^9/L for girls)											
1 (Lejeune 1991)	Prospective	243	8% (2, 22)	95% (91, 98)	LR+	1.67 (0.48, 5.78)	Serious[1]	N/A[2]	No serious	Very serious[3]	VERY LOW
					LR-	0.97 (0.87 to 1.07)	Serious[1]	N/A[2]	No serious	No serious	MODERATE

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	LRs	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
LE and protein (assessed by analyser) vs culture (10^5 cfu/ml, method of collection not reported) and microscopy (WBC > 25 x 10^9/L for boys or 50 x 10^9/L for girls)											
1 (Lejeune 1991)	Prospective	243	89% (75, 97)	95% (91, 98)	**LR+**	18.37 (9.93 , 33.98)	Serious[1]	N/A[2]	No serious	No serious	MODERATE
					LR-	0.11 (0.05, 0.29)	Serious[1]	N/A[2]	No serious	No serious	Moderate
LE and protein and nitrite (assessed by analyser) vs culture (10^5 cfu/ml, method of collection not reported) and microscopy (WBC > 25 x 10^9/L for boys or 50 x 10^9/L for girls)											
1 (Lejeune 1991)	Prospective	243	89% (75, 97)	72% (65, 78)	**LR+**	3.17 (2.48, 4.05)	Serious[1]	N/A[2]	No serious	No serious	MODERATE
					LR-	0.15 (0.06 to 0.38)	Serious[1]	N/A[2]	No serious	No serious	MODERATE

1. Evidence downgraded one levels due to unclear index and reference test blinding.
2. Inconsistency not applicable as evidence is from a single study and not pooled in a meta-analysis.
3. Evidence downgraded two levels as 95% CI cross two minimal important differences (0.5 and 2).
4.

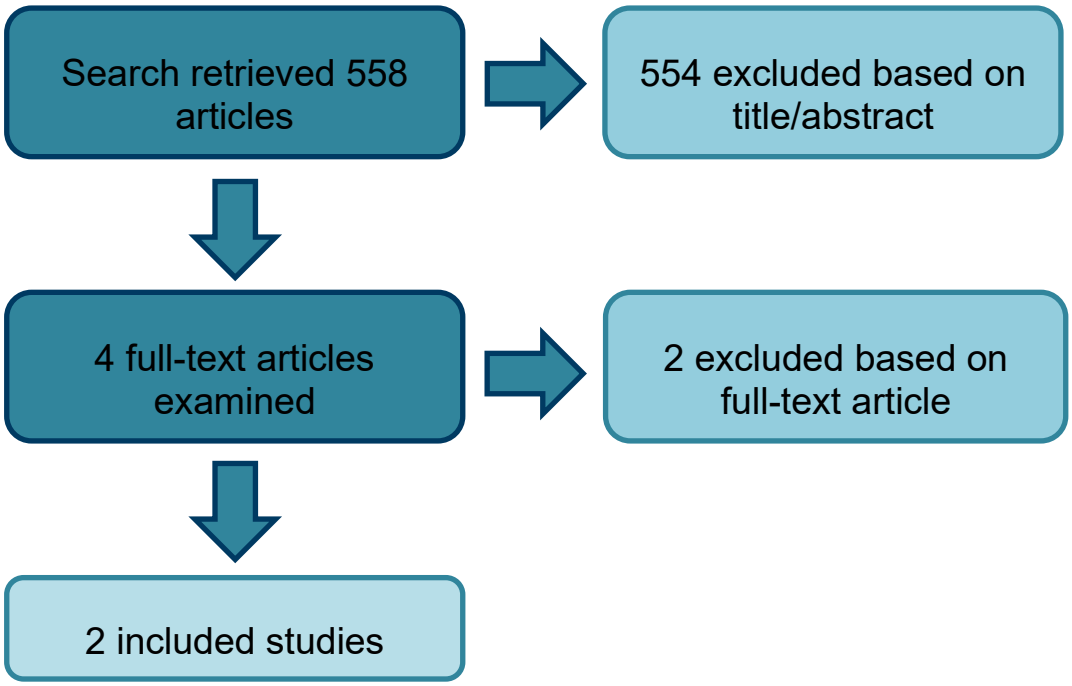
Appendix G: Sensitivity analysis

The following table highlights the results of the sensitivity analysis. Large differences from the primary analysis which were observed in the sensitivity analysis are highlighted in grey.

Primary analysis: All sampling methods		**Sensitivity analysis: Only non-invasive methods (clean catch, bag or nappy pad)**	
Number of studies, N =	**Outcomes (95% CI)**	**Number of studies, N =**	**Outcomes (95% CI)**
< 3 months: reference test culture			
Nitrite			
3, N = 613	Sens = 37% (31, 43) Spec = 96% (86, 99) LR+ = 12.09 (4.78, 30.61) LR- = 0.65 (0.58, 0.71)	1, N = 144	Sens = 34% (4, 86) Spec = 89% (83, 93) LR+ = 4.48 (0.62, 32.42) **not significant in sensitivity analysis** LR- = 0.79 (0.37, 1.66) **not significant in sensitivity analysis**
LE			
3, N = 1083	Sens = 82% (78, 85) Spec = 91% (77, 97) LR+ = 9.04 (4.21, 19.42) LR- = 0.20 (0.16, 0.23)	1, N = 144	Sens = 66% (14, 96) Spec = 87% (43, 98) LR+ = 0.43 (0.09, 1.97) **not significant in sensitivity analysis** LR- = 3.38 (1.48, 7.75)
Nitrite or LE			
5, N = 7208	Sens = 82% (70, 89) Spec = 89% (79, 95) LR+ = 8.81 (5.42, 14.33) LR- = 0.22 (0.12, 0.39)	1, N = 144	Sens = 66% (14, 96) Spec = 72% (64, 78) LR+ = 2.54 (1.16, 5.58) LR- = 0.50 (0.11, 2.29) **not significant in sensitivity analysis**
3 months or older but younger than 3 years			
Nitrite			
3, N = 3270	Sens = 50% (37, 62) Spec = 97% (88, 99) LR+ = 12.64 (3.59, 44.47) LR- = 0.60 (0.43, 0.83)	2, N = 2971	Sens = 45% (25, 67) Spec = 95% (81, 99) LR+ = 7.26 (2.30, 22.91) LR- = 0.64 (0.44,0.95)
LE			
3, N = 3313	Sens = 87% (55, 90) Spec = 85% (55, 90) LR+ = 5.54 (2.49, 12.28) LR- = 0.26 (0.11, 0.63)	1, N = 2971	Sens = 70% (48, 85) Spec = 80% (73, 85) LR+ = 3.45 (2.02, 5.91) LR- = 0.39 (0.20, 0.75)
Nitrite and LE			
2, N = 269	Sens = 14% (4, 35) Spec = 98% (95, 99)	1, N = 124	Sens = 13% (0.3, 53) Spec = 99% (95, 99)

Primary analysis: All sampling methods		Sensitivity analysis: Only non-invasive methods (clean catch, bag or nappy pad)	
	LR+ = 7.95 (1.89, 33.45) LR- = 0.88 (0.74, 1.04) **not significant in primary analysis**		LR+ = 14.50 (1.00, 210.99) **not significant in sensitivity analysis** LR- = 0.88 (0.68, 1.15) **Not significant in sensitivity analysis**
Nitrite or LE			
6, N = 8739	Sens = 77% (72, 82) Spec = 92% (81, 97) LR+ = 10.17 (4.57, 22.64) LR- = 0.25 (0.21, 0.30)	1, N = 2971	Sens = 80% (68, 88) Spec = 75% (65, 83) LR+ = 2.54 (1.16, 5.58) LR- = 0.50 (0.11, 2.29)

Appendix H: Economic evidence study selection



Appendix I: Economic evidence tables

Bibliographic reference	**Whiting, P., Westwood, M., Bojke, L., Palmer, S., Richardson, G., Cooper, J., Watt, I., Glanville, J., Sculpher, M. and Kleijnen, J., 2006. Clinical effectiveness and cost-effectiveness of tests for the diagnosis and investigation of urinary tract infection in children: a systematic review and economic model.**
Evaluation design	

Interventions	Dipstick testing (with and without culture as confirmatory test) versus microscopy and/or culture
Comparators	As above
Base-line cohort characteristics	Children under 5 years with suspected UTI, stratified by age (<1, 1-2, 2-3, and >3 years) and gender
Type of Analysis	Cost-utility
Structure	Decision tree and Markov model
Cycle length	Not specified
Time horizon	Lifetime
Perspective	NHS
Country	UK
Currency unit	GBP
Cost year	2003
Discounting	6% for costs, 2% for health benefits
Other comments	-

Results

Due to the large number of possible strategies, the strategy with the highest probability of being the most cost effective was reported for a range of thresholds, rather than reporting costs and QALYs.

At a threshold of £20,000/QALY strategies with the highest expected net monetary benefit were as follows:

Girl <1 year: Dipstick (positive for nitrite or LE), followed by confirmatory laboratory culture, followed by MCUG

Girl 1-2 years: Dipstick (positive for nitrite and LE), followed by MCUG

Girl 2-3 years: Dipstick (positive for nitrate and LE), followed by MCUG

Girl >3 years: Treat all patients with suspected UTI

Boy <1 year: Dipstick (positive for nitrite and LE) followed by MCUG

Boy 1-2 years: Dipstick (positive for nitrite and LE) followed by MCUG

Boy 2-3 years: Treat all patients with suspected UTI

Bibliographic reference	**Whiting, P., Westwood, M., Bojke, L., Palmer, S., Richardson, G., Cooper, J., Watt, I., Glanville, J., Sculpher, M. and Kleijnen, J., 2006. Clinical effectiveness and cost-effectiveness of tests for the diagnosis and investigation of urinary tract infection in children: a systematic review and economic model.**
	Boy >3 years: Treat all patients with suspected UTI
Data sources	**Base-line data**: Prevalence of UTI and presence of reflux derived from meta-analysis of RCTs. Recurrence of UTI, probability of renal scarring, and probability of ESRD were taken from individual studies of the epidemiology or natural history of UTI. **Effectiveness data**: Sensitivity and specificity of diagnostic tests were sourced from meta-analyses of clinical studies. **Cost data**: Costs were taken from standard NHS sources (BNF and PSSRU), or previous costing and economic analyses from the perspective of the NHS. **Utility data**: Due to the lack of utility values for children with UTI, utilities were taken from a single source that examined the cost effectiveness of treatment strategies for women with suspected UTIs.
Uncertainty	**One-way sensitivity analysis**: A deterministic sensitivity analysis was conducted in which strategies involving glucose testing with dipsticks were included for children >3 (these were excluded from the main analysis due to poor quality of data). Results indicated that glucose testing followed by MCUG becomes the optimal strategy for girls at thresholds ≥£24,000/QALY for girls and ≥£40,000/QALY for boys. **Probabilistic sensitivity analysis**: Probabilistic sensitivity analysis showed that, due to the large number of possible strategies, the probability of the strategy with the highest expected net monetary benefit being the most cost effective was generally low at any threshold.
Applicability	**Directly Applicable** This study is from the perspective of the NHS, and is therefore directly applicable.
Limitations	**Very serious limitations** While the overall quality of the analysis is high, the assumption is made that the accuracy of diagnostic tests does not vary with children's age. Since the review question is explicitly focused on determining whether diagnostic accuracy varies according to age (and, by extension, whether this affects cost effectiveness), this classifies as a very serious limitation.
Conflicts	**None**

Bibliographic reference	**Hay, A., Birnie, K., Busby, J., Delaney, B., Downing, H., Dudley, J., Durbaba, S., Fletcher, M., Harman, K., Hollingworth, W. and Hood, K., 2016. The Diagnosis of Urinary Tract infection in Young children (DUTY): a diagnostic prospective observational study to derive and validate a clinical algorithm for the diagnosis of urinary tract infection in children presenting to primary care with an acute illness.**

Evaluation design

Interventions	Dipstick testing (to direct initial antibiotic treatment) followed by laboratory testing in all children
Comparators	Laboratory testing (with antibiotics prescribed according to results) Presumptive treatment (antibiotics prescribed for children with suspected UTI prior to laboratory testing)
Base-line cohort characteristics	Children under 5 years at low risk of UTI (definition: GP responds yes to question 'if this child was NOT in the DUTY study would you have requested a urine sample?')
Type of Analysis	Cost-utility
Structure	Markov model
Cycle length	1 day
Time horizon	21 days
Perspective	NHS
Country	UK
Currency unit	GBP
Cost year	2011
Discounting	N/A (time horizon shorter than 1 year)
Other comments	-

Results

Strategy	Cost	QALDs	Net monetary benefit	Incremental net monetary benefit (versus no laboratory testing)
Laboratory testing	1.100	20.709	1090.44	-
Dipstick testing	1.183	20.709	1090.38	-0.05
Presumptive treatment	1.187	20.709	1090.4	-0.04

Data sources

Base-line data	Baseline data were taken from the accompanying DUTY clinical study
Effectiveness data	Effectiveness data were taken from the accompanying DUTY clinical study
Cost data	Costs were taken from standard NHS sources (PSSRU), or previous costing and economic analyses from the perspective of the NHS.
Utility data	Utilities were taken from previous economic analyses (Whiting 2006) or from studies of QoL in specific disease states.

Bibliographic reference	**Hay, A., Birnie, K., Busby, J., Delaney, B., Downing, H., Dudley, J., Durbaba, S., Fletcher, M., Harman, K., Hollingworth, W. and Hood, K., 2016. The Diagnosis of Urinary Tract infection in Young children (DUTY): a diagnostic prospective observational study to derive and validate a clinical algorithm for the diagnosis of urinary tract infection in children presenting to primary care with an acute illness.**
Uncertainty	**One-way sensitivity analysis**: N/A **Probabilistic sensitivity analysis**: Bootstrapping was used to produce 95% confidence intervals of incremental net monetary benefit. This showed that the laboratory testing strategy was significantly more cost-effective than the dipstick strategy.
Applicability	**Partially Applicable** This evaluation is based on a patient population that overlaps with, but is not identical to the population of interest (<5 years) and is therefore assessed as partially applicable.
Limitations	**Minor limitations** This evaluation is based on a high-quality clinical study and is generally of a high quality. Although the model time horizon is short (21 days), this is appropriate to the decision problem, as it investigates outcomes associated with a single episode of UTI.
Conflicts	**None**

Appendix J: Health economic analysis

Introduction

2007 NICE guidance on the diagnosis and management of urinary tract infection in under 16s recommends that children over the age of 3 years with a suspected UTI should initially be tested with a urine dipstick. However, due to lack of evidence regarding the accuracy of dipstick tests in younger children, this recommendation was not previously extended to children under the age of 3 years, for whom urgent microscopy and culture was recommended. The purpose of this economic evaluation is to determine whether dipstick testing prior to microscopy and culture is cost-effective in this younger age group, using accuracy data synthesised from the clinical review for this update.

Methods

Type of analysis

Cost-utility analysis, in which cost are measured in GBP and health effects are measured in quality-adjusted life years (QALYs).

Target population

Children with suspected UTI under the age of 3 years, stratified into two age groups:

- Infants younger than 3 months
- Children 3 months or older but younger than 3 years

Interventions

The analysis compares two major strategies:

- '**No dipstick testing':** A urine sample is sent for urgent microscopy and culture in all children with suspected UTI. Antibiotic treatment is started immediately for all children, with treatment adjusted or discontinued as appropriate when test results are received.
- **'Dipstick testing':** All children with suspected UTI are dipstick tested. For children with a positive dipstick test a urine sample is sent for urgent microscopy and culture, and antibiotic treatment is started. Children with a negative dipstick test are assumed to not have UTI, and no further testing or treatment is administered unless symptoms persist. This option consists of four sub-strategies, according to interpretation of nitrite and leukocyte esterase (LE) results:
  - Presence of nitrite alone is considered a positive test result
  - Presence of LE alone is considered a positive test result
  - Presence of nitrite or LE is considered a positive test result
  - Presence of nitrite and LE is considered a positive test result

Perspective

The analysis was conducted from the perspective of the NHS and social services (PSS).

Discounting

A discount rate of 3.5% per annum was applied to all costs and QALYs after the first year.

Model structure

As there is considerable uncertainty regarding the possible outcomes of a false negative dipstick result, the model uses a number of scenarios to explore the potential consequences of an untreated UTI.

- **Basic scenario:** A false negative test result for UTI only results in a longer duration of symptoms, after which there are no further adverse consequences
- **Scenario 1:** In addition to the basic scenario assumption, a false negative result also increases the risk of children with UTI developing septicaemia
- **Scenario 2:** In addition to the base case assumption, a false negative result also increases the risk of PRS in the future, and hence the risk of progressing to end-stage renal disease (ESRD).
- **Scenario 3:** In addition to the base case assumption, a false negative result also increases the risk of septicaemia and PRS.

Basic scenario

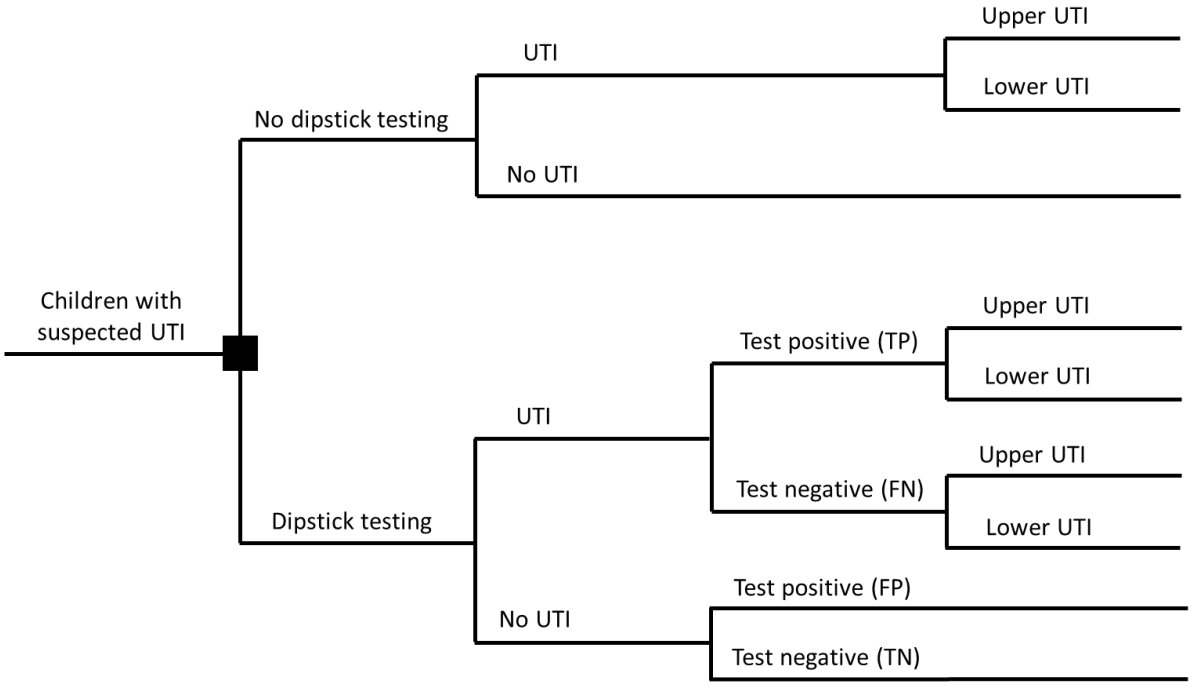
The 'basic scenario' structure consists of two elements: a short-term decision tree, which simulates children's UTI status, subsequent test results, and treatment of the UTI episode; and a long-term Markov model, which estimates lifetime cost and QALY outcomes, and captures any downstream effects of UTI.

The short-term decision tree for the base case is shown in Figure 6. At the start of the tree, the decision is made between a 'no dipstick testing' and 'dipstick testing' strategy. In the 'no dipstick testing arm of the model, all patients with suspected UTI have an underlying UTI status (either UTI or no UTI). A urine sample is sent for microscopy and culture for all children, which provides a definitive test of UTI. All children with UTI are appropriately treated with a course of antibiotics, which is assumed to resolve the infection. UTI may either take the form of upper UTI (pyelonephritis) or lower UTI, which affects the duration of symptoms.

For the 'dipstick testing' arm, all children are initially tested with dipstick. Again, all children have an underlying UTI status (UTI or no UTI) and for each of these groups dipstick test can produce a positive or negative test result, with probabilities according the sensitivity and specificity of testing. Children testing positive (both true positives and false positive) receive antibiotic treatment and a urine sample is sent for microscopy and culture. Children with a false negative result experience a delay in treatment ('untreated UTI'), after which time their true UTI positive status is discovered, antibiotic treatment is administered and a urine sample is sent for microscopy and culture. Children with a true negative result appropriately receive no further treatment or testing for UTI.

For the long-term section of the base case, a simple Markov model is used to estimate lifetime QALYs. Following the resolution of UTI, all children are assumed to return to a healthy state, and age-specific baseline mortality rates for the general population are used to estimate survival. Since, in the base case scenario, testing strategies only affect the duration of UTI, no differences in costs and QALYs occur between strategies during the long-term phase of the model. However, this element is included as it is required to capture long-term differences in health outcomes and costs when the incidence of septicaemia and PRS are included in the model.

Figure 6: Diagram of short-term decision tree section of model

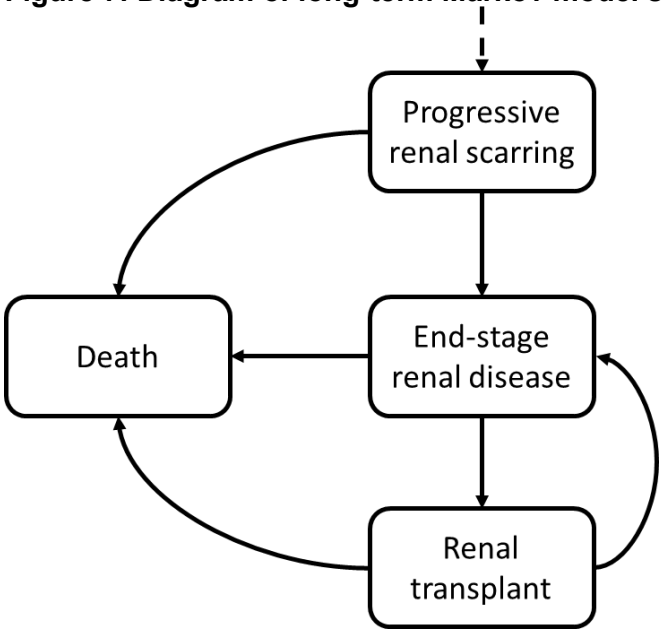


Including risk of progressive renal scarring

For the scenario in which false negative results are associated with an increased risk of PRS, all children with UTI have a baseline risk of developing PRS in the future, with differing probabilities according to whether the infection is upper or lower UTI. Children with a false negative test result have an increased risk of PRS.

For children developing PRS, the long-term Markov model simulates progress through various stages of disease, as shown in Figure 7. Patients with PRS have a probability of developing end-stage renal disease (ESRD) after a number of years. In this state, patients have an elevated annual probability of death, and also have an annual probability of receiving a renal transplant. From the renal transplant state, patients have an increased probability of death in the first year after surgery, after which mortality is assumed to return to baseline rate. Patients can also experience transplant failure, which results in a return to the end-stage renal disease state.

Figure 7: Diagram of long-term Markov model section of the model



Including risk of septicaemia

For the scenario in which false negative results are associated with an increased risk of septicaemia, all children with UTI have a baseline risk of developing septicaemia. Children with a false negative test result have an increased risk of septicaemia relative to baseline. Children who develop septicaemia also have a probability of death. In order to capture the lifetime QALY loss from septicaemia-related death, these children do not progress to the long-term Markov phase of the model.

Model inputs

Accuracy of dipstick testing

Sensitivity and specificity of dipstick tests for each interpretation of nitrite/LE results are displayed in Table 14, stratified by age group. Values were taken from a meta-analysis of studies identified in the clinical review with culture as the reference test, methodology of which is detailed in the 'methods and process' section, and full results are detailed in appendices E and F.

Table 1414: Accuracy of dipstick testing

Children under 3 months		
Dipstick interpretation	**Sensitivity (95% CI)**	**Specificity (95% CI)**
Nitrite	37% (31%-43%)	96% (86%-99%)
LE	82% (78%-85%)	91% (77%-97%)
Nitrite and LE	34% (21%-50%)	100% (98%-100%)
Nitrite or LE	82% (70%-89%)	89% (79%-95%)
Children 3 months or older but younger than 3 years		
Dipstick interpretation	**Sensitivity (95% CI)**	**Specificity (95% CI)**
Nitrite	50% (37%-62%)	97% (88%-99%)
LE	77% (55%-90%)	85% (77%-91%)
Nitrite and LE	14% (4%-35%)	98% (95%-99%)

Children under 3 months		
Nitrite or LE	77% (72%-82%)	92% (81% - 97%)

Prevalence of UTI

Inputs for the baseline prevalence of UTI in children with suspected UTI are displayed in Table 15. Values for overall prevalence were taken from a meta-analysis of UTI prevalence in children with fever (Shaikh et al, 2008[1]). As age ranges in this meta-analysis did not coincide exactly those in the model, UTI prevalence for children 6-12 months was used to inform the population of children 3 months or older but younger than 3 years.

Values for the prevalence of upper UTI in children with UTI were taken from Whiting et al (2006)[2]. For the under 3 months population, a prevalence value for children of 1 year was used (as this was the youngest age for which data were available), and for the population of children 3 months or older but younger than 3 years, a simple average of prevalence at 1 year and at 2 years was used.

Table 15: Prevalence of UTI in children with suspected UTI

Children under 3 months		
Parameter	**Value (95% CIs)**	**Source**
Prevalence of UTI	7.2% (5.8%-8.6%)	Shaikh 2008[1]
Prevalence of upper UTI in children with UTI	83% (77%-89%)	Whiting 2006[2]
Children 3 months or older but younger than 3 years		
Parameter	**Value (95% CIs)**	**Source**
Prevalence of UTI	5.4% (3.4%-7.4%)	Shaikh 2008[1]
Prevalence of upper UTI in children with UTI	67% (60%-74%)	Whiting 2006[2]

Duration of UTI

Values for the duration of lower and upper UTI were taken from Whiting et al (2006)[2] and are displayed in Table . As per the Whiting evaluation, the assumption was made that a false negative test for UTI results in symptoms being extended by 4 days.

Table 16: Duration of UTI

UTI characteristics	Duration	Source
Lower UTI – treated	3 days	Whiting 2006[2]
Lower UTI – untreated	7 days	Whiting 2006[2]
Upper UTI – treated	10 days	Whiting 2006[2]
Upper UTI – untreated	14 days	Whiting 2006[2]

Probability of septicaemia

Due to the lack of data directly relating to the relating to the incidence and mortality associated with septicaemia and sepsis, data relating to bacteraemia were used to populate the model, making the assumption that deaths that occur are due to bacteraemia which subsequently develops into sepsis.

Probabilities relating to septicaemia incidence and mortality are displayed in Table17 . The baseline probability of bacteraemia in children under 3 months was taken directly from a source in the literature (Schnadower et al, 2010[3]). However, analogous sources for children over 3 months were relatively scarce. One study (Pitetti et al, 2002[4]) was identified which reported the incidence of bacteraemia in children younger than 2 months and children between 2 months and 3 years. Therefore, in order to estimate septicaemia incidence or the older group of children, an odds ratio was calculated between the older and younger group of children in the Pitetti study, and this was applied to the probability of septicaemia in children under 3 months from the Schnadower study.

As no data were available specifically on the incidence of septicaemia in untreated UTI an estimated relative risk of 2 was initially applied to the baseline incidence of septicaemia (in both the younger and the older group). Due to the lack of evidence for the value this parameter might take in reality, a threshold analysis was carried out in order to determine the relative risk at which the decision between 'dipstick' and 'no dipstick' strategies' would change at a threshold of £20,000 per QALY. This parameter was also varied widely in one-way and probabilistic sensitivity analysis.

Probabilities of death from septicaemia were taken from a Public Health England report providing age-stratified thirty-day all-cause fatality subsequent to E. coli bacteraemia[5]. We assumed case-fatality rates for children under 1 year applied to our modelled cohort of children under 3 months and case-fatality rates for 1–14 year-olds year applied to our modelled cohort of children 3 months or older but younger than 3 years.

Table17 : Probabilities of developing septicaemia

Children under 3 months		
Parameter	**Value (SE)**	**Source**
Probability of bacteraemia – treated UTI	6.6% (0.57%)	Schnadower 2010[3]
Relative risk – probability of bacteraemia in untreated UTI versus probability of bacteraemia in treated UTI	2	Assumption
Probability of death from bacteraemia	7.7% (1.13%)	Public Health England 2015[5]
Children 3 months or older but younger than 3 years		
Parameter	**Value (SE)**	**Source**
Probability of bacteraemia <2 months	22.7% (8.93%)	Pitetti 2002[4]
Probability of bacteraemia 2 months – 3 years	3.2% (1.41%)	Pitetti 2002[4]
Odds ratio – probability of bacteraemia 2 months to 3 years versus probability of bacteraemia <2 months	0.11	Calculated
Calculated probability of bacteraemia	0.9%	Calculated
Relative risk – probability of bacteraemia in untreated UTI versus probability of bacteraemia in treated UTI	2	Assumption
Probability of death from bacteraemia	4.3% (1.39%)	Public Health England 2015[5]

Progression of PRS

Parameters relating to the incidence and progression of PRS are displayed in Table 15. The baseline probability of renal scarring in a child with UTI was calculated using values for the probability of reflux, proportion of reflux which is mild/moderate (with the remainder classified as severe), and probability of renal scarring given severe reflux (making the assumption that only patients with severe reflux are at risk of renal scarring) taken from Whiting et al (2006)[2]. As with the incidence of septicaemia, the increase in risk of PRS caused by a delay in treatment of UTI is unknown, and therefore an arbitrarily chosen relative risk of 2 was applied to the baseline probability of PRS for the base case. Due to the lack of evidence for the value this parameter might take in reality, a threshold analysis was carried out in order to determine the relative risk at which the decision between 'dipstick' and 'no dipstick' strategies' would change at a threshold of £20,000 per QALY. This parameter was also varied widely in one-way and probabilistic sensitivity analysis.

As no data were available specifically on the incidence of septicaemia in untreated UTI an estimated relative risk of 2 was initially applied to the baseline incidence of septicaemia (in both the younger and the older group). Due to the lack of evidence for the value this parameter might take in reality, a threshold analysis was carried out in order to determine the relative risk at which the decision between 'dipstick' and 'no dipstick' strategies' would change at a threshold of £20,000 per QALY. This parameter was also varied widely in one-way and probabilistic sensitivity analysis.

The lifetime probability of developing ESRD in individuals with PRS was taken from Whiting et al (2006)[2]. Mean and range for age ESRD onset were also taken from the Whiting study, from which the proportion of patients progressing to ESRD each year was calculated using a triangular distribution. The annual probability of death from ESRD in a European population was taken from Goodkin et al (2003)[6].

An annual probability of receiving a renal transplant was derived by fitting a beta distribution to the median days wait for renal transplant (taken from the NHS Annual Report on Kidney Transplantation, 2014[7]), from which an estimate of the proportion of patients receiving a transplant within 365 days was calculated. Probability of death and of renal graft failure in the first year after transplant were taken directly from the NHS Annual Report on Kidney Transplantation. The assumption was made that mortality returns to baseline from the second year after transplantation onwards. To calculate the annual probability of renal graft failure from the second year after transplant onwards the one year graft failure rate was first subtracted from the five year failure probability (again taken from the NHS Annual Report) to provide a failure probability for years 2 to 5 after transplant, which was converted to an annual probability.

Table 15: Parameters relating to the incidence and progression of PRS

Parameter	Value (accuracy)	Source
Prevalence of reflux	28.8% (SE = 5.9%)	Whiting 2006[2]
Proportion of reflux classified as mild/moderate	87.7% (SE = 17.5%)	Whiting 2006[2]
Probability of renal scarring in lower UTI in patients with severe reflux	27% (95% CI = 4.6%-60.1%)	Whiting 2006[2]
Probability of renal scarring in upper UTI in patients with severe reflux	44% (95% CI = 27.3%-68.6%)	Whiting 2006[2]
Probability of renal scarring in all children with treated lower UTI	0.43%	Calculated
Probability of renal scarring in all children with treated upper UTI	0.96%	Calculated

Parameter	Value (accuracy)	Source
Relative risk – probability of PRS in children with untreated UTI versus treated UTI	2	Assumption
Probability of progression to ESRD in children with PRS	5% (95% CI = 2.5% to 8.4%)	Whiting 2006[2]
Mean age of ESRD onset	13.67 (range = 7-24)	Whiting 2006[2]
Annual probability of death from ESRD	15.6% (95% CI = 14.2%-17.0%)	Goodkin 2003[6]
Median days wait for renal transplant	342 (95% CI = 249-342)	NHS Annual Report on Kidney Transplantation 2014[7]
Annual probability of receiving renal transplant	76.8%	Calculated
Probability of death in the first year after renal transplant	1% (95% CI = 0%-3%)	NHS Annual Report on Kidney Transplantation 2014[7]
Probability of renal graft failure in first year after transplant	4% (95% CI = 2%-7%)	NHS Annual Report on Kidney Transplantation 2014[7]
Probability of renal graft failure in first 5 years after transplant	16% (95% CI = 12%-21%)	NHS Annual Report on Kidney Transplantation 2014[7]
Annual probability of renal graft failure from year 2 after transplant onwards	3.2%	Calculated

Costs

Costs used in the model are displayed in Table along with their sources. Costs from Whiting et al 2006[2] and Kerr et al 2012[8] have been adjusted to 2015/16 values using the Health Service Cost Index, taken from PSSRU Unit Costs of Health and Social Care 2015.

The cost of antibiotic treatment is calculated using a simple mean of the pack cost for Amoxicillin 125mg/1.25ml oral suspension paediatric, Cefalexin 125mg/5ml oral suspension, Co-amoxiclav 125mg/31mg/5ml oral suspension, and Cefradine 250mg capsules, with prices taken from the NHS Drug Tariff[9]. For each of these treatments a per-pack rather than a per-day cost was used, as it was determined that the remainder of the pack was unlikely to be re-used by other patients once the treatment course had completed.

The cost of renal transplant was calculated using a weighted mean of kidney transplant procedures for patients of 18 years and younger in the NHS National Schedule of Reference Costs 2015-16[10], with all patients subsequently receiving 2 vials of basiliximab, 75% of patients receiving tacrolimus immunotherapy (150 micrograms/kg daily for a 70kg individual over 15 days)[9] and 25% of patients receiving ciclosporin immunotherapy (2mg/kg daily for a 70kg individual over 15 days)[9] Similarly, the cost of septicaemia was calculated using a weighted average of all sepsis treatment in the NHS Reference Costs 2015-16.

Table 19: Costs used to populate the model

Item	Cost	Source
Dipstick test	£0.12	Siemans Multistix 10SG Urinalysis Strips x 100 - medisave.co.uk - accessed 18/04/17
Microscopy (bacteriuria and pyuria)	£22.24	Whiting 2006[2]

Item	Cost	Source
Laboratory culture	£3.61	Whiting 2006[2]
Antibiotic treatment	£2.14	NHS Drug Tariff[9]
Additional cost of treating upper UTI	£23.98	Whiting 2006[2]
Additional cost of untreated UTI	£25.02	Whiting 2006[2]
Additional cost of untreated upper UTI	£173.72	Whiting 2006[2]
Cost of dialysis per year	£26,585.15	Kerr 2012[8]
Cost of renal transplant	£20,115.17	NHS Reference Costs[10] and NHS Drug Tariff[9]
Cost of septicaemia	£2,163.51	NHS Reference Costs[10]

Quality of life

Quality of life (QoL) values used to populate the model are displayed in Table 20. QoL scores for patients with UTI or no UTI were taken from Bermingham et al (2012)[11]. These values were utility scores for adult women with a UTI, measured using the SF-36 and mapped to the EQ-5D, which were used due to a lack of QoL values for children measured directly with the EQ-5D. These values were applied for the first 14 days of the model (the duration of an untreated upper UTI). For example, a patient with treated upper UTI would have a QoL score of 0.724 (UTI) for the first 10 days of the model, and a score of 0.922 (no UTI) for the following 4 days. The assumption was made that patients with sepsis experience a QoL equivalent to that of UTI for the entire 14 days. In reality, it is likely that the QoL associated with sepsis is lower than that of UTI. However, this assumption is unlikely to substantially affect results, as the vast majority of QALY loss associated with sepsis arises from the risk of mortality.

After the first 14 days of the model, QoL of children without ESRD was sourced from age-specific UK population EQ-5D norms (Kind et al, 1999)[12]. QoL scores for patients with ESRD and for the first year after renal transplant are sourced from Whiting et al (2006)[13]. The assumption is made that patients' QoL returns to that of the general population from the second year after transplant onwards, unless graft failure occurs.

Table 20: Quality of life values used to populate the model

State	QoL (SE)	Source
UTI	0.724 (N/A)	Bermingham 2012[11]
No UTI	0.922 (N/A)	Bermingham 2012[11]
ESRD (on dialysis)	0.604 (0.009)	Wasserfallen 2014[13]
First year after renal transplant	0.73 (0.011)	Cleemput 2004[14]

Sensitivity analysis

In order to characterise the uncertainty surrounding model results, extensive deterministic and probabilistic sensitivity analyses were carried out.

- One-way sensitivity analyses conducted on the following parameters:
- Prevalence of UTI
- Accuracy of dipstick tests
- Additional duration of untreated UTI
- Quality of life associated with UTI
- Cost of microscopy, culture and antibiotic treatment

- Baseline probability of PRS
- Relative risk of PRS in untreated versus treated UTI
- Baseline probability of septicaemia
- Probability of death from septicaemia
- Relative risk of septicaemia in untreated versus treated UTI

In addition, four deterministic scenarios were included in the one-way sensitivity analysis:

- Cost of dipstick test added to the 'no dipstick testing' strategy: This scenario was included to reflect a pathway in which all children receive a dipstick test, but a urine sample is also sent for microscopy and culture regardless of the result.
- Antibiotic adverse events included: This scenario used pessimistic estimates of the potential consequences of antibiotic treatment. This comprised a 0.05% probability of anaphylactic shock, 0.33% probability of death from anaphylactic shock, and a 1% probability of 'other adverse events', which were assumed to cause a reduction in QoL by 0.5 for 3 days.
- Probability of ESRD set to upper-bound from Round et al (2011)[15]: In order to explore the uncertainty in the progression of UTI to ESRD, a pessimistic upper-bound value of 0.65% for the probability of eventually developing ESRD from a UTI was used in the model, rather than using the probability of developing PRS as an intermediate step.
- Probability of death from septicaemia estimated from Schnadower et al (2010)[3]: As the base case analysis for scenarios 2 and 3 used probability of death from bacteraemia which was not necessarily related to UTI, a probability of death was estimated from the Schnadower study. The assumption was made that all deaths in children with UTI in this study resulted from septicaemia, thereby giving an 'upper bound' for the probability of death from septicaemia of 1.6%.

For scenarios 1 and 2, threshold analyses were conducted on the relative risk of PRS and septicaemia for untreated versus treated UTI, to determine at what value the 'no dipstick testing' strategy became cost effective at a threshold of £20,000 per QALY.

Two-way sensitivity analyses were conducted on the following pairs of parameters:

- Relative risk of PRS and baseline probability of PRS (scenario 1)
- Relative risk of septicaemia and baseline probability of septicaemia (scenario 2)
- Relative risk of septicaemia and probability of death from septicaemia (scenario 2)
- Relative risk of PRS and baseline probability of PRs (scenario 3)

For the probabilistic sensitivity analysis, all model input parameters were assigned probability distributions (rather than being expressed as point estimates) to reflect the uncertainty surrounding the available clinical and cost data. 1,000 iterations of the model were run, each drawing random values from parameter distributions.

Probability parameters were assigned beta distributions in order to account for the fact that probability values must lie between 0 and 1. In order to account for the fact that sensitivity and specificity values are typically negatively correlated, accuracy values for dipstick tests were transformed onto the log odds scale (in order to ensure that values could not lie outside of the 0 to 1 range) and assigned a normal distribution, with a Cholesky decomposition used to correlate sensitivity and specificity. The meta-analysis used to synthesise dipstick accuracy values lacked a sufficient number of studies to produce a correlation coefficient, so an assumed value of -0.5 was used.

Cost parameters for which there was uncertainty regarding the point estimate were assigned gamma distributions, to ensure that costs could not be negative. As utilities are bound at 1 but have no lower bound, these values were transformed via the formula: D = 1 – utility. The

resulting D was assigned a gamma distribution (as this value is bound at 0 with no upper limit), and subsequently transformed back into a utility value. Costs which took a specific value (such as the cost of dipsticks) were not varied probabilistically.

Where available, standard errors or 95% confidence intervals were used to inform the shape of distributions. For QOL values for UTI/no UTI a standard error of 0.02 was assumed. For cost parameters with uncertainty regarding the point estimate a standard error of 20% of the parameter mean was assumed.

Results

Infants younger than 3 months

Basic scenario

Base case results for infants younger than 3 months in the basic scenario are shown in Table. These results show that, when the assumption is made that untreated UTI only results in an extra 4 days of symptoms, the 'no dipstick' strategy results in the highest total cost and highest number of QALYs. This is because the 'no dipstick' strategy entails testing all children with microscopy and culture – a more expensive test, but also one with a higher accuracy, which avoids QALY loss due to false negative results. However, due to the relatively minor consequences of an untreated UTI in this scenario, the 'no dipstick' strategy generates a relatively small number of incremental QALYs, resulting in a high ICER of £776,964 compared to a 'dipstick testing' strategy with the presence of LE interpreted as a positive result.

Table21: Basic scenario base case results for infants younger than 3 months

Strategy	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Dipstick - LE	£19.17	25.22166	-	-	-
Dipstick - Nitrite or LE	£19.69	25.22166	£0.52	0.00000	dominated
Dipstick - Nitrite and LE	£21.27	25.22158	£2.10	-0.00007	dominated
Dipstick - Nitrite	£22.03	25.22159	£2.87	-0.00007	dominated
No dipstick	£41.02	25.22169	£21.85	0.00003	£776,964

Results of one-way sensitivity analyses for the base case scenario are shown in Table 16. These results show that, for the basic scenario, the ICER of the 'no dipstick' strategy remains well above the NICE cost-effectiveness threshold for all sensitivity analysis scenarios.

The ICER is relatively sensitive to changes in the baseline prevalence of UTI, although this parameter would have to take an extremely high value for 'no dipstick' to be considered cost-effective. Similarly, setting the additional duration of an untreated UTI, accuracy of dipstick tests, and QOL of UTI to extreme values still results in an ICER well in excess of £20,000 per QALY.

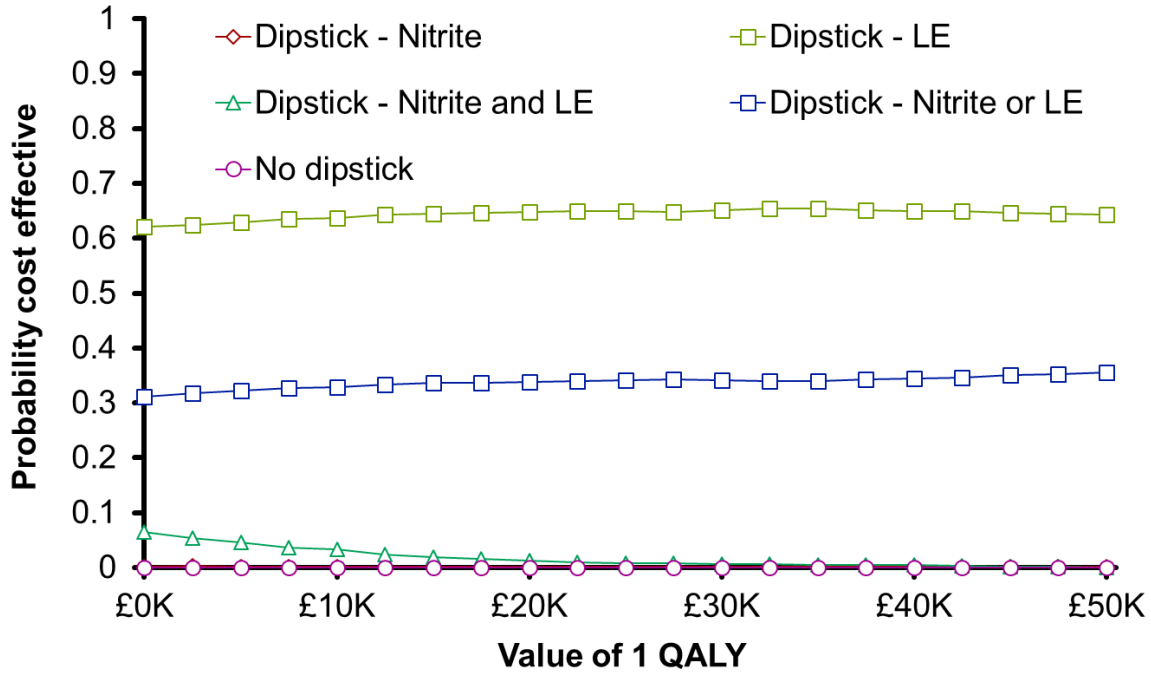
Including the cost of a dipstick test in the 'no dipstick testing' strategy (in order to represent a scenario in which all children with suspected UTI receive a dipstick test, followed by microscopy/culture regardless of the result) results in a relatively small increase in the ICER to £781,228 per QALY. Including a pessimistic estimation of adverse events associated with antibiotic treatment results in the 'no dipstick' strategy being dominated by 'dipstick LE', due to the QALY loss associated with treating all patients with antibiotics in the former scenario.

Table 16: Basic scenario one-way sensitivity analysis results for infants under 3 months

Scenario	Δ Costs – 'No dipstick' versus 'dipstick LE'	Δ QALYs – 'No dipstick' versus 'dipstick LE'	ICER – 'No dipstick' versus 'dipstick LE'
Base case	£21.85	0.00003	£776,964
UTI prevalence set to 1%	£24.86	0.00000	£6,365,712
UTI prevalence set to 25%	£13.20	0.00010	£135,159
Additional duration of untreated UTI set to 20 days	£21.85	0.00014	£155,393
Accuracy of dipstick tests set to lower 95% CI	£17.84	0.00003	£519,134
Accuracy of dipstick tests set to upper 95% CI	£23.54	0.00002	£1,369,546
Quality of life of UTI set to 0.1	£21.85	0.00004	£505,916
Cost of microscopy, culture and antibiotic treatment doubled	£45.48	0.00003	£1,617,430
'No dipstick' strategy also associated with the cost of a dipstick test	£21.97	0.00003	£781,228
Antibiotic adverse events included	£21.85	-0.00004	Dipstick dominates no dipstick

Probabilistic sensitivity analysis results are shown as a cost-effectiveness acceptability curve (CEAC) in Figure 8. These results are consistent with those of the one-way sensitivity analysis; they show that, for a scenario in which an untreated UTI only extends the duration of symptoms, the 'no dipstick' strategy has a negligible probability of being the most cost effective strategy at a threshold of £20,000 per QALY. 'Dipstick – LE' and 'Dipstick – nitrite or LE' are the strategies with the highest probability of being cost-effective across all thresholds.

Figure 8: Cost-effectiveness acceptability curve of basic scenario results for infants under 3 months



Scenario 1: Untreated UTI associated with an increased risk of PRS

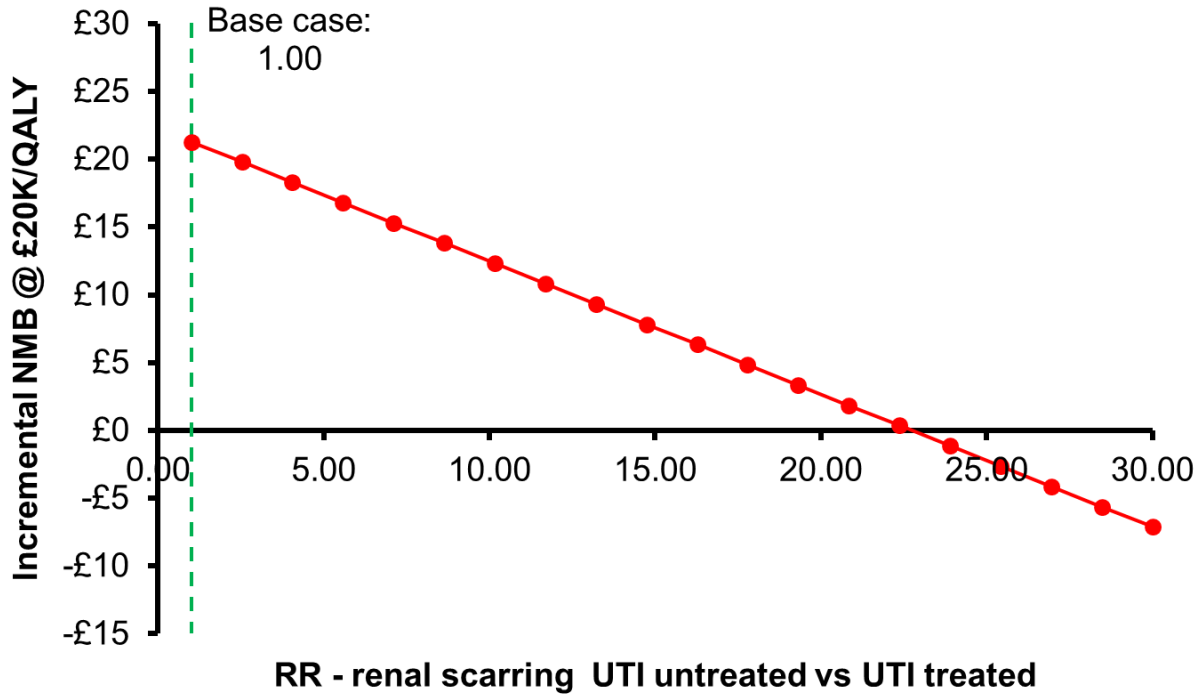
Base case results for the scenario in which untreated UTI is associated with an increased risk of PRS are shown in Table 23. The results show that including this assumption in the analysis reduces the ICER of 'no dipstick' compared to 'dipstick – LE' to a value of £334,327 per QALY.

Table 23: Scenario 1 base case results for infants under 3 months

Strategy	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Dipstick - LE	£19.42	25.22162	-	-	-
Dipstick - Nitrite or LE	£19.94	25.22162	£0.52	0.00000	dominated
Dipstick - Nitrite and LE	£22.19	25.22145	£2.77	-0.00017	dominated
Dipstick - Nitrite	£22.91	25.22146	£3.49	-0.00016	dominated
No dipstick	£41.02	25.22169	£21.60	0.00006	£334,327

A threshold analysis of the incremental net monetary benefit of 'dipstick – LE' versus 'no dipstick' over a range of values for the relative risk of PRS in children with untreated versus treated UTI is shown in Figure 9. These results show that the relative risk of PRS would have to be substantially higher – over 20 – for 'no dipstick' to be considered cost-effective in this scenario.

Figure 9: Scenario 1 threshold analysis for infants under 3 months – plotting relative risk of renal scarring against incremental net monetary benefit of 'dipstick' versus 'no dipstick'



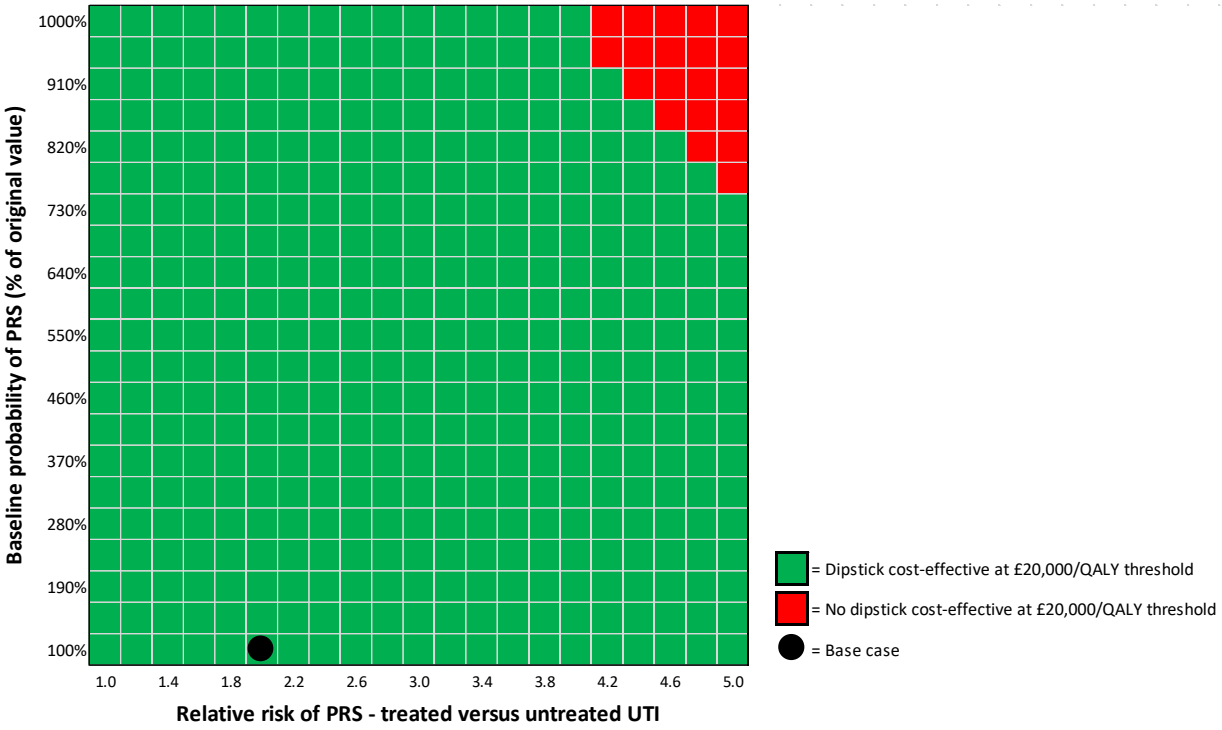
One-way sensitivity analysis results for the scenario in which untreated UTI is associated with an increased risk of PRS are shown in Table 24. These results show that varying the baseline prevalence of UTI, probability of PRS, and relative risk of PRS for untreated versus treated UTI substantially affects the ICER, although none of the analyses result in an ICER at which the 'no dipstick' strategy could be considered cost-effective at a £20,000 per QALY threshold.

Table 24: Scenario 1 one-way sensitivity analysis results for infants under 3 months

Scenario	Δ Costs – 'No dipstick' versus 'dipstick LE'	Δ QALYs – 'No dipstick' versus 'dipstick LE'	ICER – 'No dipstick' versus 'dipstick LE'
Base case	£21.60	0.00006	£334,327
UTI prevalence set to 1%	£24.83	0.00001	£2,766,974
UTI prevalence set to 25%	£12.33	0.00022	£54,965
Baseline probability of PRS halved relative to base case	£21.72	0.00005	£468,566
Baseline probability of PRS doubled relative to base case	£21.35	0.00010	£211,194
Probability of ESRD set to upper bound from Round 2012	£19.62	0.00035	£55,509
Relative risk of PRS for untreated UTI versus treated UTI set to 1.2	£21.80	0.00004	£615,483
Relative risk of PRS for untreated UTI versus treated UTI set to 4	£21.10	0.00014	£153,370
Cost of microscopy, culture and antibiotic treatment doubled	£45.23	0.00006	£700,162
Antibiotic adverse events included	£21.60	-0.00001	Dipstick dominates no dipstick

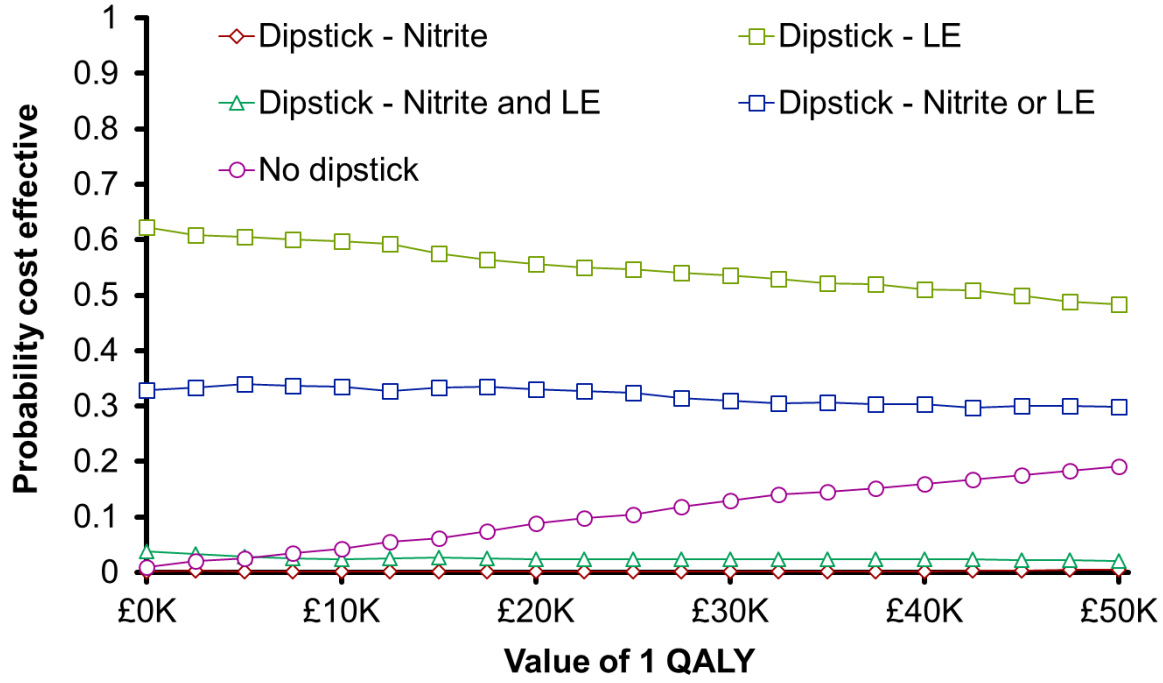
A two-way sensitivity analysis, in which both baseline probability of PRS and relative risk of PRS in untreated versus treated UTI are varied is displayed in Figure 10. This figure further shows that the incidence of PRS or added risk of PRS from an untreated UTI would need to be substantially higher for 'no dipstick' to be considered cost effective.

Figure 10: Scenario 1 two-way sensitivity analysis results for infants under 3 months – plotting relative risk of PRS against baseline probability of PRS



Probabilistic sensitivity analysis results are shown as a CEAC in Figure 11. These results show that, for this scenario, the 'no dipstick' strategy has a low probability of being cost-effective at a threshold of £20,000 per QALY.

Figure 11: Scenario 1 cost-effectiveness acceptability curve for infants under 3 months



Scenario 2: Untreated UTI associated with an increased risk of septicaemia

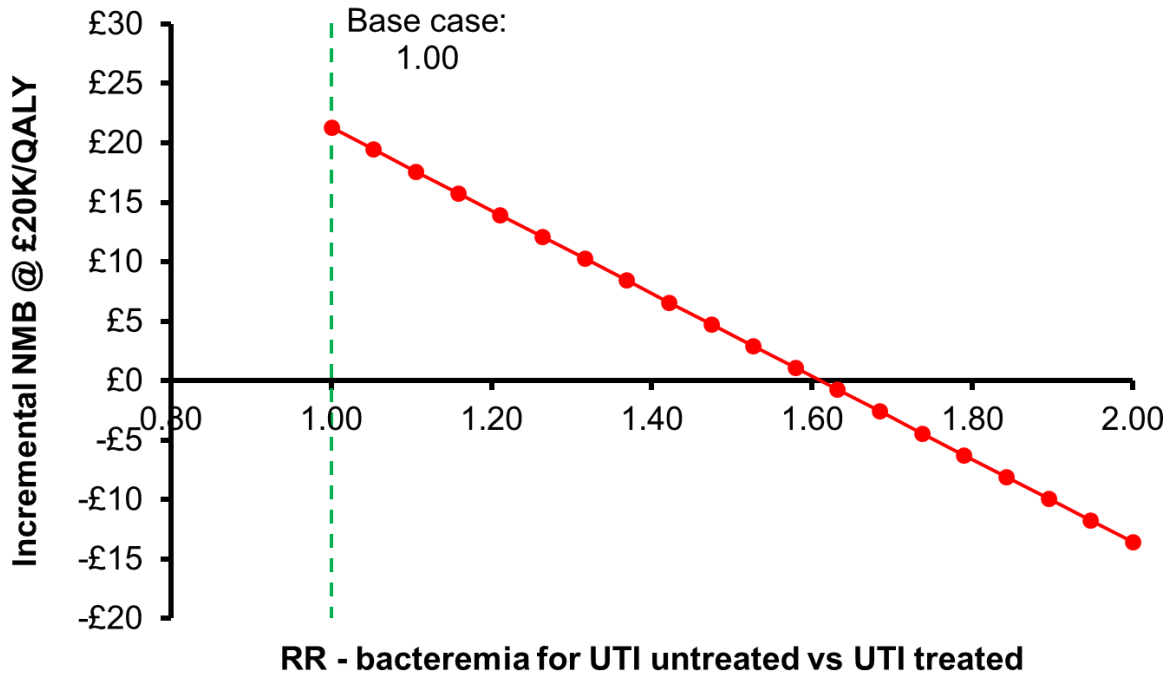
Results for the scenario in which untreated UTI is associated with an increased risk of sepsis are shown in Table 25. These results show that including this assumption in the analysis considerably reduces the ICER of 'no dipstick' compared to 'dipstick – LE' to a value of £11,914 per QALY. This is because septicaemia is associated with a risk of death in the model, meaning that the expected QALY loss associated with an untreated case of UTI is considerably higher.

Table 25: Scenario 2 base case results for infants under 3 months

Strategy	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Dipstick - LE	£21.00	25.22001	-	-	-
Dipstick - Nitrite or LE	£21.52	25.22001	£0.52	0.00000	dominated
Dipstick - Nitrite and LE	£28.00	25.21553	£7.00	-0.00448	dominated
Dipstick - Nitrite	£28.46	25.21581	£7.46	-0.00420	dominated
No dipstick	£41.02	25.22169	£20.01	0.00168	£11,914

A threshold analysis of the incremental net monetary benefit of 'dipstick – LE' versus 'no dipstick' over a range of values for the relative risk of septicaemia in children with untreated versus treated UTI is shown in Figure 12. These results show that, while the 'no dipstick' strategy is cost-effective with the arbitrarily chosen relative risk of 2 used in the base case, reducing this value to below 1.6 would result in an ICER of above £20,000 per QALY.

Figure 12: Scenario 2 threshold analysis for infants under 3 months – plotting relative risk of PRS against incremental net monetary benefit of 'dipstick' versus 'no dipstick'



One-way sensitivity analysis results for the scenario in which untreated UTI is associated with an increased risk of septicaemia are shown in Table 26. These results show that, in contrast to the previous scenario, the cost-effectiveness of the 'no dipstick' strategy at a threshold of £20,000 per QALY is sensitive to changes in parameters. Specifically, a reduction in the prevalence of UTI, baseline prevalence of bacteraemia, or relative risk of septicaemia in untreated versus treated UTI results in a considerably higher ICER.

Table 26: Scenario 2 one-way sensitivity analysis results for infants under 3 months

Scenario	Δ Costs – 'No dipstick' versus 'dipstick LE'	Δ QALYs – 'No dipstick' versus 'dipstick LE'	ICER – 'No dipstick' versus 'dipstick LE'
Base case	£20.01	0.00168	£11,914
UTI prevalence set to 1%	£24.61	0.00023	£105,471
UTI prevalence set to 25%	£6.82	0.00583	£1,170
Baseline probability of bacteraemia set to 1%	£21.57	0.00028	£76,983
Baseline probability of bacteraemia set to 20%	£16.25	0.00507	£3,205
Probability of death from bacteraemia set to 1%	£20.01	0.00024	£82,667
Probability of death from bacteraemia set to 20%	£20.02	0.00431	£4,647
Probability of death from bacteraemia estimated from Schnadower 2010 (1.6%)	£20.01	0.00037	£54,021
Relative risk of bacteraemia for untreated UTI versus treated UTI set to 1.2	£21.48	0.00036	£59,927
Relative risk of bacteraemia for untreated UTI versus treated UTI set to 4	£16.34	0.00498	£3,279
Cost of microscopy, culture and antibiotic treatment doubled	£43.65	0.00168	£25,983
Cost of septicaemia doubled	£18.18	0.00168	£10,820
Antibiotic adverse events included	£20.01	0.00161	£12,430

Two-way sensitivity analyses, in which relative risk of septicaemia is varied simultaneously with with basline probability of septicaemia, and with probability of septicaemia, are shown in Figure 13 and Figure 14. These figures demonstrate that, while the 'no dipstick' strategy is cost effective in the base case, a relatively small reduction in any of these three parameters results in the strategy no longer being cost-effective at a £20,000 per QALY threshold.

Figure 13: Scenario 2 two-way sensitivity analysis for infants under 3 months – plotting relative risk of septicaemia against baseline risk of septicaemia

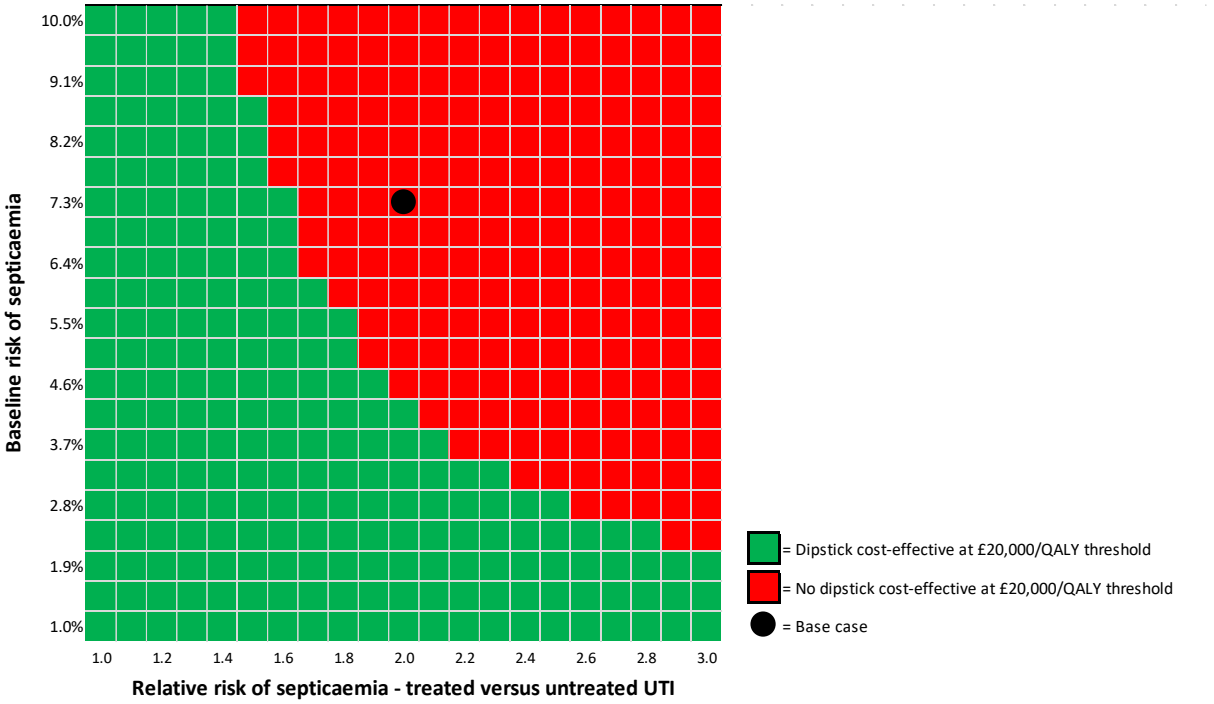
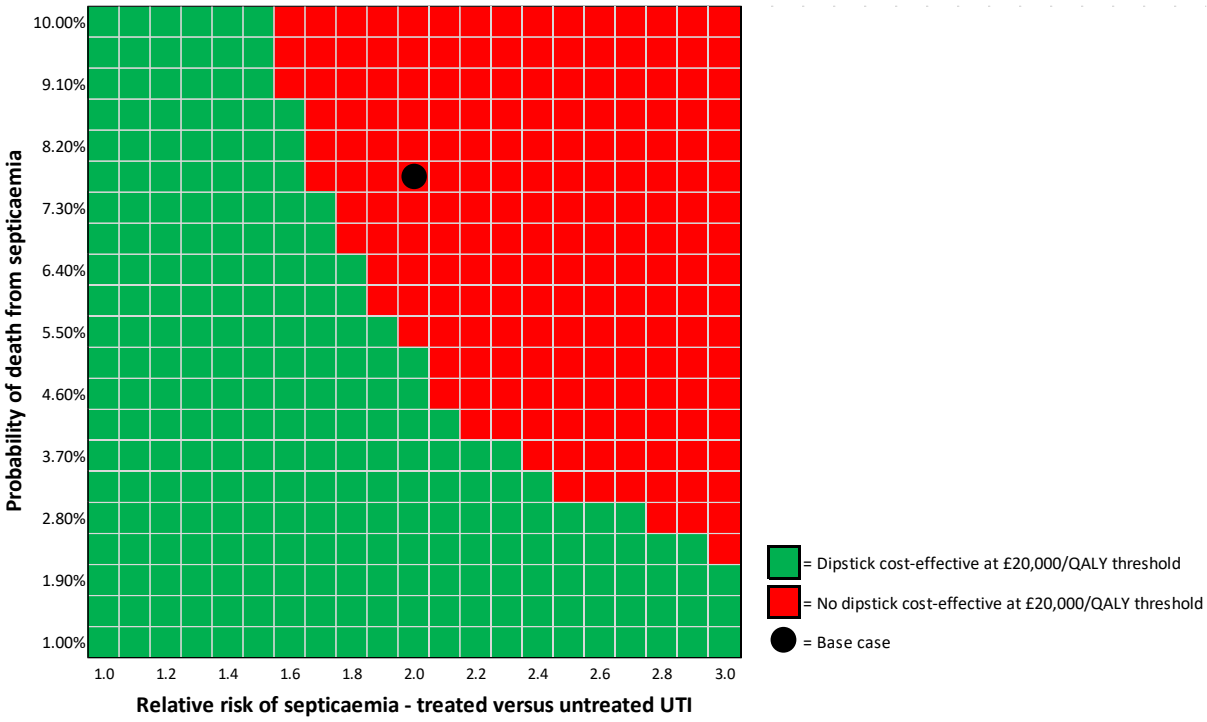
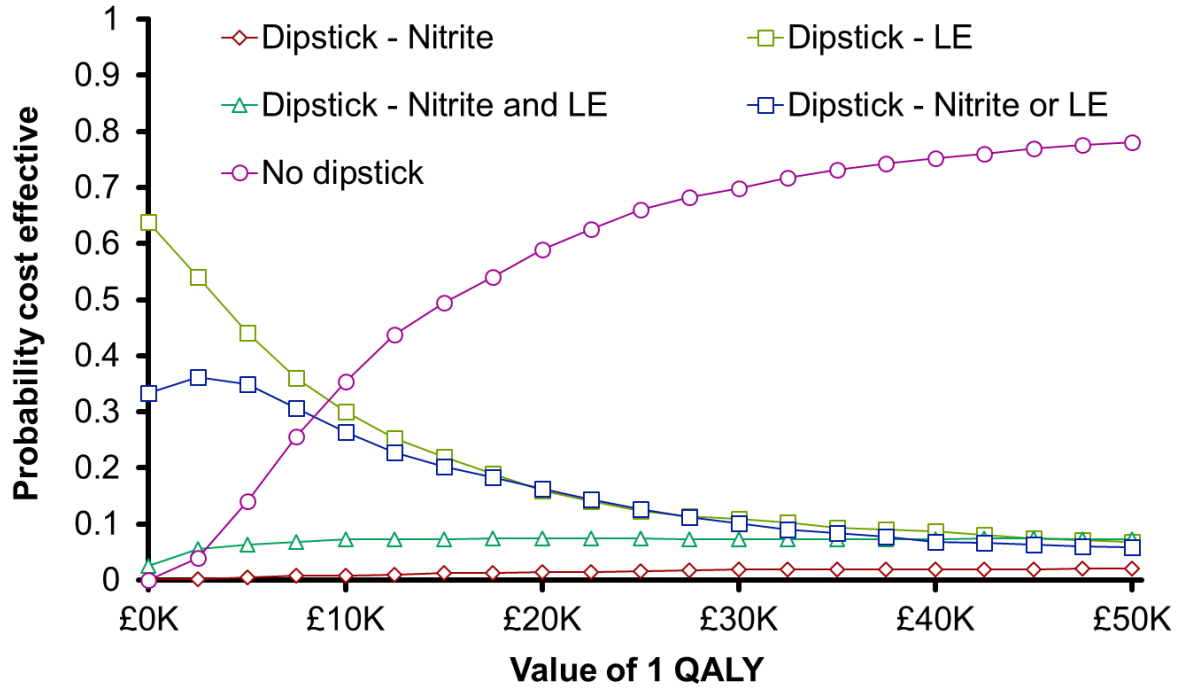


Figure 14: Scenario 2 two-way sensitivity analysis for infants under 3 months – plotting relative risks of septicaemia against probability of death from septicaemia



Probabilistic sensitivity analysis results are shown as a CEAC in Figure 15. These results show that while the 'no dipstick' strategy has the highest probability of being cost effective at a threshold of £20,000 per QALY, there is considerable uncertainty surrounding this result.

Figure 15: Scenario 2 cost-effectiveness acceptability curve for infants under 3 months



Scenario 3: Untreated UTI associated with an increased risk of septicaemia and PRS

Results for the scenario in which untreated UTI is associated with both an increased risk of sepsis and an increased risk of PRS is shown in Table 27. These results show that including both of these assumptions results in an ICER of £11,517 for 'no dipstick' compared to 'dipstick – LE'.

Table 27: Scenario 3 base case results for infants under 3 months

Strategy	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Dipstick - LE	£21.25	25.21997	-	-	-
Dipstick - Nitrite or LE	£21.77	25.21997	£0.52	0.00000	dominated
Dipstick - Nitrite and LE	£28.92	25.21539	£7.66	-0.00458	dominated
Dipstick - Nitrite	£29.33	25.21568	£8.08	-0.00429	dominated
No dipstick	£41.02	25.22169	£19.76	0.00172	£11,517

One-way sensitivity analysis results for the scenario in which untreated UTI is associated with both an increased risk of septicaemia and PRS are shown in Table 28. These results show that, as with scenario 2, the ICER is sensitive to variation in the prevalence of UTI, baseline probability of bacteraemia, and relative risk of bacteraemia in untreated versus treated UTI. The sensitivity analyses in which these parameters are lowered result in the 'no dipstick' strategy no longer being cost effective at a threshold of £20,000 per QALY. Contrastingly, the ICER is relatively insensitive to changes in the baseline incidence of PRS and relative risk of PRS in untreated versus treated UTI.

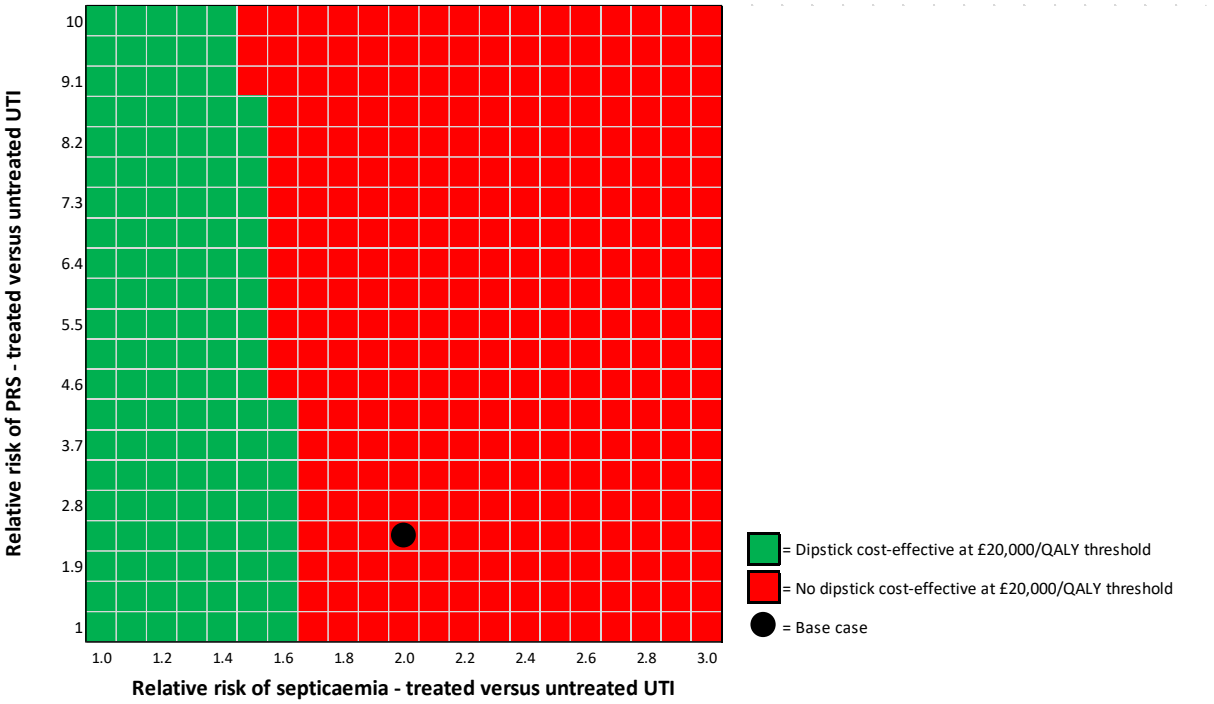
Table 28: Scenario 3 one-way sensitivity analysis results for infants under 3 months

Scenario	Δ Costs – 'No dipstick' versus 'dipstick LE'	Δ QALYs – 'No dipstick' versus 'dipstick LE'	ICER – 'No dipstick' versus 'dipstick LE'
Base case	£19.76	0.00172	£11,517
UTI prevalence set to 1%	£24.57	0.00024	£103,095
UTI prevalence set to 25%	£9.84	0.00477	£2,064
Baseline probability of bacteraemia set to 1%	£21.32	0.00032	£67,294
Baseline probability of bacteraemia set to 20%	£16.00	0.00510	£3,135
Probability of death from bacteraemia set to 1%	£19.76	0.00028	£70,905
Probability of death from bacteraemia set to 20%	£19.77	0.00434	£4,552
Probability of death from bacteraemia estimated from Schnadower 2010 (1.6%)	£19.76	0.00041	£48,549
Baseline probability of PRS halved relative to base case	£19.89	0.00170	£11,712
Baseline probability of PRS doubled relative to base case	£19.52	0.00175	£11,138
Relative risk of bacteraemia for untreated UTI versus treated UTI set to 1.2	£21.23	0.00039	£53,765
Relative risk of bacteraemia for untreated UTI versus treated UTI set to 4	£16.09	0.00502	£3,207
Relative risk of PRS for untreated UTI versus treated UTI set to 1.2	£19.96	0.00169	£11,833
Relative risk of PRS for untreated UTI versus treated UTI set to 4	£19.27	0.00179	£10,771
Cost of microscopy, culture and antibiotic treatment doubled	£43.40	0.00172	£25,289

Scenario	Δ Costs – 'No dipstick' versus 'dipstick LE'	Δ QALYs – 'No dipstick' versus 'dipstick LE'	ICER – 'No dipstick' versus 'dipstick LE'
Cost of septicaemia doubled	£17.93	0.00172	£10,446
Antibiotic adverse events included	£19.76	0.00165	£12,005

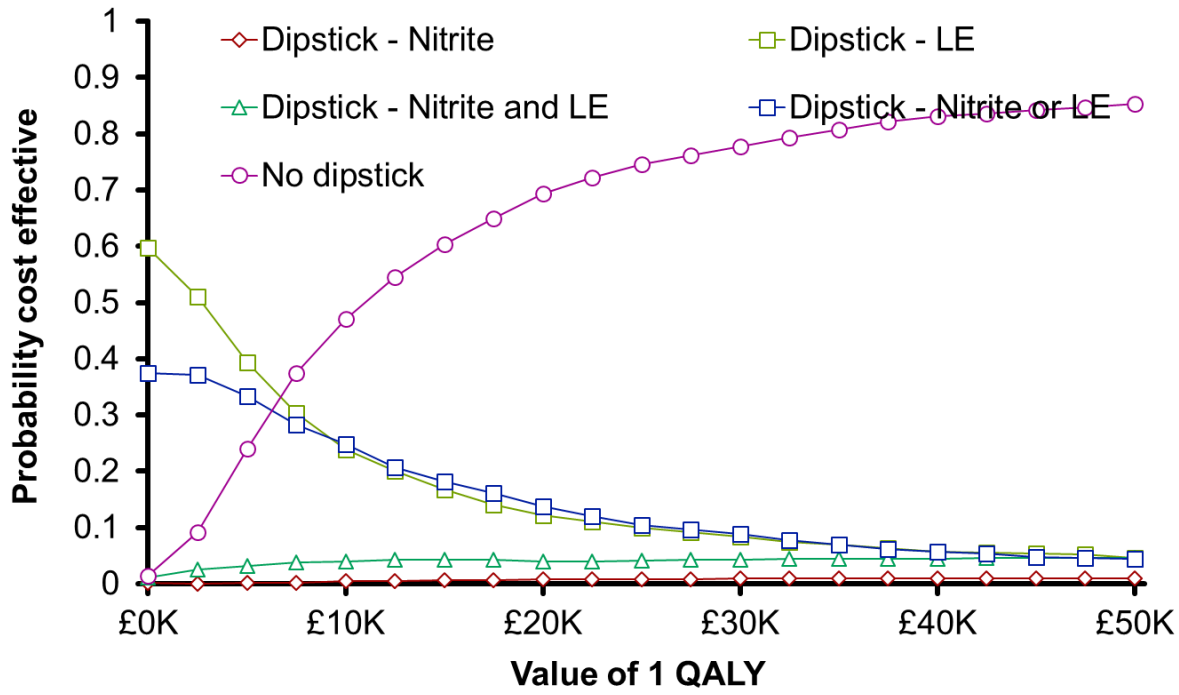
A two-way sensitivity analysis, in which both the relative risk of septicaemia and relative risk of PRS are varied simultaneously, is shown in Figure 16. These results demonstrate that the cost-effectiveness of the 'no dipstick' strategy is relatively sensitive to changes in the relative risk of septicaemia, and comparatively much less sensitive to changes in the relative risk of PRS.

Figure 16: Scenario 3 two-way sensitivity analysis for infants under 3 months – relative risk of septicaemia versus relative risk of PRS



Probabilistic sensitivity analysis results are shown as a CEAC in Figure 17. These results show that, for this scenario, the 'no dipstick' strategy is likely to be cost-effective strategy at a threshold of £20,000 per QALY.

Figure 17: Scenario 3 cost-effectiveness acceptability curve for infants under 3 months



Children 3 months or older but younger than 3 years

Basic scenario

Cost-effectiveness results for children 3 months or older but younger than 3 years in the basic scenario are shown in Table 29. As with the under 3 months population, the 'no dipstick' strategy results in the highest overall costs and highest number of QALYs, due to all children being tested with microscopy and culture – a more accurate yet more costly test. However, due to the relatively minor consequences of an untreated UTI in this scenario, the 'no dipstick' strategy is associated with a high ICER of £849,353 compared to a 'dipstick testing' strategy with the presence of nitrite or LE interpreted as a positive result.

Table 29: Basic scenario base case results for children 3 months or older but younger than 3 years

Strategy	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Dipstick - Nitrite or LE	£7.93	25.14975	-	-	-
Dipstick - Nitrite	£8.19	25.14972	£0.26	-0.00003	dominated
Dipstick - LE	£9.78	25.14975	£1.85	0.00000	dominated
Dipstick - Nitrite and LE	£10.03	25.14967	£2.11	-0.00007	dominated
No dipstick	£30.82	25.14977	£22.89	0.00003	£849,353

Results of one-way sensitivity analyses for the base case scenario are shown in Table . These results show that, although the ICER is relatively sensitive to changes in the prevalence of UTI, additional duration of untreated UTI, accuracy of dipstick tests, and QOL of patients with UTI, the ICER of the 'no dipstick' strategy remain well above the threshold of £20,000 per QALY in all scenarios.

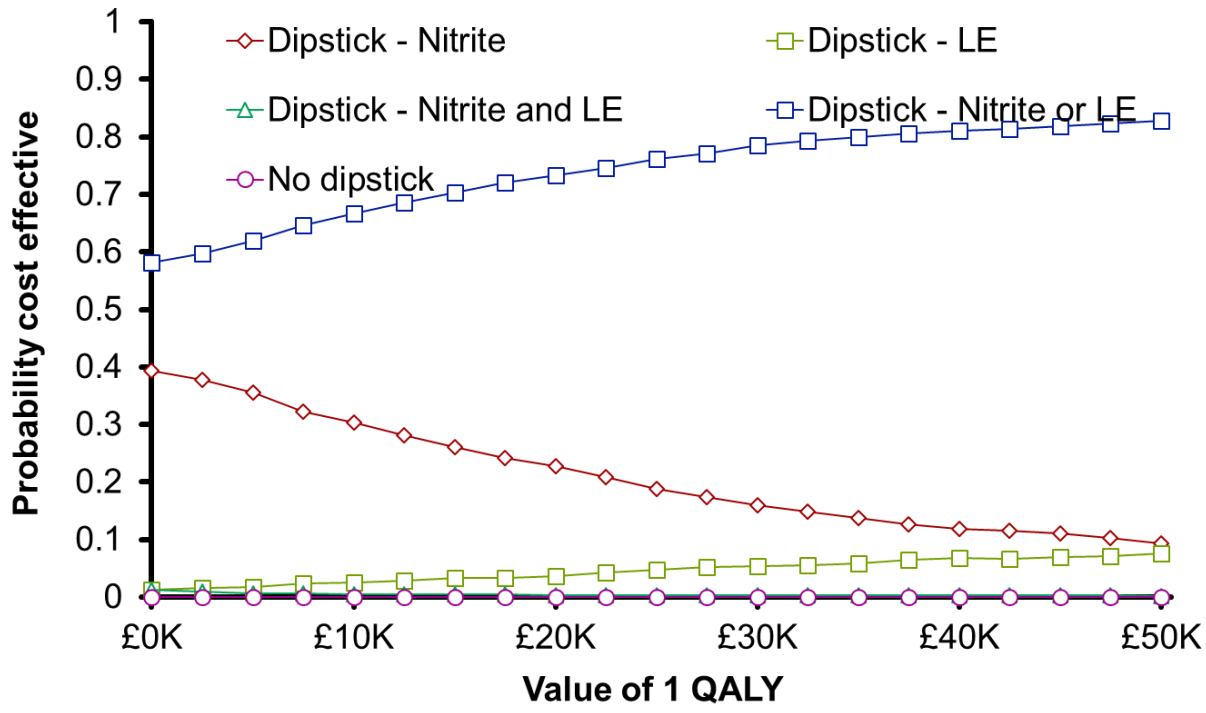
As with the population under 3 months, adding the cost of a dipstick test to the 'no dipstick' strategy (to reflect a scenario in which all children are tested with dipstick followed by microscopy/culture regardless of the result) does not substantially affect the ICER.

Table 30: Basic scenario one-way sensitivity analysis results for children 3 months or older but younger than 3 years

Scenario	Δ Costs – 'No dipstick' versus 'dipstick nitrite or LE'	Δ QALYs – 'No dipstick' versus 'dipstick nitrite or LE'	ICER – 'No dipstick' versus 'dipstick nitrite or LE'
Base case	£22.89	0.00003	£849,353
UTI prevalence set to 1%	£25.12	0.00000	£5,033,701
UTI prevalence set to 25%	£12.95	0.00012	£103,778
Additional duration of untreated UTI set to 20 days	£22.89	0.00013	£169,871
Accuracy of dipstick tests set to lower 95% CI	£19.68	0.00003	£599,976
Accuracy of dipstick tests set to upper 95% CI	£23.39	0.00001	£1,995,969
Quality of life of UTI set to 0.1	£22.89	0.00005	£416,349
Cost of microscopy, culture and antibiotic treatment doubled	£47.25	0.00003	£1,753,196
'No dipstick' strategy also associated with the cost of a dipstick test	£23.01	0.00003	£853,802
Antibiotic adverse events included*	£22.89	-0.00004	Dipstick dominates no dipstick

Results of probabilistic sensitivity analysis (displayed in Figure 18) show that the 'no dipstick' strategy has a negligible probability of being cost-effective at a threshold of £20,000 per QALY. 'Dipstick – nitrite or LE' has the highest probability of being cost-effective at this threshold, followed by 'dipstick – Nitrite'.

Figure 18: Basic scenario cost-effectiveness acceptability curve for children 3 months or older but younger than 3 years



Scenario 1: Untreated UTI associated with an increased risk of PRS

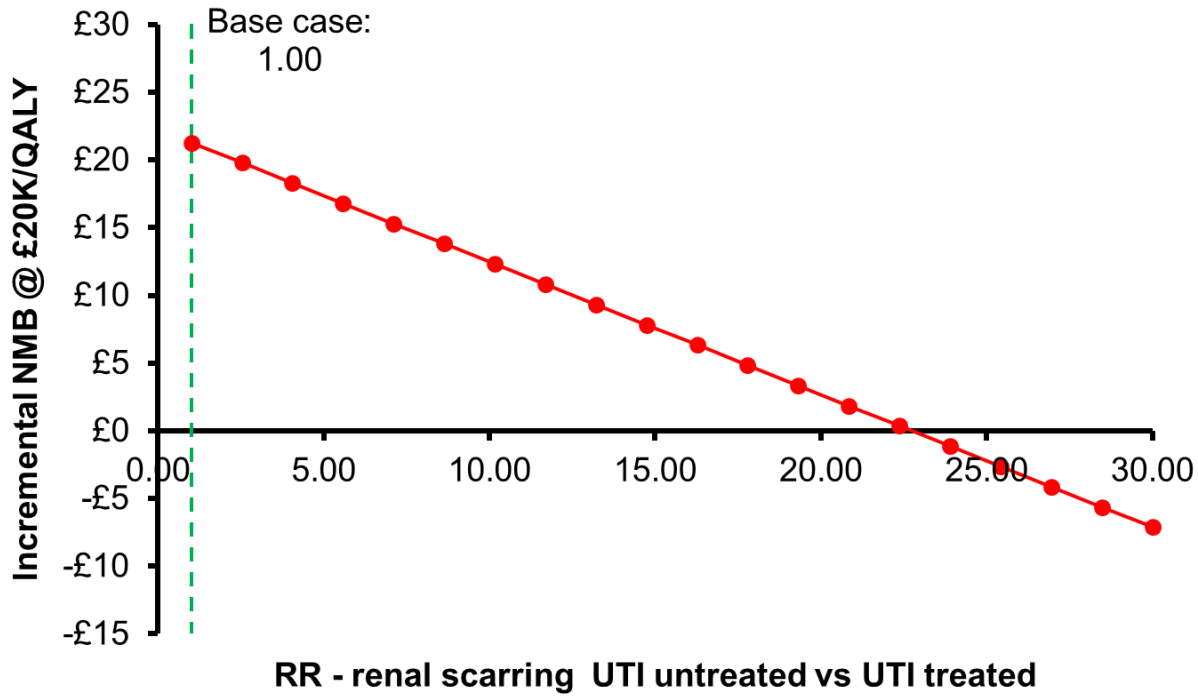
Results for the scenario in which untreated UTI is associated with an increased risk of PRS are shown in Table 31. The results show that including this assumption in the analysis reduces the ICER of 'no dipstick' compared to 'dipstick – LE' to a value of £364,766 per QALY.

Table 31: Scenario 1 base case results for children 3 months or older but younger than 3 years

Strategy	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Dipstick - Nitrite or LE	£8.17	25.14971	-	-	-
Dipstick - Nitrite	£8.71	25.14964	£0.54	-0.00007	dominated
Dipstick - LE	£10.02	25.14971	£1.85	0.00000	dominated
Dipstick - Nitrite and LE	£10.93	25.14954	£2.76	-0.00017	dominated
No dipstick	£30.82	25.14977	£22.65	0.00006	£364,766

A threshold analysis of the incremental net monetary benefit of 'dipstick – LE' versus 'no dipstick' over a range of values for the relative risk of PRS in children with untreated versus treated UTI is shown in Figure 19. These results show that the relative risk of PRS would have to be substantially higher – over 20 – for 'no dipstick' to be considered cost-effective in this scenario.

Figure 19: Scenario 1 threshold analysis for children 3 months or older but younger than 3 years – relative risk of PRS versus incremental net monetary benefit of 'dipstick' compared to 'no dipstick'



One-way sensitivity analysis results for the scenario in which untreated UTI is associated with an increased risk of PRS are shown in Table 32. These results show that varying the baseline prevalence of UTI, probability of PRS, and relative risk of PRS for untreated versus treated UTI substantially affects the ICER, although none of the analyses result in an ICER at which the 'no dipstick' strategy could be considered cost-effective at a £20,000 per QALY threshold.

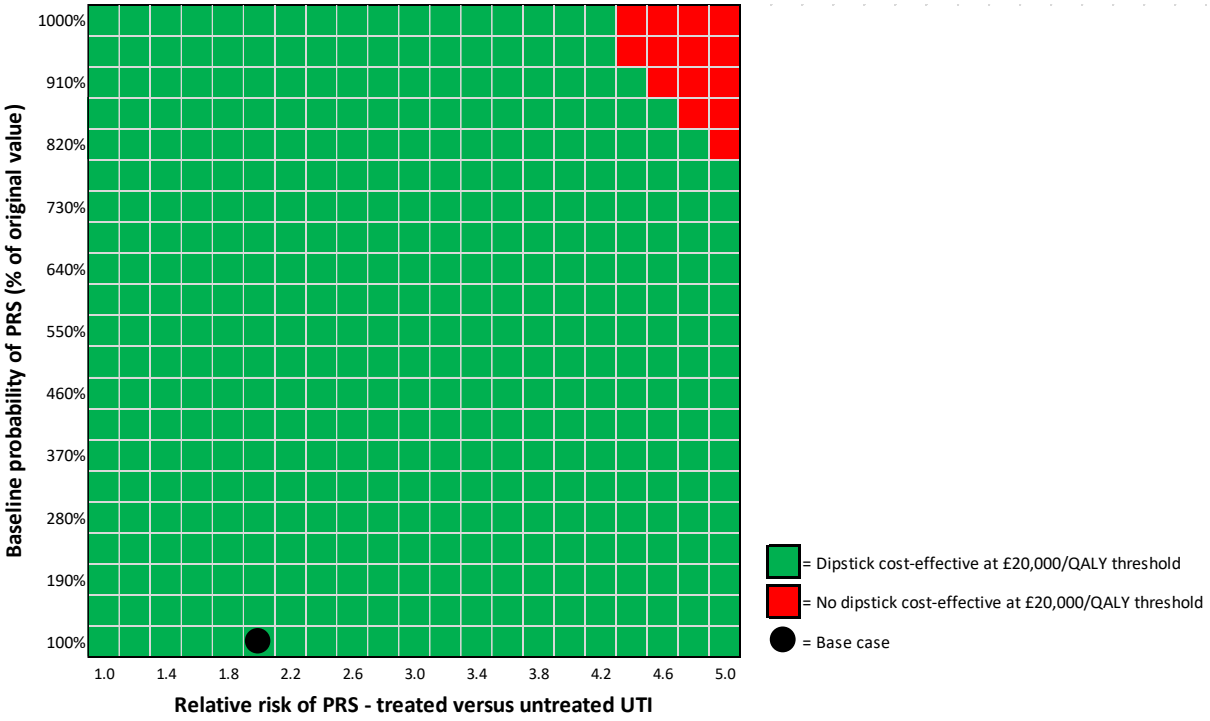
Table 32: Scenario 1 one-way sensitivity analysis results for children 3 months or older but younger than 3 years

Scenario	Δ Costs – 'No dipstick' versus 'dipstick nitrite or LE'	Δ QALYs – 'No dipstick' versus 'dipstick nitrite or LE'	ICER – 'No dipstick' versus 'dipstick nitrite or LE'
Base case	£22.65	0.00006	£364,766
UTI prevalence set to 1%	£25.08	0.00001	£2,180,882
UTI prevalence set to 25%	£11.83	0.00029	£41,168
Baseline probability of PRS halved relative to base case	£22.77	0.00004	£511,433
Baseline probability of PRS doubled relative to base case	£22.41	0.00010	£230,458
Probability of ESRD set to upper bound from Round 2012	£20.59	0.00036	£56,797
Relative risk of PRS for untreated UTI versus treated UTI set to 1.2	£22.84	0.00003	£672,245

Scenario	Δ Costs – 'No dipstick' versus 'dipstick nitrite or LE'	Δ QALYs – 'No dipstick' versus 'dipstick nitrite or LE'	ICER – 'No dipstick' versus 'dipstick nitrite or LE'
Relative risk of PRS for untreated UTI versus treated UTI set to 4	£22.17	0.00013	£167,460
Cost of microscopy, culture and antibiotic treatment doubled	£47.01	0.00006	£757,058
Antibiotic adverse events included	£22.65	-0.00001	Dipstick dominates no dipstick

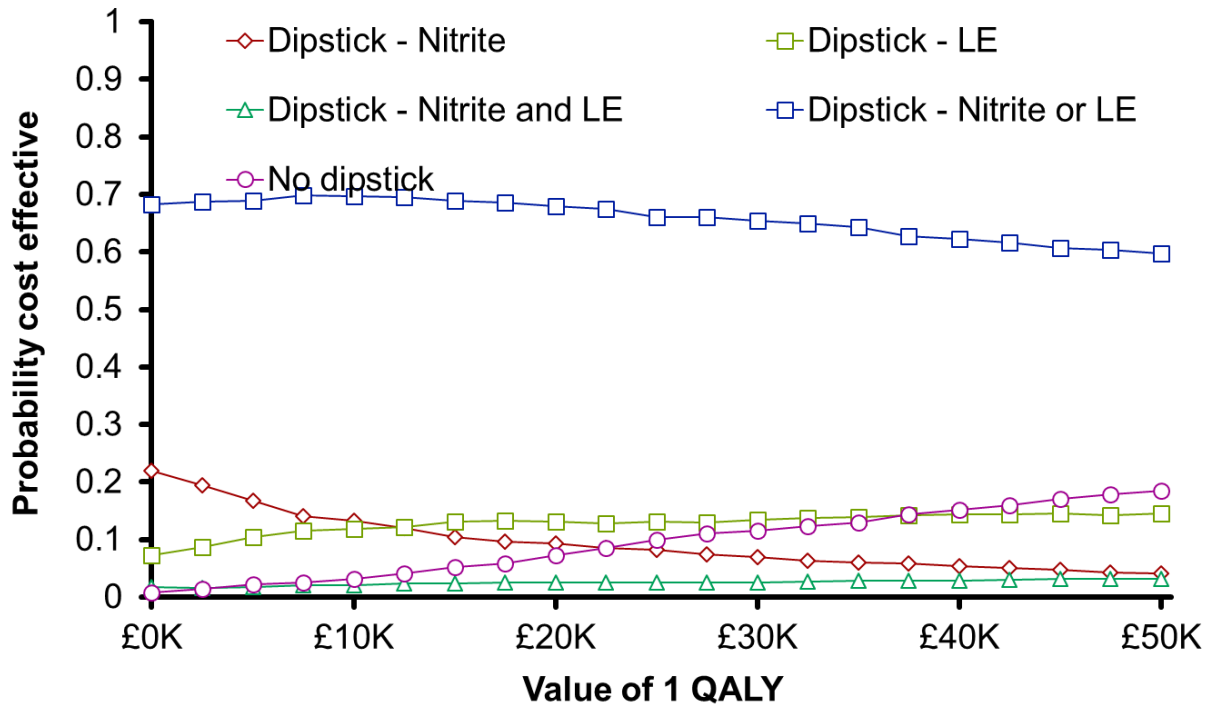
A two-way sensitivity analysis, in which both baseline probability of PRS and relative risk of PRS in untreated versus treated UTI are varied is displayed in Figure 20. This figure further shows that the incidence of PRS or added risk of PRS from an untreated UTI would need to be substantially higher for 'no dipstick' to be considered cost effective.

Figure 20: Scenario 1 two-way sensitivity analysis for children 3 months or older but younger than 3 years – relative risk of PRS versus baseline probability of PRS



Results of probabilistic sensitivity analysis (displayed in Figure 21) show that the 'no dipstick' strategy has a low probability (less than 5%) of being cost-effective at a threshold of £20,000 per QALY, with 'dipstick – nitrite or LE' showing the highest probability of being cost-effective.

Figure 21: Scenario 1 cost-effectiveness acceptability curve for children 3 months or older but younger than 3 years



Scenario 2: Untreated UTI associated with an increased risk of septicaemia

Results for the scenario in which untreated UTI is associated with an increased risk of sepsis are shown in Table 33. These results show that including this assumption in the analysis considerably reduces the ICER of 'no dipstick' compared to 'dipstick – LE' to a value of £172,917 per QALY. This is because septicaemia is associated with a risk of death in the model, meaning that the expected QALY loss associated with an untreated case of UTI is considerably higher.

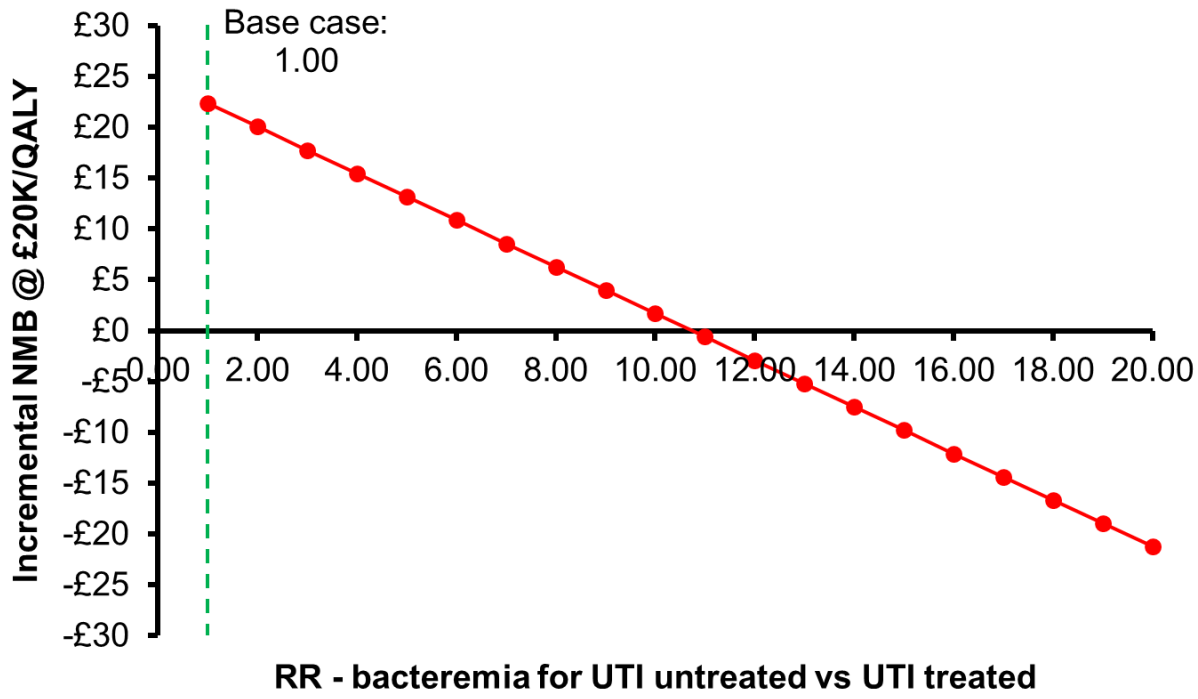
Table 33: Scenario 2 base case results for children 3 months or older but younger than 3 years

Strategy	Costs	QALYs	Δ Costs – 'No dipstick' versus 'dipstick nitrite or LE'	Δ QALYs – 'No dipstick' versus 'dipstick nitrite or LE'	ICER – 'No dipstick' versus 'dipstick nitrite or LE'
Dipstick - Nitrite or LE	£8.14	25.14964	-	-	-
Dipstick - Nitrite	£8.64	25.14949	£0.51	-0.00015	dominated
Dipstick - LE	£9.99	25.14964	£1.85	0.00000	dominated
Dipstick - Nitrite and LE	£10.82	25.14928	£2.68	-0.00036	dominated
No dipstick	£30.82	25.14977	£22.68	0.00013	£172,917

A threshold analysis of the incremental net monetary benefit of 'dipstick – LE' versus 'no dipstick' over a range of values for the relative risk of PRS in children with untreated versus treated UTI is shown in Figure 22. These results show that, in contrast to the equivalent scenario in children under 3 months, the relative risk of bacteraemia would have to be substantially higher (over 10) for 'no dipstick' to be considered cost-effective.

These results show that the relative risk of PRS would have to be substantially higher – over 20 – for 'no dipstick' to be considered cost-effective in this scenario.

Figure 22: Scenario 2 threshold analysis for children 3 months or older but younger than 3 years – relative risk of septicaemia versus incremental net monetary benefit of 'dipstick' compared to 'no dipstick'



One-way sensitivity analysis results for the scenario in which untreated UTI is associated with an increased risk of septicaemia are shown in Table 34. These results show that, unlike in the population of infants younger than 3 months in most scenarios, the ICER of 'no dipstick' remains well above the threshold of £20,000 per QALY, due to the lower baseline prevalence of bacteraemia in children 3 months or older but younger than 3 years.

In two scenarios the ICER of 'no dipstick' is below or close to the £20,000 per QALY threshold: the scenario in which UTI prevalence is set to 25% and the scenario in which the baseline probability of bacteraemia is set to 20%. However, even when the relative risk of bacteraemia for untreated UTI versus treated UTI is set to 4 the ICER still remains well in excess of the threshold.

Table 34: Scenario 2 one-way sensitivity analysis results for children 3 months or older but younger than 3 years

Scenario	Δ Costs – 'No dipstick' versus 'dipstick nitrite or LE'	Δ QALYs – 'No dipstick' versus 'dipstick nitrite or LE'	ICER – 'No dipstick' versus 'dipstick nitrite or LE'
Base case	£22.68	0.00013	£172,917
UTI prevalence set to 1%	£25.08	0.00002	£1,032,696
UTI prevalence set to 25%	£11.97	0.00061	£19,720
Baseline probability of bacteraemia set to 0.1%	£22.86	0.00004	£567,977
Baseline probability of bacteraemia set to 20%	£16.25	0.00507	£3,205
Probability of death from bacteraemia set to 1%	£22.68	0.00005	£441,398

Scenario	Δ Costs – 'No dipstick' versus 'dipstick nitrite or LE'	Δ QALYs – 'No dipstick' versus 'dipstick nitrite or LE'	ICER – 'No dipstick' versus 'dipstick nitrite or LE'
Probability of death from bacteraemia set to 20%	£22.68	0.00052	£43,989
Probability of death from bacteraemia estimated from Schnadower 2010 (1.6%)	£22.68	0.00007	£343,422
Relative risk of bacteraemia for untreated UTI versus treated UTI set to 1.2	£22.85	0.00005	£478,072
Relative risk of bacteraemia for untreated UTI versus treated UTI set to 4	£22.26	0.00034	£65,548
Cost of microscopy, culture and antibiotic treatment doubled	£47.04	0.00013	£358,634
Cost of septicaemia doubled	£22.47	0.00013	£171,312
Antibiotic adverse events included	£22.68	0.00006	£382,290

Two-way sensitivity analyses, in which relative risk of septicaemia is varied simultaneously with with basline probability of septicaemia, and with probability of septicaemia, are shown in Figure 23 and Figure 24. These figures demonstrate tha a substantial increase in these paramters from base case values would be required in order for 'no dipstick' to be cost-effective.

Figure 23: Scenario 2 two-way sensitivity analysis results for children 3 months or older but younger than 3 years – relative risk of septicaemia versus baseline probability of septicaemia

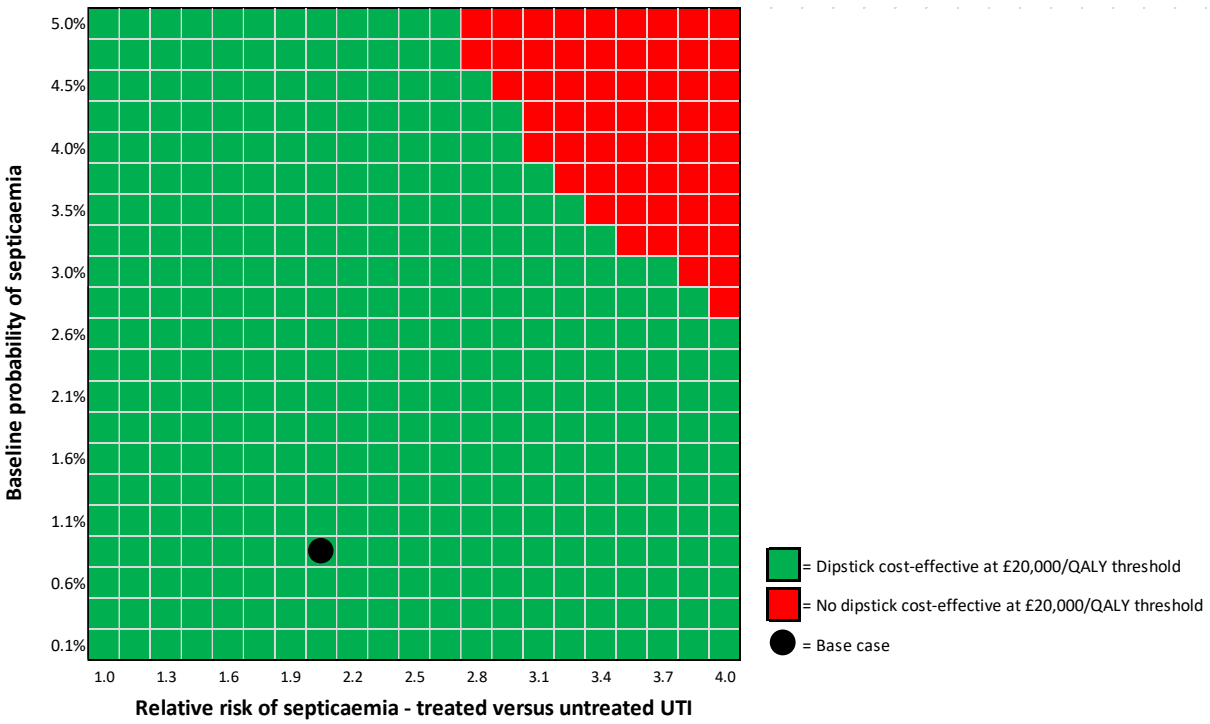
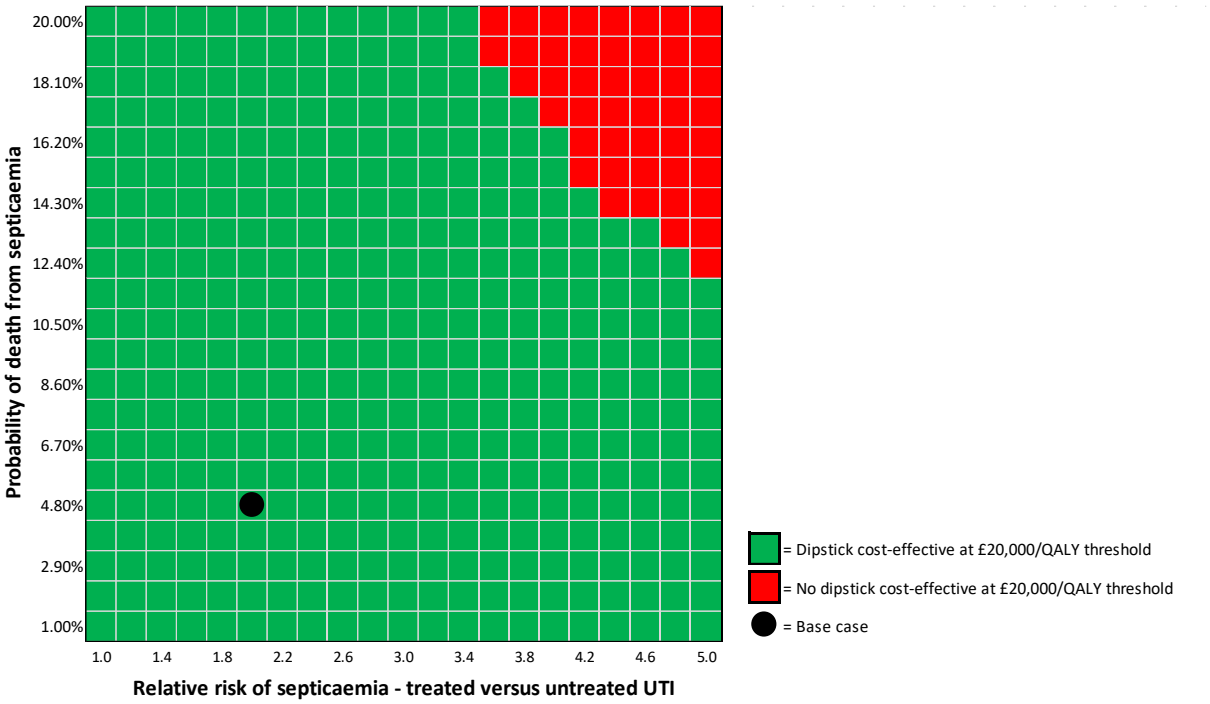
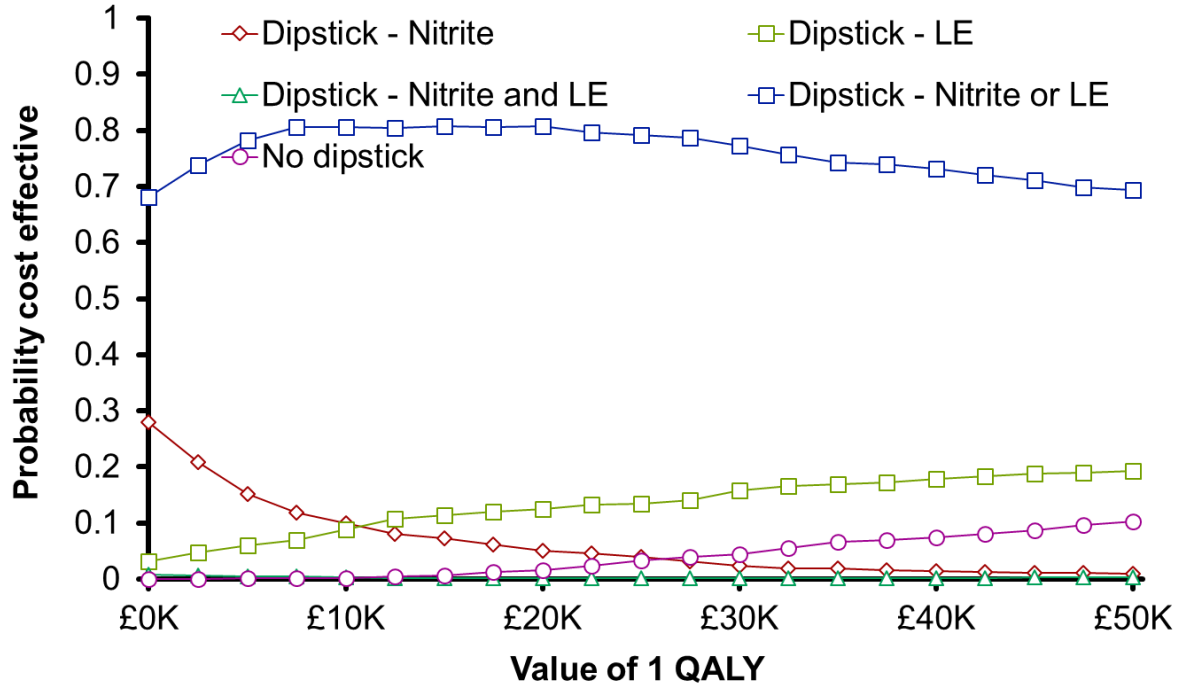


Figure 24: Scenario 2 two-way sensitivity analysis for children 3 months or older but younger than 3 years – relative risk of septicaemia versus probability of death from septicaemia



Results of probabilistic sensitivity analysis (displayed in Figure 25) show that the 'no dipstick' strategy has a low probability (less than 5%) of being cost-effective at a threshold of £20,000 per QALY, with 'dipstick – nitrite or LE' having the highest probability of being cost-effective.

Figure 25: Scenario 2 cost-effectiveness acceptability curve for children 3 months to 3 years



Scenario 3: Untreated UTI associated with an increased risk of septicaemia and PRS

Results for the scenario in which untreated UTI is associated with both an increased risk of sepsis and an increased risk of PRS is shown in Table 35. These results show that including both of these assumptions results in an ICER of £134,939 for 'no dipstick' compared to 'dipstick – LE'.

Table 35: Scenario 3 base case results for children 3 months to 3 years

Strategy	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Dipstick - Nitrite or LE	£8.38	25.14961	-	-	-
Dipstick - Nitrite	£9.17	25.14941	£0.79	-0.00020	dominated
Dipstick - LE	£10.23	25.14961	£1.85	0.00000	dominated
Dipstick - Nitrite and LE	£11.72	25.14915	£3.34	-0.00046	dominated
No dipstick	£30.82	25.14977	£22.44	0.00017	£134,939

One-way sensitivity analysis results for the scenario in which untreated UTI is associated with both an increased risk of septicaemia and PRS are shown in Table 36. These results show that, as with the population of infants younger than 3 months, the ICER is relatively sensitive to changes in the prevalence of UTI and parameters relating to the incidence of bacteraemia, but relatively insensitive to parameters relating to the incidence of PRS.

However, unlike the younger population, only two scenarios result in an ICER below £20,000 per QALY: the scenario in which the prevalence of UTI is set to 25% and the scenario in which the baseline probability of bacteraemia is set to 20%.

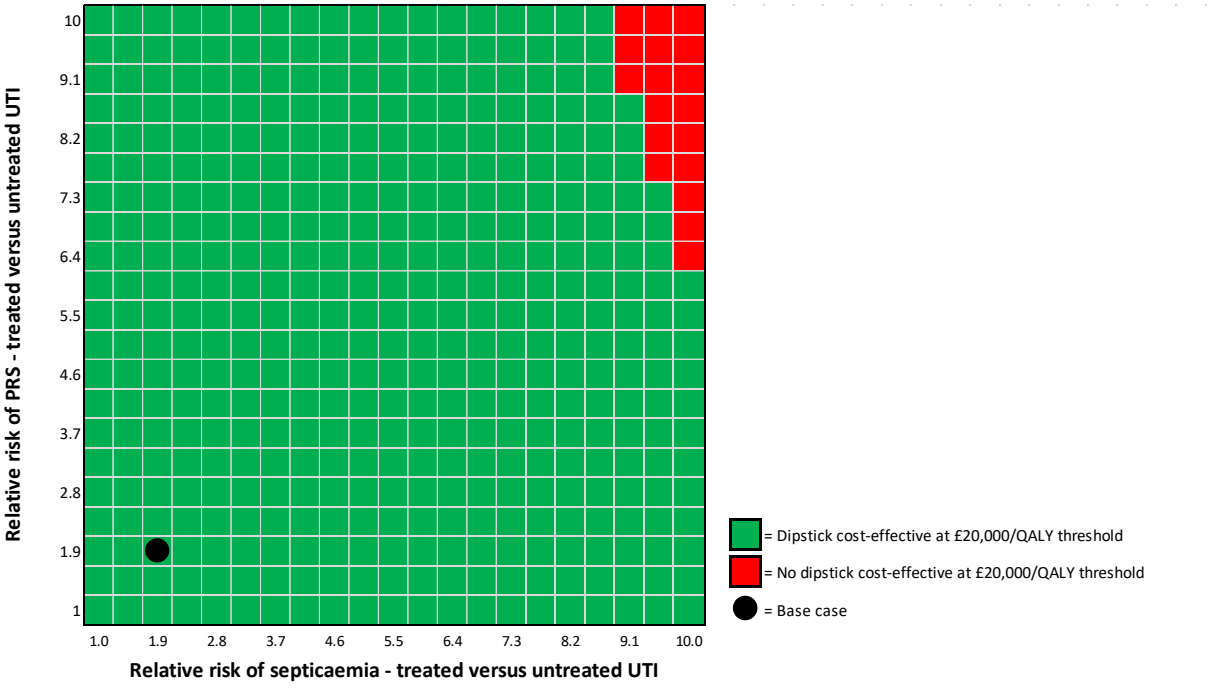
Table 36: Scenario 3 one-way sensitivity analysis results for children 3 months or older but younger than 3 years

Scenario	Δ Costs – 'No dipstick' versus 'dipstick nitrite or LE'	Δ QALYs – 'No dipstick' versus 'dipstick nitrite or LE'	ICER – 'No dipstick' versus 'dipstick nitrite or LE'
Base case	£22.44	0.00017	£134,939
UTI prevalence set to 1%	£25.04	0.00003	£813,077
UTI prevalence set to 25%	£10.86	0.00077	£14,107
Baseline probability of bacteraemia set to 0.1%	£22.62	0.00008	£300,011
Baseline probability of bacteraemia set to 20%	£17.28	0.00272	£6,348
Probability of death from bacteraemia set to 1%	£22.44	0.00009	£259,318
Probability of death from bacteraemia set to 20%	£22.44	0.00055	£40,754
Probability of death from bacteraemia estimated from Schnadower 2010 (1.6%)	£22.44	0.00010	£221,759
Baseline probability of PRS halved relative to base case	£22.56	0.00015	£151,679
Baseline probability of PRS doubled relative to base case	£22.20	0.00020	£110,216
Relative risk of bacteraemia for untreated UTI versus treated UTI set to 1.2	£22.61	0.00008	£272,600
Relative risk of bacteraemia for untreated UTI versus treated UTI set to 4	£22.02	0.00037	£58,765
Relative risk of PRS for untreated UTI versus treated UTI set to 1.2	£22.63	0.00014	£163,776

Scenario	Δ Costs – 'No dipstick' versus 'dipstick nitrite or LE'	Δ QALYs – 'No dipstick' versus 'dipstick nitrite or LE'	ICER – 'No dipstick' versus 'dipstick nitrite or LE'
Relative risk of PRS for untreated UTI versus treated UTI set to 4	£21.96	0.00024	£92,825
Cost of microscopy, culture and antibiotic treatment doubled	£46.80	0.00017	£281,421
Cost of septicaemia doubled	£22.23	0.00017	£133,673
Antibiotic adverse events included	£22.44	0.00009	£237,559

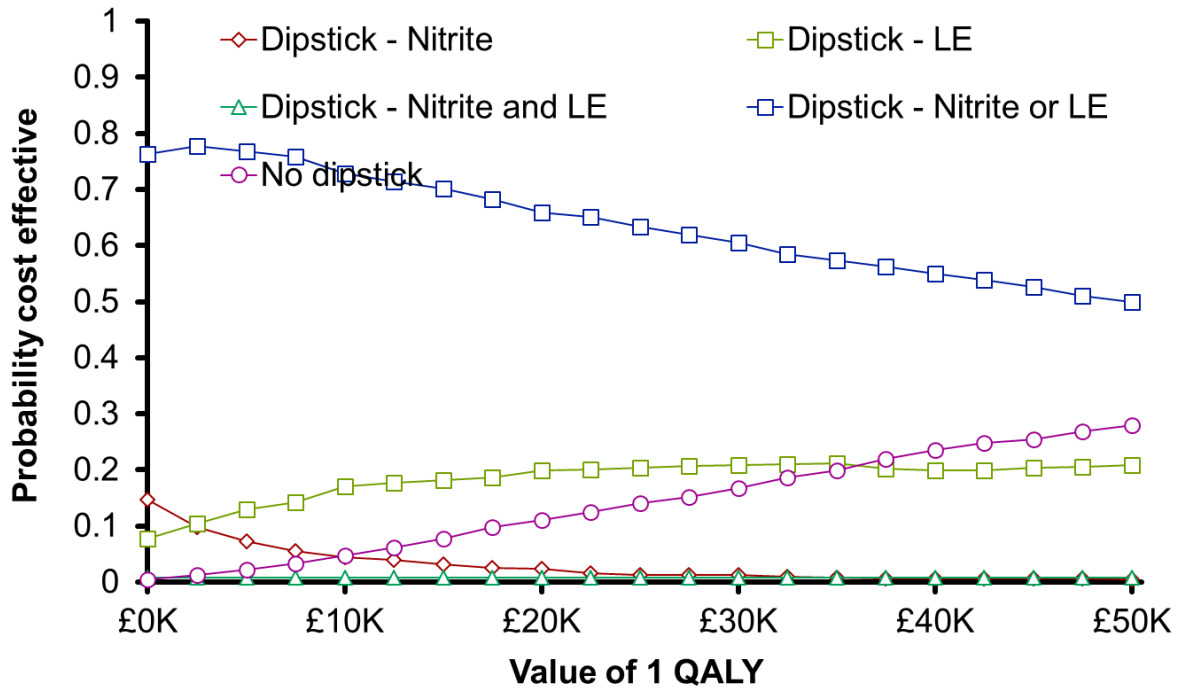
A two-way sensitivity analysis, in which the relative risk of septicaemia and relative risk of PRS are varied simultaneously, is shown in Figure 26. These results show that a substantial increase in these paramters from base case values would be required in order for the 'no dipstick' strategy to be cost-effective.

Figure 26: Scenario 3 two-way sensitivity analysis for children 3 months or older but younger than 3 years – relative risk of septicaemia versus relative risk of PRS



Results of probabilistic sensitivity analysis (displayed in Figure 27) show that the 'no dipstick' strategy has a relatively low probability of being cost-effective at a threshold of £20,000 per QALY, with 'dipstick – nitrite or LE' showing the highest probability of being cost-effective.

Figure 27: Scenario 3 cost-effectiveness acceptability curve for children 3 months or older but younger than 3 years



Discussion

Overall, the results of this analysis show that, in the majority of scenarios, and for the majority of sensitivity analyses, a strategy in which all children are treated with antibiotics and a sample is sent for laboratory testing is not cost-effective compared to a strategy in which initial dipstick testing is used to determine which children should receive treatment and further tests. This is primarily for three key reasons. First, the prevalence of UTI in children with suspected UTI is relatively low. Second, the sensitivity and specificity of dipstick testing in children under 3 is relatively high, particularly for 'LE' and 'nitrite or LE' strategies. These first two factors mean that a relatively small proportion of children with suspected UTI have a false negative test result. Third, in the majority of scenarios, the consequences of a false negative result are relatively insignificant. In the basic scenario, an extra 4 days of UTI symptoms results in a very small absolute QALY loss. For scenario 1, while the consequences of ESRD are severe, the probability of an individual case of UTI resulting in PRS, and the probability of PRS progressing to ESRD are small, meaning that the QALY loss associated with a false negative test result are, at the cohort level, relatively small.

Only in scenarios in which an increased risk of septicaemia associated with an untreated UTI is assumed is 'no dipstick' potentially a cost-effective strategy. This is because both the baseline probability of septicaemia and probability of death associated with septicaemia in the model are both relatively high, meaning that an untreated UTI results in a much higher expected QALY loss than in other scenarios. In the base-case scenario for children under 3 months, using an arbitrarily chosen relative risk of septicaemia of 2, the 'no dipstick' strategy has an ICER of £11,914 per QALY. However, the cost-effectiveness of this strategy is reliant on a number of parameters, most importantly a high baseline probability of UTI, probability of septicaemia, probability of death from septicaemia, and relative risk of septicaemia in untreated versus treated UTI. Sensitivity analyses have shown that a relatively small reduction in any of these parameters results in the ICER of the 'no dipstick' strategy exceeding £20,000 per QALY.

In children 3 months or older but younger than 3 years, unlike in the younger cohort, the base case results for scenarios including a risk of septicaemia still show an ICER of well in excess of £20,000 per QALY for the 'no dipstick' strategy. This is due to the considerably lower baseline risk of bacteraemia for the older group of patients. Sensitivity analyses for these scenarios show that the prevalence of UTI, baseline incidence of bacteraemia, probability of death from bacteraemia, or the relative risk of septicaemia would have to be substantially higher for the 'no dipstick' strategy to be cost-effective.

Of the four possible interpretations of dipstick results, the 'LE' and 'nitrite or LE' strategies are consistently more cost-effective than 'nitrite' and 'nitrite and LE strategies'. This is principally because the former two strategies have a far higher sensitivity, while retaining a relatively high specificity. For the large majority of scenarios, the 'LE' interpretation is the most cost-effective for children under 3 months, while the 'nitrite or LE' interpretation is the most cost-effective for children 3 months or older but younger than 3 years. However, it is likely that this dichotomy is an artefact of random variation in the results of the meta-analyses used to synthesise accuracy data. The sensitivity and specificity of 'LE' and 'nitrite or LE' interpretations are relatively close to one another, with confidence intervals indicating that there is considerable overlap in plausible values.

A key limitation of this analysis is the considerable uncertainty surrounding the consequences of a false negative test result for UTI. This problem is addressed by exploring a wide range of possible outcomes of untreated UTI, ranging from fairly mild (4 extra days of symptoms) to severe (risk of death from septicaemia). However, these scenarios are highly speculative, and in some cases may not fully reflect clinical reality. For example, in scenarios which include an increased risk of PRS associated with false negative results, the assumption is made that there is a direct link between a single untreated UTI event and the development of PRS. In reality, PRS is likely to develop over a longer period of time, and is a cumulative result of several UTI incidents. Therefore, quantifying the added risk of delaying treatment of a single infection for a short period is a highly speculative process. For the scenarios including an increased risk of septicaemia, the assumption is made that UTI has the possibility of progressing to septicaemia, the probability of which increases if antibiotic treatment is delayed. However, the evidence that septicaemia occurs secondary to UTI is inconclusive; the two conditions are often coincident, but the order of causality is not clear.

Another limitation of the analysis is that the full complexity of potentially overlapping symptoms and conditions which may occur in children with possible UTI is not captured. Particularly in infants under 3 months, symptoms of UTI are frequently non-specific, and therefore children in the age group are typically referred to secondary care regardless of initial testing. This means that, for such children, a single test alone may not be sufficient to determine that a child is in no need of further investigation for other causes, and therefore the model assumption that children without a UTI are otherwise healthy is potentially unrealistic.

In the base case, the analysis uses values for the prevalence of UTI in febrile children taken from a meta-analysis primarily based on studies which were conducted in an emergency department setting. It is possible that these values provide an overestimate of the true prevalence of UTI in the population of interest, as it is likely that symptoms are more severe in children presenting to an emergency department. However, sensitivity analyses indicate that this could only affect decision making in scenarios where untreated UTI is associated with an increased risk of septicaemia in children under 3 months, as a lower prevalence of UTI results in an increase in the ICER of laboratory testing compared to dipstick testing.In conclusion, this analysis shows that, in the majority of exploratory scenarios, a strategy in which all children with suspected UTI are prescribed antibiotics and a urine sample sent for microscopy and culture is not cost-effective compared to a scenario in which initial dipstick testing is used to determine which children should receive treatment and further testing. Only

in scenarios in which a substantial added risk of septicaemia is assumed to result from untreated UTI is a ‘no dipstick’ strategy potentially cost-effective.

References

1. Shaikh, N., Morone, N.E., Bost, J.E. and Farrell, M.H., 2008. Prevalence of urinary tract infection in childhood: a meta-analysis. The Pediatric infectious disease journal, 27(4), pp.302-308.
2. Whiting, P., Westwood, M., Bojke, L., Palmer, S., Richardson, G., Cooper, J., Watt, I., Glanville, J., Sculpher, M. and Kleijnen, J., 2006. Clinical effectiveness and cost-effectiveness of tests for the diagnosis and investigation of urinary tract infection in children: a systematic review and economic model.
3. Schnadower, D., Kuppermann, N., Macias, C.G., Freedman, S.B., Baskin, M.N., Ishimine, P., Scribner, C., Okada, P., Beach, H., Bulloch, B. and Agrawal, D., 2010. Febrile infants with urinary tract infections at very low risk for adverse events and bacteremia. Pediatrics, pp.peds-2010.
4. Pitetti, R.D. and Choi, S., 2002. Utility of blood cultures in febrile children with UTI. The American journal of emergency medicine, 20(4), pp.271-274.
5. MRSA, MSSA and E. coli bacteraemia and C. difficile infection: 30 day all-cause fatality. Available at: https://www.gov.uk/government/publications/mrsa-mssa-and-e-coli-bacteraemia-and-c-difficile-infection-30-day-all-cause-fatality [accessed 21/04/17]
6. Goodkin, D.A., Bragg-Gresham, J.L., Koenig, K.G., Wolfe, R.A., Akiba, T., Andreucci, V.E., Saito, A., Rayner, H.C., Kurokawa, K., Port, F.K. and Held, P.J., 2003. Association of comorbid conditions and mortality in hemodialysis patients in Europe, Japan, and the United States: the Dialysis Outcomes and Practice Patterns Study (DOPPS). Journal of the American Society of Nephrology, 14(12), pp.3270-3277.
7. NHS Blood and Transplant. Annual Report on Kidney Transplantation. Report for 2013/14. Available at: http://www.odt.nhs.uk/pdf/organ_specific_report_kidney_2014.pdf [accessed 21/04/17]
8. Kerr, M., Bray, B., Medcalf, J., O'donoghue, D.J. and Matthews, B., 2012. Estimating the financial cost of chronic kidney disease to the NHS in England. Nephrology Dialysis Transplantation, p.gfs269.
9. NHS Drug Tariff April 2017. Available at: https://www.nhsbsa.nhs.uk/sites/default/files/2017-04/April_2017.pdf [accessed 21/04/17]
10. NHS Reference Costs 2015 to 2016. Available at: https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016 [accessed 21/04/17]
11. Bermingham, S.L. and Ashe, J.F., 2012. Systematic review of the impact of urinary tract infections on health-related quality of life. BJU international, 110(11c), pp.E830-E836.
12. Kind, P., Hardman, G. and Macran, S., 1999. UK population norms for EQ-5D (Vol. 172). York: Centre for Health Economics, University of York.
13. Wasserfallen, J.B., Halabi, G., Saudan, P., Perneger, T., Feldman, H.I., Martin, P.Y. and Wauters, J.P., 2004. Quality of life on chronic dialysis: comparison between haemodialysis and peritoneal dialysis. Nephrology Dialysis Transplantation, 19(6), pp.1594-1599.
14. Cleemput, I., Kesteloot, K., Moons, P., Vanrenterghem, Y., Van Hooff, J.P., Squifflet, J.P. and De Geest, S., 2004. The construct and concurrent validity of the EQ-5D in a renal transplant population. Value in Health, 7(4), pp.499-509.
15. Round, J., Fitzgerald, A.C., Hulme, C., Lakhanpaul, M. and Tullus, K., 2012. Urinary tract infections in children and the risk of ESRF. Acta Paediatrica, 101(3), pp.278-282.

Appendix K: Excluded studies

K.1 Clinical studies

Short Title	Title	Reason for exclusion
Amir (2013)	Urinary tract infection in children	Dipstick testing not included.
Anacleto (2009)	Bedside diagnosis of outpatient childhood urinary tract infection using three-media dipslide culture test	Only patients with positive finding on dipstick were included.
Antwi (2008)	Urine dipstick as a screening test for urinary tract infection	Only selected patients received urine culture.
Ayazi (2007)	Comparison of urine culture and urine dipstick analysis in diagnosis of urinary tract infection	Indirect population: age ranged from 15 days to 11 years and no subgroup analysis for those < 3 years old is possible.
Ayazi (2013)	Diagnostic Accuracy of the Quantitative C-Reactive Protein, Erythrocyte Sedimentation Rate and White Blood Cell Count in Urinary Tract Infections among Infants and Children	No dipstick testing.
Ayazi (2013)	Diagnostic accuracy of the quantitative C-reactive protein, erythrocyte sedimentation rate and white blood cell count in urinary tract infections among infants and children	Dipstick testing not included.
Bachur (2001)	Reliability of the urinalysis for predicting urinary tract infections in young febrile children.	Classification of index test (dipstick positive and/or microscopy positive) is not included in review protocol.
Baumer (2005)	Managing urinary tract infections in young children	Narrative review.
Bellino (2013)	Urinary tract infections in sows in Italy: Accuracy of urinalysis and urine culture against histological findings	Study population - animals.
Bereket (2013)	Use of urinalysis as a screening tool for asymptomatic infants	Only selected patients received urine culture.
Berger (2006)	Diagnosing urinary tract infections in young children	Letter to the editor.
Bin (2010)	Duration of fever affects the likelihood of a positive bag urinalysis or catheter culture in young children	Only patients testing positive on dipstick received urine culture.
Bonadio (2014)	Urinary tract infection in outpatient febrile infants younger than 30 days of age: a 10-year evaluation	Not all patients received a dipstick test.
Bonsu (2007)	Leukocyte counts in urine reflect the risk of concomitant sepsis in bacteriuric infants: a retrospective cohort study	Definition of index test positive not reported.
Cantey (2015)	Lack of clinical utility of urine gram stain for suspected urinary tract infection in pediatric patients	Indirect population: median age 4 years (IQR = 10 months - 10 years) and no subgroup analysis for those < 3 years old is possible.
Clyne (2014)	Paediatrics: dipstick adequate for febrile UTI test	Overview of an included study (Glissmeyer 2014).
Coulthard (2010)	Point-of-care diagnostic tests for childhood urinary-tract infection: phase-contrast	Indirect population: median age 6.2 years (range = 12 weeks –

Short Title	Title	Reason for exclusion
	microscopy for bacteria, stick testing, and counting white blood cells	17.7 years) and no subgroup analysis for those < 3 years old is possible.
Deville (2004)	The urine dipstick test useful to rule out infections. A meta-analysis of the accuracy	Meta-analysis on dipstick testing. Included studies were assessed against the review protocol.
Dilek (2014)	Validity of urine and blood tests for detection of urinary tract infections in children	Indirect population: range = 0.5 – 12 years and no subgroup analysis for those < 3 years old is possible.
Downing (2012)	The diagnosis of urinary tract infections in young children (DUTY): protocol for a diagnostic and prospective observational study to derive and validate a clinical algorithm for the diagnosis of UTI in children presenting to primary care with an acute illness	Study protocol for an included study (Hay 2016).
Elhassanien (2013)	Fever without source in infants and young children: Dilemma in diagnosis and management	Dipstick testing not included.
Eliacik (2016)	A Comparison of Bladder Catheterization and Suprapubic Aspiration Methods for Urine Sample Collection from Infants with a Suspected Urinary Tract Infection	Only all positive culture results tested with dipstick.
Emamghorashi (2008)	Evaluation of urinary tract infection in newborns with jaundice in south of Iran	No outcomes in protocol available and case control study design.
Galetto-Lacour (2010)	Validation of a laboratory risk index score for the identification of severe bacterial infection in children with fever without source	Diagnostic accuracy of dipstick testing not reported.
Geurts (2014)	Impact analysis of an evidence-based guideline on diagnosis of urinary tract infection in infants and young children with unexplained fever	Not all patients received dipstick and of those only selected patients received urine culture and
Ghaemi (2007)	Late onset jaundice and urinary tract infection in neonates	Dipstick testing not included.
Gupta (2015)	Profile of urinary tract infections in paediatric patients	No dipstick testing.
Hay (2016)	Improving the Diagnosis and Treatment of Urinary Tract Infection in Young Children in Primary Care: Results from the DUTY Prospective Diagnostic Cohort Study	Insufficient data available to calculate 2 by 2 diagnostic accuracy table for under 3 years.
Hiraoka (1994)	Rapid dipstick test for diagnosis of urinary tract infection.	Indirect population aged 1 month to 15 years and no subgroup analysis for those aged < 3 years old.
Hoberman (1997)	Urinary tract infections in young febrile children.	Insufficient evidence available to calculate diagnostic accuracy 2x2 table.
Hosseini (2009)	Urine culture obtained from bag specimens and suprapubic aspiration in neonates	Dipstick testing not included.
Jafari (2015)	Urinary screening in primary school children in yazd, iran	Indirect population: Primary school age children included.

Short Title	Title	Reason for exclusion
Karacan (2010)	Evaluation of urine collection methods for the diagnosis of urinary tract infection in children	Dipstick testing not included.
Kiddoo (2015)	Randomized Crossover Trial of Single Use Hydrophilic Coated vs Multiple Use Polyvinylchloride Catheters for Intermittent Catheterization to Determine Incidence of Urinary Infection	No diagnostic accuracy data provided, does not include < 3 years old subgroup.
Kjolvmark (2012)	Elevated urine levels of heparin-binding protein in children with urinary tract infection	Indirect population: mean age 6 years (range = 0 - 18 years) and no subgroup analysis for those < 3 years old is possible.
Kocer (2015)	Diagnostic Accuracy of a New Urinalysis System, DongJiu, for Diagnosis of Urinary Tract Infection	Mean age 16 years.
Krahenbuhl (2012)	Evaluation of a novel in-vitro diagnostic device for the detection of urinary tract infections in diaper-wearing children	Inadequate data.
Lertdumrongluk (2015)	Diagnostic accuracy of urine heparin binding protein for pediatric acute pyelonephritis	Indirect age group (0.3 to 6.4 years) with no < 3 years subgroup. Case-control study design.
Lertdumrongluk (2014)	Diagnostic accuracy of urine heparin binding protein for pediatric acute pyelonephritis	Duplicate.
Lunn (2010)	Automated microscopy, dipsticks and the diagnosis of urinary tract infection	Only selected patients positive on dipstick testing received urine culture.
Mori (2010)	Diagnostic performance of urine dipstick testing in children with suspected UTI: a systematic review of relationship with age and comparison with microscopy	Systematic review of dipstick and microscopy. Included studies were assessed against the review protocol.
Ojha (2014)	Profile of children with urinary tract infection and the utility of urine dipstick as a diagnostic tool	Indirect population: age range = 2 months – 13 years) and no subgroup analysis for those < 3 years old is possible.
Pugia (2004)	Near-patient testing for infection using urinalysis and immuno-chromatography strips	Age of participants not reported.
Ramlakhan (2011)	Dipstick urinalysis for the emergency department evaluation of urinary tract infections in infants aged less than 2 years	Not all patients received reference test - only a proportion (all positive dipsticks and few negatives) selected for culture.
Reed (1995)	Urinary tract infection in malnourished rural African children.	Indirect population of 0 to 5 years, subgroup for < 3 years not available.
Schroeder (2005)	Choice of urine collection methods for the diagnosis of urinary tract infection in young, febrile infants	Inconsistent reporting of results.
Schroeder (2015)	Diagnostic accuracy of the urinalysis for urinary tract infection in infants <3 months of age	Study selects children with positive bacteriuria and a random sample of children with negative urine culture. Children with negative urine culture did not receive dipstick testing.

Short Title	Title	Reason for exclusion
St John (2006)	The use of urinary dipstick test to exclude urinary tract infection: A systematic review of the literature	Only patients testing positive on dipstick received urine culture.
Unal (2011)	Comparison of different urinalysis techniques in the diagnosis of urinary tract infection among febrile children without an apparent origin of fever	Full article unavailable.
Velasco (2015)	Febrile young infants with altered urinalysis at low risk for invasive bacterial infection. a Spanish Pediatric Emergency Research Network's Study	urine culture only performed if dipstick was abnormal
Watson (2016)	Evaluation of novel urinary tract infection biomarkers in children	Indirect age group (up to 3.9 years) with catheterisation.
Weems (2014)	Urinary tract infections in a neonatal intensive care unit	Cases were their own control (samples with negative culture compared with sample with positive culture).
Weisz (2010)	The presence of urinary nitrites is a significant predictor of pediatric urinary tract infection susceptibility to first- and third-generation cephalosporins	Diagnostic outcomes not provided and not calculable.
Westwood (2005)	Further investigation of confirmed urinary tract infection (UTI) in children under five years: a systematic review (Structured abstract)	Meta-analysis on urine testing in under 5 year olds. Included studies were assessed against the review protocol.
Whiting (2005)	Rapid tests and urine sampling techniques for the diagnosis of urinary tract infection (UTI) in children under five years: a systematic review	Meta-analysis on urine testing in under 5 year olds. Included studies were assessed against the review protocol.
Whiting (2006)	Clinical effectiveness and cost-effectiveness of tests for the diagnosis and investigation of urinary tract infection in children: a systematic review and economic model	Systematic review in under 5 years. Included studies were assessed against the review protocol.
Williams (2010)	Absolute and relative accuracy of rapid urine tests for urinary tract infection in children: a meta-analysis	Systematic review in children. Included studies were assessed against the review protocol.
Wu (2005)	Auditing the management of childhood urinary tract infections in a regional hospital	Audit. Only select patients (a proportion of those positive on dipstick or microscopy) received urine culture.

K.2 Economic studies

Reference	Reason for exclusion
Sekhar, D.L., Wang, L., Hollenbeak, C.S., Widome, M.D. and Paul, I.M., 2010. A cost-effectiveness analysis of screening urine dipsticks in well-child care. Pediatrics, 125(4), pp.660-663.	Analysis of screening in healthy children, rather than in children with suspected UTI
Little, P., Turner, S., Rumsby, K., Warner, G., Moore, M., Lowes, J.A., Smith, H., Hawke, C., Turner, D., Leydon, G.M. and Arscott, A., 2009. Dipsticks and diagnostic algorithms in urinary tract infection: development and validation, randomised trial, economic analysis, observational cohort and qualitative study. Health Technol Assess, 13(19), pp.1-73.	Patient population is adult women

Appendix L: DUTY study included data

The following contains the data obtained from the DUTY study authors and included in this evidence review.

Diagnostic Accuracy of Dipstick Tests in the Diagnosis of Urinary Tract infection in Young children (DUTY) Study

Alastair D Hay,[1] Kate Birnie,[2] John Busby,[2] Brendan Delaney,[3] Harriet Downing,[1] Jan Dudley,[4] Stevo Durbaba,[5] Margaret Fletcher,[6,7] Kim Harman,[1] William Hollingworth,[2] Kerenza Hood,[8] Robin Howe,[9] Michael Lawton,[2] Catherine Lisles,[8] Paul Little,[10] Alasdair MacGowan,[11] Kathryn O'Brien,[12] Timothy Pickles,[8] Kate Rumsby,[10] Jonathan AC Sterne,[2] Emma Thomas-Jones,[8] Judith van der Voort,[13] Cherry-Ann Waldron,[8] Penny Whiting,[2] Mandy Wootton[9] and Christopher C Butler[12,14]

1 Centre for Academic Primary Care, National Institute for Health Research (NIHR) School of Primary Care Research, School of Social and Community Medicine, University of Bristol, Bristol, UK
2 School of Social and Community Medicine, University of Bristol, Bristol, UK
3 Department of Primary Care and Public Health Sciences, National Institute for Health Research (NIHR) Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London, London, UK
4 Bristol Royal Hospital for Children, University Hospitals Bristol NHS Foundation Trust, Bristol, UK
5 Department of Primary Care and Public Health Sciences, Division of Health and Social Care Research, King's College London, London, UK
6 Centre for Health and Clinical Research, University of the West of England, Bristol, UK
7 South West Medicines for Children Local Research Network, University Hospitals Bristol NHS Foundation Trust, Bristol, UK
8 South East Wales Trials Unit (SEWTU), Institute for Translation, Innovation, Methodology and Engagement, School of Medicine, Cardiff University, Cardiff, UK
9 Specialist Antimicrobial Chemotherapy Unit, Public Health Wales Microbiology Cardiff, University Hospital Wales, Cardiff, UK
10 Primary Care and Population Sciences Division, University of Southampton, Southampton, UK
11 Southmead Hospital, North Bristol NHS Trust, Bristol, UK
12 Cochrane Institute of Primary Care & Public Health, School of Medicine, Cardiff University, Cardiff, UK
13 Department of Paediatrics and Child Health, University Hospital of Wales, Cardiff, UK
14 Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

Background

In February 2017, the Diagnosis of Urinary Tract infection in Young children (DUTY) study received an invitation from Omnia Abdulrazeg, a Technical Analyst at the National Institute for Health and Care Excellence (NICE). The DUTY study was asked to provide 2 by 2 diagnostic accuracy tables as sensitivity and specificity of dipstick leukocytes and nitrites for children aged <3 years. This information could be used to inform an update of NICE guideline CG54, Urinary tract infection in under 16s: diagnosis and management, for a section relating to the use of dipstick tests to diagnose UTI in this age group.

Methods

The DUTY study was a multicentre, prospective, diagnostic cohort study. Children were eligible if they were aged <5 years, presented to primary care with any acute illness episode of <28 days duration and had constitutional or urinary symptoms associated with their acute

illness. Children were excluded if they were not constitutionally unwell, had a neurogenic or surgically reconstructed bladder, used a urinary catheter, presented with trauma, or had taken antibiotics within the past week.

UTI was defined as pure (single) or predominant growth of a uropathogen (*Enterobacteriaceae*) at ≥105 colony-forming units (CFU)/mL. We defined predominant growth as $\geq 10^5$ CFU/mL of a uropathogen with a 3-log10 (1,000-fold) or greater difference between the growth of this and the next species.

The dipstick nitrite variable had two categories: negative or positive. The leukocytes variable had the following five categories: negative, trace, 1+, 2+, 3+. For this analysis we dichotomised the leukocytes variable into negative/trace vs. positive, and also as an alternative coding of negative vs. trace/positive. We also created a variable for either nitrite or leukocytes positive (according to the original and alternative coding of the dichotomised leukocytes variable).

Age was stratified according to the following categories: <3 months, 3 months to <3 years, and 3 to 5 years. The 3 to 5 years age category was included in the frequency tables for completeness.

Analyses were stratified by sampling method (i.e. clean catch and nappy pad samples), because of the difference in contamination rates.

For more details on study methods see: Hay AD, Birnie K, Busby J, Delaney B, Downing H, Dudley J, et al. The Diagnosis of Urinary Tract infection in Young children (DUTY): a diagnostic prospective observational study to derive and validate a clinical algorithm for the diagnosis of urinary tract infection in children presenting to primary care with an acute illness. Health Technol Assess 2016;20(51).

Results

The frequency of UTI by dipstick result, stratified by age group, in nappy pad samples are shown in Table 17. The resulting sensitivity and specificity, with 95% confidence intervals (CI) are shown in Table 38 for children aged: <3 months and 3 months to <3 years. The frequencies for clean catch samples are shown in Table 39, with sensitivity and specificity values in Table .

Discussion

We have responded to the specific question asked i.e. to provide the diagnostic parameters of dipstick testing for children aged under 3 years, but in clinical practice the diagnostic parameters might differ since some of the diagnostic value of the dipsticks might be 'used up' by the clinical symptoms and signs used to select for urine sample testing.

These are a portion of the DUTY study results. Full results have been published in the HTA monograph series, please see:

Hay AD, Birnie K, Busby J, Delaney B, Downing H, Dudley J, et al. The Diagnosis of Urinary Tract infection in Young children (DUTY): a diagnostic prospective observational study to derive and validate a clinical algorithm for the diagnosis of urinary tract infection in children presenting to primary care with an acute illness. Health Technol Assess 2016;20(51)

https://dx.doi.org/10.3310/hta20510

Table 17: Frequency of UTI by dipstick results, stratified by age group in NAPPY PAD samples

Age <3 months	UTI negative	UTI positive	Total	Age 3 months to <3 years	UTI negative	UTI positive	Total	Age 3 to 5 years	UTI negative	UTI positive	Total
Nitrites											
Negative	113	1	114	Negative	1,729	12	1,741	Negative	61	0	61
Positive	14	0	14	Positive	318	17	335	Positive	6	0	6
Total	127	1	128	Total	2,047	29	2,076	Total	67	0	67
Leukocytes											
Negative/trace	106	0	106	Negative/trace	1,706	14	1,720	Negative/trace	58	0	58
Positive	21	1	22	Positive	341	15	356	Positive	9	0	9
Total	127	1	128	Total	2,047	29	2,076	Total	67	0	67
Leukocytes - alternative coding											
Negative	96	0	96	Negative	1,596	13	1,609	Negative	54	0	54
Trace/positive	31	1	32	Trace/positive	451	16	467	Trace/positive	13	0	13
Total	127	1	128	Total	2,047	29	2,076	Total	67	0	67
Either positive											
Negative/trace	96	0	96	Negative/trace	1,474	6	1,480	Negative/trace	53	0	53
Positive	31	1	32	Positive	573	23	596	Positive	14	0	14
Total	127	1	128	Total	2,047	29	2,076	Total	67	0	67
Either positive - alternative coding											
Negative	90	0	90	Negative	1,392	6	1,398	Negative	49	0	49
Trace/positive	37	1	38	Trace/positive	655	23	678	Trace/positive	18	0	18
Total	127	1	128	Total	2,047	29	2,076	Total	67	0	67

For nappy pad samples, there were 6 children with missing data for the dipstick tests. These were excluded from the analysis.

Table 38: Sensitivity and Specificity (95% CI) for NAPPY PAD samples

	Age <3 months		Age 3 months to <3 years	
	Sensitivity	Specificity	Sensitivity	Specificity
Nitrites	0.0 (0.0, 97.5)	89.0 (82.2, 93.8)	58.6 (38.9, 76.5)	84.5 (82.8, 86.0)
Leukocytes	100.0 (2.5, 100.0)	83.5 (75.8, 89.5)	51.7 (32.5, 70.6)	83.3 (81.7, 84.9)
Leukocytes - alternative coding	100.0 (2.5, 100.0)	75.6 (67.2, 82.8)	55.2 (35.7, 73.6)	78.0 (76.1, 79.7)
Either positive	100.0 (2.5, 100.0)	75.6 (67.2, 82.8)	79.3 (60.3, 92.0)	72.0 (70.0, 73.9)
Either positive - alternative coding	100.0 (2.5, 100.0)	70.9 (62.1, 78.6)	79.3 (60.3, 92.0)	68.0 (65.9, 70.0)

Table 39: Frequency of UTI by dipstick results, stratified by age group in CLEAN CATCH samples

Age <3 months	UTI negative	UTI positive	Total	Age 3 months to <3 years	UTI negative	UTI positive	Total	Age 3 to 5 years	UTI negative	UTI positive	Total
Nitrites											
Negative	16	0	16	Negative	726	12	738	Negative	1,881	23	1,904
Positive	0	0	0	Positive	23	10	33	Positive	26	15	41
Total	16	0	16	Total	749	22	771	Total	1,907	38	1,945
Leukocytes											
Negative/trace	14	0	14	Negative/trace	672	7	679	Negative/trace	1,717	16	1,733
Positive	2	0	2	Positive	77	15	92	Positive	190	22	212
Total	16	0	16	Total	749	22	771	Total	1,907	38	1,945
Leukocytes - alternative coding											
Negative	13	0	13	Negative	635	4	639	Negative	1,607	13	1,620
Trace/positive	3	0	3	Trace/positive	114	18	132	Trace/positive	300	25	325
Total	16	0	16	Total	749	22	771	Total	1,907	38	1,945
Either positive											
Negative/trace	14	0	14	Negative/trace	660	5	665	Negative/trace	1,706	11	1,717
Positive	2	0	2	Positive	89	17	106	Positive	202	27	229
Total	16	0	16	Total	749	22	771	Total	1,908	38	1,946

Age <3 months	UTI negative	UTI positive	Total	Age 3 months to <3 years	UTI negative	UTI positive	Total	Age 3 to 5 years	UTI negative	UTI positive	Total
Either positive - alternative coding											
Negative	13	0	13	Negative	625	3	628	Negative	1,600	11	1,611
Trace/positive	3	0	3	Trace/positive	124	19	143	Trace/positive	308	27	335
Total	16	0	16	Total	749	22	771	Total	1,908	38	1,946

For clean catch samples, there were 8 children with missing data for the dipstick tests. These were excluded from the analysis.

Table 40: Sensitivity and Specificity (95% CI) for CLEAN CATCH samples

	Age <3 months		Age 3 months to <3 years	
	Sensitivity	Specificity	Sensitivity	Specificity
Nitrites	-	100 (79.4, 100.0)	45.5 (24.4, 67.8)	96.9 (95.4, 98.0)
Leukocytes	-	87.5 (61.7, 98.5)	68.2 (45.1, 86.1)	89.7 (87.3, 91.8)
Leukocytes - alternative coding	-	81.3 (54.4, 96.0)	81.8 (59.7, 94.8)	84.8 (82.0, 87.3)
Either positive	-	87.5 (61.7, 98.5)	77.3 (54.6, 92.2)	88.1 (85.6, 90.3)
Either positive - alternative coding	-	81.3 (54.4, 96.0)	86.4 (65.1, 97.1)	83.4 (80.6, 86.0)

- Sensitivity cannot be calculated due to 0 events