ACTA PÆDIATRICA
NURTURING THE CHILD

Outpatient urine collection methods for paediatric urinary tract infections: systematic review of diagnostic accuracy studies

Journal:	*Acta Paediatrica*
Manuscript ID	SPAE-2021-0474.R1
Manuscript Type:	Review Article
Date Submitted by the Author:	07-Jul-2021
Complete List of Authors:	Boon, Hanne Ann; KU Leuven, Public Health and Primary Care Lenaerts, Wouter; KU Leuven, Public Health and Primary Care Van Aerde, Cedric; KU Leuven, Public Health and Primary Care Verbakel, Jan; KU Leuven, Public Health and Primary Care; University of Oxford Nuffield Department of Primary Care Health Sciences
Keywords:	systematic review, child, urinary tract infection, urine collection, primary health care

SCHOLARONE™
Manuscripts

Outpatient urine collection methods for paediatric urinary tract infections: systematic review of diagnostic accuracy studies (Title)

Hanne Ann Boon (MD)[1], Wouter Lenaerts (MD)[1], Cedric Van Aerde (MSc)[1], Jan Y Verbakel (MD, PhD)[1,2] (Authors)

[1]EPI-centre, Academic Centre for General Practice, KU Leuven, Kapucijnenvoer 7, 3000 Leuven, BELGIUM

[2]Nuffield Department of Primary Care Health Sciences, University of Oxford, Radcliffe Observatory Quarter, Woodstock Rd, Oxford OX2 6GG, UK

Short title: Urine collection methods in outpatient children

Corresponding author: Professor Dr. Jan Y Verbakel, EPI-centre, Academic Centre for General Practice, KU Leuven, Kapucijnenvoer 7, 3000 Leuven, BELGIUM, jan.verbakel@kuleuven.be, Tel +32 (0)16 37 66 21,

Word count (including Abbreviation list, Keynotes, Abstract, Introduction, Methods, Results, Discussion, Statements, References): 4610

Manuscript pages: 21

Abstract word count: 203

References:39 **Tables:** 2 **Figures:** 3 **Supplemental Content:** 4

For Peer Review Only

ABSTRACT

Aim

To investigate the diagnostic test accuracy of urine collection methods for urinary tract infections in outpatient children.

Methods

A systematic literature review until April 2021 (Medline, Web Of Science, Embase, Cinahl) to examine the diagnostic test accuracy of urine culture on collection methods for urinary tract infection in outpatient children below 18 years. Contamination rates were studied as secondary outcome. The risk of bias was assessed using the QUADAS-2 criteria. Two-by-two tables were extracted in duplicate to calculate sensitivities, specificities, likelihood ratios with 95% confidence intervals.

Results

The search identified seven relevant studies. Clean-catch compared to catheterization in children less than 90 days showed a sensitivity and specificity of 97% and 89% (95%CI 84-100% and 67-99%), respectively. Adhesive bags compared to catheterization showed a sensitivity of 83% (95%CI 75-90%) and specificity of 91% (95%CI 83-96%). There was a similar diagnostic accuracy when comparing urine sampling by means of adhesive bags versus nappy pads. The contamination rate was 5% for clean-catch, 30-80% for adhesive bags and 64% for nappy pads.

Conclusion

Nappy pads and adhesive bags are easy to use with comparable accuracy, but are extremely prone to contamination. Clean-catch urine sampling might be an accurate alternative in young infants in ambulatory care.

Keywords (Mesh terms): Urinary Tract Infections, Child, Primary Health Care, Urine Specimen Collection/methods, Systematic Review [Publication Type]

For Peer Review Only

KEY NOTES

- Early diagnosis of urinary tract infection is challenging in the outpatient setting, because there are few convenient collection methods and data on the diagnostic test accuracy is rare.
- This systematic literature review identified seven studies with paired data on collection methods, which show that clean-catch urine sampling might be useful for ruling out urinary tract infection (LR- of 0.03).
- The contamination rate was highest for adhesive bags and nappy pads and the LR- for bags was 0.18.

ABBREVIATION LIST

cfu/mL= colony-forming units per millilitre

LR+ = positive likelihood ratio

LR- = negative likelihood ratio

UTI= urinary tract infection

95%CI= 95% confidence intervals

INTRODUCTION

Urinary tract infections (UTI), comprising cystitis and pyelonephritis, occur in 6% and 0.6% of acutely ill children presenting to primary care, respectively.[1,2] UTI diagnosis is important, as adequate and timely treatment can alleviate short-term distress of the child, prevent renal scarring,[3] and progression to a more serious systemic infection.[4]

Diagnosis is challenging, as only a minority of children with UTI present with typical UTI symptoms. In a previous study, we found that changes in urine appearance, such as cloudy urine, malodorous urine and hematuria, are alarming signs for UTI, while only few features (e.g. circumcision, stridor) lower the probability of UTI considerably.[7]

For Peer Review Only

Because the clinical presentation is often unclear, primary care physicians need appropriate collection methods for urine testing to rule out UTI in children efficiently. In most outpatient settings, it is not feasible or acceptable to perform invasive, sterile methods such as suprapubic aspiration or catheterization, as recommended in guidelines.[5,6,8] Therefore, primary care physicians use other simple and non-invasive methods such as urine collection bags, nappy pads or clean-catch urine sampling in children that are not yet potty-trained.

Urine collection bags are adhesive plastic bags that are applied around the genitalia of the child. Parents feel that bags are easy to use, but common problems are leaking of the bag, distress on removal and red marks of the glue.[9,10]

Nappy or diaper pads are absorbent pads positioned in the diaper, after which urine is aspirated from the pad with a syringe.[11] Replacing the pad every 30 minutes might reduce contamination.[12]

Clean-catch urine sampling is a direct-catch collection of a first-stream specimen when the child urinates spontaneously. However, parents prefer bags and pads to the clean-catch method, because it is time-consuming and messy.[10] Some studies, therefore, propose clean-catch urine with additional bladder stimulation or cold application to accelerate urination.[13-15]

The aim of this study was to investigate which collection methods are accurate for diagnosing UTI in children presenting to the outpatient setting.

PATIENTS AND METHODS

A systematic review on the diagnostic test accuracy of urine collection methods for urinary tract infections in outpatient children.

Protocol and registration

For Peer Review Only

The protocol for this study was registered a priori on the International prospective register of systematic reviews (PROSPERO, ID number CRD42019139142) and this study is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses guidelines (Appendix S1).

The primary outcome was UTI, defined as growth of a single pathogen on urine culture, in accordance with the criteria per collection method, as used in the European Society for Paediatric Urology guidelines.[5] The secondary outcomes were (1.) the contamination rate (%), (2.) successful catch rate (%), (3.) time to obtain a sample, (4.) complication rate (%), (5.) level of subjective pain during collection, and (6.) parental and clinical satisfaction, when reported. Contamination rates were reported in accordance with the definitions that were used in the primary studies. Successful catch rate was defined as the percentage of samples obtained after a first attempt and complication as any condition occurring directly following urine collection by any method and reported as such.

Information sources and search strategy

A uniform and comprehensive search strategy was used for Medline (through Pubmed), Embase, Web of Science and Cinahl on June 16th 2019. (Appendix S2) Studies were deduplicated in Endnote X.9.2 (Clarivate Analytics, USA). Two investigators (WL and CVA) independently selected articles in duplicate, first based on title and abstract and consequently on the full text, using Covidence online software (Veritas Health Innovation, Australia). One investigator (HB) resolved the conflicts. References of the final articles were searched for other relevant publications. An update of the search was performed on April 6th 2021 following the same protocol by two investigators (HB and WL).

Eligibility criteria

All primary observational studies conducted in ambulatory care settings that compared the diagnostic accuracy of urine culture on paired samples obtained through two or more collection methods in the same patient were eligible for inclusion. Both non-invasive

For Peer Review Only

(adhesive bags, nappy pads or clean-catch) and invasive (catheterization or suprapubic aspiration) collection methods were eligible for the comparison. When possible, culture from invasive methods was used as the reference standard. Ambulatory care was defined as all healthcare settings where patients were not admitted to the hospital. Consequently, studies on children admitted to the pediatric ward or (neonatal) intensive care unit were excluded. Only studies on children below or equal to 18 years were eligible.

Studies with a sample size lower than 50 or on patients older than 18 years were excluded. Randomized controlled trials were excluded, because these studies do not provide paired comparisons of collection methods per patient. Case-control studies with differential sample schemes for cases and controls were also excluded. Studies that only evaluated the diagnostic accuracy of urine tests (e.g. not urine collection methods) were not eligible.

Data collection

Two-by-two tables (true positives, false positives, false negatives, true negatives) were extracted and imported in duplicate to Excel (Microsoft, 2016) by WL and CVA. For zero cells, we used a 0.5 continuity correction. We excluded studies that provided duplicate data.

Risk of bias and applicability assessment

The risk of bias and applicability was assessed with the Quality Assessment of Diagnostic Accuracy Studies tool (QUADAS-2) using RevMan version 5.4 (Cochrane Collaboration Review Manager, Copenhagen). WL assessed the risk of bias and applicability and disagreements were resolved during a consensus meeting (WL, HB, JV). We referred to the European Association of Urology guideline to assess the reference standard bias for interpretation of urine culture positivity threshold (per collection method).[5]

Data analysis

Sensitivity, specificity, positive (LR+) and negative (LR-) likelihood ratios were calculated using Review Manager (RevMan 5.3). To calculate the 95% confidence intervals (95%CI) for

For Peer Review Only

the likelihood ratios, we used an online test calculator by Schwartz *et al.* (http://araw.mede.uic.edu/cgi-bin/testcalc.pl?DT=0&Dt=0&dT=0&dt=0&2x2=Compute) and MedCalc (https://www.medcalc.org/calc/diagnostic_test.php). LR+'s above or equal to five were considered useful for diagnosis and LR-'s lower or equal to 0.20 were considered useful for ruling out UTI.[16] Forest plots of the likelihood ratios were developed using the package ggplot2 (version 3.3.2; Wickham, 2020) in R version 4.0.2 (R Foundation, Austria).[17,18]

RESULTS

Study selection

We screened 7497 articles by title and abstract of which we evaluated 217 articles by full text (Figure 1). Ultimately, seven articles on studies performed in outpatient healthcare were selected for inclusion.[19-25]

Most of the articles were excluded because of an inappropriate study design (55/217, 25%), or because the full text was simply not available (78/217, 36%). Most of these studies did not compare two urine collection methods in one patient and were therefore excluded. Samples of less than 50 children were less frequent (7/217, 3%). Finally, we excluded 29 studies (30/217, 14%) because of a setting different from our review question, e.g. on children admitted to the hospital.

Study characteristics

All study characteristics are listed in Table 1. Seven prospective, cross-sectional studies were included, comprising 1158 children. Six studies were conducted at the emergency department and one study was conducted at an outpatient department of a hospital.[24] None of the studies took place in general practice. Most of the studies (n=5) included children below three years of age, and two studies were on children below five years of age. The

For Peer Review Only

mean age ranged from 40 days to 9 months.[19,20,22,25] In the study by Pylkkänen *et al.,* age was not reported, however specimens were obtained through either suprapubic aspiration, adhesive bags or by clean-catch sampling, assuming children were not yet toilet-trained. Consequently, this study was scored as a high concern for selection bias.[24]

The sample size ranged from 50 to 477 children. There were 379 girls (60.1%) and 252 (39.9%) boys. The following collection methods were used as index tests at least once: adhesive bags, nappy pads, clean-catch, and catheterization (Appendix S3).

Risk of bias and applicability assessment

Overall, there was a low risk of bias and no concerns on the applicability (Figure 2). A detailed overview of the risk of bias assessment is provided in Appendix S4. Risk of bias was caused mainly by the reference standard (n=3), particularly for the inappropriate interpretation of urine culture positivity thresholds, compared to the guidelines. A blind interpretation of both index and reference standard was not possible in two studies and unclear in one. Secondly, an unclear risk of bias was reported for patient selection in one article due to missing information.

Diagnostic accuracies per collection method

Table 1 describes the sensitivities, specificities, and likelihood ratios, while Figure 3 shows forest plots of the likelihood ratios per collection method.

Clean-catch versus urethral catheterization

Clean-catch samples showed 5.6% (n=1/18) false negatives and 5.9% (n=2/34) false positives, compared to urethral catheterization in children less than 90 days old.[22] This provided a sensitivity of 97.0% (95%CI 84.2-99.9%) and specificity of 89.5% (95%CI 66.9-98.7%), corresponding with a LR+ of 9.2 (95%CI 2.5-34.2) and LR- of 0.03 (95%CI 0.00-0.23). All boys included in the study had been circumcised. Both the clean-catch samples and the catheterized sample were obtained within one hour after assessment of the patient.

For Peer Review Only

Adhesive bags versus catheterization

One study reported that adhesive bags gave 8.6% (n=8/93) false positive results, compared to catheterization.[20] Seventeen percent of the negative bag samples (n=17/99), gave a positive result on catheterization. The study authors reported that catheterization was successful in 92.3% of the attempts.[20]

A second study used an adapted collection bag, called a 'midstream collector', which divides urine directly into two compartments, the first-stream going into the main part of the bag and the end-stream going into a test tube.[25] Both urine samples were cultured and compared to catheterization. No false positives were reported by either first- (bag) or end-stream urine samples (tube). Bag specimens gave 19% (n=9/47) false negatives compared to catheterization, and tube specimens 22% (n=11/49) false negatives. Bags were left in place with no time limit in this study, as long as no stool contamination occurred. The influence of urine collection in an upright position, which limits urine contact with the surrounding skin and genitalia, was insignificant.

Nappy pads versus adhesive bags

Feasey et al. compared urine nappy pads to adhesive bags and found no false positives or false negatives.[21] The average time to obtain a sample was 70 minutes when using bags and 93 minutes when using nappy pads. One child needed replacement of the adhesive bag compared to seven children who needed replacement pads. Pads were considered less invasive than bags because they gave less skin irritation and less discomfort for the child.

First-stream versus late-stream catheterization

For catheterization, one study reported no false positive results and 1% (n=1/81) false negative results when first-stream urine was compared to late-stream urine, while assuming unequivocal results (mixed growth or low amount of bacteria) to be negative.[19] This provided

For Peer Review Only

a sensitivity of 85.7% (95%CI 42.1-99.6%) and a specificity of 100.0% (95%CI 95.5-100.0%), corresponding with a LR+ of 138.0 (95%CI 8.6-2224.0) and a LR- of 0.14 (95%CI 0.02-0.88).

A second study included 199 paired urine cultures and showed 33% (n=8/24) false positives and 1% (n=2/175) false negatives of first-stream compared to late-stream.[23] These results corresponded with a LR+ of 20.1 (95%CI 10.0-40.4) and a LR- of 0.12 (95%CI 0.03-0.43).

Adhesive bags and clean-catch versus suprapubic aspiration

One study compared both clean-catch samples or bag specimens to suprapubic aspiration samples, which showed 15.6% (n=50/320) false positives and 33.1% (n=52/157) false negatives.[24] This provided a LR+ of 2.6 (95%CI 2.1-3.3) and a LR- of 0.24 (95%CI 0.18-0.31).

Contamination rate per collection method

The percentage of contaminated samples as defined in the primary studies is shown in Table 2.

Noninvasive methods

Clean-catch samples showed 5% contamination,[22] while other non-invasive collection methods gave 30 to 80% (adhesive bags),[20,21] 64% (nappy pads)[21] and 10% (midstream collector)[25] contamination.

Invasive methods

Catheterization gave contamination rates between 0 and 9%. A first-stream sample culture gave a contamination rate between 5% to 7%, compared to 0 to 2% for late-stream sample cultures.[19,23] For suprapubic aspiration, contamination rate was 0.03%.[24]

DISCUSSION

For Peer Review Only

Principal findings

These results suggest that clean-catch urine might be very useful for ruling out UTI compared to more invasive, sterile methods, such as catheterization. At a LR- of 0.03 (95%CI 0.00 to 0.23), assuming a pre-test probability of UTI of 6%, post-test probability would decrease to 0.2% when urine culture is negative on a clean-catch obtained sample. The contamination rate of clean-catch was 5%, which was comparable to catheterization. However, caution is needed in the interpretation because these results are based on only one study with small sample size in young infants (<90 days old). Contamination rate might have been low, given the study authors used narrow criteria to define contamination, boys were all circumcised and urine had been collected under standardized conditions.

Nappy pads offer a comparable diagnostic accuracy for UTI to adhesive bags but were considered less invasive and easier to implement. However, they both showed a high contamination rate. Contamination might be explicable because of contact between the collected urine with surrounding skin and genitalia, incorrect application, or late removal. Negative bag samples show a LR- of 0.16 (95%CI 0.09 to 0.30) and 0.18 (95%CI 0.12 to 0.18) compared to catheterization, and therefore might be useful for ruling out UTI, however false negative samples remain possible. When urine culture is negative on bag obtained-samples, post-test probability of UTI decreases from 6% to 1%.

If catheterization is performed, urine culture should be performed on a late-stream sample, instead of a first-stream sample.

Limitations and strengths

We performed a comprehensive search strategy to obtain all available information on collection methods for outpatient settings. Studies on hospitalized children were not included, because these results are not applicable for outpatient settings where the prevalence of UTI is relatively low. Only seven studies with an appropriate cross-sectional design (at least 2

For Peer Review Only

matched samples per patient) were included. There were no studies that were performed in general practice.

Secondly, the analyses were restricted to descriptive statistics, because less than three studies were available per collection method and therefore, a meta-analysis was not possible. The study on nappy pads used a suboptimal reference standard, providing only the diagnostic accuracy of nappy pads compared to adhesive bags.

Additionally, we merely reported the values of the contamination rate per method as defined in the primary studies and therefore, accurate comparison between different methods was not possible.

Comparison with existing literature

Diviney *et al.* also proposed clean-catch sampling as the preferred method, because the contamination rate was comparable to invasive methods.[27] Other studies found contamination rates of clean-catch between 4.5% to of 27%,[28-35] whereas Herreros *et al.*[22] reported a contamination rate of 5%. In addition, the proportion of contamination can be influenced by the criteria used to define contamination, the person performing urine collection, techniques used to reduce contamination or stimulate urination, patient characteristics and setting. Prior cleansing of the genitalia gives a reduction in contamination ranging between 8 to 22%,[26] but can also induce false negatives.[36]

The contamination rate for adhesive bags in this study ranged from 30 to 80%, which is comparable with previous findings that showed a pooled percentage of 47% (95%CI 36-58%, n=15) contamination with 61% (95%CI 38-82%, n=12) false positive samples.[37] The study by Etoubleau *et al.* found a relatively low contamination rate of 30%, however only included bag-obtained samples after positive urinalysis and defined contamination by mixed bacterial growth $\geq 10^5$ cfu/mL only, which could be considered a narrow definition.[20] The study by Feasey et al. used a broader definition of contamination rate, which might explain why the contamination rate was higher. For both studies, bags were applied by study nurses and not

For Peer Review Only

by parents. There was prior cleansing and urination was checked every 20 to 30 minutes. In practice however, parents often apply the bag themselves without prior cleansing, and therefore contamination rate might be even higher in routine clinical care.

Etoubleau et al. reported a high number of false negative bag specimens (17 out of 99),[20] compared to catheterized specimens, which is in contract with most recommendations reporting that bag specimens are very unreliable for diagnosis (due to contamination), but can be helpful for ruling out UTI when negative.[5,38] The American guidelines recommend against any use of bag specimens and two recent studies also reported suboptimal sensitivities of bag specimens compared to sterile methods.[8,37,39] False negative bag samples in this study might have been caused by long duration between collection of the bag sample versus the catheter sample, increased hydration during urine collection, usage of an antimicrobial agent before application of the bag or a low positivity threshold for catheterized specimens (>1000 cfu/mL).

The Quick-Wee method is a clean-catch method where urine voiding is triggered through a cold stimulus to improve efficiency. Based on one randomized controlled trial, urine collection is more successful within five minutes (30% versus 9%, p<0.001) and results in greater parental and clinical satisfaction.[13] Although this method might be very relevant for primary care, we did not find a single outpatient study describing its diagnostic accuracy.

Guidelines recommend obtaining urine through catheterization or suprapubic aspiration in children, as these are the most reliable methods.[5,8] However, since low-threshold screening with noninvasive methods is essential for ruling out UTI in low prevalence settings, alternative guidelines for the initial workup in outpatient settings are necessary.

Implications for research and practice

For Peer Review Only

In young infants, clean-catch urine might be very useful for ruling out UTI when urine culture is negative, for settings where more invasive methods are not feasible nor acceptable to perform. In practice, choosing between clean-catch, bags, catheterization or suprapubic aspiration also depends on the seriousness of infection, setting characteristics and experience with catheterization. More information on the time required for obtaining a clean-catch sample and feasibility to implement such a method is needed, as previous studies indicated that parents prefer collection bags and nappy pads.

Other non-invasive collection methods should be developed as an alternative for adhesive bags or nappy pads, with novel techniques to minimize the risk of contamination, such as technology to indicate when the child has urinated, a collection device that minimizes contact between urine and genitalia or collection methods that directly incorporate a diagnostic test. The performance and feasibility of other non-invasive interventions that might accelerate the time to urination, such as suprapubic tapping or cold stimulus, should be further investigated.

ACKNOWLEDGEMENTS

We would like to thank Elien Van Begin for her help in preparing the protocol and the KU Leuven for providing funding for this research.

DECLARATIONS

Funding

This study did not receive any specific funding. HB is funded by a KU Leuven Starting Grant (No ERX-D5331-STG/18/008). The financial sponsor played no role in the design, execution, analysis and interpretation of data, nor in the writing of the study.

Ethics and consent not applicable for this study

For Peer Review Only

Conflict of interest

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Data sharing: Full datasets can be obtained from the corresponding author at jan.verbakel@kuleuven.be

JV affirms that the manuscript is an honest, accurate and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

REFERENCES

1. O'Brien K, Stanton N, Edwards A, Hood K, Butler CC. Prevalence of urinary tract infection (UTI) in sequential acutely unwell children presenting in primary care: exploratory study. *Scand J Prim Health Care.* 2011;29(1):19-22.
2. Verbakel JY, Lemiengre MB, De Burghgraeve T, et al. Validating a decision tree for serious infection: diagnostic accuracy in acutely ill children in ambulatory care. *BMJ Open.* 2015;5(8):e008657.
3. Shaikh N, Mattoo TK, Keren R, et al. Early Antibiotic Treatment for Pediatric Febrile Urinary Tract Infection and Renal Scarring. *JAMA Pediatr.* 2016;170(9):848-854.
4. Dayan PS, Hanson E, Bennett JE, Langsam D, Miller SZ. Clinical course of urinary tract infections in infants younger than 60 days of age. *Pediatr Emerg Care.* 2004;20(2):85-88.
5. Stein R, Dogan HS, Hoebeke P, et al. Urinary tract infections in children: EAU/ESPU guidelines. *Eur Urol.* 2015;67(3):546-558.
6. NICE guideline CG54: Urinary tract infection in under 16s Evidence reviews for UTI diagnosis in under 3 years 2017.
7. Boon H. A., Van den Bruel A, Struyf T., Gillemot A., Bullens D, Verbakel JY. Clinical features to diagnose paediatric urinary tract infections in children. *Accepted for publication in Annals of Family Medicine.* 2021.
8. Roberts KB. Revised AAP guideline on UTI in febrile infants and young children. *Am Fam Physician.* 2012;86(10):940-946.
9. Al-Orifi F, McGillivray D, Tange S, Kramer MS. Urine culture from bag specimens in young children: are the risks too high? *J Pediatr.* 2000;137(2):221-226.

For Peer Review Only

10. Liaw LC, Nayar DM, Pedler SJ, Coulthard MG. Home collection of urine for culture from infants by three methods: survey of parents' preferences and bacterial contamination rates. *BMJ.* 2000;320(7245):1312-1313.
11. Butler CC, Sterne JA, Lawton M, et al. Nappy pad urine samples for investigation and treatment of UTI in young children: the 'DUTY' prospective diagnostic cohort study. *Br J Gen Pract.* 2016;66(648):e516-524.
12. Rao S, Bhatt J, Houghton C, Macfarlane P. An improved urine collection pad method: a randomised clinical trial. *Arch Dis Child.* 2004;89(8):773-775.
13. Kaufman J, Fitzpatrick P, Tosif S, et al. Faster clean catch urine collection (Quick-Wee method) from infants: randomised controlled trial. *Bmj-British Medical Journal.* 2017;357.
14. Morris L. PURLs: An easy approach to obtaining clean-catch urine from infants. *J Fam Pract.* 2018;67(3):166-169.
15. Demonchy D, Ciais C, Fontas E, et al. Evaluation of bladder stimulation as a non-invasive technique for urine collection to diagnose urinary tract infection in infants under 6 months: a randomized multicenter study ("EE-Sti.Ve.N"). *Trials.* 2019;20(1).
16. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA.* 1994;271(9):703-707.
17. *R: A Language and Environment for Statistical Computing* [computer program]. Vienna, Austria2020.
18. Wickham H, Averick M, Bryan J, et al. Welcome to the tidyverse. *Journal of Open Source Software.* 2019;4(43):1686.
19. Dayan PS, Chamberlain JM, Boenning D, Adirim T, Schor JA, Klein BL. A comparison of the initial to the later stream urine in children catheterized to evaluate for a urinary tract infection. *Pediatr Emerg Care.* 2000;16(2):88-90.
20. Etoubleau C, Reveret M, Brouet D, et al. Moving from bag to catheter for urine collection in non-toilet-trained children suspected of having urinary tract infection: a paired comparison of urine cultures. *J Pediatr.* 2009;154(6):803-806.
21. Feasey S. Are Newcastle urine collection pads suitable as a means of collecting specimens from infants? *Paediatr Nurs.* 1999;11(9):17-21.
22. Herreros ML, Tagarro A, Garcia-Pose A, Sanchez A, Canete A, Gili P. Accuracy of a new clean-catch technique for diagnosis of urinary tract infection in infants younger than 90 days of age. *Paediatr Child Health.* 2015;20(6):e30-32.
23. Megged O, Zilberstein L, Ben Shalom E, Erlichman M. Contamination of urinary cultures in initial-stream versus later-stream urine in children undergoing bladder catheterization for the diagnosis of urinary tract infection. *Eur J Emerg Med.* 2017;24(6):e17-e20.
24. Pylkkanen J, Vilska J, Koskimies O. Diagnostic value of symptoms and clean-voided urine specimen in childhood urinary tract infection. *Acta Paediatr Scand.* 1979;68(3):341-344.
25. Verliat-Guinaud J, Blanc P, Garnier F, Gajdos V, Guigonis V. A midstream urine collector is not a good alternative to a sterile collection method during the diagnosis of urinary tract infection. *Acta Paediatr.* 2015;104(9):e395-400.
26. LaRocco MT, Franek J, Leibach EK, et al. Effectiveness of Preanalytic Practices on Contamination and Diagnostic Accuracy of Urine Cultures: a Laboratory Medicine Best Practices Systematic Review and Meta-analysis. *Clin Microbiol Rev.* 2016;29(1):105-147.
27. Diviney J, Jaswon MS. Urine collection methods and dipstick testing in non-toilet-trained children. *Pediatr Nephrol.* 2020.
28. Tosif S, Baker A, Oakley E, Donath S, Babl FE. Contamination rates of different urine collection methods for the diagnosis of urinary tract infections in young children: an observational cohort study. *J Paediatr Child Health.* 2012;48(8):659-664.
29. Rivas-Garcia A, Lorente-Romero J, Lopez-Blazquez M, Rodriguez-Jimenez C, Castro-Rodriguez C, Miguez-Navarro MC. Contamination in Urine Samples Collected Using Bladder Stimulation

For Peer Review Only

and Clean Catch Versus Urinary Catheterization in Infants Younger Than 90 Days. *Pediatr Emerg Care.* 2020.

30. Mulcrone AE, Parikh M, Ahmad FA. Reducing infant catheterization in the emergency department through clean-catch urine collection. *J Am Coll Emerg Physicians Open.* 2020;1(6):1533-1541.
31. Teo S, Cheek JA, Craig S. Improving clean-catch contamination rates: A prospective interventional cohort study. *Emerg Med Australas.* 2016;28(6):698-703.
32. Alam MT, Coulter JB, Pacheco J, et al. Comparison of urine contamination rates using three different methods of collection: clean-catch, cotton wool pad and urine bag. *Ann Trop Paediatr.* 2005;25(1):29-34.
33. Labrosse M, Levy A, Autmizguine J, Gravel J. Evaluation of a New Strategy for Clean-Catch Urine in Infants. *Pediatrics.* 2016;138(3).
34. Kaufman J, Fitzpatrick P, Tosif S, et al. Faster clean catch urine collection from infants: The quick-wee randomised controlled trial. *Arch Dis Child.* 2017;102:A108-A109.
35. Ho IV, Lee CH, Fry M. A prospective comparative pilot study comparing the urine collection pad with clean catch urine technique in non-toilet-trained children. *Int Emerg Nurs.* 2014;22(2):94-97.
36. Lee JW., SJ. L. Comparison of ultrasound-guided suprapubic aspiration with urethral catheterization in infants. *J Korean Soc Pediatr Nephrol.* 2007;11:59-64.
37. Ochoa Sangrador C, Pascual Terrazas A. [Systematic review of the validity of urine cultures collected by sterile perineal bags]. *An Pediatr (Barc).* 2016;84(2):97-105.
38. Whiting P, Westwood M, Bojke L, al. e. Clinical effectiveness and cost-effectiveness of tests for the diagnosis and investigation of urinary tract infection in children: a systematic review and economic model. *Health Technol Assess.* 2006;10(36).
39. Kim GA, Koo JW. Validity of bag urine culture for predicting urinary tract infections in febrile infants: a paired comparison of urine collection methods. *Korean J Pediatr.* 2015;58(5):183-189.

For Peer Review Only

TABLES

Table 1: Study characteristics and diagnostic accuracies per collection method comparison

Study	Age	Prev (n/N)	Setting	cfu/mL (I)	cfu/mL (R)	Sens (%) (95%CI)	Spec (%) (95%CI)	LR+ (95%CI)	LR- (95%CI)
Clean-catch (I) versus catheterization (R)									
Herreros *et al.* 2015	≤90d	33/52	ED	10^5	10^4	97.0 (84.2-99.9)	89.5 (66.9-98.7)	9.2 (2.5-34.2)	0.03 (0.00-0.23)
Adhesive bag (I) versus catheterization (R)									
Etoubleau *et al.* 2009	<3y	102/192	ED	10^5	10^3	83.3 (74.7-90.0)	91.1 (83.2-96.1)	9.4 (4.8-18.3)	0.18 (0.12-0.28)
Midstream collector bag (I) versus catheterization (R)									
Verliat-Guinaud *et al.* 2015	<3y	56/94	ED	10^5	10^3	83.9 (71.7-92.4)	100.0 (90.8-100.0)	65.0 (4.1-1017.0)	0.16 (0.09-0.30)
Midstream collector tube (I) versus catheterization (R)									
Verliat-Guinaud *et al.* 2015	<3y	56/94	ED	10^5	10^3	80.4 (67.6-89.8)	100.0 (90.8-100.0)	62.0 (3.9-975.0)	0.20 (0.12-0.34)
Nappy pad (I) versus adhesive bag (R)									
Feasey *et al.*	≤5y	2/50	ED	10^5	10^5	100.0	100.0	78.0	0.20

For Peer Review Only

1999						(15.8-100.0)	(92.6-100.0)	(4.6-1310.0)	(0.02-2.41)
Early stream catheterization (I) versus late stream catheterization (R)									
Dayan *et al.* 2000	≤4y	7/87	ED	$5x10^4$	$5x10^4$	85.7 (42.1-99.6)	100.0 (95.5-100.0)	138.0 (8.6-2224.0)	0.14 (0.02-0.88)
Megged *et al.* 2017	≤2y	18/199	ED	$5x10^4$	$5x10^4$	88.9 (65.3-98.6)	95.6 (91.5-98.1)	20.1 (10.0-40.4)	0.12 (0.03-0.43)
Clean-catch urine or adhesive bag (I) versus suprapubic aspiration (R)									
Pylkkänen *et al.* 1979	NR	322/477	OD	10^5	NR	83.9 (79.4-87.7)	67.7 (59.8-75.0)	2.6 (2.1-3.3)	0.24 (0.18-0.31)

Prev= prevalence, n= number of UTI cases; N= sample size, cfu/mL= colony-forming-units per milliliter, I= index test, R= reference standard, Sens= sensitivity, Spec= Specificity, LR+= positive likelihood ratio, LR-= negative likelihood ratio, 95%CI= 95% confidence intervals, d= days, y= years, ED= emergency department, OD= outpatient department

For Peer Review Only

Table 2: Reported % contamination rate for each collection method

Author, year	Definition of contamination*	contamination rate* (%, n/N)
CLEAN-CATCH		
Herreros *et al.,* 2015	Mixed bacterial growth or growth of one or more nonpathogenic bacteria.	5.0% (n=3/60)
ADHESIVE BAGS		
Etoubleau *et al.*, 2009	Mixed bacterial growth, $\geq 10^5$ cfu/mL	30.2% (n=58/192)
Feasey *et al.,* 1999	Mixed bacterial growth or growth of less than 10^5 cfu/mL of a single organism.	80.0% (n=40/50)
MIDSTREAM COLLECTOR		
Verliat-Guinaud *et al.,* 2015	Bacteriuria $\geq 10^3$ cfu/mL with ≥2 pathogens	9.6% (n=9/94) in bag and tube
NAPPY PADS		
Feasey *et al.,* 1999	Mixed bacterial growth or growth of less than 10^5 cfu/mL of a single organism.	64.0% (n=32/50)
CATHETERIZATION		
Dayan *et al.,* 2000	Growth of a single pathogen between 10^2 to $5x10^4$ cfu/mL or $>5x10^4$ cfu/mL, which is negative in a second sample.	4.6% (n=4/87), first-stream samples
Etoubleau *et al.*, 2009	Mixed bacterial growth, $\geq 10^3$ cfu/mL	8.3% (n=16/192)
Herreros *et al.,* 2015	Mixed bacterial growth or	8.3% (n= 5/60)

For Peer Review Only

	growth of one or more nonpathogenic bacteria.	
Megged *et al.,* 2017	Bacterial growth of $<10^4$cfu/mL	7.0% (n=14/199), first-stream samples 2.0% (n=4/199), late-stream samples
Verliat-Guinaud *et al.,* 2015	Bacteriuria of ≥2 pathogens ≥10^3 cfu/mL	8.5% (n=8/94)
SUPRAPUBIC ASPIRATION		
Pylkkänen *et al.,* 1979	1 or 2 occasional colonies on blood agar plates.	0.03% (6/200)

*as defined in the primary study; NR= not reported, n= number of samples contaminated, N= total number of samples

FIGURE Legends

Figure 1: PRISMA flowchart of included studies

PRISMA= Preferred Reporting Items for a Systematic review and Meta-Analysis of Diagnostic Test Accuracy Studies

Figure 2: Risk of bias and applicability assessment

Review authors' judgements about each domain presented as percentages across included studies

For Peer Review Only

Figure 3: Forest plots of likelihood ratios for the collection methods

CC= clean-catch urine; CATH= urinary catheterization; AB= adhesive bags; MS bag= midstream collector bag; MS tube= midstream collector tube; NP= nappy pad; SPA= suprapubic aspiration

For Peer Review Only

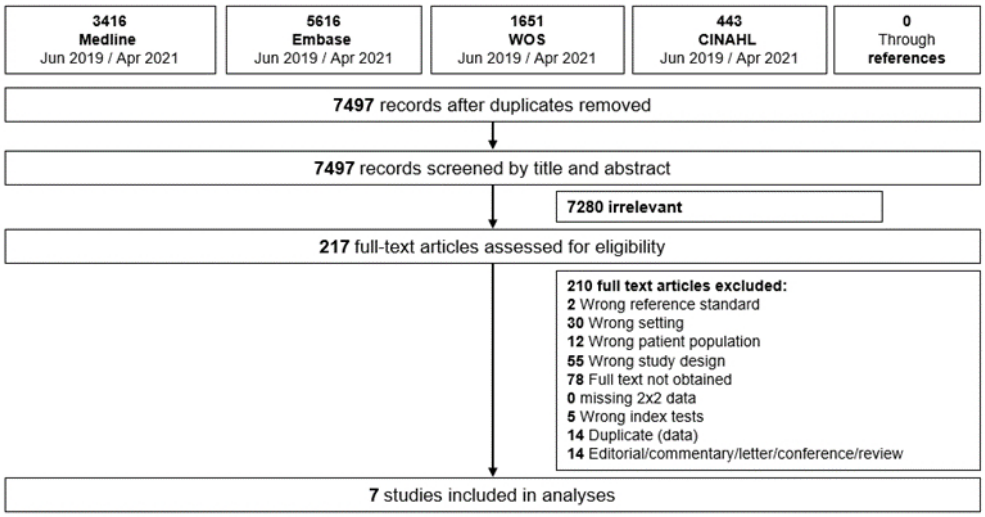


Figure 1: PRISMA flowchart of included studies/ PRISMA= Preferred Reporting Items for a Systematic review and Meta-Analysis of Diagnostic Test Accuracy Studies

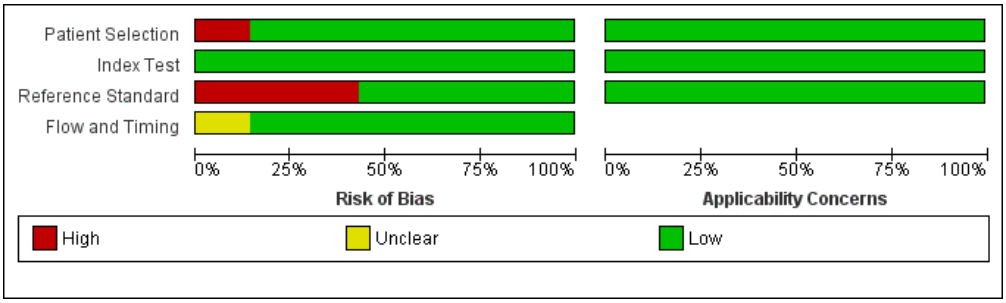


Figure 2: Risk of bias and applicability assessment/ Review authors' judgements about each domain presented as percentages across included studies

239x70mm (72 x 72 DPI)

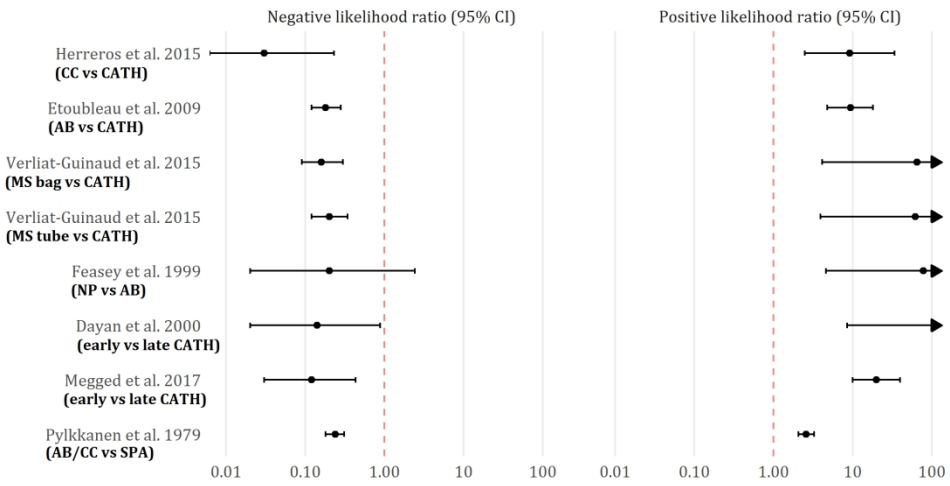


Figure 3: Forest plots of likelihood ratios for the collection methods/ CC= clean-catch urine; CATH= urinary catheterization; AB= adhesive bags; MS bag= midstream collector bag; MS tube= midstream collector tube; NP= nappy pad; SPA= suprapubic aspiration

APPENDICES

Appendix S1: PRISMA guidelines

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
TITLE / ABSTRACT			
Title	1	Identify the report as a systematic review (+/- meta-analysis) of diagnostic test accuracy (DTA) studies.	1
Abstract	2	Abstract: See PRISMA-DTA for abstracts.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3-4
Clinical role of index test	D1	State the scientific and clinical background, including the intended use and clinical role of the index test, and if applicable, the rationale for minimally acceptable test accuracy (or minimum difference in accuracy for comparative design).	3-4
Objectives	4	Provide an explicit statement of question(s) being addressed in terms of participants, index test(s), and target condition(s).	4
METHODS			

For Peer Review Only

Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	5
Eligibility criteria	6	Specify study characteristics (participants, setting, index test(s), reference standard(s), target condition(s), and study design) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5-6
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search	8	Present full search strategies for all electronic databases and other sources searched, including any limits used, such that they could be repeated.	5
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5-6
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6
Definitions for data extraction	11	Provide definitions used in data extraction and classifications of target condition(s), index test(s), reference standard(s) and other characteristics (e.g. study design, clinical setting).	5-6
Risk of bias and applicability	12	Describe methods used for assessing risk of bias in individual studies and concerns regarding the applicability to the review question.	6

For Peer Review Only

Diagnostic accuracy measures	13	State the principal diagnostic accuracy measure(s) reported (e.g. sensitivity, specificity) and state the unit of assessment (e.g. per-patient, per-lesion).	6-7
Synthesis of results	14	Describe methods of handling data, combining results of studies and describing variability between studies. This could include, but is not limited to: a) handling of multiple definitions of target condition. b) handling of multiple thresholds of test positivity, c) handling multiple index test readers, d) handling of indeterminate test results, e) grouping and comparing tests, f) handling of different reference standards	6-7
RESULTS			
Study selection	17	Provide numbers of studies screened, assessed for eligibility, included in the review (and included in meta-analysis, if applicable) with reasons for exclusions at each stage, ideally with a flow diagram.	7
Study characteristics	18	For each included study provide citations and present key characteristics including: a) participant characteristics (presentation, prior testing), b) clinical setting, c) study design, d) target condition definition, e) index test, f) reference standard, g) sample size, h) funding sources	7-8
Risk of bias and applicability	19	Present evaluation of risk of bias and concerns regarding applicability for each study.	8
Results of individual studies	20	For each analysis in each study (e.g. unique combination of index test, reference standard, and positivity threshold) report 2x2 data (TP, FP, FN, TN) with estimates of diagnostic accuracy and confidence intervals, ideally with a forest or receiver operator characteristic (ROC) plot.	8-10
Synthesis of results	21	Describe test accuracy, including variability; if meta-analysis was done, include results and confidence intervals.	8-10

For Peer Review Only

Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression; analysis of index test: failure rates, proportion of inconclusive results, adverse events).	8-10
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence.	10-11
Limitations	25	Discuss limitations from included studies (e.g. risk of bias and concerns regarding applicability) and from the review process (e.g. incomplete retrieval of identified research).	11-12
Conclusions	26	Provide a general interpretation of the results in the context of other evidence. Discuss implications for future research and clinical practice (e.g. the intended use and clinical role of the index test).	11-14
FUNDING			
Funding	27	For the systematic review, describe the sources of funding and other support and the role of the funders.	14

Adapted From: McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, The PRISMA-DTA Group (2018). Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA. 2018 Jan 23;319(4):388-396. doi: 10.1001/jama.2017.19163.

For more information, visit: **www.prisma-statement.org**.

For Peer Review Only

Appendix S2: Search strategy (Medline through Pubmed*) * The search strategies for Embase, Web of Science and Cinahl are available from the corresponding author on request.

Concept 1: children

"Child"[Mesh] OR "Infant"[Mesh] OR "Adolescent"[Mesh] OR “Minors”[Mesh] OR adolescen*[tiab] OR baby[tiab] OR babies[tiab] OR child*[tiab] OR boy*[tiab] OR girl*[tiab] OR infant*[tiab] OR infancy[tiab] OR junior*[tiab] OR juvenile*[tiab] OR kid*[tiab] OR minor*[tiab] OR newborn*[tiab] OR neonat*[tiab] OR neo-nat*[tiab] OR offspring*[tiab] OR perinatal*[tiab] OR peri-natal*[tiab] OR premature*[tiab] OR preterm*[tiab] OR pre-term*[tiab] OR pediatric*[tiab] OR paediatric*[tiab] OR preschool*[tiab] OR pre-school*[tiab] OR postnatal*[tiab] OR post-natal*[tiab] OR premature*[tiab] OR preterm*[tiab] OR pre-term*[tiab] OR preteen*[tiab] OR pre-teen*[tiab] OR prepubertal*[tiab] OR prepubescent*[tiab] OR pubescent*[tiab] OR puberty[tiab] OR pueril*[tiab] OR schoolchild*[tiab] OR sophomore*[tiab] OR student*[tiab] OR teen*[tiab] OR toddler*[tiab] OR underage*[tiab] OR youth*[tiab] OR young*[tiab]

AND

Concept 2: urine collection methods

"Urine Specimen Collection"[Mesh] OR “absorbent pad*”[tiab] OR absorbent-pad*[tiab] OR bagspecimen*[tiab] OR bag-urine[tiab] OR bladder-aspiration[tiab] OR bladder-catheter*[tiab] OR bladder-stimulation[tiab] OR bladder-puncture[tiab] OR catheter-specimen*[tiab] OR cleancatch[tiab] OR clean-voided[tiab] OR collection-bag*[tiab] OR diaper*[tiab] OR first-stream*[tiab] OR first-void*[tiab] OR midstream[tiab] OR “mid-stream”[tiab] OR nappy[tiab] OR nappies[tiab] OR quick-wee[tiab] OR specimen-collection*[tiab] OR sterile-bag*[tiab] OR sterile-sampl*[tiab] OR sterile-specim*[tiab] OR sterile-urin*[tiab] OR suprapubic-aspiration[tiab] OR spa[tiab] OR suprapubic-stimulation[tiab] OR tubc[tiab] OR urethral-catheter*[tiab] OR urinary-catheter[tiab] OR urinary-bag*[tiab] OR urine-bag*[tiab] OR urine-capt*[tiab] OR urine-collect*[tiab] OR urinesampl*[tiab] OR urine-specimen*[tiab]

For Peer Review Only

AND

Concept 3: urinary tract infection / contamination rate

"Urinary tract infections"[Mesh] OR "infection/urine"[Mesh] OR "cystitis"[Mesh] OR "Pyelitis"[Mesh] OR "Pyonephrosis"[Mesh] OR "bacteruria"[Mesh] OR "pyuria"[Mesh] OR bacteriuria[tiab] OR bladder-infection*[tiab] OR bladder-inflammation*[tiab] OR ((contaminat*[tiab]) AND (bacterium[tiab] OR bacterial[tiab] OR sample*[tiab] OR specimen*[tiab] OR urine[tiab])) OR cystitis[tiab] OR cystitides[tiab] OR cystopyelitis[tiab] OR kidney-infection*[tiab] OR kidneyinflammation*[tiab] OR pyelitis[tiab] OR pyelocystitis[tiab] OR pyelonephritis[tiab] OR pyonephrosis[tiab] OR pyuria[tiab] OR urinary-tract-infection*[tiab] OR urinary-tractinflammation*[tiab] OR UTI[tiab] OR urosepsis[tiab] OR ((infection*[tiab] OR inflam*[tiab]) AND (bladder[tiab] OR kidney[tiab] OR "urinary tract"[tiab]))

For Peer Review Only

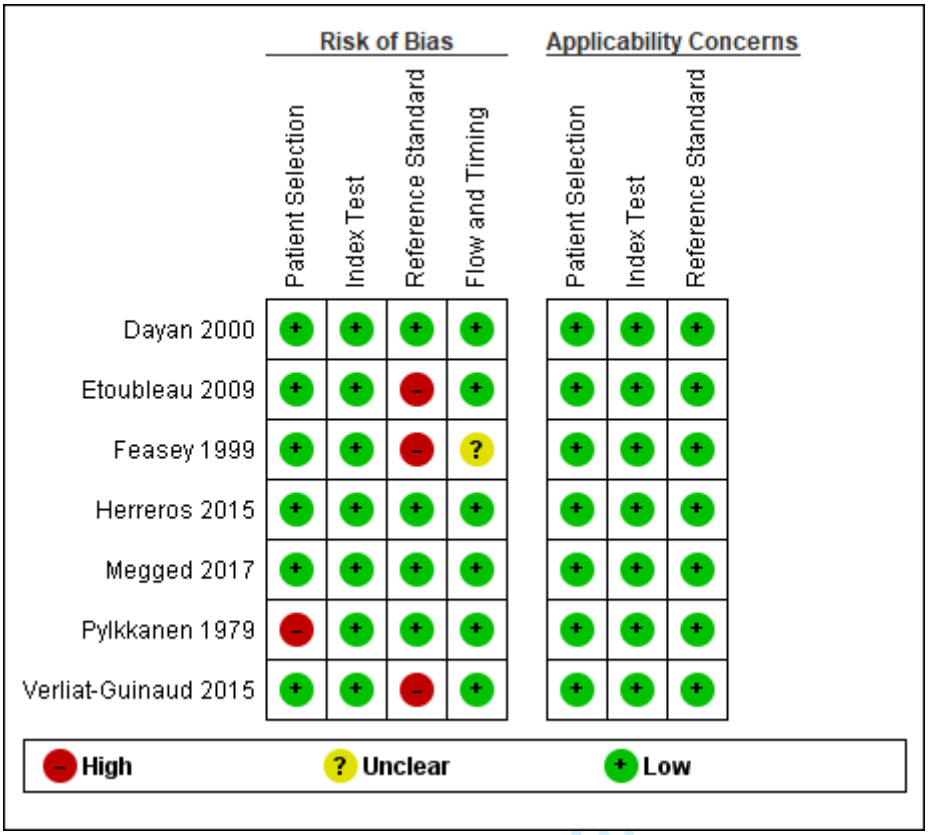
Appendix S3. Overview of the collection methods

		Reference standard			
		UC	SPA	NP	BS
Index test	CC	Herreros et al, 2018	Pylkkänen et al, 1979		
	BS	Etoubleau et al, 2009			
	MC	Verliat-Guinaud et al, 2015			
	NP				Feasey et al, 2009
	UC	Dayan et al, 2000 Megged et al, 2017			

CC = clean-catch; BS = bag specimen; MC: midstream collector; NP = nappy pad; UC = urinary catheterization; SPA = suprapubic aspiration

For Peer Review Only

Appendix S4. Detailed overview of the risk of bias and applicability assessment



Review authors' judgements about each domain for each included study

For Peer Review Only