EDITORIAL COMMENTARY

Defining urinary tract infection by bacterial colony counts: a case for 100,000 colonies/ml as the best threshold

Malcolm G. Coulthard[1]

Why is it important to get the diagnosis of UTIs right?

We need to diagnose urinary tract infections quickly

Urinary tract infections (UTIs) are common in childhood and cause a considerable health burden. Though many children have mild symptoms and are easily treated, some present severely unwell with urosepsis. UTIs may be associated with renal scarring, and in severe cases with hypertension, and renal impairment. They should be treated promptly to relieve symptoms and ideally within three days in infants < 2 years of age to reduce their risk of developing permanent kidney scars [1, 2]. Virtually all authors, including the American Academy of Pediatrics (AAP) [3, 4] and the UK National Institute for Health and Clinical Excellence (NICE) [5], emphasise the importance of rapid antibiotic treatment in the very young to reduce sequelae.

And we need to diagnose them accurately

Whilst the case for identifying children with UTIs is clear, it is also important not to falsely overdiagnose them. Not every child with dysuria has a UTI; some may simply have vulvitis or balanitis, and others may have a febrile illness and a poor fluid intake resulting in them passing concentrated urine that stings. Diagnosing these children as having UTIs will lead to the misuse of antibiotics and may result in children having unnecessary investigations, which is burdensome and wasteful.

Why can diagnosing UTIs be difficult?

Making any clinical diagnosis involves a complex integration of information, including the child's prior clinical history, known risk factors such as age and sex, and the probability of a particular symptom or symptom complexes being caused by a particular illness, as well as the results of laboratory tests. Indeed, the power with which a positive or negative laboratory test result can rule in or rule out a diagnosis depends on its prior probability as well as the test's sensitivity and specificity, as illustrated by leaf plots [6]. False diagnoses can only be confidently ruled out by tests that are highly sensitive, and true diagnoses can only be confidently confirmed by tests that are highly specific. An ideal test is one which can be calibrated to be both highly sensitive and highly specific.

The possible urine tests for UTIs can be divided into those that detect the presence of bacterial themselves and those which detect associated changes, such as increased white blood cell (WBC) numbers or nitrite concentrations. Unfortunately it has always been known that urinary WBC numbers have low diagnostic power [7]; they are neither sufficiently sensitive (they may disappear rapidly [8]) nor specific (they are often present in many children with fever from other causes [9]), and recently it has been shown that urine nitrite sticks miss 77% of UTIs in infants [10]. We are therefore left dependent upon identifying bacteria in urine to make a reliable laboratory diagnosis of UTI in children. First, I will deal with urine culture and then with point-of-care microscopy.

How to create a robust laboratory definition of UTI by bacterial culture?

The primary problem of contamination

There would be no difficulty in diagnosing UTIs by urine culture if uropathogens were rarely found elsewhere and if

✉ Malcolm G. Coulthard
  malcolm.coulthard@nuth.nhs.uk

1 Great North Children's Hospital, Queen Victoria Road, Newcastle upon Tyne NE1 4LP, UK

urine was a fastidious culture medium because then any bacterial colonies identified could be assumed to be pathological. However, this is far from the case, and it is only too common for urine from uninfected children to become contaminated with peri-urethral or skin bacteria. Kass introduced diagnostic bacterial quantification 60 years ago as a strategy to distinguish contamination from urine infection [11, 12], but the best ways to collect children's urine samples and interpret their culture results remain controversial.

How to define sterile urine and rule out a UTI?

It seems obvious that if a urine sample from a child who is not taking an antibiotic does not grow any uropathogenic bacteria (including coliforms, *Escherichia coli*, *Enterococcus*, *Proteus*, or *Klebsiella*), then it can be concluded that they do not have a UTI so long as the collection was not made using inappropriate skin antisepsis. In standard clinical laboratory practice, 1 μl of urine is inoculated onto a petri dish, so "sterile" urine with no bacterial growth will be reported if the urine contains $< 10^3$ colony forming units (cfu) per millilitre.

It follows that if a child has two or more urine samples collected in immediate sequence and at least one of them was sterile, they do *not* have a UTI. It further follows that if another urine did grow bacteria, either the child had developed a UTI at that very instant (extremely unlikely) or the specimen was contaminated. Therefore, in publications which report discordant serial cultures, I conclude that one sterile sample excludes a UTI, and the positives indicate contamination regardless of the bacterial count, species, or urine collection method. Similarly, two serially collected urine samples containing different species are evidence that both were contaminated.

Many authors have interpreted data without using this and consistent approach and have drawn illogical conclusions according to their prior expectations. For example, there are six studies of children where paired samples collected by suprapubic aspiration (SPA) and clean voiding have shown discordant results in both directions, where logic indicates that some of the SPA urines must have become lightly contaminated. However, in each case, the authors (who appear to believe that SPA samples cannot become contaminated) have argued that the pairs with void-positive/SPA-negative results demonstrate urethral contamination, whilst SPA-positive/void-negative pairs indicated that the "UTI had been missed" in the urethral specimen [13–18]. Similarly, when only the first flow from a urethral catheter sample (where the authors expected contamination might occur) grew organisms, it was called contamination, but when only the subsequent stream was culture-positive, it is assumed that the first flow sterility was due to "non-uniform bacterial excretion in different urine phases" or the "potential randomness of growth of some bacteria" [19].

How to define a highly probable UTI?

Because it is possible for apparently scrupulously collected urine samples from children to become contaminated, there is no scientific way to determine whether any particular single specimen that grows a pure uropathogen indicates a UTI or contamination. However, the probability that it is a genuine UTI increases sharply if exactly the same result is found in a second sample, because the chances of a random false-positive error will be reduced by its square, for example, from 10 to 1% or from 5 to 0.25%. Although it is clearly not possible to reach 100% certainty, imposing a diagnostic criterion of ≥2 identical cultures from published reports greatly increases the confidence (in these examples to 99% or 99.75%) of a genuine UTI, so I have used this "gold standard" method to analyse the literature and have excluded less rigorous studies.

How to determine the best diagnostic threshold for childhood UTIs?

By accepting two identical positive cultures as evidence of a UTI and rejecting all other results, it is possible to independently determine the diagnostic sensitivity and specificity for a range of possible colony count thresholds and to see if this varies with the urine collection method. This approach avoids the bias inherent in most studies where each child has a single sample taken and a particular culture cut-off is pre-determined, and its validity is then judged by seeing how well it seems to separate children according to the clinician's prior assessment (or worse, where different pre-determined thresholds are allocated to different urine collection methods).

It is important to recognise that routine quantitative culture testing has a limited range which does not reach the true bacterial numbers in most UTIs, which was shown to be between 10^7 and 10^8 bacteria/ml in adults by Kass [11]. This means that a typical adult sample bottle will contain around *one billion* organisms and that 1 μl of urine inoculated onto a petri dish will contain approximately 50,000 viable organisms, where the nutrient agar is only sufficient to sustain about 500 colonies, so only about 1% will grow. Thus, without a pre-culture dilution step, it is only possible to measure bacterial concentrations between 10^3/ml and about 5×10^5/ml (usually reported as $\geq 10^5$ cfu/ml). Kass warned that contamination during voiding frequently produced urinary colony counts of $\geq 10^5$/ml [11, 12] and would therefore generate false-positive results in adults. Some of the published paediatric studies did use pre-dilution to detect high bacterial concentrations, and some inoculated larger urine volumes to identify very low counts.

Performing meta-analyses to determine the best colony count UTI diagnostic thresholds for voided, SPA, and catheter urine samples

Collecting the material

I searched for papers where children had paired urine samples cultured quantitatively by reviewing those referenced in the AAP and NICE childhood UTI guidelines [3–5], by undertaking a MEDLINE (1946 to March 2019) search for "urinary tract infection or bacteriuria" in children (≤ 18 years) and by following up earlier papers not included in computerised databases. I excluded children already known to have structural urinary tract abnormalities and included studies which compared urine collection methods in healthy individuals as well as those with a clinical suspicion of UTI. I also reanalysed the data our group has previously published on bacterial quantitation [20] to look for sex or age effects.

Analysis methods

I used the urine culture "gold standard" described to define children *with a UTI* as having ≥ 2 urine cultures with the same pure uropathogen and those *without a UTI* if they had at least one urine culture with $< 10^3$ cfu/ml. I have plotted colony counts on $\log_{10}$ axes, used <> symbols to denote concentrations only reported in ranges, and compared their geometric means using unpaired t tests. The standard clinical colony count ranges of 10^3 to $\geq 10^5$ cfu/ml are shaded in coded colours, higher values obtained by prior dilution are shown in grey, and counts of $< 10^3$ cfu/ml produced by culturing 100 μl of urine are shown below the main plots. I either used χ^2 or Fisher's exact test to test for differences in the pooled data, according to the dataset size. I plotted leaf plots to determine the effects of their sensitivity and specificity data on the predictive values of positive or negative test results for all levels of pre-test probability [6].
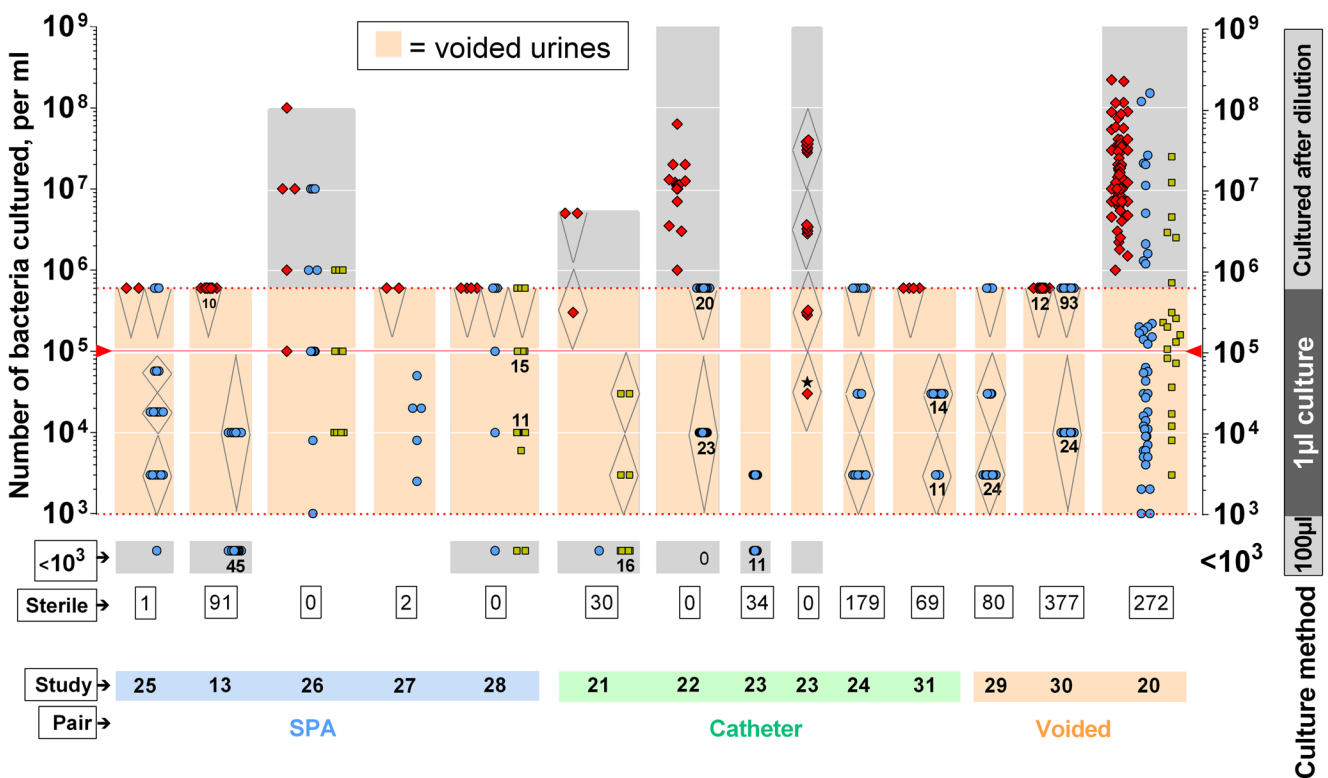

Fig. 1 Concentrations of uropathogens cultured from voided urines, plotted on a log scale. Red symbols = both urines grew the same organism (UTI). Blue circles = sample grew a uropathogen, but its pair was sterile (contaminant). Yellow squares = mixed bacteria. The salmon boxes indicate standard laboratory culture count limits, and the grey boxes show the extended limits in that laboratory. The <> symbols indicate data that was reported in concentration ranges (e.g., $> 10^5$/ml or 10^{3-4}/ml). When more than 10 data points fall together, the number is indicated below. The number of sterile urines is shown in a box at the bottom. The starred data point was the first sample from a patient whose subsequent colony counts were $> 10^7$ cfu/ml. The technique used to collect the paired urine for each study is indicated at the bottom

Meta-analysis 1: voided urines

High-quality quantitative data for voided urines

There were 13 valid papers [13, 20–31] including one with two sub-studies [23], giving 14 reports on 1270 children, of whom 106 had UTIs and 1164 did not (Table 1). Most groups studied children suspected of having UTIs, one collected a clean catch sample before catheterising children for micturating cystograms [31], one screened for UTIs in healthy school children [22], and two compared different voided collection techniques in healthy children in hospital [30] or at home [29]. Some studies were rejected as they had insufficiently detailed data [14–18, 32–34].

Children with UTIs

The quantitative culture results for voided urines are plotted on salmon pink columns in Fig. 1, with children with UTIs shown by red diamonds. The three right-hand plots show their total study data because these groups only collected voided sample pairs. In the other studies, one sample was voided and its pair was collected by either SPA or urethral catheter as indicated below, and these results are plotted in Fig. 2. The urine colony counts for children with UTIs were similar regardless of the urine collection method, so I have analysed them together. Of the 212 infected urines, 210 (99%) had colony counts of $\geq 10^5$ cfu/ml and two were in the range $10^{4–5}$ cfu/ml, which gave a sensitivity of 1.00 at a threshold of 10^4 cfu/ml and of 0.99 at 10^5 cfu/ml. In the 150 samples where higher concentrations could be counted after dilution, the range was much higher, giving a sensitivity of 0.93 at a threshold of 10^6 cfu/ml and 0.58 at 10^7 cfu/ml. The two children whose first samples contained $10^{4–5}$ cfu/ml both had $> 10^7$/ml in their subsequent specimens [23], consistent with them both developing UTIs. A reanalysis of our previously published colony count data [20] showed that the bacterial concentrations were unaffected by sex or age, comparing infants < 2 years, children aged 2 to 9 years, and older children.
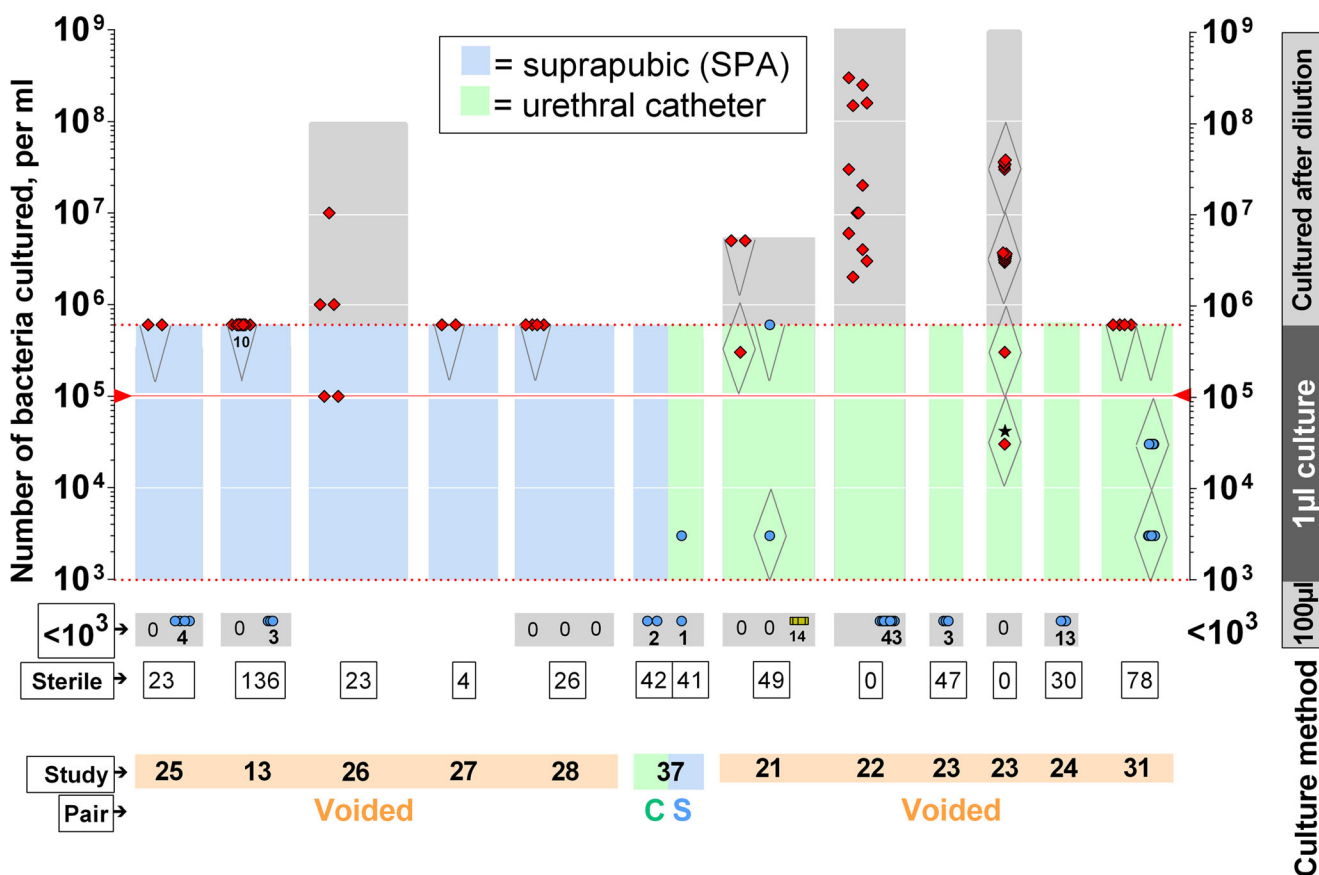

Fig. 2 Concentrations of uropathogens cultured by SPA (blue columns) or catheter sampling (green columns). Red symbols = both urines grew the same organism (UTI). Blue circles = sample grew a uropathogen, but its pair was sterile (contaminant). Yellow squares = mixed bacteria. The blue and green boxes indicate standard laboratory culture count limits, and the grey boxes show the extended limits in that laboratory. The <> symbols indicate data that was reported in concentration ranges (e.g., > 10^5/ml or $10^{3–4}$/ml). For clarity, the number of data points is indicated below in some cases. The number of sterile urines is shown in boxes. The starred data point was the first sample from a patient whose subsequent colony counts were > 10^7 cfu/ml. The technique used to collect the paired urine for each study is indicated at the bottom (mostly voided)

Children without UTIs

Most of the 1583 voided urine samples from children without UTIs were sterile or had mixed organisms, but 303 (19.1%) were contaminated with a single uropathogen (Fig. 1, blue circles). These would be indistinguishable from genuine UTIs if just one sample was collected per child, giving a specificity of 0.809. Many had high counts, giving the following false-positive rates (specificities): 13.8% (0.862) at 10^4 cfu/ml, 9.3% (0.907) at 10^5 cfu/ml, and 1.0% (0.990) at 10^6 cfu/ml. Urines contaminated with mixed bacterial species including a uropathogen (Fig. 1, yellow squares) had similar colony counts.
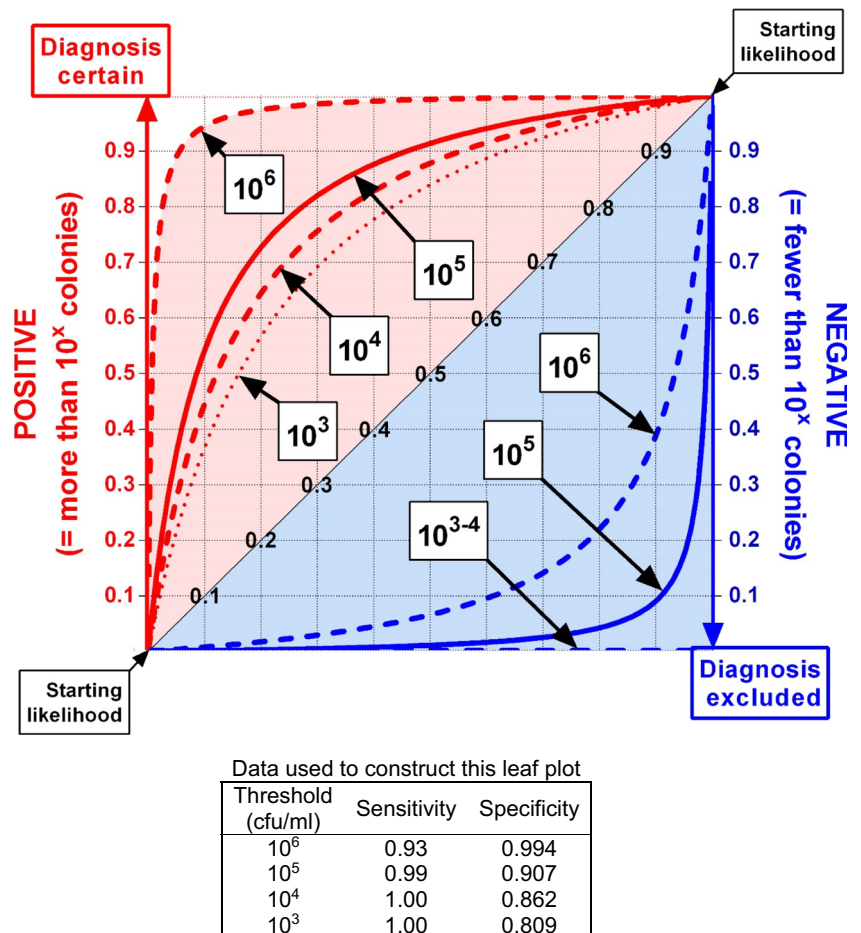
Defining the best clinical diagnostic threshold for a single voided urine sample

Setting an ideal diagnostic threshold depends upon the clinical circumstances and typically involves compromise—higher cut-offs risk missing genuine UTIs, whilst lower ones risk making false diagnoses in children with contaminated samples. In addition, the diagnostic impact of test results is greatly influenced by the clinical probability that the child had a UTI in the first place. These judgements are made easier to visualise by using a
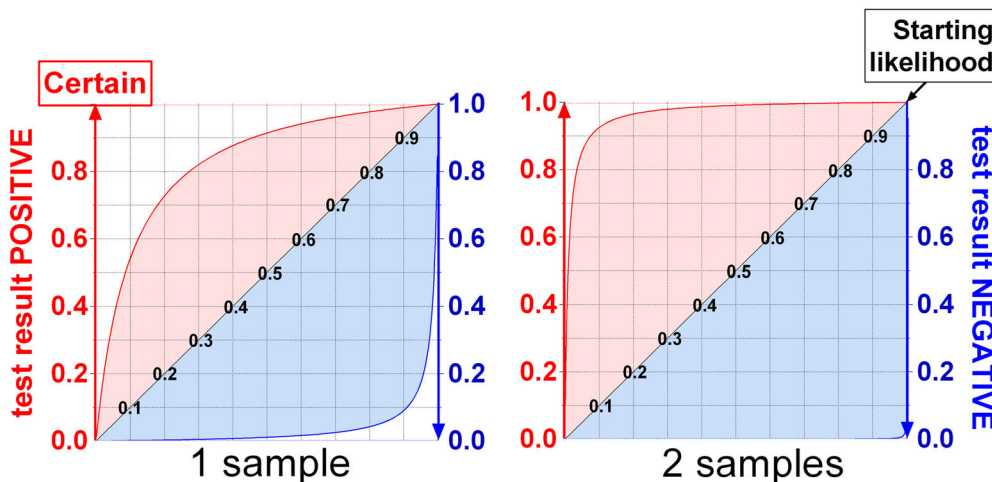
leaf plot [6] to compare the impacts of using different thresholds at all levels of prior diagnostic probabilities. Figure 3 shows that at a diagnostic threshold of 10^3 or 10^4 cfu/ml, a negative result completely excludes a UTI at all levels of clinical suspicion, but a positive one leaves a lot of doubt, especially when the child's clinical circumstances are not compelling and overdiagnosis will be common. Raising the threshold to 10^5 cfu/ml reduces the false-positive rate a little, but increases the risk of missing genuine UTIs among children whose clinical likelihood of having one is considered fairly high.

Raising the cut-off to 10^6 cfu/ml would dramatically increase the power to rule in UTIs at all levels of prior probability, though it would also further increase the risk of missing genuine cases compared with using a lower threshold. However, this option would require modification of standard laboratory methods, so the best available cut-off to deliver diagnostic balance is $\geq 10^5$ cfu/ml, which delivers a sensitivity of 0.99 and a specificity of 0.907. Routinely collecting two samples would improve these to 1.000 and 0.991, respectively, which produces a marked increase in diagnostic efficacy as shown by comparing their leaf plots (Fig. 4, top row), but would cause considerable extra inconvenience in clinical practice. This analysis indicates that lower targets substantially
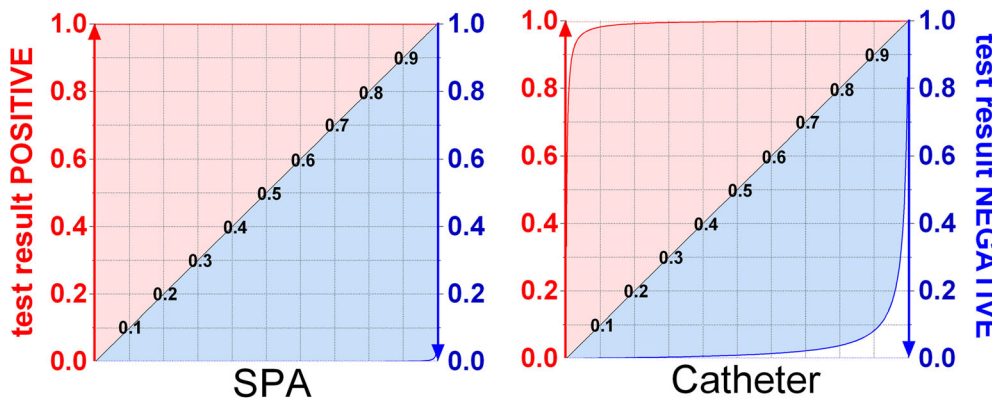

Fig. 3 Leaf plot showing the diagnostic values of using colony count thresholds between 10^3 and 10^6 cfu/ml on a child's single voided urine culture result (calculated from the sensitivity and sensitivity data shown). The vertical increase in height of the red line above the diagonal (vein of the "leaf") shows the impact that a positive culture result has on the probability of that child having a UTI. The vertical drop from the "vein" to the blue line indicates the effect of a negative urine culture on excluding a UTI
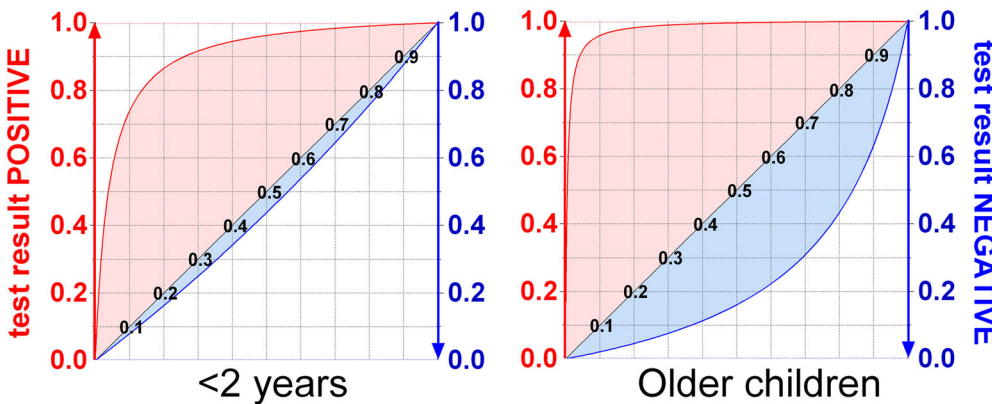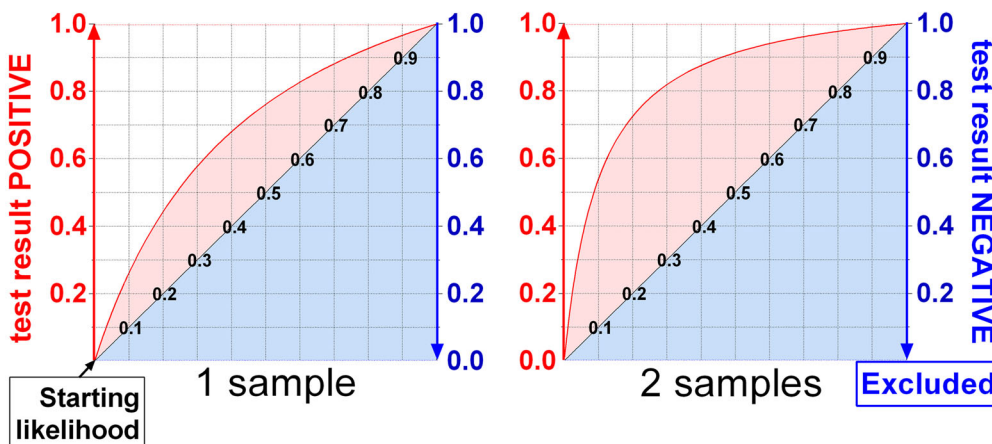
Data used to construct this leaf plot

Threshold (cfu/ml)	Sensitivity	Specificity
10^6	0.93	0.994
10^5	0.99	0.907
10^4	1.00	0.862
10^3	1.00	0.809

◄ Fig. 4 Leaf plots comparing the efficacy of various urine testing strategies at all levels of pre-test probability (see Fig. 3 for guidance to interpretation). Top row, clean-voided urine collections of one or two samples. Second row, SPA and catheter collection methods. Third row, comparing the usefulness of nitrite stick testing in infants < 2 years and older children. Bottom row, phase-contrast microscopy screening of one or two samples

worsen the false-positive diagnosis rates and rejects the notion that they would be preferable [4, 5, 35].

Meta-analysis 2: SPA sampling

Background to SPA usage

Low-level bacterial contamination was recognised to occur from the first use of SPA sampling in women [36] and children [16, 37], so colony count thresholds were originally set above 10^3 cfu/ml to avoid false-positive results. Despite this, Hellerstein cited these publications [38] as if they supported the notion that "any growth" of Gram-negative bacilli from an SPA in a child indicates a UTI, and his "table of recommendations" is the "evidence" that the AAP and NICE guidelines use to justify including this policy in their guidelines [3, 5].

High-quality quantitative data for SPA urines

Twelve studies [13–17 25–28, 32, 33, 37] included 709 children, of whom 114 had UTIs and 595 did not (Table 1), but half of these only reported how many cases exceeded various pre-determined colony counts, leaving six that provided sufficient quantitative data for full analysis [13, 25–28, 37]. Nearly half the subjects and two-thirds of those with UTIs were aged < 2 years. Two groups only enrolled infant boys [25, 33].

Children with UTIs

All 23 children with UTIs who had quantitative SPA data [13, 25–28, 37] had colony counts $\geq 10^5$ cfu/ml (Fig. 2), and three of the five children tested for higher bacterial concentrations had counts $\geq 10^6$ cfu/ml.

Children without UTIs

Most of the 594 children without UTIs had sterile urines, despite some using detection limits of just 10 [13, 25, 37] or 100 cfu/ml [16, 28, 33]. Unlike voided urines, there were no cases with high-concentration contamination, but 33 (5.5%) specimens had small numbers (< 2000/ml) of single uropathogens [13, 17, 25, 32, 37]. The SPA specificity is therefore 1.00 at a 10^5/ml threshold, but would fall and produce false positives by AAP and NICE criteria (Figs. 1 and 2).

How could SPA samples become contaminated?

The fact that blood cultures are commonly lightly contaminated [39] suggests that skin bacteria may be carried forward on the tip of a sampling needle, and the fact that the flora that contaminated the SPA samples were typically skin commensals [15, 25, 37, 40] and faecal organisms [15, 17, 40] which colonise the nappy area suggests that they may have become contaminated in the same way. Estimates of the bacterial numbers that could fit onto the leading edge of a needle tip are consistent with the reported urine colony counts.

Defining the best clinical diagnostic threshold for SPAs

SPA samples are 99% sensitive and 100% specific for diagnosing UTIs at the same 10^5 cfu/ml threshold that is the best for voided urines, so it is sensible to use the same values. Choosing to perform a single SPA would therefore avoid the approximately 10% false-positive risk seen with voided urines and almost guarantees to correctly diagnose a UTI, as shown in a leaf plot (Fig. 4).

It is unclear how the assertion that "any growth" in SPA urine ($\geq 10^3$ cfu/ml in standard laboratories) is diagnostic of a UTI was adopted without any scientific basis. This concept has been sustained by studies with biased designs and illogical analytic methods [13–17] and continues to enjoy wide support [3–5, 35, 38].

Meta-analysis 3: urethral catheter sampling

Background to urethral catheter sampling

Like SPAs, the AAP recommends urethral catheterisation to avoid bacterial contamination [3], but in this case they advise using a 10^5 cfu/ml threshold of a single uropathogen to diagnose UTIs [38]. Four other thresholds have been advocated on little evidence, namely 5×10^4 cfu/ml [19], 10^4 cfu/ml [19, 34, 41], 10^3 cfu/ml [18, 35, 42–45], and 10 cfu/ml [46]. In addition, the AAP now also endorses discarding the first drops of catheter urine [4] with little evidence.

High-quality quantitative data for catheter urines

Six eligible publications [21–24, 31, 37] containing seven studies enrolled 508 children (Table 1), and all but one [22] cultured the "mid-stream" urine flow after discarding the first few millilitres. I rejected 15 studies because they introduced a selection bias [41–44, 46–48], used discrepant culture thresholds for catheter and non-catheter samples not providing quantitative data [18, 34, 43, 46, 49–51], or provided incomplete data [52].

Children with UTIs Forty-nine of the 50 children with UTIs grew $\geq 10^5$ cfu/ml (Fig. 2); the one child whose first catheter sample grew 10^{4-5} cfu/ml and next voided sample grew $> 10^7$/ml has already been described. As with SPA and voided urines, most catheter samples tested for higher bacterial concentration ranges had counts $> 10^6$ cfu/ml.

Children without UTIs The catheter urines from the uninfected children grew fewer contaminants than the voided specimens (Fig. 1), but more than the SPA samples (Fig. 2). The specificity for 473 specimens from 458 children was 0.829 at a threshold of 10^2 cfu/ml, 0.960 at 10^3 cfu/ml, 0.983 at 10^4 cfu/ml, and 0.998 at 10^5 cfu/ml.

Should the initial catheter urine stream be discarded?

Two groups showed that the first drops of catheter urine were more likely to be contaminated than the subsequent stream [37, 53], but these cultures were clinically irrelevant skin commensals and very light growths of uropathogens (all $< 10^4$/ml). A more recent study that recommends discarding the first flow of catheter urine had multiple shortcomings including only reporting culture positivity rates at fixed threshold ranges and using biased analysis [19].

Defining the best clinical diagnostic threshold for catheter samples

For catheter samples (collected without discarding the first flow of urine), the threshold of 10^5 cfu/ml delivers 99% sensitivity and 99.8% specificity, so again it would be sensible to use the same diagnostic cut-off for all sampling methods. The leaf plot in Fig. 4 shows that a single catheter sample is marginally better at ruling UTIs in than collecting two voided samples and marginally worse than using an SPA, but no better at ruling them out than a single voided urine sample.

Table 1 Paediatric studies of paired-urine cultures that could be used for three meta-analyses of voided, suprapubic (SPA), and catheter urine collection methods, presented in date order

Reference and year		Reason for study	Urine collection methods	Age range	Children's UTI status			Meta-analysis		
					Total	UTI	No UTI	Voided	SPA	Catheter
[21]	1959	?UTI	MSU and catheter	3–12 years	58	3	55	Yes		Yes
[37]	1959	Study	Catheter and SPA	0.3–10 years	42	0	42		Yes	Yes
[22]	1960	Screening	Clean-catch and catheter	6–18 years	203	15	188	Yes		Yes
[23]	1961	Study	Clean-catch and catheter	< 2 years	50	0	50	Yes		Yes
		?UTI	Clean-catch and catheter	< 2 years	13	13	0	Yes		Yes
[24]	1969	Screening study	Clean-catch and catheter	1–31 days	44	0	44	Yes		Yes
			Repeated clean-catch		154	0	154			
[25]	1965	?UTI	Clean-catch and SPA	"Preterm"	25	2	23	Yes	Yes	
[13]	1965	?UTI	Clean-catch or bag and SPA	0.5–12 years	154	10	144	Yes	Yes	
[26	1970	?UTI	Clean-catch or bag and SPA	0–11 years	28	5	23	Yes	Yes	
[14]	1972	?UTI	Bag and SPA	< 2 years	63	12	51		Yes	
		?UTI	Clean-catch and SPA	2–8 years	45	11	34		Yes	
[15]	1972	?UTI	Clean-catch and SPA	4–6 days	14	9	5		Yes	
[16]	1973	?UTI	Bag and SPA	< 1.5 years	86	20	66		Yes	
			Clean-catch and SPA	3–12 years	34	10	24		Yes	
[27]	1973	?UTI	Clean-catch and SPA	1–11 years	6	2	4	Yes	Yes	
[28]	1976	?UTI	Clean-catch and bag and SPA	< 3 years	30	4	26		Yes	
[32]	1979	?UTI	Clean-catch or bag and SPA	"Children"	73	0	73		Yes	
[33]	1993	?UTI	Clean-catch and SPA	< 0.4 year	51	13	38		Yes	
[17]	1999	?UTI	Clean-catch and SPA	< 2 years	58	16	42		Yes	
[29]	2000	Study	Pad and clean-catch and bag	0–1.5 years	44	0	44	Yes		
[30]	2005	?UTI	Pad and clean-catch and bag	< 3 years	160	12	148	Yes		
[31]	2007	MCUG	Clean-catch and catheter	< 2 years	98	4	94	Yes		Yes
[20]	2010	?UTI	Repeat pad or clean-catch	< 18 years	203	36	167	Yes		
				Totals	1736	197	1539			

Conclusions about which diagnostic colony count to use

Sensitivity When I only analysed high-quality paired-sample publications without making prior assumptions, rather than relying on single-sample cultures in case series that also relied upon clinical opinion, a clear picture emerged that children with UTIs have similarly high concentrations of bacteria in their urine as Kass showed for adults over 60 years ago [11, 12]. Most had between 10^5 and 10^8 bacteria per ml, which translates to $\geq 10^5$ cfu/ml in standard laboratory reports. In over 200 reported high-quality childhood UTI cases, only two children had lower values (10^{4-5}), which subsequently rose to 10^7/ml, giving a sensitivity of 0.99 regardless of the urine sampling method used. This study shows that adopting a universal diagnostic threshold of 10^5 cfu/ml will correctly diagnose almost all genuine UTIs and that there is no evidence to support the need for using lower thresholds under any circumstances, such as for SPAs [3–5].

Specificity The false-positive rate of childhood urine culture is markedly dependent upon the collection technique used, with the specificity being lowest for single voided urine samples, making it more difficult to confidently rule out UTIs and thereby potentially leading to children being treated and investigated unnecessarily. Leaf plots demonstrate that trying to increase the specificity by lowering the diagnostic colony count below 10^5/ml will lead to children with genuine UTIs being missed. How else can the false-positive rate be reduced? We have shown that this can be achieved by using more invasive urine collection methods or by collecting paired samples; a third approach could be by integrating point-of-care diagnostic screening. Here, I consider these in turn.

Invasive urine collection Both SPA and catheter samples (without any need to discard the first drops) are relatively easy to collect in hospital settings and have excellent specificity when a diagnostic threshold of 10^5 cfu/ml is used, at 1.00 and 0.998, respectively. Hence, meaningful answers can be almost guaranteed from a single invasively collected specimen (Fig. 4). These methods allow samples to be obtained promptly so long as the child has some urine present in the bladder, which may be important in ill babies who need an urgent infection screen before antibiotic therapy is commenced. Urethral catheterisation can be challenging in very low–birthweight babies, but using firmer feeding tubes or umbilical arterial lines may be helpful [40, 47], and SPAs are often hard to obtain without ultrasound guidance, making them less convenient to perform. In very small babies, failed catheter sampling is sometimes followed by successful SPA collections, and vice-versa [40, 47]. However, invasive collection methods are impractical

outside of clinical settings, and many paediatricians prefer not to use them because they may be traumatic [54].

Collecting paired voided samples Although routinely collecting a second voided urine sample greatly improves diagnostic efficacy (as shown in Fig. 4), practical considerations mean it is unlikely to be adopted widely (though it is helpful that gentle suprapubic stimulation with a cold fluid–soaked gauze makes babies void more promptly [55]). However, we have found that advice to collect a second sample in selected cases after immediate point-of-care screening is consistently followed.

Incorporating point-of-care screening Unfortunately, nitrite sticks, which are the most convenient point-of-care urine test to detect UTI in older children, miss three-quarters of cases in children aged < 2 years, so they cannot be used to exclude the diagnosis in the most vulnerable children [10]. Figure 4 shows leaf plots for both age groups.

In Newcastle, specially trained paediatric nurses undertake immediate phase-contrast microscopy which reliably identifies bacteria in fresh unspun urine with 100% sensitivity but a lower specificity of 0.686 [56] (Fig. 4, bottom row). In non-emergency situations, we collect a single voided mid-stream or nappy pad urine sample according to age and use phase-contrast microscopy to rule out two-thirds of the uninfected children at once, which enables us to discard those samples (with cost-savings) and inform the families of this result immediately. We then collect second samples from those children where bacteria were present, some due to contamination and some due to UTIs (these typically have hundreds of identical organisms per high-power field, equivalent to about 10^{6-7}/ml), which most families accept as worthwhile because the first result is uncertain. This increases the specificity to 0.90. Children with bacteria in both specimens can then be presumed to have a UTI and started on antibiotics. If we remain suspicious of contamination in both samples because of scant organisms or an excess of epithelial cells, we may collect further samples after very careful washing.

Future research For future research in this area to add value, it is critical that studies must use rigorous methods to define UTIs, whether those involve invasive or paired sampling. They must also culture all their samples identically and publish raw quantitative data to allow independent analysis and not choose different diagnostic colony count thresholds for different collection methods. Even very large studies cannot contribute knowledge if they use single voided samples and a diagnostic threshold of 10^3 cfu/ml in some cases and count some mixed growths as positive [57, 58].

Summary By only reviewing publications with robust, paired, quantitative urine culture data, I have identified that some widely recommended

guidelines on diagnosing childhood UTIs are not secure [3–5]. Three key evidence-based points are as follows: (1) using a diagnostic culture threshold of 10^5 cfu/ml is 99% sensitive for UTIs however they are collected, (2) samples collected by SPA or urethral catheterisation (without discarding the first drops) are also highly specific, and (3) repeating voided urine collections improve specificity, especially if it is integrated with point-of-care phase-contrast microscopy. Introducing lower diagnostic thresholds for selected situations is both illogical and unhelpful.

Compliance with ethical standards

Conflict of interest The author declares that he has no conflicts of interest.

References

1. Coulthard MG, Lambert HJ, Vernon SJ, Hunter EW, Keir M, Matthews JNS (2014) Does prompt treatment of urinary tract infection in preschool children prevent renal scarring: mixed retrospective and prospective audits. Arch Dis Child 99:342–347
2. Karavanaki KA, Soldatou A, Koufadaki AM, Tsentidis C, Haliotis FA, Stefanadis CJ (2017) Delayed treatment of the first febrile urinary tract infection in early childhood increased the risk of renal scarring. Acta Paediatr 106:149–154
3. American Academy of Pediatrics (1999) Practice parameter: the diagnosis, treatment and evaluation of the initial urinary tract infection in febrile infants and young children. Pediatr 103:834–852
4. American Academy of Pediatrics (2011) Urinary tract infection: clinical practice guideline for the diagnosis and management of the initial UTI in febrile infants and children 2 to 24 months. Pediatr 128:595–610
5. National Institute for Health and Clinical Excellence (NICE) (2007) Urinary tract infection in children. (http://guidance.nice.org.uk/CG054). Accessed 14/16/2019
6. Coulthard MG, Coulthard T (2019) The leaf plot: a novel way of presenting the value of tests. Br J Gen Pract 69:205–206
7. Pryles CV, Eliot CR (1965) Pyuria and bacteriuria in infants and children. Am J Dis Child 110:628–635
8. Stansfeld JM (1962) The measurement and meaning of pyuria. Arch Dis Child 37:257–262
9. Turner GM, Coulthard MG (1995) Fever can cause pyuria in children. BMJ 311:924
10. Coulthard MG (2019) Using urine nitrite sticks to test for urinary tract infection in children aged <2 years: a meta-analysis. Pediatr Nephrol 34:1283–1288
11. Kass EH (1956) Asymptomatic infections of the urinary tract. Trans Assoc Am Phys 69:56–65
12. Kass EH (1957) Bacteriuria and the diagnosis of infections of the urinary tract. JAMA Intern Med 100:709–714
13. Saccharow L, Pryles CV (1969) Further experience with the use of percutaneous suprapubic aspiration of the urinary bladder. Pediatr 43:1018–1024
14. Bergwall B, Ljungblom B-A (1972) Suprapubic aspiration of the urinary bladder for accurate diagnosis of urinary tract infections in children. Acta Paediatr Scand 61:390
15. Abbott GD (1972) Neonatal bacteriuria: a prospective study in 1, 460 infants. BMJ 1(5795):267–369
16. Aronson AS, Gustafson B, Svenningsen NW (1973) Combined suprapubic aspiration and clean-voided urine examination in infants and children. Acta Paediatr Scand 62:396–400
17. Ramage IJ, Chapman JP, Hollman AS, Elabassi M, McColl JH, Beattie TJ (1999) Accuracy of clean-catch urine collection in infancy. J Pediatr 135:765–767
18. Verliat-Guignaud J, Blanc P, Garnier F, Gajdos V, Guigonis V (2015) A midstream urine collector is not a good alternative to a sterile collection method during the diagnosis of urinary tract infection. Acta Paediatr 104:e395–e400
19. Megged O, Zilberstein L, Shalom EB, Erlichman M (2017) Contamination of urinary cultures in initial-stream versus later-stream urine in children undergoing bladder catheterisation for the diagnosis of urinary tract infection. Eur J Emerg Med 24:e17–e20
20. Coulthard MG, Kalra M, Lambert HJ, Nelson A, Smith T, Perry JD (2010) Redefining urinary tract infections by bacterial colony counts. Pediatr 125:335–341
21. Pryles CV, Steg LN (1959) Specimens of urine obtained from young girls by catheter vs voiding. Pediatr 23:441–452
22. Kunin CM, Southall I, Paquin AJ (1960) Epidemiology of urinary-tract infections: a pilot study of 3057 school children. NEJM 263:817–823
23. Pryles CV, Luders D, Alkan MK (1961) Colony counts in paired specimens of urine obtained by catheter versus voiding from normal infants and infants with urinary tract infection. Pediatr 27:17–28
24. McCarthy JM, Pryles CV (1963) Clean voided and catheter neonatal urine specimens. Am J Dis Child 106:473–478
25. Nelson JD, Peters PC (1965) Suprapubic aspiration of urine in premature and term infants. Pediatr 36:132–134
26. Conn NK (1970) A study of some of the methods of urinary collection in children. J Clin Pathol 23:81–84
27. Sharma SK, Singh SM, Bapna BC (1973) Comparison of the bacterial flora and colony counts of the paired urine samples obtained by suprapubic aspiration and clean voided mid-stream at the All-India Institute of Medical Sciences Hospital, New Delhi. Indian J Med Res 61:179–183
28. Hardy JD, Furnell PM, Brumfitt W (1976) Comparison of sterile bag, clean catch and suprapubic aspiration in the diagnosis of urinary infection in early childhood. Br J Urol 48:279–283
29. Liaw LCT, Nayar DM, Pedler SJ, Coulthard MG (2000) Home collection of urine for culture from infants by three methods: survey of parents' preferences and bacterial contamination rates. BMJ 320:1312–1313
30. Alam MT, Coulter JBS, Pacheo J, Correia JB, Ribeiro MGB, Coelho MFC, Bunn JEG (2005) Comparison of urine contamination rates using three different methods of collection: clean-catch, cotton wool pad and urine pad. Ann Trop Paediatr 25:29–34
31. Lau AY, Wong S-N, Yip K-T, Fong K-W, Li SP-S, Que T-L (2007) A comparative study on bacterial cultures of urine samples obtained by clean-void technique versus urethral catheterization. Acta Paediatr 96:432–436
32. Pylkkanen J, Vilska J, Koskimies O (1979) Diagnostic value of symptoms and clean-voided urine specimen in childhood urinary tract infection. Acta Paediatr Scand 68:341–344
33. Amir J, Ginzburg M, Staussberg R, Varsano I (1993) The reliability of midstream urine culture from circumcised male infants. Am J Dis Child 147:969–970
34. Herreros ML, Tagarro A, Garcia-Pose A, Sanchez A, Canete A, Gili P (2018) Performing a urine dipstick test with a clean-catch urine sample is an accurate screening method for urinary tract infections in young infants. Acta Paediatr 107:145–150
35. Tullus K (2016) Low urinary bacterial counts: do they count? Pediatr Nephrol 31:171–174
36. Stamey TA (1972) Urinary tract infections. Williams & Williams, Baltimore, pp 1–29
37. Pryles CV, Atkin MD, Morse TS, Welch KJ (1959) Comparative bacteriologic study of urine obtained from children by percutaneous suprapubic aspiration of the bladder and by catheter. Pediatr 24:983–991
38. Hellerstein S (1982) Recurrent urinary tract infections in children. Pediatr Infect Dis J 1:271–281
39. Dawson S (2014) Blood culture contaminants. J Hosp Infect 87:1–10

40. Tobiansky R, Evans N (1998) A randomized controlled trial of two methods for collection of sterile urine in neonates. J Paediatr Child Health 34:460–462

41. Karacan C, Erkek N, Senel S, Gunduz SA, Catli G (2010) Evaluation of urine collection methods for the diagnosis of urinary tract infection in children. Med Princ Pract 19:188–191

42. Etoubleau C, Reveret M, Brouet D, Badier I, Brosset P, Fourcarde L, Bahans C, Garnier F, Blanc P, Guigonis V (2009) 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 154:803–806

43. McGillivray D, Mok E, Mulrooney E, Kramer MS (2005) A head-to-head comparison: "clean-void" bag versus catheter urinalysis in the diagnosis of urinary tract infection in young children. J Pediatr 147:451–456

44. Al-Orifi F, McGillivray D, Tange S, Kramer MS (2000) Urine culture from bag specimens in young children: are the risks too high? J Pediatr 137:221–226

45. Swerkersson S, Jodal U, Ahren C, Sixt R, Stoklanad E, Hansson S (2016) Urinary tract infection in infants: the significance of low bacterial count. Pediatr Nephrol 31:239–245

46. Tosif S, Baker A, Oakley E, Donath S, Babl FE (2012) 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 48:659–664

47. Austin BJ, Bollard C, Gunn TR (1999) Is urethral catheterisation a successful alternative to suprapubic aspiration in neonates? J Paediatr Child Health 35:34–36

48. Pollack CV, Pollack ES, Andrew ME (1994) Suprapubic bladder aspiration versus urethral catheterisation in ill infants: success, efficiency, and complication rates. Ann Emerg Med 23:225–230

49. Li PS, Ma LC, Wong SN (2002) Is bag urine culture useful in monitoring urinary tract infection in infants? J Paediatr Child Health 38:377–381

50. Armengol CE, Hendley JO, Schlager TA (2001) Should we abandon standard microscopy when screening for urinary tract infections in young children. Pediatr Infect Dis J 20:1176–1177

51. Shaw KN (1998) Screening for urinary tract infection in infants in the emergency department: which test is best? Pediatr 101:e1065

52. Gilljam B-M, Svensson M-L (2008) In-out catheterization of young children with suspected urinary tract infection: a retrospective journal study. Pediatr Nurs 34:241–245

53. Dayan PS, Chamberlain JM, Boenning D, Adirim T, Schor J, Klein BL (2000) A comparison of the initial to the later stream urine in children catheterized to evaluate for a urinary tract infection. Pediatr Emerg Care 16:88–90

54. Coutinho K, Stensland K, Akhavan A, Jayadevan R, Stock JA (2014) Pediatrician noncompliance with the American Academy of Pediatrics guidelines for the workup of UTI in infants. Clin Pediatr 53:1139–1148

55. Kaufman J, Fitzpatrick P, Tosif S, Hopper SM, Donath SM, Bryant PA, Babl FE (2017) Faster clean catch urine collection (Quick-Wee method) from infants; randomised controlled trial. BMJ 357:j1341

56. Coulthard MG, Nelson A, Smith T, Perry JD (2010) Point-of-care diagnostic tests for childhood urinary tract infection: phase-contrast microscopy for bacteria, stick-testing, and counting white blood cells. J Clin Pathol 63:823–829

57. Downing H, Thomas-Jones E, Gal M, Waldron C-A, Sterne J, Hollingworth W, Hood K, Delaney B, Little P, Howe R, Wooton M, Macgowan A, Butler CC, Hay AD (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. BMC Infect Dis 12: 158

58. Birnie K, Hay AD, Wootton M, Howe R, MacGowan A, Whiting P, Lawton M, Delaney B, Downing H, Dudley J, Hollingworth W, Lisles C, Little P, O'Brien K, Pickles T, Rumsby K, Thomas-Jones E, Van der Voort J, Waldron C-A, Harman K, Hood K, Butler CC, Sterne JAC (2017) Comparison of microbiological diagnosis of urinary tract infection in young children by routine health service laboratories and a research laboratory: diagnostic cohort study. PLoS One 12. https://doi.org/10.1371/journal.pone.017113