

Improving the Diagnosis and Treatment of Urinary Tract Infection in Young Children in Primary Care: Results from the DUTY Prospective Diagnostic Cohort Study

Alastair D. Hay, FRCGP¹
Jonathan A. C. Sterne, PhD²
Kerenza Hood, PhD³
Paul Little, FMedSci⁴
Brendan Delaney, MD⁵
William Hollingworth, PhD²
Mandy Wootton, PhD⁶
Robin Howe, FRCPath⁶
Alasdair MacGowan, MD⁷
Michael Lawton, MPhil²
John Busby, MSc²
Timothy Pickles, BSc³
Kate Birnie, PhD²
Kathryn O'Brien, PhD⁸
Cherry-Ann Waldron, PhD³
Jan Dudley, PhD⁹
Judith Van Der Voort, FRCPCH¹⁰
Harriet Downing, MPhil¹
Emma Thomas-Jones PhD³
Kim Harman, DHealth⁴
Catherine Lises, MSc³
Kate Rumsby, MSc⁴
Stevó Durbaba, MSc¹¹
Penny Whiting, PhD¹²
Christopher C. Butler, FRCGP¹³



Conflicts of interest: authors report none.

CORRESPONDING AUTHOR

Alastair D. Hay, FRCGP
Centre for Academic Primary Care
School of Social and Community Medicine
Canyng Hall
University of Bristol
39 Whatley Rd
Clifton Bristol, BS8 2PS, UK
alastair.hay@bristol.ac.uk

ABSTRACT

PURPOSE Up to 50% of urinary tract infections (UTIs) in young children are missed in primary care. Urine culture is essential for diagnosis, but urine collection is often difficult. Our aim was to derive and internally validate a 2-step clinical rule using (1) symptoms and signs to select children for urine collection; and (2) symptoms, signs, and dipstick testing to guide antibiotic treatment.

METHODS We recruited acutely unwell children aged under 5 years from 233 primary care sites across England and Wales. Index tests were parent-reported symptoms, clinician-reported signs, urine dipstick results, and clinician opinion of UTI likelihood (clinical diagnosis before dipstick and culture). The reference standard was microbiologically confirmed UTI cultured from a clean-catch urine sample. We calculated sensitivity, specificity, and area under the receiver operator characteristic (AUROC) curve of coefficient-based (graded severity) and points-based (dichotomized) symptom/sign logistic regression models, and we then internally validated the AUROC using bootstrapping.

RESULTS Three thousand thirty-six children provided urine samples, and culture results were available for 2,740 (90%). Of these results, 60 (2.2%) were positive: the clinical diagnosis was 46.6% sensitive, with an AUROC of 0.77. Previous UTI, increasing pain/crying on passing urine, increasingly smelly urine, absence of severe cough, increasing clinician impression of severe illness, abdominal tenderness on examination, and normal findings on ear examination were associated with UTI. The validated coefficient- and points-based model AUROCs were 0.87 and 0.86, respectively, increasing to 0.90 and 0.90, respectively, by adding dipstick nitrites, leukocytes, and blood.

CONCLUSIONS A clinical rule based on symptoms and signs is superior to clinician diagnosis and performs well for identifying young children for noninvasive urine sampling. Dipstick results add further diagnostic value for empiric antibiotic treatment.

Ann Fam Med 2016;14:325-336. doi: 10.1370/afm.1954.

INTRODUCTION

The accurate and timely diagnosis of urinary tract infection (UTI) in children is important to alleviate short-term suffering¹ and prevent the possible long-term consequences, such as renal scarring, impaired renal growth, recurrent pyelonephritis, impaired glomerular function, hypertension, end-stage renal disease, and preeclampsia.²⁻⁴ Guidelines universally recommend microbiological confirmation using urine samples collected by clean catch (preferred in Europe),⁵ catheterization, or suprapubic aspiration for unwell children where clean catch is not immediately available (preferred in the United States⁶ and Australia⁷).

There are 3 possible explanations why one-half of UTIs are not diagnosed at the earliest opportunity in UK primary care.⁸ First, relevant evidence in primary care regarding which children should be suspected of having a UTI is scarce. Guidelines, which emphasize the importance

of fever,^{6,7,9} are largely informed by studies conducted in emergency departments.¹⁰⁻¹² Second, the symptoms and signs of UTI are often nonspecific, especially in very young children. Finally, obtaining an uncontaminated sample can be challenging, time-consuming, and, for invasive catheter- and suprapubic aspiration-sampling methods, painful¹³ and frightening,¹⁴ and they may induce infection.¹⁵

We report a large, prospective cohort study designed to investigate the diagnostic features of UTI in young children in primary care. Our aim was to develop and internally validate a 2-step clinical rule: step 1 used symptoms and signs to select children for urine sampling, and step 2 (once urine was obtained) used symptoms, signs, and dipstick testing to guide empiric antibiotic treatment. Coefficient- and points-based clinical rules were developed for use with and without computer assistance.

METHODS

Design

Diagnosis of Urinary Tract Infection in Young children (DUTY) was a multicenter, prospective, diagnostic cohort study that recruited children seeking care at National Health Service (NHS) primary care sites. General practitioners, nurses, and children's emergency department physicians (from here on clinicians) working in primary care sites (general practice clinics, emergency departments, and walk-in centers) are those who provide primary care for children. Primary care site clinicians were recruited, and the staff were trained by 4 United Kingdom center hubs (Bristol, Cardiff, London, and Southampton).

Ethical approval was granted by the South West Southmead Research Ethics Committee, ref #09/H0102/64.

Participants

Children were eligible if aged under 5 years and with complaints of any acute (less than 28 days) illness episode, where the illness was associated with (1) at least 1 constitutional symptom or sign identified by the National Institute of Health and Clinical Excellence (NICE)⁵ as a potential marker for UTI (fever, vomiting, lethargy/malaise, irritability, poor feeding, and failure to thrive), and/or (2) at least 1 urinary tract symptom identified by NICE⁵ as a potential marker of UTI (abdominal pain, jaundice in children aged younger than 3 months, hematuria, offensive urine odor, cloudy urine, loin pain, frequency, apparent pain on passing urine, and changes to continence). As a result, constitutionally unwell children consulting with an apparently obvious cause for their symptoms (such as acute otitis

media or bronchiolitis, without a urinary symptom) were included. Children were excluded if they were constitutionally well (eg, acute conjunctivitis only), had a neurogenic or surgically reconstructed bladder, had a permanent or intermittent urinary catheter, reported trauma as the main problem, or had been taking antibiotics within 7 days. Clinicians, aided by research nurses where available, were asked to recruit consecutive eligible children, and where it was not possible, to log children's age and sex (Supplemental Appendix, Figure 1, <http://www.annfammed.org/content/14/4/325/suppl/DC1>).

Index Tests and Urine Collection

Subsequent to consent, results of 107 index tests (symptoms, signs, and dipstick results, Supplemental Appendix, <http://www.annfammed.org/content/14/4/325/suppl/DC1>) were recorded on a standardized case report form by qualified clinicians blinded to the reference standard. Parent-reported items included the child's medical history and symptoms. Signs gathered from a full clinical examination included clinicians' global illness severity impression and abdominal tenderness. Before urine dipstick testing, clinicians recorded their rating of UTI likelihood (clinical diagnosis).

Our preferred urine collection method was clean catch. For toilet-trained children, we used a sterile bowl that the parent could hold under the child or put in a potty. For other children, the parent cleaned the diaper area using water alone and sat the child on their knee with the bowl placed under their perineum. If it was not possible to obtain a sample at the site, the parent was given equipment and advice on taking the sample at home. Where a clean catch was not feasible, we used NICE-recommended sterile urine collection pads (Newcastle SteriSet sterile urine collection packs) placed inside the diaper,⁵ but because of differences in contamination rates and children's ages between clean-catch and diaper pad samples, the current analysis focuses on clean-catch samples.

Urine samples were tested at the site using Siemens/Bayer multistix 8SG dipsticks (Siemens Healthineers). Urine samples were split into 2 fractions for microbiological analysis. The priority fraction was sent to the site's usual laboratory. When at least 1 mL was left, the remainder was sent in boric acid monovettes using first-class postal Safeboxes (Royal Mail Group, Ltd) to the Public Health Wales Microbiology Specialist Antimicrobial Chemotherapy Unit in Cardiff, UK (the research laboratory).

Reference Standard

The reference standard was determined at the research laboratory, which spiral plated (Don Whitley Sci-

entific) 50 μL of urine onto chromogenic agar and standard blood agar. Quantitative total counts were recorded for up to 6 organisms, and the presence of antimicrobial substances was measured. Two staff members processed the samples using a single, standardized procedure. According to UK guidelines,¹⁶ our microbiological definition of UTI was either the pure (single) or predominant growth of a uropathogen (*Enterobacteriaceae*) at $\geq 10^5$ colony-forming units (CFU)/mL. We defined predominant growth as $\geq 10^5$ CFU/mL of a uropathogen with a 3- \log_{10} (1,000-fold) or greater difference between the growth of this and the next species. For comparison, we used the US definition⁶ of a pure uropathogen growth: $>50,000$ CFU/mL with a leukocyte count of $\geq 25/\text{mm}^3$ on microscopy or leukocyte positive (threshold at nil/trace) on dipstick.

Sample Size Calculation

We assumed a candidate predictor prevalence of 10% and UTI prevalence of 2%.¹⁷ With 80% power and a 2-sided α of .05, 3,000 urine results were required to detect an odds ratio of 2.4, whereas 3,100 results would give a clinical rule with 80% sensitivity and a 95% confidence interval width of 10%. We originally proposed to recruit 4,000 children with a target of recovering urine samples from at least 3,100 (77.5%) for clinical rule derivation and a further 2,000 children for external validation. We did not anticipate, however, the need to stratify analyses by clean-catch and diaper pad collection methods. We therefore decided to use all available clean-catch results to derive the models, with internal bootstrap validation instead of external validation.

Summary of Statistical Analysis

A full description of statistical methods is given in the Supplemental Appendix (<http://www.annfam.org/content/14/4/325/suppl/DC1>). In summary, we first compared the age and sex of the children who were recruited with those children whose parents declined to participate. Then, using logistic regression, we estimated associations of index tests with positive urine cultures in 2 steps (reflecting clinical procedure): in step 1 we used symptoms and signs to select children for urine sampling, and in step 2 (once urine was obtained) we used symptoms, signs, and dipstick testing to guide empiric antibiotic treatment. Coefficient- and points-based clinical rules were developed for use with and without computer assistance. We quantified diagnostic accuracy using the area under the receiver operating characteristic (AUROC) curve with 95% confidence intervals and internally validated coefficient-based models using the bootstrap procedure described by Steyerberg et al.¹⁸ Because children

with fever of unknown origin is a group of particular clinical interest, we investigated the presence of UTI among children identified as having fever without symptoms or signs suggestive of another source. To assess the added value of dipsticks in addition to symptoms and signs alone, we (1) quantified the change in AUROC, and (2) used a simulation procedure to calculate the change in the probability of UTI associated with addition of dipstick results. Finally, because UTI laboratory criteria differ between the United Kingdom and the United States, we calculated the prevalence- and bias-adjusted k statistic to assess agreement between UK and US UTI definitions.¹⁹ We also used crude and adjusted odds ratios and the AUROC to assess strength of association and diagnostic utility of index tests identified as diagnostic using the UK UTI definition.

RESULTS

Between April 2010 and April 2012, 516 staff (61 research nurses, 182 general practitioners, and 273 site nurses) recruited participants from 233 primary care sites (225 general practices, 4 walk-in centers, and 4 children's emergency departments) across England and Wales. Of 14,724 children screened for eligibility, 4,390 (43%) were ineligible, 1,276 (12.6%) declined participation, 1,880 (18.5%) could not be recruited for other reasons, and 15 (0.15%) withdrew. We had valid consent and index test data for the remaining 7,163 (Supplemental Appendix Figure 2, <http://www.annfam.org/content/14/4/325/suppl/DC1>), of whom 6,797 (94.9%) were recruited in general practice sites, 284 (4.0%) in children's emergency departments, and 82 (1.1%) in walk-in centers. Urine was collected using diaper pads from 3,205 (reported in a separate article)²⁰ and by clean catch from 3,036 children, with reference standard (research laboratory) results available for 2,740 (90%) of clean-catch samples. Of these 2,740 children, 2,561 (93.5%) were aged 2 years or older, and 1,473 (54%) were female (Table 1). The most common working diagnoses were upper respiratory tract infection (28%), viral illness (15%), otitis media (10%), and gastroenteritis (3.6%). One-third of samples arrived at the research laboratory within 24 hours of collection. Transit time did not affect culture performance.

Sixty (2.2%) children had urine samples that met the laboratory definition for UTI: 50 (83.3%) *Escherichia coli*, 5 (8.3%) *Proteus* species, 3 (5.0%) *Klebsiella* species, 1 (1.7%) *Morganella morganii*, and 1 (1.7%) *Citrobacter farmeri*. Of these samples, 2,627 (96%) were provided within 24 hours of the index test measurement. Urinary antimicrobial substances were found in

128 (4.5%) samples and in 4 (6.7%) of the UTI-positive samples. A clinical diagnosis of UTI before urine dipstick testing was made in 168 (6.1%) children, of whom 28 (16.7%) were UTI positive. Clinical diagnosis achieved 46.6% sensitivity, 94.7% specificity, and an AUROC of 0.77 (95% CI, 0.71-0.83). Missing data and not known responses were infrequent (Table 1).

Step 1. Symptoms and Signs

The parent-reported index tests associated with UTI in crude (Table 1) and adjusted (Table 2) analyses were pain/crying while passing urine, smelly urine, previous UTI, and absence of severe cough. For the pain/crying while passing urine and smelly urine, there was a graded association with increasing symptom severity. Clinician-reported index tests associated with UTI were increasing illness severity (graded association), abdominal tenderness, and absence of ear abnormalities. None of the other index tests (Table 1 of the Supplemental Appendix, <http://www.annfammed.org/content/14/4/325/suppl/DC1>) met our criteria for model inclusion, and there was no evidence of association for fever of unknown origin (Table 3 of the Supplemental Appendix, <http://www.annfammed.org/content/14/4/325/suppl/DC1>) or prior illness duration (data not shown).

The multiple imputation-based AUROC for the coefficient-based step 1 model was 0.90 (95% CI, 0.85-0.95, internally validated AUROC of 0.88, Table 2, Figure 1). Table 3 (upper portion) shows diagnostic test characteristics (sensitivity, specificity, positive and negative predictive values, percentage of children defined as positive) according to a range of cut points for the symptoms and signs model (for any combination of symptoms and signs, the linear predictor is obtained by adding the coefficients corresponding to those categories to the constant term in Table 6 of the Supplemental Appendix, <http://www.annfammed.org/>

Table 1. Children's Characteristics and Crude Odds Ratio for Associations With Urinary Tract Infection (N = 2,740 Children)

Demographics and Index Test	Total No. (%) ^a	UTI Positive No. (%) ^b	Crude OR (95% CI)
Sex			
Male	1,267 (46.2)	13 (1.0)	1 [reference]
Female	1,473 (53.8)	47 (3.2)	3.18 (1.71,5.90)
Age of child			
<6 mo	34 (1.2)	1 (2.9)	1.13 (0.15,8.77)
6 to <12 mo	54 (2.0)	3 (5.6)	2.19 (0.62,7.77)
1 to <2 y	91 (3.3)	2 (2.2)	0.84 (0.19,3.70)
2 to <3 y	612 (22.3)	16 (2.6)	1 [reference]
3 to <4 y	1,073 (39.2)	21 (2.0)	0.74 (0.39,1.44)
≥4 y	876 (32.0)	17 (1.9)	0.74 (0.37,1.47)
Clinician diagnosis before dipstick			
Not UTI certain / very certain	1,149 (41.9)	6 (0.5)	0.28 (0.12,0.69)
Not UTI fairly certain / uncertain	1,417 (51.7)	26 (1.8)	1 [reference]
UTI fairly to very certain	168 (6.1)	28 (16.7)	10.75 (6.13,18.8)
Missing	6 (0.2)	0 (0.0)	
Pain/crying when passing urine ^c			
No problem	2,234 (81.5)	22 (1.0)	1 [reference]
Slight problem	182 (6.6)	6 (3.3)	2.97 (1.21,7.29)
Moderate problem	128 (4.7)	12 (9.4)	9.01 (4.45,18.2)
Severe problem	51 (1.9)	15 (29.4)	36.30 (17.81,74.0)
Missing/not known	145 (5.3)	5 (3.4)	
Smelly urine ^c			
No problem	2,108 (76.9)	20 (0.9)	1 [reference]
Slight problem	174 (6.4)	10 (5.7)	5.87 (2.76,12.5)
Moderate problem	179 (6.5)	16 (8.9)	9.46 (4.93,18.2)
Severe problem	51 (1.9)	10 (19.6)	23.5 (10.6,52.3)
Missing/not known	228 (8.3)	4 (1.8)	
Cough ^c			
No problem	773 (28.2)	24 (3.1)	1 [reference]
Slight problem	556 (20.3)	16 (2.9)	0.93 (0.48,1.76)
Moderate problem	829 (30.3)	17 (2.1)	0.66 (0.35,1.23)
Severe problem	579 (21.1)	3 (0.5)	0.16 (0.05,0.54)
Missing/not known	3 (0.1)	0 (0.0)	
Previous UTI ^c			
No	2,449 (89.4)	43 (1.8)	1 [reference]
Yes	177 (6.5)	12 (6.8)	3.81 (1.99,7.31)
Missing/not known	114 (4.2)	5 (4.4)	
Clinician global impression of illness severity grade ^d			
0-1	989 (36.1)	14 (1.4)	1 [reference]
2	739 (27.0)	14 (1.9)	1.35 (0.64,2.85)
3	531 (19.4)	14 (2.6)	1.89 (0.89,4.00)
4-5	363 (13.2)	12 (3.3)	2.39 (1.09,5.21)
6 or more	115 (4.2)	6 (5.2)	3.85 (1.45,10.21)
Missing	3 (0.1)	0 (0.0)	

continued

[content/14/4/325/suppl/DC1](http://www.annfammed.org/content/14/4/325/suppl/DC1)). To achieve sensitivities of 70% (all children with linear predictor of -2.729 or greater) or 85% (linear predictor -3.717 or greater) with the step 1 model would require urine sampling in 6.8% to 17.6% of children; with corresponding posi-

Table 1. Children's Characteristics and Crude Odds Ratio for Associations With Urinary Tract Infection (N = 2,740 Children)

(continued)

Demographics and Index Test	Total No. (%) ^a	UTI Positive No. (%) ^b	Crude OR (95% CI)
Abdominal examination: any tenderness ^c			
No	2,237 (81.6)	46 (2.1)	1 [reference]
Yes	63 (2.3)	8 (12.7)	7.34 (3.33,16.19)
Missing	440 (16.1)	6 (1.4)	
Ear examination: any acute abnormality ^c			
No	1,783 (65.1)	50 (2.8)	1 [reference]
Yes	635 (23.2)	4 (0.6)	0.23 (0.08,0.64)
Missing	322 (11.8)	6 (1.9)	
Dipstick: leukocytes ^c			
Negative	2,272 (82.9)	17 (0.7)	1 [reference]
Trace	154 (5.6)	6 (3.9)	5.40 (2.10,13.9)
1+	110 (4.0)	2 (1.8)	2.47 (0.56,10.8)
2+	148 (5.4)	19 (12.8)	19.61 (9.95,38.6)
3+	48 (1.8)	16 (33.3)	66.6 (30.9,143.3)
Missing	8 (0.3)	0 (0.0)	
Dipstick: nitrites ^c			
Negative	2,658 (97.0)	35 (1.3)	1 [reference]
Positive	74 (2.7)	25 (33.8)	38.4 (21.4,68.9)
Missing	8 (0.3)	0 (0.0)	
Dipstick: blood ^c			
Negative	2,297 (83.8)	29 (1.3)	1 [reference]
Non-heme	246 (9.0)	8 (3.3)	2.64 (1.19,5.84)
Heme trace	50 (1.8)	6 (12.0)	10.70 (4.23,27.08)
Heme 1+	67 (2.4)	4 (6.0)	4.98 (1.70,14.60)
Heme 2+ or 3+	72 (2.6)	13 (18.1)	17.29 (8.56,34.94)
Missing	8 (0.3)	0 (0.0)	

OR = odds ratio; UTI = urinary tract infection.

^a Total children column gives the percentage of observations within that category.^b Children with UTI column gives the percentage of positives relative to the number of observations within that category.^c Index tests independently associated with UTI in multivariable models. Missing values were assigned to the modal category for crude odds ratio.^d Graded on a scale from 1 to 10, where higher scores indicate greater severity.

tive predictive values of 22.6% to 10.6% and specificities of 94.6% to 83.9%. Although the points-based model AUROC of 0.86 (95% CI, 0.81-0.90; validated AUROC = 0.85) (Table 2 of the Supplemental Appendix, <http://www.annfammed.org/content/14/4/325/suppl/DC1>) was similar to the coefficient-based model, other diagnostic parameters were inferior: using an 85.0% sensitivity (a 3-point or greater cutoff) only increased the posttest probability to 6.9% with a lower specificity (74.4%) and higher (26.9%) urine collection rate (upper portion of Table 4, Supplemental Appendix, <http://www.annfammed.org/content/14/4/325/suppl/DC1>). Using a 5-point or greater cutoff (any 3 of 5 symptoms and signs) increased the posttest UTI probability to 17.7%, with increased specificity

(94.6%) and reduced urine collection rate (6.4%), but at the expense of reduced sensitivity (51.7%, upper portion of Table 4, Supplemental Appendix, <http://www.annfammed.org/content/14/4/325/suppl/DC1>).

Urine samples were available for 88, 91, and 612 children aged less than 12 months, 12 to 23 months, and 24 to 35 months, respectively, with laboratory confirmed UTI in 4, 2, and 16 of these children, respectively (Table 1). The coefficient model performed well in children younger than 3 years, with similar estimated odds and AUROC (Table 5 of the Supplemental Appendix, <http://www.annfammed.org/content/14/4/325/suppl/DC1>).

Step 2. Symptoms, Signs, and Dipstick Testing

Dipstick leukocytes, nitrites, and blood were strongly associated with UTI (Tables 1 and 2). The coefficient-based, multiple imputation model AUROC was 0.93 (95% CI, 0.90-0.97, validated 0.90), an increase of 0.03 ($P = 0.01$) when dipstick results were added to symptoms and signs (Table 2, Figure 1). If all children had a urine sample and dipstick test, the dipstick test results could maintain sensitivity at 80% while improving specificity from 88.3% to 93.8% and reducing the percentage of children treated with antibiotics from 13.2% to 7.8%, assuming immediate antibiotic use

(Table 3). The points-based model AUROC was 0.90 (95% CI, 0.85-0.95, validated 0.89), and increased (by 0.04, $P < 0.01$) when dipstick results were added to symptoms and signs (Table 2). Table 7 of the Supplemental Appendix (<http://www.annfammed.org/content/14/4/325/suppl/DC1>) shows there was a substantial impact on the probability of UTI associated with addition of dipstick test results to the step 1, with a median change in post-dipstick test probability of UTI of 9.9% (95% CI, 1.4%-55.5%).

Serious Adverse Events

Seventy-nine (1.1%) of the 7,163 recruited children were hospitalized; 3 related to dipstick testing (2 with UTI and 1 with diabetes).

Table 2. Coefficient-Based Models for Symptoms and Signs; Symptoms, Signs and Dipstick Results, Including Results Based on Multiple Imputation

Index Tests	Symptom and Sign Model		Symptom, Sign, and Dipstick Model	
	Adjusted OR ^a (95% CI) ^b	MI Adjusted OR ^a (95% CI)	Adjusted OR ^a (95% CI) ^b	MI Adjusted OR ^a (95% CI)
Pain/crying when passing urine				
No problem	1 [reference]	1 [reference]	1 [reference]	1 [reference]
Slight problem	1.56 (0.68-3.61)	1.73 (0.73-4.06)	1.01 (0.37-2.80)	1.16 (0.41-3.24)
Moderate problem	4.58 (2.27-9.25)	4.80 (2.30-10.04)	2.68 (1.16-6.18)	2.87 (1.21-6.82)
Severe problem	14.32 (6.81-30.11)	15.81 (7.37-33.89)	9.64 (3.92-23.69)	10.33 (4.11-25.96)
Smelly urine				
No problem	1 [reference]	1 [reference]	1 [reference]	1 [reference]
Slight problem	4.08 (2.00-8.33)	4.28 (2.02-9.05)	2.97 (1.29-6.85)	3.16 (1.32-7.59)
Moderate problem	5.00 (2.64-9.48)	5.14 (2.60-10.19)	4.16 (2.02-8.57)	4.34 (2.00-9.39)
Severe problem	8.49 (3.74-19.26)	8.76 (3.76-20.41)	4.13 (1.51-11.31)	4.44 (1.57-12.54)
Previous UTI				
No	1 [reference]	1 [reference]	1 [reference]	1 [reference]
Yes	2.71 (1.39-5.27)	2.66 (1.34-5.26)	2.39 (1.12-5.11)	2.36 (1.10-5.03)
Cough				
No problem	1 [reference]	1 [reference]	1 [reference]	1 [reference]
Slight problem	1.28 (0.67-2.45)	1.32 (0.68-2.55)	1.27 (0.59-2.72)	1.30 (0.60-2.81)
Moderate problem	1.31 (0.69-2.48)	1.38 (0.72-2.68)	1.95 (0.95-4.00)	2.04 (0.98-4.22)
Severe problem	0.28 (0.08-0.93)	0.29 (0.09-0.97)	0.36 (0.09-1.48)	0.36 (0.09-1.51)
Clinician global impression of illness severity ^c				
0-1	1 [reference]	1 [reference]	1 [reference]	1 [reference]
2	1.97 (0.95-4.12)	1.98 (0.93-4.19)	2.14 (0.93-4.91)	2.13 (0.92-4.97)
3	2.66 (1.28-5.54)	2.72 (1.28-5.81)	2.65 (1.16-6.07)	2.63 (1.13-6.14)
4-5	3.57 (1.61-7.91)	3.87 (1.72-8.73)	2.96 (1.18-7.42)	3.24 (1.28-8.24)
6 or more	6.84 (2.52-18.56)	7.24 (2.59-20.25)	5.80 (1.81-18.60)	6.28 (1.92-20.61)
Abdominal examination: any tenderness				
No	1 [reference]	1 [reference]	1 [reference]	1 [reference]
Yes	2.40 (1.03-5.61)	2.24 (0.95-5.25)	1.34 (0.40-4.45)	1.18 (0.35-3.94)
Ear examination: any acute abnormality				
No	1 [reference]	1 [reference]	1 [reference]	1 [reference]
Yes	0.30 (0.11-0.83)	0.27 (0.10-0.74)	0.46 (0.18-1.22)	0.40 (0.15-1.09)
Dipstick: leukocytes				
Negative	1 [reference]	1 [reference]
Trace	1.81 (0.68-4.81)	1.78 (0.66-4.78)
1+	0.70 (0.16-3.13)	0.66 (0.14-3.12)
2+	5.27 (2.52-11.04)	5.19 (2.45-10.98)
3+	10.45 (4.11-26.53)	10.36 (3.94-27.26)
Dipstick: nitrites				
Negative	1 [reference]	1 [reference]
Positive	5.25 (2.56-10.77)	5.37 (2.58-11.19)
Dipstick: blood				
Negative	1 [reference]	1 [reference]
Non-heme	0.88 (0.36-2.17)	0.89 (0.35-2.21)
Heme trace	4.16 (1.34-12.85)	4.08 (1.28-13.05)
Heme 1+	2.65 (0.87-8.03)	2.84 (0.92-8.79)
Heme 2+ or 3+	1.71 (0.65-4.51)	1.74 (0.64-4.73)
AUROC curve (95% CI) ^d	0.89 (0.84-0.94)	0.90 (0.85-0.95)	0.93 (0.89-0.96)	0.93 (0.90-0.97)
Validated AUROC curve ^e	0.87	0.88	0.90	0.90
Calibration slope ^e	0.89	0.87	0.83	0.83

AUROC = area under the receiver operating characteristic; MI = multiple imputation; OR = odds ratio; UTI = urinary tract infection.

^a Calculated using shrunken estimates from the bootstrap internal validation calibration slope.

^b Missing values coded to modal category.

^c Graded on a scale from 1 to 10, where higher scores indicate greater severity.

^d Calculated without internal validation.

^e Calculated from the bootstrap internal validation.

Effects of Replacing US With UK Definition of UTI

Data were available for all 2,740 (100%) children, 35 (1.3%) of whom had findings that were UTI positive using the US definition of UTI. We found good agreement (prevalence- and bias-adjusted $\kappa = 0.98$), and crude and adjusted odds ratios were similar, comparing US and UK definitions of UTI, showing the same graded associations, except for severe cough US (adjusted OR = 0.74 (95% CI, 0.23-2.37) compared with 0.29 UK (95% CI, 0.09-.97), data available from the authors). Step 1 and step 2 diagnostic utilities were stable based on the US definition, with validated AUROCs of 0.88 and 0.93 respectively.

DISCUSSION

Summary of Findings

In a large cohort of young children with acute illness in primary care, 2.2% of clean-catch urine samples met the UK microbiological criteria for UTI. Based on data obtained from clean-catch samples, we developed novel coefficient (for computer use) and points-based clinical rules to help clinicians select children for urine

sampling and antibiotic treatment with high diagnostic utility. For step 1, the coefficient-based rule was diagnostically superior to the points-based rule, which in turn was superior to clinical diagnosis. For step 2, dipstick testing was diagnostically superior to symptoms and signs alone (both coefficient and points-based rules), and was not diagnostically useful in children with the lowest UTI probability, for whom step 1 would not result in urine collection.

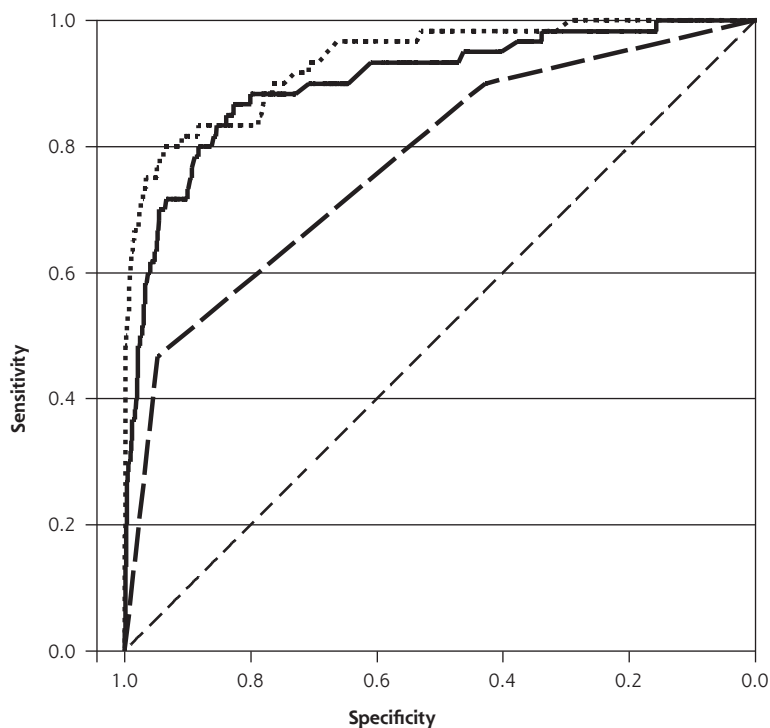
Strengths and Limitations

To our knowledge, this study is the largest and most rigorous investigation of diagnostic accuracy of UTI in children aged under 5 years in primary care. The generalizability of the final sample is supported by the similarity between participating children and those invited but declining. We achieved high levels of data completeness across a large number of primary care sites and maintained blinding of recruiting staff to the reference standard. Index tests were measured according to routine clinical practice using standardized reporting forms and equipment, and nearly all tests were completed within 24 hours of urine sample retrieval, minimizing disease progression bias. The

low number of samples with antimicrobial substances minimized treatment paradox. Our reference standard was specific to common uropathogens and excluded index tests. Two members of the laboratory staff, blinded to all index tests but not age, performed the microbiological cultures and interpreted results using a standardized process in a single laboratory. Our broad eligibility criteria allowed us to identify previously unidentified clinical features useful for both increasing (smelly urine) and decreasing (absence of severe cough, normal ear examination) UTI probability, as well as to demonstrate the absence of diagnostic utility of other features (such as fever, fever of unknown origin, vomiting, lethargy, irritability, and poor feeding) widely believed to be diagnostically useful.^{5,6} Our results are stable using the more conservative US definition of UTI.

The main limitation is the relatively small number of UTI

Figure 1. ROC curve for multiple imputation, coefficient-based models for clinician diagnosis of UTI (dashed line), symptoms and signs only (solid line), and symptoms, signs and dipstick (dotted line).



ROC = receiver operating characteristic; UTI = urinary tract infection.

Table 3. Diagnostic Test Characteristics of Coefficient-Based Models for Sensitivity Cut Points, Using Symptom and Sign Model for Urine Sampling and Antibiotic Treatment and the Symptom, Sign, and Dipstick Model for Antibiotic Treatment

Linear Predictor Cut Point, Shrunken Coefficients (\geq) ^a	Linear Predictor Cut Point, Unshrunk Coefficients (\geq) ^a	Sensitivity (95% CI)	Specificity (95% CI)	Positive Predictive Value (95% CI)	Negative Predictive Value (95% CI)	Children Clinical Rule Positive % (95% CI)
Symptom and sign model						Urine Sampled, Antibiotic Treated^b
-0.504	-0.195	20.0 (11.7-32.0)	99.8 (99.5-99.9)	66.7 (42.9-84.2)	98.2 (97.7-98.7)	0.7 (0.4-1.0)
-1.092	-0.87	30.0 (19.8-2.7)	99.5 (99.1-99.7)	56.3 (39.0-72.1)	98.4 (97.9-98.9)	1.2 (0.8-1.6)
-1.813	-1.698	40.0 (28.5-52.8)	98.2 (97.6-98.6)	32.9 (23.1-44.4)	98.7 (98.1-99.0)	2.7 (2.1-3.3)
-2.059	-1.98	50.0 (37.6-62.4)	97.5 (96.9-98.1)	31.3 (22.8-41.2)	98.9 (98.4-99.2)	3.5 (2.9-4.3)
-2.372	-2.34	60.0 (47.2-71.5)	96.3 (95.5-97.0)	26.7 (19.9-34.7)	99.1 (98.6-99.4)	4.9 (4.2-5.8)
-2.729	-2.75	70.0 (57.3-80.2)	94.6 (93.7-95.4)	22.6 (17.1-29.1)	99.3 (98.9-99.6)	6.8 (5.9-7.8)
-3.396	-3.515	80.0 (68.0-88.3)	88.3 (87.0-89.4)	13.3 (10.1-17.2)	99.5 (99.1-99.7)	13.2 (12.0-14.5)
-3.717	-3.884	85.0 (73.6-92.0)	83.9 (82.4-85.2)	10.6 (8.1-13.6)	99.6 (99.2-99.8)	17.6 (16.2-19.1)
-4.567	-4.86	93.3 (83.5-97.5)	61.0 (59.1-62.8)	5.1 (3.9-6.6)	99.8 (99.4-99.9)	40.2 (38.4-42.0)
-5.299	-5.7	96.7 (87.6-99.2)	37.8 (35.9-39.6)	3.4 (2.6-4.3)	99.8 (99.2-100.0)	63.0 (61.2-64.8)
-6.138	-6.664	100	15.7 (14.4-17.1)	2.6 (2.0-3.3)	100	84.6 (83.2-85.9)
Symptom, sign, and dipstick model						Antibiotic Treated^c
0.801	1.43	20.0 (11.7-32.0)	99.9 (99.7-100.0)	85.7 (57.3-96.4)	98.2 (97.7-98.7)	0.5 (0.3-0.9)
-0.122	0.321	40.0 (28.5-52.8)	99.9 (99.7-100.0)	88.9 (70.7-96.4)	98.7 (98.2-99.0)	1.0 (0.7-1.4)
-1.346	-1.15	60.0 (47.2-71.5)	99.3 (98.8-99.5)	64.3 (51.0-75.7)	99.1 (98.7-99.4)	2.0 (1.6-2.6)
-3.114	-3.275	80.0 (68.0-88.3)	93.8 (92.9-94.7)	22.5 (17.4-28.6)	99.5 (99.2-99.7)	7.8 (6.8-8.8)
-3.700	-3.98	83.3 (71.7-90.8)	88.3 (87.0-89.5)	13.8 (10.6-17.7)	99.6 (99.2-99.8)	13.2 (12.0-14.6)
-4.746	-5.237	96.7 (87.6-99.2)	66.3 (64.5-68.1)	6.0 (4.7-7.7)	99.9 (99.6-100.0)	35.0 (33.3-36.8)
-5.235	-5.825	98.3 (89.1-99.8)	53.1 (51.2-54.9)	4.5 (3.5-5.7)	99.9 (99.5-100.0)	48.1 (46.2-49.9)
-5.955	-6.69	100	29.5 (27.8-31.2)	3.1 (2.4-3.9)	100	71.2 (69.4-72.8)

Note: Results based on models using multiple imputation to deal with missing values. For comparison, clinician diagnosis sensitivity was 46.6% and specificity was 94.7%.

^a Derived from the coefficient-based models using multiple imputation where the coefficients are listed within the Supplemental Appendix, Table 5 (<http://www.annfammed.org/content/14/4/325/suppl/DC1>).

^b Percentage of children who would be at or above this threshold assuming all children had a urine sample.

^c Percentage of children who would be at or above this threshold assuming all children had a urine sample and dipstick test.

diagnoses, especially in the youngest children, which had an impact on 3 areas. First, we were not able to externally validate our rules. Although external validation is desirable before clinical application, bootstrap validation takes account of model overoptimism. By analogy, it is reasonable for clinical practice to change on the basis of a single, high-quality, well-conducted

randomized trial, though replication is desirable. That said, because we recruited from first-point-of-contact, primary care sites, we consider it necessary to evaluate the rule's performance before use in secondary care. Second, our rule development breached the widely quoted "10 events per candidate predictor." This rule of thumb, however, has little theoretical justification

and has been shown to be conservative,²¹ and the consequences of variable selection are strongly dependent on the strength of association of candidate predictors with the outcome. Here, predictors of UTI are biologically plausible, and associations are substantial. Finally, children aged 2 years or younger are underrepresented in these analyses because of the difficulty of obtaining clean-catch samples in this age-group. We found our rule, however, to be diagnostically accurate in children aged under 3 years who are able to provide a clean-catch urine sample. Our secondary care experience, as well as a recent report describing a bladder stimulation technique for infants,²² suggest that when sufficient time, space, and personnel are available, clean-catch sampling is possible in most young children.

We mitigated the impact of false-positive urine cultures (arising as a result of asymptomatic bacteriuria²³ or contamination) using 3 design features. First, children were eligible only if experiencing constitutional and/or urinary tract symptoms; second, the rule was developed using only clean-catch samples; and third, we used a single research laboratory, which used methods superior to NHS laboratories to distinguish contaminated urine. Incorporation bias could have inflated the AUROC for step 2 using the US definition of UTI, because dipstick leukocytes were used as both an index test and within the reference standard definition.

Results in Context with Other Studies

One systematic review of 8 primary studies (7,892 children),²⁴ and 5 primary studies^{10,11,25-27} of a further 18,796 children (with only one study²⁵ conducted in general practice surgeries) assessed UTI prevalence and the diagnostic value of clinical symptoms and signs in children aged under 5 years.²⁸ These studies found similar UTI prevalence and showed abdominal pain, back pain, dysuria, frequency, and new-onset urinary incontinence to be positively associated with UTI.²⁴ Stridor, audible wheeze, circumcision, temperature less than 39°C with a source, abnormal chest sounds, chest crackles, aged under 3 years, not feeling hot, and rapid breathing were inversely associated with UTI. The largest study, which included 16,000 children coming to children's emergency departments, derived a complex model based on 27 symptoms and signs, with an AUROC of 0.80 (95% CI, 0.78-0.82).²⁶ The only previous study to recruit from general practice surgeries found that younger age, urinary frequency, and pain/crying on passing urine were associated with UTI, but it had an insufficient number of children with UTI to develop a clinical rule.²⁵ Previous investigation of malodorous urine has shown conflicting results,^{27,29} but our study strongly supports its diagnostic value. Dipstick testing has been considered diagnostically unhelpful in young children.⁵

Clinical and Research Implications

In keeping with recently updated US guidelines,⁶ our results support a risk-based approach to the identification of children for investigation of UTI. Pain or crying while passing urine, smelly urine, previous UTI, absence of severe cough, severe illness, abdominal tenderness, and absence of ear abnormalities can be used for deciding which children for whom a urine sample (step 1) and dipstick results would improve specificity for antibiotic treatment (step 2). For both steps, increasing diagnostic sensitivity can be achieved by increasing urine-sampling rates, which may not be feasible or affordable. Precisely how these results are used is likely to depend on clinician preference. Some clinicians may wish to use these as risk factors to feed into clinical judgment. Others may prefer to use a checklist approach and use the points-based clinical rules (Figure 2, which can be used without computer assistance), which focus attention on predictive factors rather than those (such as fever) with poor diagnostic utility. Clinicians concerned about overdiagnosis and treatment could select a higher specificity threshold, whereas higher sensitivity thresholds would reduce underdiagnosis. When the rule is used, it should supplement, not replace, clinical judgment.

Further research is needed to distinguish pathogens from contamination and asymptomatic bacteriuria.²³ Given the expense of an external validation study and the low rates of routine urine sampling (which render routine data sets unsuitable), we consider the most cost-effective future research strategy would be to assess the impact of the DUTY clinical rule on clinical behavior and patient outcome in a randomized trial. The strongest design would integrate the presentation of the coefficient-based clinical rule within routine clinical care, probably via electronic medical records.

Our rule can be used to enhance current clinical practice in the identification of young children for noninvasive urine sampling in primary care. Fever should not be used to stratify UTI probability, and dipstick testing can be used to improve specificity for empiric antibiotic treatment in this population.

To read or post commentaries in response to this article, see it online at <http://www.annfamned.org/content/14/4/325>.

Key words: urinary tract infections; primary health care; pediatrics; diagnosis; anti-bacterial agents

Submitted November 20, 2015; submitted, revised, March 24, 2016; accepted April 7, 2016.

Author affiliations: Centre for Academic Primary Care, School of Social and Community Medicine, University of Bristol, Clifton, Bristol, United Kingdom (Hay); School of Social and Community Medicine, University of Bristol, Clifton, Bristol, United Kingdom (Sterne, Hollingworth, Law-

Figure 2. DUTY (Diagnosis of Urinary Tract infection in Young children) clean-catch criteria.**How to use the DUTY Clean-Catch Urine Criteria**

1. The *DUTY Clean-Catch Urine Criteria* are for children in whom a clean-catch sample is possible.
2. Urinary tract infection (UTI) was defined as $\geq 10^5$ colony-forming units (CFU)/mL of a single or predominant uropathogen cultured from a clean-catch urine specimen.
3. Table A: Use the symptoms and signs to decide whether a clean-catch urine sample should be collected/antibiotics given (≥ 5 points or can be operationalized as 'any 3 of the 5' symptoms and signs has been shown to be cost-effective). Clinicians concerned about overdiagnosis and treatment can select a higher specificity (at least 6 points) threshold. Higher sensitivity thresholds (eg, ≥ 3 points or ≥ 4 points) would reduce underdiagnosis, but these thresholds have not been shown to be cost-effective.
4. It is not clear which of the following possible antibiotic treatment strategies is most cost-effective: (i) immediate presumptive treatment of all sampled children; (ii) immediate dipstick-guided treatment; or (iii) laboratory-guided (delayed) treatment.
5. For children with urine samples at the ≥ 5 -point threshold, the probability of UTI will be 18% (Supplemental Appendix, Table 4, <http://www.annfamned.org/content/14/4/325/suppl/DC1> [upper portion]). Although not demonstrably cost-effective, dipstick testing can raise or lower this probability (see Table B).
6. Table B: refer to Supplemental Appendix, Table 3, <http://www.annfamned.org/content/14/4/325/suppl/DC1> (lower portion) for probability of UTI with total score.
7. Consider advising the parents of all children (with and without urine samples) to seek medical advice if their child gets worse.
8. The DUTY Clean-Catch Urine Criteria are designed to supplement and not replace clinical judgment.

Table A: Should I get a urine sample?

Clinical Characteristic (Present/Absent) ^a	Points ^b
<i>Symptoms and signs</i>	<i>To guide urine collection</i>
Pain/crying passing urine ^c	2
Smelly urine ^c	2
Previous UTI ^c	1
Absence of severe cough ^d	2
Severe illness present ^e	2
<i>Collect clean-catch urine if symptoms and signs points total ≥ 5 "any 3 of the 5"</i>	

Table B: Should I give antibiotic treatment

Clinical Characteristic (Present/Absent) ^a	Points ^b
<i>Symptoms and signs</i>	<i>To guide antibiotic treatment</i>
Pain/crying passing urine ^c	2
Smelly urine ^c	2
Previous UTI ^c	1
Absence of severe cough ^d	2
Severe illness present ^e	2
Dipstick: Leukocytes positive	2
Dipstick: Nitrites positive	3
Dipstick: Blood positive	1

UTI = urinary tract infection.

^a Clinical characteristic wording as used in study Case Report Form and reported by parent/clinician unless stated otherwise.

^b Refer to the Supplemental Appendix, Table 3, <http://www.annfamned.org/content/14/4/325/suppl/DC1> (upper portion), for probability of UTI with total score.

^c Parents were asked to report presence/absence.

^d Parents were asked to grade presence of cough as no problem, slight problem, moderate problem or severe problem.

^e Score of ≥ 6 on the clinician global illness severity scale with range 0 (child completely well) to 10 (child extremely unwell).

ton, Busby, Birnie); South East Wales Trials Unit (SEWTU), Centre for Trials Research, Cardiff University, Heath Park, Cardiff, United Kingdom (Hood, Pickles, Waldron, Lises); Primary Care and Population Science, Faculty of Medicine, University of Southampton, Aldermoor Health Centre, Aldermoor Close, Southampton, United Kingdom (Little, Harman, Rumsby); Guys' and St Thomas' Charity Chair in Primary Care Research, NIHR Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London, Department of Primary Care and Public Health Sciences, London, United Kingdom (Delaney); Specialist Antimicrobial Chemotherapy Unit, Public Health Wales Microbiology Cardiff, University Hospital Wales, Heath Park, Cardiff, United Kingdom (Wootton, Howe); North Bristol NHS Trust, Southmead Hospital, Westbury-on-Trym, Bristol, United Kingdom (MacGowan); Division of Population Medicine, School of Medicine, Cardiff University, Heath Park, Cardiff, United Kingdom (O'Brien); Bristol Royal Hospital for Children, University Hospitals Bristol, NHS Foundation Trust, Bristol, United Kingdom (Dudley); Department of Paediatrics and Child Health, University Hospital of Wales, Heath Park, Cardiff, United Kingdom (Van Der Voort); King's College London, Division of Health and Social Care Research, Department of Primary Care and Public Health Sciences, London, United Kingdom (Durbaba); NIHR CLAHRC West, University Hospitals Bristol NHS Foundation Trust, Bristol, United Kingdom (Whit-

ing); Nuffield Department of Primary Care Health Sciences, University of Oxford, New Radcliffe House, Radcliffe Observatory Quarter, Oxford, United Kingdom, and General Practitioner, Cwm Taf University Health Board, Wales, United Kingdom (Butler).

Author contributions: Hay and Butler (co-chief investigators) were responsible for overall study design, management, and data interpretation. Hay led the writing of, and approved, the final manuscript. Butler also made substantial contributions to writing and reviewed the final manuscript. Sterne was lead study statistician and made substantial contributions to overall study design and statistical analysis, and writing, and reviewed the final manuscript. Hood contributed to overall study design, data analysis, and study management, and contributed to and reviewed the final manuscript. Little contributed to study design, supervised recruitment, contributed to and reviewed the final manuscript. Delaney contributed to overall study and data collection instrument design, supervised recruitment, and contributed to and reviewed the final manuscript. Hollingworth was lead health economist and made substantial contributions to study design, data analysis, and writing, and reviewed the final manuscript. Wootton was responsible for microbiology data and contributed to study design and data analysis, and contributed to and reviewed the final manuscript. Howe and MacGowan were co-

lead microbiologists and contributed to study design and data analysis, and contributed to and reviewed the final manuscript. Lawton conducted data analysis and clinical rule development, contributed to the writing, and reviewed the final manuscript. Busby conducted economic analysis, contributed to the writing, and reviewed the final manuscript. Pickles contributed to the data management and statistical analysis, and contributed to and reviewed the final manuscript. Birnie contributed to data analysis, and contributed to and reviewed the final manuscript. O'Brien contributed to the overall study design and data interpretation, and contributed to and commented on the final manuscript. Waldron contributed to study management, and contributed to and reviewed the final manuscript. Dudley and Van Der Voort provided specialist pediatric nephrology advice, contributed to study design, and contributed to and reviewed the final manuscript. Downing and Thomas-Jones were co-lead study managers, and contributed to and reviewed the final manuscript. Harman contributed to study management, and contributed to writing and reviewed the final manuscript. Lisles contributed to data management and data analysis, and contributed to and reviewed the final manuscript. Rumsby contributed to study management, and contributed to and reviewed the final manuscript. Durbaba developed Web-based data collection systems, and contributed to and reviewed the final manuscript. Whiting contributed to study design, literature review, and writing, and reviewed the final manuscript.

All authors had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis

Hay is guarantor for the study and 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 have been explained.

Funding support: The study was commissioned and funded by the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) program (project number 08/66/01) and a longer, more comprehensive version will be published in full in Health Technology Assessment (both in print and online). AH is funded by NIHR Research Professorship (NIHR-RP-02-12-012) and JS by NIHR Senior Investigator Award (NF-SI-0611-10168). Jonathan Sterne is funded by a National Institute for Health Research Senior Investigator award (NF-SI-0611-10168), and Alastair Hay by a NIHR Research Professorship (NIHR RP-R2-12-012).

Disclaimer: The views and opinions expressed by the authors do not necessarily reflect those of the National Health Service (NHS), NIHR, HTA, or the Department of Health.

Data Sharing Statement: The full data set will be made available when all studies described within the protocol are complete and published. Application for the data to be released should be made in writing to Professor Alastair Hay (co-Chief Investigator) via the Freedom of Information Officer at the University of Bristol.

Acknowledgments: We wish to thank the children, their families, the general practices and other recruitment sites. We also wish to thank the providers of nursing/clinical study officer support from the Primary Care Research Networks Greater London; Kent and Medway; Sussex; Surrey; Thames Valley; the Western; Peninsular; Cumbria and Lancashire; Northumberland and Tyne and Wear; and the National Institute for Social Care and Health Research, Clinical Research Centre in Wales (NISCHR-CRC). Additionally, we wish to acknowledge the support given by the South East Wales Trials Unit (funded by NISCHR), The Wales School of Primary Care Research (funded by NISCHR), the Comprehensive Local Research Networks of Central and East London, Western, Peninsula, Hampshire, and Isle of Wight, and the NIHR Biomedical Research and Development Department, Guy's and St Thomas' NHS Foundation Trust.

Supplementary materials: Available at <http://www.AnnFamMed.org/content/14/4/325/suppl/DC1/>.

References

- Butler C, O'Brien K, Pickles T, Hood K, Wootton M, Howe R, et al. Childhood urinary tract infection in primary care: a prospective observational study of prevalence, diagnosis, treatment, and recovery. *Br J Gen Pract*. 2015;65(633):e217-223.
- Coulthard MG, Lambert HJ, Vernon SJ, Hunter EW, Keir MJ, Matthews JN. Does prompt treatment of urinary tract infection in preschool children prevent renal scarring: mixed retrospective and prospective audits. *Arch Dis Child*. 2014;99(4):342-347.
- Jacobson SH, Eklöf O, Eriksson CG, Lins LE, Tidgren B, Winberg J. Development of hypertension and uraemia after pyelonephritis in childhood: 27 year follow up. *BMJ*. 1989;299(6701):703-706.
- Farnham SB, Adams MC, Brock JW III, Pope JC IV. Pediatric urological causes of hypertension. *J Urol*. 2005;173(3):697-704.
- National Collaborating Centre for Women's and Children's Health. *Urinary Tract Infection in Children: Diagnosis, Treatment and Long Term Management*. London, England: RCOG Press; 2007.
- Roberts KB; Subcommittee on Urinary Tract Infection, Steering Committee on Quality Improvement and Management. Urinary tract infection: clinical practice guideline for the diagnosis and management of the initial UTI in febrile infants and children 2 to 24 months. *Pediatrics*. 2011;128(3):595-610.
- Ammenti A, Cataldi L, Chimenz R, et al; Italian Society of Pediatric Nephrology. Febrile urinary tract infections in young children: recommendations for the diagnosis, treatment and follow-up. *Acta Paediatr*. 2012;101(5):451-457.
- Coulthard MG, Vernon SJ, Lambert HJ, Matthews JN. A nurse led education and direct access service for the management of urinary tract infections in children: prospective controlled trial. *BMJ*. 2003;327(7416):656.
- Victorian Paediatric Clinical Network. Clinical Practice Guidelines - Urinary Tract Infection. The Royal Children's Hospital Melbourne. http://www.rch.org.au/clinicalguide/guideline_index/Urinary_Tract_Infection_Guideline/. Updated April 2015.
- Gorelick MH, Shaw KN. Clinical decision rule to identify febrile young girls at risk for urinary tract infection. *Arch Pediatr Adolesc Med*. 2000;154(4):386-390.
- Gorelick MH, Hoberman A, Kearney D, Wald E, Shaw KN. Validation of a decision rule identifying febrile young girls at high risk for urinary tract infection. *Pediatr Emerg Care*. 2003;19(3):162-164.
- Shaw KN, Gorelick M, McGowan KL, Yakscoe NM, Schwartz JS. Prevalence of urinary tract infection in febrile young children in the emergency department. *Pediatrics*. 1998;102(2):e16.
- Kozer E, Rosenbloom E, Goldman D, Lavy G, Rosenfeld N, Goldman M. Pain in infants who are younger than 2 months during suprapubic aspiration and transurethral bladder catheterization: a randomized, controlled study. *Pediatrics*. 2006;118(1):e51-e56.
- Merritt KA, Ornstein PA, Spicker B. Children's memory for a salient medical procedure: implications for testimony. *Pediatrics*. 1994;94(1):17-23.
- Lohr JA, Downs SM, Dudley S, Donowitz LG. Hospital-acquired urinary tract infections in the pediatric patient: a prospective study. *Pediatr Infect Dis J*. 1994;13(1):8-12.
- Public Health England. *UK Standards for Microbiology Investigations: Investigation of Urine (B41)*. London, England: Public Health England; 2014.
- 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.

18. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. 2001;54(8):774-781.
19. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol*. 1990;43(6):543-549.
20. Butler CCSJ, Lawton M, O'Brien K, et al. 'Nappy pad' urine samples for guiding investigation and treatment of urinary tract infection (UTI) in young children: Findings from the 'DUTY' prospective diagnostic cohort study. *Br J Gen Pract*. In press.
21. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. *Am J Epidemiol*. 2007;165(6):710-718.
22. Herreros Fernández ML, González Merino N, Tagarro García A, et al. A new technique for fast and safe collection of urine in newborns. *Arch Dis Child*. 2013;98(1):27-29.
23. Wettergren B, Jodal U, Jonasson G. Epidemiology of bacteriuria during the first year of life. *Acta Paediatr Scand*. 1985;74(6):925-933.
24. Shaikh N, Morone NE, Lopez J, et al. Does this child have a urinary tract infection? *JAMA*. 2007;298(24):2895-2904.
25. O'Brien K, Edwards A, Hood K, Butler CC. Prevalence of urinary tract infection in acutely unwell children in general practice: a prospective study with systematic urine sampling. *Br J Gen Pract*. 2013;63(607):e156-e164.
26. Craig JC, Williams GJ, Jones M, et al. The accuracy of clinical symptoms and signs for the diagnosis of serious bacterial infection in young febrile children: prospective cohort study of 15 781 febrile illnesses. *BMJ*. 2010;340:c1594.
27. Gauthier M, Gouin S, Phan V, Gravel J. Association of malodorous urine with urinary tract infection in children aged 1 to 36 months. *Pediatrics*. 2012;129(5):885-890.
28. Hay AD, Whiting P, Butler CC. How best to diagnose urinary tract infection in preschool children in primary care? *BMJ*. 2011;343:d6316.
29. Struthers S, Scanlon J, Parker K, Goddard J, Hallett R. Parental reporting of smelly urine and urinary tract infection. *Arch Dis Child*. 2003;88(3):250-252.

Get the *Annals of Family Medicine* by E-mail

Make sure you see every new issue while it's fresh; have the table of contents sent to you by e-mail for easy access to articles of interest.

Don't miss important research. Request the e-mail table of contents at http://www2.highroadsolution.com/aafp_annals_preference_center/search.aspx

ANNALS OF FAMILY MEDICINE
Indexed in the MEDLINE and MEDLARS databases WWW.ANNFAMMED.ORG JANUARY/FEBRUARY 2016 VOL. 14, NO. 1

Now Available:
The Wonder and the Mystery
Annals anthology of personal reflections and innovative ideas

The full text of the journal is available online at <http://www.annfammed.org> and through various aggregators, including PubMed Central, EBSCO, and MDCConsult. The *Annals* is indexed in the MEDLINE and MEDLARS, Science Citation Index Expanded, Current Contents/Clinical Medicine, PsycINFO, EMBASE, and CINHAI databases.

EDITORIALS

In This Issue: Size Matters
Kurt C. Stange

Achieving PCMH Status May Not Be Meaningful for Small Practices
Kelley K. Glancey; James G. Kennedy

The Paradox of Size: How Small, Independent Practices Can Thrive in Value-Based Care
Farzad Mostashari

ORIGINAL RESEARCH

Solo and Small Practices: A Vital, Diverse Part of Primary Care
Winston R. Liaw; Anuradha Jetty; Stephen Petterson; Lars E. Peterson; Andrew W. Bazemore
Family physicians in solo and small practices outnumber those in larger practices.

Large Independent Primary Care Medical Groups
Lawrence P. Casalino; Melinda A. Chen; C. Todd Staub; Matthew J. Press; Jayme L. Mendelssohn; John T. Lynch; Yesenia Miranda
Large physician-owned groups have the potential to make primary care attractive to physicians and improve patient care.

Primary Care Physician Panel Size and Quality of Care: A Population-Based Study in Ontario, Canada
Simone Dahroug; William Hogg; Jaime Younger; Elizabeth Muggah; Grant Russell; Richard H. Glazier
In Ontario, larger patient panel sizes do not decrease quality of care, but cancer screening rates are slightly lower.

Willingness to Exchange Health Information via Mobile Devices: Findings From a Population-Based Survey
Katrina J. Serrano; Mandi Yu; William T. Riley; Vaishali Patel; Penelope Hughes; Kathryn Marchesini; Audie A. Atienza
Willingness to exchange health information via mobile devices varies with the sensitivity of the