Supplemental materials for:

Hay AD, Sterne J, Hood K, et al. Improving the diagnosis and treatment of urinary tract infection in young children in primary care: results from the DUTY prospective diagnostic cohort study. *Ann Fam Med*. 2016;14(4):325-336.

This supplemental material has been provided by the author and not edited by Annals of Family Medicine

Supplemental Appendix: Documents in support of the DUTY prospective diagnostic cohort study article

Methods: statistical analysis (full text version)

We compared the age and gender of the children who were recruited with those children whose parents declined to participate. We used logistic regression to estimate associations of index tests with urine culture positivity. Where categorical variables had one category with very few observations, we examined the frequency of symptom and sign categories blind to association with urine culture results and merged the least frequent categories prior to analyses. P values were derived using likelihood ratio tests. For ordinal variables, both heterogeneity and trend p values were derived. A heterogeneity p value tests the null hypothesis that prevalence of UTI does not differ between categories, while a trend p value tests the null hypothesis that there is no linear increase in the log odds of UTI per category. Continuous variables were divided into quintiles and trend p values were derived using the median within categories. We examined plots of the log odds of culture positivity against the median within quintiles for evidence of non-linearity. We used two methods for dealing with missing data, including "don't know" responses. First, missing data were coded as the modal non-missing value. Second, we repeated multivariable analyses using the chained equations approach to multiple imputation: estimates and Wald p values[1] based on 50 imputed datasets were derived using Rubin's rules.[2] A complete case analysis was not feasible due to the reduction in sample size.

Step 1 - symptoms and signs

We derived 'coefficient-based' models in two stages. First, we selected symptoms and signs with either trend or heterogeneity univariable p values <0.01. Second, we derived models from selected symptoms and signs using backwards stepwise selection and an exclusion criterion of heterogeneity p value >0.1. We investigated using more liberal p value thresholds of 0.1 and 0.2 at the first stage, and found no important diagnostic utility differences of the final models (results available on request).

We generated 'points-based' models (easy to calculate without a computer) by dichotomising parent-reported symptom variables to 'present/absent', except for cough which was dichotomised at 'severe/all other categories' and clinicians' global illness severity impression at ≥6 threshold. We removed other physical examination variables as these contributed least to the models. We derived the points by dividing each coefficient by the smallest coefficient in the model and rounding to the nearest integer.

We quantified diagnostic accuracy using the area under the receiver operating characteristic (AUROC) curve with 95% confidence interval. We internally validated coefficient-based models using the bootstrap procedure described by Steyerberg:[3] we calculated a validated AUROC and a calibration slope (shrinkage factor) by which we multiplied model coefficients in order to derive internally validated odds ratios. Because 'points-based' models have fixed coefficients such internal validation is not possible: instead we internally validated these models before rounding the coefficients. For each model, we selected linear predictor cut-points corresponding to a range of values for sensitivity, and then calculated the corresponding specificity, negative and positive predictive values, and proportion of children classified positive, with 95% confidence intervals. Model diagnostic parameters were compared against 'clinical diagnosis' of UTI (where clinicians considered UTI to be 'fairly' or 'very' certain). In a sensitivity analysis we fitted the coefficient models in data restricted to children under three years of age.

Since children presenting 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. We used three 'fever' variables (parent reported 'fever now or in the past 24 hours', parent reported 'fever at any time during this illness' and temperature ≥38°C on examination) combined with symptoms and signs regarded as evidence of a non-UTI illness (rash, diarrhoea, blocked/runny nose, cough, wheeze, shortness of breath, chest pain, earache, sore throat, oxygen saturation <94%, throat abnormality, ear abnormality, and chest abnormality).

Step 2 - symptoms, signs and dipstick testing

We used the model development processes described in step 1, extending the symptoms and signs models to include dipstick results, with the points-based model dipstick results dichotomised at the 'negative /positive' threshold. To assess the added value of dipsticks over symptoms and signs alone we first quantified the change in AUROC and second, used a simulation approach based on the step 1 points-based model together with multinomial logit models in which dipstick results were predicted by the dichotomised symptoms and signs as predictors. The aim of the simulation procedure was to quantify the additional value of dipstick results, by calculating the change in the probability of UTI associated with addition of dipstick results to the symptoms and signs model. Under this procedure we: (i) sampled coefficient values from the multivariate normal distribution of the multinomial logit parameter estimates; (ii) randomly generated a set of dipstick results based on the sampled coefficients; and (iii) computed the corresponding probability of UTI based on the shrunken coefficients for the symptoms, signs and dipstick points-based model. For each combination of symptoms and signs we generated 10,000 samples and calculated the probability of

UTI with and without the dipstick results and the change in probability of UTI after accounting for the dipstick results. One of the dipstick combinations was dropped since it was observed in only three individuals and led to numerical instability.

Effects of replacing US with UK UTI definition

We calculated the prevalence and bias adjusted kappa statistic to assess agreement between UK and US UTI definitions[4] and 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: Supplemental Appendix, Table 7 (full text version)

Web Table 7 shows change in the probability of UTI associated with addition of dipstick test results to the step 1 symptoms and signs model, based on the simulation study and using dichotomised symptoms and signs. There was a clear trend towards increased diagnostic value of dipstick results (change in probability of UTI) as the probability of UTI based on symptoms and signs increased. Based on these results, we defined two groups of children according to whether they step 1 probability of UTI was greater than 5% (group 1, points score ≥5) or less than 5% (group 2). In group 2 the median change in post-test UTI probability when dipstick results were additionally used for prediction was 0.3%. In group 2 (children with step 1 UTI probability >5%), the dipstick results had a substantial impact on the probability of UTI (median change in post-dipstick test probability of UTI 9.9%, 95% range 1.4% to 55.5%).

Supplemental Appendix, Table 1. Full list of index tests from the DUTY study case record form.

Parent reported symptoms

Numerical response

"How many days has your child been unwell?" "Please rate your overall impression of your child's illness when at its worst"[a], "Not counting today approximately how many times have you consulted a doctor (for this illness)", "Approximately how many nappies/pull ups has your child used in the last 24 hours?" "How many times do you usually bath or shower your child in a normal week?"

Scaled

"Compared to yesterday is your child the same, better, worse?"

Severity scale (No, slight, moderate, severe, don't know/NA)

"Child Not themselves", "Confusion/disorientation", "Disturbed sleep", "Fever now or past 24 hours", "Fever at any time during this illness", "Chills or shivering", "New generalised rash with this illness", "Nappy rash or similar", "Muscle aches or pains all over", "Headaches", "Refused feeds/eating less than normal", "Poor weight gain/weight loss (in the last month)", "Vomiting", "Diarrhoea" (both at any time and in last 24 hours), "Constipation in last week", "Abdominal pain/tummy ache/pulling legs up", "Passing urine more often", "Changes in urine appearance"; if Yes choices of Darker, Cloudy, Smelly, Bloody, Other, "Pain/crying passing urine", "Day or bed wetting when previously dry", "Blocked or runny Nose"," Cough"," Wheeze", "Short of breath, difficulty breathing or grunting", "Chest pains", "Earache/holding ear/s", "Sore throat", "More unwell compared to previous illnesses"

Past medical and family history (Yes, No)

"Does your child have any ongoing health problems?" Asthma, Diabetes, Heart disease, High blood pressure, Learning difficulty, other (specify), "What the pregnancy full term for your child?" Born late, Born early, if early estimate weeks, "Was the child breast fed?" If Yes for how long exclusively < 3months, ≥ 3 months, "Were you ever told your child's kidney, bladder or urinary system was abnormal in any way after a pregnancy ultrasound scan?" , "Has your child or member of your family ever been diagnosed with vesico-ureteric or 'kidney' reflux?", "Has your child been circumcised?" (boys only)

(Yes, No, Don't know)

"Has your child or member of your family ever been treated for urine infections?"

Child, Mother, Father, Sibling (tick box)

"Does your child or any member of your family have any other renal/urinary problem?" urethral/ureter/bladder/kidney problem

Child/parent socio-demographic characteristics

Child "Age"," Gender", "Ethnicity", ***Parent*** "Highest qualification", "Cost of living question"

Clinician reported signs on examination

"Temperature" (absolute with method), "Oxygen saturation" %, "Pulse rate" (absolute), "Respiratory rate" (absolute), "Capillary refill time" (<2 sec, 2-5 sec > 5sec), "Global impression of child" (0 -10)[a]

Scaled or Not examined; "Hydration level" , "Conscious level" , "General examination"

Normal, abnormal, Not examined (and detail) with choices and other; "Throat examination", "Ear examination", "Chest examination", "Abdomen examination". Abdomen includes any of suprapubic, loin or other abdominal tenderness. "Are you aware of the dipstick result?" Yes No "What is your working diagnosis (pre dipstick result)?" Choice of URTI, chest infection, bronchitis, bronchiolitis, pneumonia, exacerbation of asthma, tonsillitis, otitis media, pharyngitis, UTI, gastroenteritis, viral illness, other (specify), "How certain are you of your diagnosis?" Choice of uncertain, fairly, certain, very

Before the dipstick results

"Were you planning on treating this child with antibiotics?"– if Yes; Yes UTI, Yes other, immediate or delayed, "Would you be referring this child to a paediatrician or admitted them?" If Yes; Yes for UTI, Yes for other, "If the child was Not in DUTY would you have requested a urine sample?" Yes No

Dipstick tests; Leukocytes, Protein, Ketones, Blood, specific gravity, pH, Glucose, Nitrites.

Diagnosis post reviewing dipstick results; "Has your diagnosis changed?" Yes No if Yes list as working diagnosis above, "Was the child referred for same day urgent assessment at hospital?" (GP surgery) or "Admitted to hospital for this illness (ED)"? No, Yes UTI, Yes other, N/A.

[a] 0 completely well with no constitutional upset, 10 showing life threatening symptoms or signs

This supplemental material has been provided by the author and not edited by Annals of Family Medicine

Supplemental Appendix, Table 2. Model fit for dichotomised coefficient and points-based models, using multiple imputation.

	Symptom and signs model		Symptom, signs and dipstick model	
Dichotomised variables	**Adjusted OR[a] (95% CI)**	**Points**	**Adjusted OR[a] (95% CI)**	**Points**
Pain/crying passing urine	4.99 (2.88,8.64)	2	3.05 (1.65,5.66)	2
Smelly urine	6.17 (3.50,10.88)	2	4.72 (2.55,8.74)	2
Previous UTI	2.58 (1.31,5.06)	1	2.07 (0.94,4.57)	1
Absence of severe cough	4.30 (1.34,13.75)	2	2.93 (0.89,9.64)	2
Severe illness[c]	4.22 (1.72,10.35)	2	4.13 (1.56,10.93)	2
Dipstick: Leukocytes positive			4.00 (2.12,7.56)	2
Dipstick: Nitrites positive			7.15 (3.60,14.18)	3
Dipstick: Blood positive			2.03 (1.11,3.73)	1
AUROC (95% CI)[b]	0.860 (0.810,0.910)	0.856 (0.808,0.903)	0.902 (0.855,0.950)	0.900 (0.853,0.948)
Validated[d] AUROC	0.849	[e]	0.892	[e]
Calibration slope[d]	0.947		0.942	

[a] Odds ratios calculated using shrunken estimates from the bootstrap internal validation calibration slope; [b] Calculated without validation ; [c] A score of 6 or more on the global impression scale; [d] calculated from the bootstrap internal validation; [e] Points-based models cannot be internally validated because coefficients are fixed

Supplemental Appendix, Table 3. Fever of unknown origin and laboratory positivity.

			Complete cases			Modal imputation	
Symptom	**Category**	**UTI +ve / Total**	**OR (95% CI)**	**p-value**	**UTI +ve / Total**	**OR (95% CI)**	**p-value**
Parent reported fever now in or in past 24 hours, unknown origin	No	28/1283	1 (ref)	0.464	59/2698	1 (ref)	0.876
	Yes	1/20	2.36 (0.31,18.24)		1/39	1.18 (0.16,8.72)	
Parent reported fever at any time during this illness, unknown origin	No	27/1274	1 (ref)	0.163	58/2683	1 (ref)	0.484
	Yes	2/29	3.42 (0.77,15.12)		2/54	1.74 (0.41,7.32)	

Supplemental Appendix, Table 4. Diagnostic test characteristics (95% CI) for a range of sensitivity cut points for the points-based model, using symptoms and signs model (upper portion of table for urine collection and antibiotic treatment), and symptoms, signs and dipstick (lower portion of table for antibiotic treatment).

Points' cut point (≥)	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Percentage of children clinical rule positive
Symptom and sign model					*Percentage urine sampled/ antibiotic treated*[a]
7	8.3% (3.5%, 18.5%)	99.7% (99.4%, 99.8%)	35.7% (15.7%, 62.4%)	98.0% (97.4%, 98.4%)	0.5% (0.3%, 0.9%)
6	43.3% (31.5%, 56.0%)	96.4% (95.6%, 97.0%)	21.1% (14.8%, 29.2%)	98.7% (98.2%, 99.1%)	4.5% (3.8%, 5.3%)
5	51.7% (39.2%, 63.9%)	94.6% (93.7%, 95.4%)	17.7% (12.7%, 24.1%)	98.9% (98.4%, 99.2%)	6.4% (5.5%, 7.4%)
4	80.0% (68.0%, 88.3%)	78.1% (76.5%, 79.6%)	7.6% (5.7%, 9.9%)	99.4% (99.0%, 99.7%)	23.2% (21.6%, 24.8%)
3	85.0% (73.6%, 92.0%)	74.4% (72.7%, 76.0%)	6.9% (5.3%, 9.0%)	99.6% (99.1%, 99.8%)	26.9% (25.3%, 28.6%)
2	98.3% (89.1%, 99.8%)	16.9% (15.5%, 18.4%)	2.6% (2.0%, 3.3%)	99.8% (98.5%, 100.0%)	83.4% (82.0%, 84.8%)
1	100%	15.9% (14.5%, 17.3%)	2.6% (2.0%, 3.3%)	100%	84.5% (83.1%, 85.8%)
Symptom, sign and dipstick model					*Percentage antibiotic treated*[b]
13	5.0% (1.6%, 14.4%)	100% (.%,.%)	100% (.%,.%)	97.9% (97.3%, 98.4%)	0.1% (0.0%, 0.3%)
12	18.3% (10.5%, 30.2%)	99.9% (99.7%, 100.0%)	78.6% (50.6%, 92.9%)	98.2% (97.6%, 98.6%)	0.5% (0.3%, 0.9%)
11	26.7% (17.0%, 39.2%)	99.8% (99.6%, 99.9%)	76.2% (54.0%, 89.7%)	98.4% (97.8%, 98.8%)	0.8% (0.5%, 1.2%)
10	33.3% (22.6%, 46.1%)	99.6% (99.3%, 99.8%)	66.7% (48.4%, 81.0%)	98.5% (98.0%, 98.9%)	1.1% (0.8%, 1.6%)
9	48.3% (36.1%, 60.8%)	99.0% (98.5%, 99.3%)	51.8% (38.9%, 64.5%)	98.8% (98.4%, 99.2%)	2.0% (1.6%, 2.6%)
8	50.0% (37.6%, 62.4%)	98.1% (97.5%, 98.5%)	36.6% (26.9%, 47.5%)	98.9% (98.4%, 99.2%)	3.0% (2.4%, 3.7%)
7	66.7% (53.9%, 77.4%)	95.4% (94.6%, 96.2%)	24.7% (18.7%, 31.9%)	99.2% (98.8%, 99.5%)	5.9% (5.1%, 6.9%)
6	78.3% (66.2%, 87.0%)	90.7% (89.5%, 91.8%)	15.9% (12.1%, 20.5%)	99.5% (99.1%, 99.7%)	10.8% (9.7%, 12.0%)
5	81.7% (69.8%, 89.5%)	85.0% (83.6%, 86.3%)	10.9% (8.3%, 14.1%)	99.5% (99.1%, 99.7%)	16.4% (15.1%, 17.9%)
4	86.7% (75.5%, 93.2%)	68.9% (67.1%, 70.6%)	5.9% (4.5%, 7.6%)	99.6% (99.1%, 99.8%)	32.3% (30.6%, 34.1%)
3	91.7% (81.5%, 96.5%)	59.1% (57.3%, 61.0%)	4.8% (3.7%, 6.2%)	99.7% (99.2%, 99.9%)	42.0% (40.1%, 43.8%)
2	100%	15.1% (13.8%, 16.5%)	2.6% (2.0%, 3.3%)	100%	85.3% (83.9%, 86.5%)
1	100%	13.4% (12.1%, 14.7%)	2.5% (2.0%, 3.2%)	100%	86.9% (85.6%, 88.1%)
0	100%	0%	2.2% (1.7%, 2.8%)	100%	100%

[a] Percentage of children who would be at or above this threshold assuming that all children had a urine sample

[b] Percentage of children who would be at or above this threshold assuming that all children had a urine sample and dipstick test

For comparison, 'clinician diagnosis' sensitivity 46.6% and specificity 94.7%

Supplemental Appendix, Table 5. Coefficient models based on symptoms and signs; and on symptoms, signs and dipstick results, including models based on multiple imputation for children <3 years only.

	Symptom and sign model		Symptom, sign and dipstick model	
Index tests	**Adjusted OR (95 % CI)[a]**	**MI[b] adjusted OR (95% CI)**	**Adjusted OR (95 % CI)[a]**	**MI[b] adjusted OR (95% CI)**
Pain/crying when passing urine				
No problem	1 (ref)	1 (ref)	1 (ref)	1 (ref)
Slight problem	2.66 (0.56,12.54)	3.59 (0.68,18.93)	1.07 (0.08,14.45)	3.11 (0.12,80.17)
Moderate problem	3.92 (0.86,17.93)	4.44 (0.85,23.30)	1.79 (0.19,17.06)	4.67 (0.24,90.26)
Severe problem	15.60 (2.61,93.25)	20.23 (3.07,133.2)	7.68 (0.34,172.83)	40.02 (0.43,3750.1)
Smelly urine				
No problem	1 (ref)	1 (ref)	1 (ref)	1 (ref)
Slight problem	9.28 (2.20,39.17)	12.42 (2.33,66.27)	21.42 (2.47,186.05)	92.43 (2.70,3160.2)
Moderate problem	12.54 (3.56,44.12)	16.40 (3.56,75.50)	28.37 (4.54,177.20)	85.43 (3.64,2006.2)
Severe problem	9.62 (1.91,48.43)	12.93 (2.17,77.09)	12.11 (0.67,220.38)	15.50 (0.30,803.34)
Previous UTI				
No	1 (ref)	1 (ref)	1 (ref)	1 (ref)
Yes	4.84 (1.26,18.59)	4.40 (1.09,17.80)	14.90 (1.42,156.77)	17.28 (0.85,351.15)
Cough				
No problem	1 (ref)	1 (ref)	1 (ref)	1 (ref)
Slight problem	1.83 (0.51,6.58)	1.86 (0.48,7.22)	5.19 (0.81,33.34)	7.77 (0.73,82.68)
Moderate problem	1.50 (0.40,5.65)	1.69 (0.42,6.82)	13.66 (1.62,115.02)	26.60 (1.68,420.31)
Severe problem	0.81 (0.13,5.11)	0.89 (0.13,6.16)	2.11 (0.12,36.79)	2.95 (0.11,77.75)
Clinician global impression of illness severity (0-10)				
0-1	1 (ref)	1 (ref)	1 (ref)	1 (ref)
2	2.57 (0.57,11.59)	2.84 (0.56,14.31)	2.60 (0.30,22.58)	3.26 (0.22,48.95)
3	0.97 (0.16,6.01)	1.07 (0.15,7.32)	0.64 (0.05,7.72)	0.71 (0.04,13.50)
4-5	5.49 (1.20,25.13)	7.13 (1.40,36.29)	2.19 (0.23,20.94)	4.08 (0.27,61.42)
6 or more	12.63 (1.86,85.94)	19.56 (2.44,157.1)	20.97 (1.23,357.23)	89.34 (1.53,5208.1)
Abdominal exam: any tenderness				
No	1 (ref)	1 (ref)	1 (ref)	1 (ref)
Yes	2.79 (0.63,12.40)	2.92 (0.64,13.40)	1.25 (0.05,28.62)	1.11 (0.03,47.48)
Ear exam: any acute abnormality				
No	1 (ref)	1 (ref)	1 (ref)	1 (ref)
Yes	0.42 (0.08,2.14)	0.33 (0.06,1.84)	0.84 (0.11,6.28)	0.90 (0.07,11.15)
Dipstick: leukocytes				
Negative			1 (ref)	1 (ref)
Trace			24.54 (2.11,285.10)	42.29 (1.85,964.9)
+			7.47 (0.32,172.62)	15.03 (0.19,1193.5)
++			22.83 (2.97,175.71)	36.01 (2.88,449.7)
+++			49.03 (5.11,470.02)	86.91 (4.73,1596.7)
Dipstick: nitrites				
Negative			1 (ref)	1 (ref)
Positive			9.83 (1.59,60.63)	15.06 (1.36,167.11)
Dipstick: blood				
Negative			1 (ref)	1 (ref)
Non-haem			0.12 (0.00,5.76)	0.13 (0.00,9.17)
Haem trace			10.45 (0.86,126.60)	14.84 (0.65,336.33)
Haem +			7.03 (0.66,74.54)	12.50 (0.69,226.05)
Haem ++ or +++			7.83 (0.87,70.18)	12.90 (0.77,216.82)
Area under ROC curve (95% CI)	0.890 (0.81,0.98)	0.923 (0.86,0.99)	0.969 (0.94, 0.99)	0.983 (0.96,1)

[a] Missing values coded to model category [b]MI: multiple imputation

Supplemental Appendix, Table 6. Coefficients to calculate the linear predictors (displayed in the first two columns of Table 3)

	Shrunken coefficients		Unshrunken coefficients	
	Symptom and sign model	Symptom, sign and dipstick model	Symptom and sign model	Symptom, sign and dipstick model
Constant	-5.5732	-6.2139	-6.0152	-7.0011
Pain/crying when passing urine				
No problem	0	0	0	0
Slight problem	0.5455	0.1443	0.6263	0.1734
Moderate problem	1.5686	1.0531	1.8009	1.2658
Severe problem	2.7605	2.3348	3.1693	2.8062
Smelly urine				
No problem	0	0	0	0
Slight problem	1.4531	1.1513	1.6683	1.3838
Moderate problem	1.6379	1.4669	1.8805	1.7631
Severe problem	2.1708	1.4913	2.4923	1.7924
Previous UTI				
No	0	0	0	0
Yes	0.9779	0.8569	1.1227	1.0299
Cough				
No problem	0	0	0	0
Slight problem	0.2749	0.2594	0.3156	0.3118
Moderate problem	0.3254	0.7106	0.3736	0.8541
Severe problem	-1.2476	-1.0228	-1.4324	-1.2293
Clinician global impression of illness severity (0-10)				
0-1	0	0	0	0
2	0.6809	0.7574	0.7818	0.9103
3	1.0024	0.9662	1.1509	1.1613
4-5	1.3538	1.1771	1.5543	1.4148
6 or more	1.9800	1.8379	2.2732	2.209
Abdominal exam: any tenderness				
No	0	0	0	0
Yes	0.8052	0.1653	0.9244	0.1987
Ear exam: any acute abnormality				
No	0	0	0	0
Yes	-1.3171	-0.9085	-1.5122	-1.0919
Dipstick: leukocytes				
Negative		0		0
Trace		0.5744		0.6904
+		-0.4087		-0.4912
++		1.6466		1.9791
+++		2.3382		2.8103
Dipstick: nitrites				
Negative		0		0
Positive		1.6810		2.0204
Dipstick: blood				
Negative		0		0
Non-haem		-0.1219		-0.1465
Haem trace		1.4062		1.6902
Haem +		1.0447		1.2556
Haem ++ or +++		0.5523		0.6638

Supplemental Appendix, Table 7. Additional diagnostic value of dipstick testing and the prevalence of each combination of symptoms and signs (shrunken estimates), based on results of the simulation study.

Symptoms and signs[a]	Probability of UTI based on symptoms and signs model	Group based on 5% threshold for pre-dipstick probability of UTI[c]	Points total based on symptoms and signs points model	UTI/Total in multiple imputation dataset[b]	Median (2.5th and 97.5th percentiles) change in probability of UTI following dipstick testing
+,+,+,+,+	0.690	1	9	0/0	0.249 (0.045, 0.411)
+,+,+,+,-	0.463	1	8	0/1	0.306 (0.035, 0.452)
+,+,+,-,+	0.345	1	7	7/17	0.259 (0.072, 0.500)
+,+,-,+,+	0.341	1	7	0/1	0.224 (0.005, 0.544)
+,-,+,+,+	0.308	1	7	0/0	0.196 (0.029, 0.573)
-,+,+,+,+	0.265	1	7	0/1	0.189 (0.018, 0.562)
+,+,+,-,-	0.170	1	6	19/101	0.127 (0.017, 0.555)
+,+,-,+,-	0.167	1	6	0/0	0.107 (0.036, 0.621)
+,-,+,+,-	0.148	1	6	0/7	0.090 (0.037, 0.489)
-,+,+,+,-	0.123	1	6	1/9	0.085 (0.014, 0.575)
+,+,-,-,+	0.109	1	5	0/4	0.078 (0.004, 0.542)
+,-,+,-,+	0.096	1	5	2/23	0.066 (0.014, 0.546)
+,-,-,+,+	0.094	1	5	0/1	0.052 (0.013, 0.460)
-,+,+,-,+	0.079	1	5	2/22	0.059 (0.005, 0.458)
-,+,-,+,+	0.077	1	5	0/0	0.050 (0.023, 0.543)
-,-,+,+,+	0.067	1	5	0/6	0.041 (0.016, 0.368)
+,+,-,-,-	0.045	2	4	0/6	0.030 (0.013, 0.428)
+,-,+,-,-	0.039	2	4	8/202	0.025 (0.010, 0.259)
+,-,-,+,-	0.039	2	4	0/4	0.018 (0.002, 0.336)
-,+,+,-,-	0.032	2	4	6/198	0.023 (0.005, 0.326)
-,+,-,+,-	0.032	2	4	0/0	0.018 (0.005, 0.247)
-,-,+,+,-	0.027	2	4	3/69	0.014 (0.002, 0.243)
+,-,-,-,+	0.024	2	3	0/5	0.014 (0.003, 0.055)
-,+,-,-,+	0.019	2	3	0/4	0.013 (0.006, 0.143)
-,-,+,-,+	0.017	2	3	3/81	0.010 (0.004, 0.034)
-,-,-,+,+	0.017	2	3	0/1	0.007 (0.002, 0.053)
+,-,-,-,-	0.009	2	2	2/75	0.004 (0.001, 0.030)
-,+,-,-,-	0.008	2	2	0/21	0.004 (0.001, 0.018)
-,-,+,-,-	0.007	2	2	6/1424	0.003 (0.000, 0.018)
-,-,-,+,-	0.006	2	2	1/14	0.002 (0.002, 0.028)
-,-,-,-,+	0.004	2	1	0/25	0.002 (0.001, 0.014)
-,-,-,-,-	0.002	2	0	0/418	0.000 (0.000, 0.007)

Group based on 5% threshold	Median probability of UTI		
1	0.170	31/193	0.099 (0.014,0.555)
2	0.007	29/2547	0.003 (0.000,0.040)

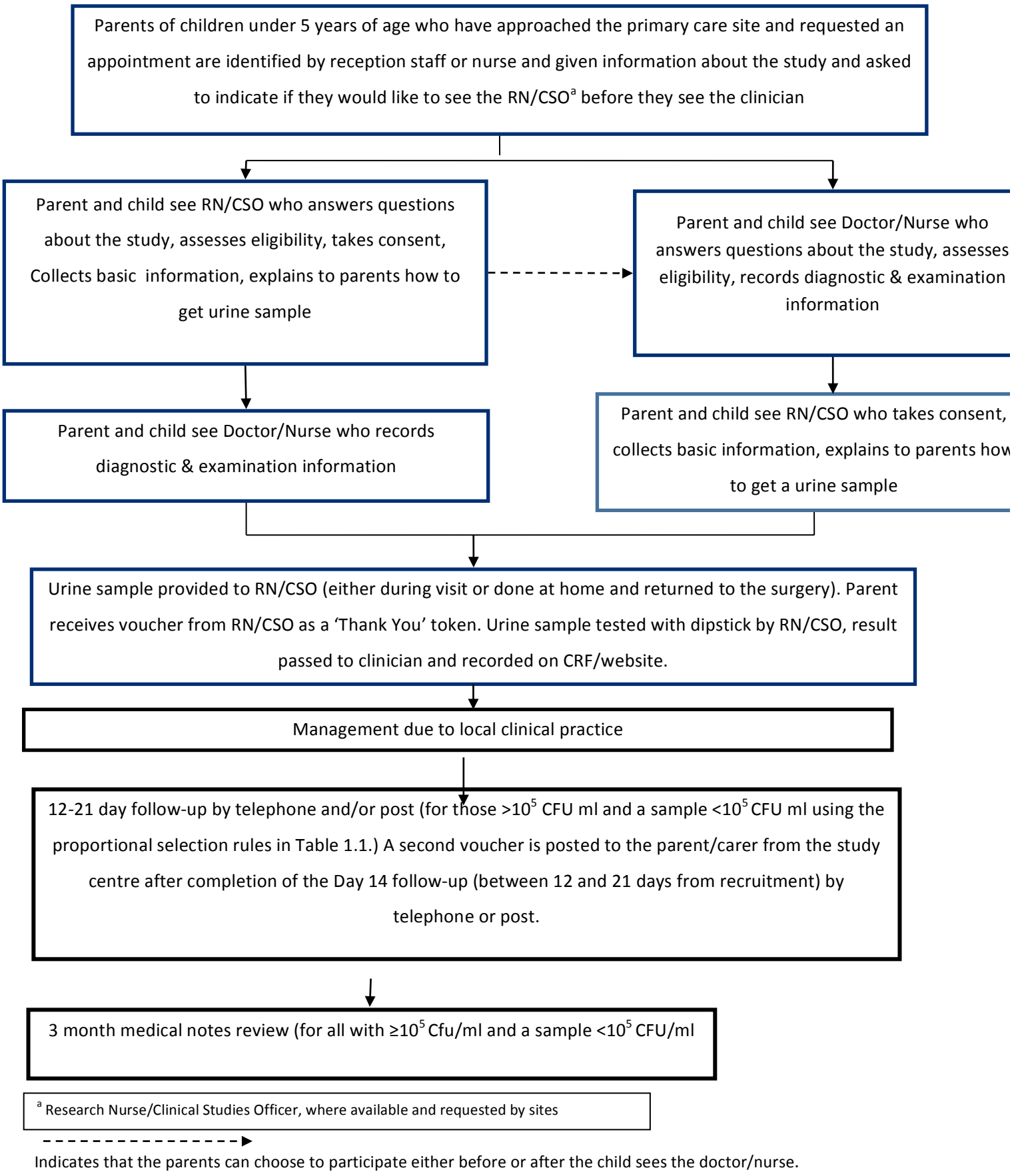
[a] The first plus if smelly urine is present, second plus if pain passing urine is present, third plus if severe cough is absent, fourth plus if there is severe illness, fifth plus if there is previous UTI

[b] Estimated number of UTI and total observations for symptom/sign combination as multiple imputation used for missing symptoms/signs

[c] 1 = Pre-dipstick probability of UTI >5%

[c] 2 = Pre-dipstick probability of UTI <5%

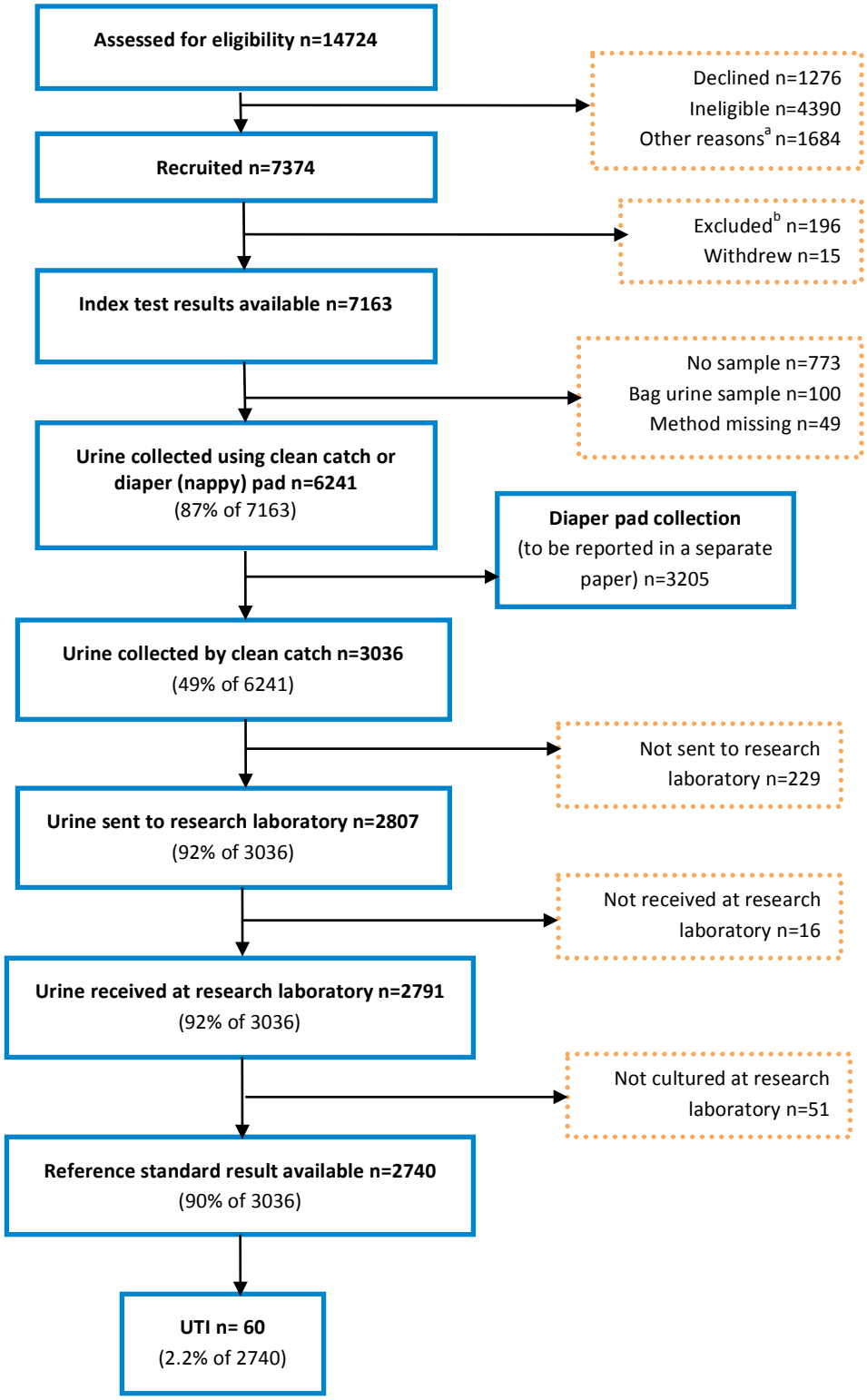
Supplemental Appendix, Figure 1. Participant recruitment methods.



[a] Research Nurse/Clinical Studies Officer, where available and requested by sites

Indicates that the parents can choose to participate either before or after the child sees the doctor/nurse.

Supplemental Appendix, Figure 2. Flow diagram of DUTY study recruitment.



[a] Includes left prior to invitation (n=811), no consent (n=214), language barrier (n=112)
[b] Excluded as retrospectively ineligible/duplicate recruitment/invalid consent (n=141) and poor data quality (n=55)

References

1. Wood AM, White IR, Royston P. How should variable selection be performed with multiply imputed data? *StatMed* 2008; 27(17): 3227-46.
2. Rubin D. Inference and missing data. *Biometrika* 1976; 63(3): 581-92.
3. 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): 774-81.
4. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990; 43(6): 543-9.