

## The Diagnosis of Urinary Tract infection in Young children (DUTY): a diagnostic prospective observational study to derive and validate a clinical algorithm for the diagnosis of urinary tract infection in children presenting to primary care with an acute illness

*Alastair D Hay, Kate Birnie, John Busby, Brendan Delaney, Harriet Downing, Jan Dudley, Stevo Durbaba, Margaret Fletcher, Kim Harman, William Hollingworth, Kerenza Hood, Robin Howe, Michael Lawton, Catherine Lises, Paul Little, Alasdair MacGowan, Kathryn O'Brien, Timothy Pickles, Kate Rumsby, Jonathan AC Sterne, Emma Thomas-Jones, Judith van der Voort, Cherry-Ann Waldron, Penny Whiting, Mandy Wootton and Christopher C Butler on behalf of the DUTY team*



**National Institute for  
Health Research**



# **The Diagnosis of Urinary Tract infection in Young children (DUTY): a diagnostic prospective observational study to derive and validate a clinical algorithm for the diagnosis of urinary tract infection in children presenting to primary care with an acute illness**

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

<sup>1</sup>Centre for Academic Primary Care, National Institute for Health Research (NIHR) School of Primary Care Research, School of Social and Community Medicine, University of Bristol, Bristol, UK

<sup>2</sup>School of Social and Community Medicine, University of Bristol, Bristol, UK

<sup>3</sup>Department of Primary Care and Public Health Sciences, National Institute for Health Research (NIHR) Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London, London, UK

<sup>4</sup>Bristol Royal Hospital for Children, University Hospitals Bristol NHS Foundation Trust, Bristol, UK

<sup>5</sup>Department of Primary Care and Public Health Sciences, Division of Health and Social Care Research, King's College London, London, UK

<sup>6</sup>Centre for Health and Clinical Research, University of the West of England, Bristol, UK

<sup>7</sup>South West Medicines for Children Local Research Network, University Hospitals Bristol NHS Foundation Trust, Bristol, UK

<sup>8</sup>South East Wales Trials Unit (SEWTU), Institute for Translation, Innovation, Methodology and Engagement, School of Medicine, Cardiff University, Cardiff, UK

<sup>9</sup>Specialist Antimicrobial Chemotherapy Unit, Public Health Wales Microbiology Cardiff, University Hospital Wales, Cardiff, UK

<sup>10</sup>Primary Care and Population Sciences Division, University of Southampton, Southampton, UK

<sup>11</sup>Southmead Hospital, North Bristol NHS Trust, Bristol, UK

<sup>12</sup>Cochrane Institute of Primary Care & Public Health, School of Medicine, Cardiff University, Cardiff, UK

<sup>13</sup>Department of Paediatrics and Child Health, University Hospital of Wales, Cardiff, UK

<sup>14</sup>Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

\*Corresponding author



**Declared competing interests of authors:** Paul Little is a member of the NIHR Journals Library Board, although he was not involved in the editorial processes for this report, and has provided consultancy work to Bayer Pharmaceuticals.

**Published July 2016**

**DOI: 10.3310/hta20510**

This report should be referenced as follows:

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

*Health Technology Assessment* is indexed and abstracted in *Index Medicus/MEDLINE*, *Excerpta Medica/EMBASE*, *Science Citation Index Expanded (SciSearch®)* and *Current Contents®/Clinical Medicine*.



ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.058

*Health Technology Assessment* is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) ([www.publicationethics.org/](http://www.publicationethics.org/)).

Editorial contact: [nhredit@southampton.ac.uk](mailto:nhredit@southampton.ac.uk)

The full HTA archive is freely available to view online at [www.journalslibrary.nihr.ac.uk/hta](http://www.journalslibrary.nihr.ac.uk/hta). Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: [www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)

## Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

## HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: <http://www.nets.nihr.ac.uk/programmes/hta>

## This report

The research reported in this issue of the journal was funded by the HTA programme as project number 08/66/01. The contractual start date was in January 2010. The draft report began editorial review in July 2013 and was accepted for publication in June 2014. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

© Queen's Printer and Controller of HMSO 2016. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library ([www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)), produced by Prepress Projects Ltd, Perth, Scotland ([www.prepress-projects.co.uk](http://www.prepress-projects.co.uk)).

## Health Technology Assessment Editor-in-Chief

**Professor Hywel Williams** Director, HTA Programme, UK and Foundation Professor and Co-Director of the Centre of Evidence-Based Dermatology, University of Nottingham, UK

## NIHR Journals Library Editor-in-Chief

**Professor Tom Walley** Director, NIHR Evaluation, Trials and Studies and Director of the EME Programme, UK

## NIHR Journals Library Editors

**Professor Ken Stein** Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

**Professor Andree Le May** Chair of NIHR Journals Library Editorial Group (EME, HS&DR, PGfAR, PHR journals)

**Dr Martin Ashton-Key** Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

**Professor Matthias Beck** Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

**Professor Aileen Clarke** Professor of Public Health and Health Services Research, Warwick Medical School, University of Warwick, UK

**Dr Tessa Crilly** Director, Crystal Blue Consulting Ltd, UK

**Dr Eugenia Cronin** Senior Scientific Advisor, Wessex Institute, UK

**Dr Peter Davidson** Director of NETSCC, HTA, UK

**Ms Tara Lamont** Scientific Advisor, NETSCC, UK

**Professor Elaine McColl** Director, Newcastle Clinical Trials Unit, Institute of Health and Society, Newcastle University, UK

**Professor William McGuire** Professor of Child Health, Hull York Medical School, University of York, UK

**Professor Geoffrey Meads** Professor of Health Sciences Research, Health and Wellbeing Research and Development Group, University of Winchester, UK

**Professor John Norrie** Health Services Research Unit, University of Aberdeen, UK

**Professor John Powell** Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

**Professor James Raftery** Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

**Dr Rob Riemsma** Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

**Professor Helen Roberts** Professor of Child Health Research, UCL Institute of Child Health, UK

**Professor Jonathan Ross** Professor of Sexual Health and HIV, University Hospital Birmingham, UK

**Professor Helen Snooks** Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

**Professor Jim Thornton** Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

**Professor Martin Underwood** Director, Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of members of the NIHR Journals Library Board:  
[www.journalslibrary.nihr.ac.uk/about/editors](http://www.journalslibrary.nihr.ac.uk/about/editors)

**Editorial contact:** [nihredit@southampton.ac.uk](mailto:nihredit@southampton.ac.uk)



# Abstract

## The Diagnosis of Urinary Tract infection in Young children (DUTY): a diagnostic prospective observational study to derive and validate a clinical algorithm for the diagnosis of urinary tract infection in children presenting to primary care with an acute illness

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

<sup>1</sup>Centre for Academic Primary Care, National Institute for Health Research (NIHR) School of Primary Care Research, School of Social and Community Medicine, University of Bristol, Bristol, UK

<sup>2</sup>School of Social and Community Medicine, University of Bristol, Bristol, UK

<sup>3</sup>Department of Primary Care and Public Health Sciences, National Institute for Health Research (NIHR) Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London, London, UK

<sup>4</sup>Bristol Royal Hospital for Children, University Hospitals Bristol NHS Foundation Trust, Bristol, UK

<sup>5</sup>Department of Primary Care and Public Health Sciences, Division of Health and Social Care Research, King's College London, London, UK

<sup>6</sup>Centre for Health and Clinical Research, University of the West of England, Bristol, UK

<sup>7</sup>South West Medicines for Children Local Research Network, University Hospitals Bristol NHS Foundation Trust, Bristol, UK

<sup>8</sup>South East Wales Trials Unit (SEWTU), Institute for Translation, Innovation, Methodology and Engagement, School of Medicine, Cardiff University, Cardiff, UK

<sup>9</sup>Specialist Antimicrobial Chemotherapy Unit, Public Health Wales Microbiology Cardiff, University Hospital Wales, Cardiff, UK

<sup>10</sup>Primary Care and Population Sciences Division, University of Southampton, Southampton, UK

<sup>11</sup>Southmead Hospital, North Bristol NHS Trust, Bristol, UK

<sup>12</sup>Cochrane Institute of Primary Care & Public Health, School of Medicine, Cardiff University, Cardiff, UK

<sup>13</sup>Department of Paediatrics and Child Health, University Hospital of Wales, Cardiff, UK

<sup>14</sup>Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

\*Corresponding author [alastair.hay@bristol.ac.uk](mailto:alastair.hay@bristol.ac.uk)

**Background:** It is not clear which young children presenting acutely unwell to primary care should be investigated for urinary tract infection (UTI) and whether or not dipstick testing should be used to inform antibiotic treatment.

**Objectives:** To develop algorithms to accurately identify pre-school children in whom urine should be obtained; assess whether or not dipstick urinalysis provides additional diagnostic information; and model algorithm cost-effectiveness.

**Design:** Multicentre, prospective diagnostic cohort study.

**Setting and participants:** Children < 5 years old presenting to primary care with an acute illness and/or new urinary symptoms.

**Methods:** One hundred and seven clinical characteristics (index tests) were recorded from the child's past medical history, symptoms, physical examination signs and urine dipstick test. Prior to dipstick results clinician opinion of UTI likelihood ('clinical diagnosis') and urine sampling and treatment intentions ('clinical judgement') were recorded. All index tests were measured blind to the reference standard, defined as a pure or predominant uropathogen cultured at  $\geq 10^5$  colony-forming units (CFU)/ml in a single research laboratory. Urine was collected by clean catch (preferred) or nappy pad. Index tests were sequentially evaluated in two groups, stratified by urine collection method: parent-reported symptoms with clinician-reported signs, and urine dipstick results. Diagnostic accuracy was quantified using area under receiver operating characteristic curve (AUROC) with 95% confidence interval (CI) and bootstrap-validated AUROC, and compared with the 'clinician diagnosis' AUROC. Decision-analytic models were used to identify optimal urine sampling strategy compared with 'clinical judgement'.

**Results:** A total of 7163 children were recruited, of whom 50% were female and 49% were < 2 years old. Culture results were available for 5017 (70%); 2740 children provided clean-catch samples, 94% of whom were  $\geq 2$  years old, with 2.2% meeting the UTI definition. Among these, 'clinical diagnosis' correctly identified 46.6% of positive cultures, with 94.7% specificity and an AUROC of 0.77 (95% CI 0.71 to 0.83). Four symptoms, three signs and three dipstick results were independently associated with UTI with an AUROC (95% CI; bootstrap-validated AUROC) of 0.89 (0.85 to 0.95; validated 0.88) for symptoms and signs, increasing to 0.93 (0.90 to 0.97; validated 0.90) with dipstick results. Nappy pad samples were provided from the other 2277 children, of whom 82% were < 2 years old and 1.3% met the UTI definition. 'Clinical diagnosis' correctly identified 13.3% positive cultures, with 98.5% specificity and an AUROC of 0.63 (95% CI 0.53 to 0.72). Four symptoms and two dipstick results were independently associated with UTI, with an AUROC of 0.81 (0.72 to 0.90; validated 0.78) for symptoms, increasing to 0.87 (0.80 to 0.94; validated 0.82) with the dipstick findings. A high specificity threshold for the clean-catch model was more accurate and less costly than, and as effective as, clinical judgement. The additional diagnostic utility of dipstick testing was offset by its costs. The cost-effectiveness of the nappy pad model was not clear-cut.

**Conclusions:** Clinicians should prioritise the use of clean-catch sampling as symptoms and signs can cost-effectively improve the identification of UTI in young children where clean catch is possible. Dipstick testing can improve targeting of antibiotic treatment, but at a higher cost than waiting for a laboratory result. Future research is needed to distinguish pathogens from contaminants, assess the impact of the clean-catch algorithm on patient outcomes, and the cost-effectiveness of presumptive versus dipstick versus laboratory-guided antibiotic treatment.

**Funding:** The National Institute for Health Research Health Technology Assessment programme.

# Contents

<b>List of tables</b>	<b>xv</b>
<b>List of figures</b>	<b>xxi</b>
<b>List of boxes</b>	<b>xxv</b>
<b>Glossary</b>	<b>xxvii</b>
<b>List of abbreviations</b>	<b>xxix</b>
<b>Plain English summary</b>	<b>xxxii</b>
<b>Scientific summary</b>	<b>xxxiii</b>
<b>Chapter 1 Introduction</b>	<b>1</b>
Prevalence of urinary tract infection	1
Importance of diagnosis	2
Asymptomatic bacteriuria	2
Longer-term adverse consequences	3
Missed diagnosis	3
Frequency of paediatric urine sampling in primary care	4
Methods of urine sampling in primary care	4
Which children should have their urine sampled?	5
Existing evidence for the diagnostic value of symptoms and signs	5
<i>Other risk factors</i>	6
<i>Summary of symptoms and signs</i>	6
Dipstick tests	7
Microbiological diagnosis of urinary tract infection in the laboratory	7
<i>Bacteriuria threshold</i>	7
<i>Culture techniques and uropathogenic organisms</i>	8
<i>Contamination</i>	8
Economic considerations	9
Summary of DUTY study design	9
<i>Summary of the research brief</i>	9
<i>Summary of the justification for the DUTY study design</i>	9
Aim of the DUTY study	10
<i>Research objectives</i>	10
Changes to the original funding proposal	10
<i>Planned change to age inclusion criterion</i>	10
<i>Planned change to research objective 5</i>	10
<i>Unanticipated changes to the validation process and reference standard</i>	11
<i>Other changes from the detailed project description</i>	11
<b>Chapter 2 Overall study methods</b>	<b>13</b>
Summary of study design	13
Ethics and research and development approvals	14
Parental contributions to the study	14

Site recruitment	15
<i>Recruitment period and locations</i>	15
<i>Selection of primary care sites</i>	15
<i>Site recruitment approach</i>	15
<i>The role of the local research networks</i>	15
<i>Recruitment models</i>	16
Study staff	16
<i>Site and staff training</i>	16
(Local) NHS laboratory recruitment	17
Service support costs	17
Participant selection	18
<i>Eligibility criteria</i>	18
<i>Widening participation</i>	18
Participant recruitment	19
<i>Registration and informed consent</i>	20
<i>Non-registration</i>	20
Urine sample collection	20
<i>Rationale for urine collection methods</i>	20
<i>Maximising urine retrieval rates</i>	21
<i>Urine dipstick methods</i>	21
<i>Dispatching urines to the laboratories</i>	22
Study thank you vouchers	22
Data collection	22
<i>Clinical case report form</i>	22
Patient follow-up	24
<i>Telephone follow-up at day 14</i>	24
<i>Three-month notes review process</i>	25
<i>Day-14 and 3-month selection</i>	25
Withdrawal and loss to follow-up	25
Discrepant laboratory results and patient safety	25
<i>Serious adverse events</i>	26
Processing of urine samples by a 'typical' local NHS laboratory	26
<i>Specimen collection</i>	26
<i>Sample processing</i>	26
<i>Urine microscopy</i>	27
<i>Urine culture</i>	27
<i>Bacterial identification</i>	27
<i>Susceptibility testing</i>	28
<i>Reporting procedure and storage of organisms</i>	28
<i>Quality control</i>	28
Processing of urine samples by the research laboratory	28
<i>Data processing</i>	28
<i>Urine microscopy</i>	29
<i>Urine culture</i>	29
<i>Colony counting</i>	29
<i>Identification of isolates</i>	29
<i>Susceptibility testing</i>	29
<i>Detection of antimicrobial substance in urine</i>	30
<i>Storage of urine and cultured organisms</i>	30
<i>Data entry to website</i>	30
<i>Quality control</i>	30
<i>Returning positive isolates to the research laboratory after the study</i>	30

Electronic data entry	30
<i>Data entry in primary care sites</i>	31
<i>Data entry in the NHS and research laboratories</i>	31
<i>Follow-up data entry in research centres</i>	31
Data management	31
<i>Data quality</i>	31
<i>Data cleaning</i>	31
<i>Data storage and retention</i>	32
Statistical methods	32
<i>Research objective algorithm</i>	32
<i>Sample size calculation</i>	32
<i>Determining prevalence rates</i>	33
Economic evaluation	33
<i>Other analyses</i>	33
Preventing and minimising bias	34
<b>Chapter 3 General results</b>	<b>35</b>
Site recruitment	35
<i>Types of primary care site</i>	36
<i>Recruitment models used by primary care sites</i>	36
NHS laboratories	37
Participant recruitment	37
<i>Eligibility assessment</i>	37
<i>Comparison of children recruited to DUTY with those not recruited</i>	37
<i>Recruitment accruals</i>	39
<i>Exclusions</i>	41
<i>Withdrawals</i>	41
Serious adverse events	41
<i>Positive unintended consequences of the DUTY study</i>	42
Urine collection	42
<i>Urine collection methods</i>	42
Microbiological culture results	43
Participant follow-up	44
<i>Proportional selection for follow-up</i>	44
<i>Fourteen-day follow-up</i>	45
<i>Three-month notes review</i>	45
Participant characteristics: descriptive statistics	46
<i>Description of study population</i>	46
<i>Description of presenting symptoms and signs</i>	47
<i>Description of clinical examination findings</i>	54
Study monitoring	61
<b>Chapter 4 Microbiological diagnosis of urinary tract infection by NHS and research laboratories</b>	<b>63</b>
Introduction	63
Methods	63
<i>Urine collection and sample processing</i>	63
<i>Statistical analysis</i>	63
Results	65

Discussion	73
<i>Summary of findings</i>	73
<i>Results in context with other studies</i>	73
<i>Clinical and microbiological implications</i>	74
<i>Implications for the DUTY study</i>	74
<b>Chapter 5 Derivation and validation of a clinical algorithm to guide diagnosis and treatment of urinary tract infection in pre-school children in primary care</b>	<b>75</b>
Introduction	75
Methods	75
<i>Participants</i>	75
<i>Index tests and urine collection</i>	75
<i>Reference standard</i>	75
<i>Statistical analysis</i>	76
<i>Fever of unknown origin</i>	77
<i>Points-based models</i>	77
<i>Added value of dipstick testing</i>	77
Results	78
<i>Participants</i>	78
<i>Clean-catch models</i>	78
<i>Nappy pad models</i>	89
Discussion	101
<i>Summary of findings</i>	101
<i>Strengths and limitations</i>	102
<i>Results in context with other studies</i>	102
<i>Clinical and research implications</i>	103
<i>Conclusions</i>	104
<b>Chapter 6 Health economic analysis and modelling of diagnostic strategies</b>	<b>105</b>
Introduction	105
Methods	105
<i>Overview</i>	105
<i>Model structure</i>	106
<i>Model inputs</i>	115
<i>Assumptions</i>	125
<i>Analytical methods</i>	126
<i>Sensitivity analysis</i>	126
Results	127
<i>Comparison of sampling strategies</i>	127
<i>Medium- and long-term outcomes</i>	131
<i>Comparison of testing and treatment strategies for the clean-catch DUTY5% strategy</i>	136
<i>Comparison of the full DUTY risk score with points-based algorithm</i>	137
<i>Deterministic sensitivity analysis</i>	141
Discussion	142
<i>Summary of findings</i>	142
<i>Strengths and limitations</i>	142
<i>Results in context with other studies</i>	143
<i>Clinical and research implications</i>	143
<i>Conclusions</i>	144

<b>Chapter 7 Determinants of urinary contamination</b>	<b>145</b>
Introduction	145
Methods	145
<i>Participants, urine collection and microbiological methods</i>	145
Definition of contamination	146
Statistical analysis	147
Results	147
<i>Microbiological predictors of contamination</i>	149
<i>Clinical predictors of contamination</i>	151
Discussion	153
<i>Summary of main results</i>	153
<i>Comparison with existing literature</i>	153
<i>Clinical implications</i>	154
<b>Chapter 8 Other results</b>	<b>155</b>
Introduction	155
Microbiological results (sensitivities and susceptibilities)	155
Follow-up outcomes	157
Generalisability and potential bias	162
<i>Verification bias</i>	162
<i>Incorporation bias</i>	166
Prevalence in urinary tract infection and its variation	166
Discussion	168
<b>Chapter 9 Discussion</b>	<b>169</b>
Summary of objectives and main results	169
Strengths	171
Weaknesses	171
<i>Design weaknesses</i>	171
<i>Analytic weaknesses</i>	172
Results in the context of research literature	173
Clinical and research implications	173
<b>Acknowledgements</b>	<b>175</b>
<b>References</b>	<b>187</b>
<b>Appendix 1 National Institute for Health Research Health Technology Assessment brief</b>	<b>199</b>
<b>Appendix 2 Systematic review (update) for the DUTY study: accuracy of symptoms and signs and dipstick tests for diagnosing UTI in children &lt; 5 years old in primary care and choice of urine sampling method</b>	<b>201</b>
<b>Appendix 3 DUTY study protocol</b>	<b>217</b>
<b>Appendix 4 Case report forms</b>	<b>233</b>
<b>Appendix 5 Day-14 data collection forms</b>	<b>247</b>
<b>Appendix 6 Three-month follow-up data collection form</b>	<b>265</b>
<b>Appendix 7 Laboratory data entry forms</b>	<b>271</b>

<b>Appendix 8</b> Laboratory methods	<b>277</b>
<b>Appendix 9</b> Health economic analysis and modelling of diagnostic strategies	<b>281</b>
<b>Appendix 10</b> STROBE checklist	<b>291</b>
<b>Appendix 11</b> STARD checklist	<b>293</b>



# List of tables

<b>TABLE 1</b> DUTY eligibility criteria	<b>18</b>
<b>TABLE 2</b> Selection rules used to select children for day-14 and 3-month follow-up	<b>24</b>
<b>TABLE 3</b> Summary of NHS laboratory methods: microscopy	<b>26</b>
<b>TABLE 4</b> Summary of NHS laboratory methods: culture	<b>26</b>
<b>TABLE 5</b> Reporting of significant isolates	<b>27</b>
<b>TABLE 6</b> Recruitment of primary care sites involved in DUTY across study centres	<b>35</b>
<b>TABLE 7</b> Recruitment models used by primary care sites involved in DUTY across study centres	<b>36</b>
<b>TABLE 8</b> Children screened but not recruited by the four study centres	<b>39</b>
<b>TABLE 9</b> Comparison of ages and sex of children recruited with those declined to participate	<b>39</b>
<b>TABLE 10</b> Frequency of urine sample growth in each of the research and NHS laboratory result categories	<b>43</b>
<b>TABLE 11</b> Proportions of cases selected for follow-up in each of the urine sample result categories and numbers completed	<b>44</b>
<b>TABLE 12</b> Day-14 follow-up interviews completed for the four study centres	<b>45</b>
<b>TABLE 13</b> Three-month note reviews completion by study centres	<b>45</b>
<b>TABLE 14</b> Demographic data (from section 2 of the CRF) for the study population overall	<b>46</b>
<b>TABLE 15</b> Comparison of DUTY recruits with national census data	<b>47</b>
<b>TABLE 16</b> Symptoms and signs	<b>48</b>
<b>TABLE 17</b> Clinical examination findings for all 7163 recruited participants	<b>55</b>
<b>TABLE 18</b> Parent-reported medical history for the 7163 recruited participants	<b>57</b>
<b>TABLE 19</b> Parent-reported medical history of previous urinary, bladder or renal problems	<b>58</b>
<b>TABLE 20</b> Clinical observations recorded for recruited participants	<b>59</b>
<b>TABLE 21</b> Treating clinicians' working diagnosis	<b>59</b>
<b>TABLE 22</b> Antibiotic treatment prescribed by clinical diagnosis	<b>60</b>

<b>TABLE 23</b> Characteristics of children and urine samples collected via clean catch or nappy pads, for the 4808 children with both NHS laboratory and research laboratory results	<b>66</b>
<b>TABLE 24</b> Extent of agreement between NHS and research laboratory result	<b>68</b>
<b>TABLE 25</b> Associations of signs, symptoms and urine dipstick tests with separate NHS and research laboratory results, and 'agree UTI positive'	<b>68</b>
<b>TABLE 26</b> Areas under the ROC curve (95% CI) from sensitivity analyses using six-variable logistic regression models	<b>72</b>
<b>TABLE 27</b> Clean-catch samples: children's characteristics and crude ORs for index tests associated with UTI	<b>80</b>
<b>TABLE 28</b> Clean-catch samples: 'full' coefficient models based on symptoms and signs and on symptoms, signs and dipstick results, including results based on MI	<b>82</b>
<b>TABLE 29</b> Clean-catch diagnostic test characteristics (95% CI) for a range of cut-points, using both the symptoms and signs model (upper part of table) and the symptoms, signs and dipstick model (lower part of table)	<b>84</b>
<b>TABLE 30</b> Parameters to calculate risk score for the clean-catch 'full' coefficient-based models	<b>85</b>
<b>TABLE 31</b> Fever and fever of unknown origin	<b>86</b>
<b>TABLE 32</b> Clean-catch points-based models using multiple imputation	<b>87</b>
<b>TABLE 33</b> Clean-catch rule of points-based models using multiple imputation for under-3-year-olds	<b>88</b>
<b>TABLE 34</b> Clean-catch diagnostic test characteristics (95% CI) for a range of cut-points for the points-based model, using both symptoms and signs model (upper part of table) and symptoms, signs and dipstick (lower part of table)	<b>88</b>
<b>TABLE 35</b> Outcome for illustrative 1000 children $\geq 2$ years presenting unwell to primary care using $\geq 5$ points (sensitivity = 51.7%, specificity = 94.6%)	<b>89</b>
<b>TABLE 36</b> Clean-catch diagnostic test characteristics (95% CI) for all possible cut-points of the coefficient-based points-based model, using the symptoms and signs model	<b>90</b>
<b>TABLE 37</b> Additional diagnostic value of dipstick testing and also the prevalence of each combination of symptoms and signs (clean-catch samples)	<b>91</b>
<b>TABLE 38</b> Nappy pad samples: children's characteristics and crude ORs for index tests associated with UTI	<b>93</b>
<b>TABLE 39</b> Nappy pad samples: models based on symptoms and signs, and on symptoms, signs and dipstick results, including results based on multiple imputation	<b>95</b>

<b>TABLE 40</b> Parameters to calculate risk score for the nappy pad ‘full’ coefficient-based models	97
<b>TABLE 41</b> Relationship between fever of unknown origin and UTI in nappy pad samples	97
<b>TABLE 42</b> Nappy pad points-based dichotomised models with and without dipstick test	98
<b>TABLE 43</b> Nappy pad diagnostic test characteristics (95% CI) for a range of cut-points using both the symptoms and signs model (upper part of table) and the symptoms, signs and dipstick model (lower part of table)	99
<b>TABLE 44</b> Nappy pad diagnostic test characteristics (95% CI) for a range of cut-points of the points-based model using both the symptoms and signs model (upper part of the table) and the symptoms, signs and dipstick model (lower part of the table)	100
<b>TABLE 45</b> Nappy pad diagnostic test characteristics (95% CI) for a range of cut-points of the coefficient-based points-based model, using the symptoms and signs model (upper part of table)	101
<b>TABLE 46</b> Prevalence of UTI in children presenting to primary care	115
<b>TABLE 47</b> Risk stratification for CJ and full DUTY risk scores	117
<b>TABLE 48</b> Risk stratification for DUTY points-based risk scores	118
<b>TABLE 49</b> Short-term model probabilities	119
<b>TABLE 50</b> Diagnostic test performance	120
<b>TABLE 51</b> Short-term Markov model probabilities	120
<b>TABLE 52</b> Symptom duration, costs and utilities in each treatment pathway during acute illness	121
<b>TABLE 53</b> Medium-term Markov model probabilities	122
<b>TABLE 54</b> Long-term probabilities	122
<b>TABLE 55</b> Costs (£)	123
<b>TABLE 56</b> Utilities	124
<b>TABLE 57</b> Diagnostic pathway	128
<b>TABLE 58</b> Diagnostic accuracy	129
<b>TABLE 59</b> Treatment pathway	130
<b>TABLE 60</b> Short-term costs and benefits	132

<b>TABLE 61</b> Medium- and long-term outcomes	<b>135</b>
<b>TABLE 62</b> Comparison of testing and treatment strategies for clean-catch DUTY5% algorithm	<b>136</b>
<b>TABLE 63</b> Comparison of full and points-based models for clean-catch samples	<b>138</b>
<b>TABLE 64</b> Sensitivity analysis, short-term parameters	<b>141</b>
<b>TABLE 65</b> Sensitivity analysis, long-term parameters	<b>141</b>
<b>TABLE 66</b> Comparing other research definitions of contamination (in date order)	<b>146</b>
<b>TABLE 67</b> Numbers of contaminated urines by different definition (research laboratory)	<b>149</b>
<b>TABLE 68</b> Number of contaminated urines by collection method (clean catch/ nappy pad)	<b>149</b>
<b>TABLE 69</b> Organisms prevalent in clean-catch contaminated urines compared with UTI-positive urines	<b>149</b>
<b>TABLE 70</b> Organisms prevalent in nappy pad contaminated urines compared with UTI-positive urines	<b>150</b>
<b>TABLE 71</b> Organisms prevalent in contaminated urines	<b>150</b>
<b>TABLE 72</b> Numbers of urines from nappy pad and clean-catch collection methods with $\geq 10$ SECs	<b>151</b>
<b>TABLE 73</b> Predictors of contamination in clean-catch urines	<b>151</b>
<b>TABLE 74</b> Predictors of contamination in nappy pad urines	<b>152</b>
<b>TABLE 75</b> Antimicrobial sensitivities for the 94 organisms considered positive for UTI at the research laboratory	<b>156</b>
<b>TABLE 76</b> NHS local laboratory antimicrobial sensitivities for the 94 organisms considered positive for UTI at the research laboratory	<b>158</b>
<b>TABLE 77</b> Early outcomes index visit and first 2 weeks	<b>159</b>
<b>TABLE 78</b> Outcomes up to 3 months	<b>163</b>
<b>TABLE 79</b> Comparison of those with a UTI status with those without a UTI status	<b>164</b>
<b>TABLE 80</b> Comparison of days unwell prior to consulting for those with a UTI status with those without a UTI status	<b>166</b>
<b>TABLE 81</b> Variation in UTI prevalence (clean-catch samples)	<b>167</b>
<b>TABLE 82</b> Variation in UTI prevalence (nappy pad samples)	<b>168</b>

<b>TABLE 83</b> Inclusion criteria	<b>201</b>
<b>TABLE 84</b> Diagnostic utility of symptoms, signs and urine dipstick testing for UTI in children < 5 years	<b>204</b>
<b>TABLE 85</b> Symptoms and signs and dipstick test results helpful for ruling out UTI in children < 5 years	<b>209</b>
<b>TABLE 86</b> Typical results recorded for a calibrated loop method	<b>277</b>
<b>TABLE 87</b> Dilution of urine	<b>277</b>
<b>TABLE 88</b> Summary of research laboratory methods	<b>277</b>
<b>TABLE 89</b> Identification of isolates at the research laboratory	<b>278</b>
<b>TABLE 90</b> Time for activities (minutes)	<b>286</b>
<b>TABLE 91</b> Fourteen-day resource use	<b>288</b>
<b>TABLE 92</b> Fixed and variable costs of infection excluding antibiotic treatment or imaging scans	<b>289</b>
<b>TABLE 93</b> Comparison of mean TAPQOL domain scores across GP diagnosis	<b>289</b>
<b>TABLE 94</b> STROBE statement: checklist of items that should be included in reports of observational studies	<b>291</b>
<b>TABLE 95</b> STARD checklist for reporting of studies of diagnostic accuracy (version January 2003)	<b>293</b>



# List of figures

<b>FIGURE 1</b> DUTY scheme (see <i>Figure 4</i> for numbers recruited)	<b>14</b>
<b>FIGURE 2</b> DUTY study participant flow diagram (the numbers of children recruited are reported in <i>Figure 4</i> )	<b>19</b>
<b>FIGURE 3</b> Recruitment of primary care sites	<b>35</b>
<b>FIGURE 4</b> DUTY study participant flow diagram	<b>38</b>
<b>FIGURE 5</b> Histograms of the ages of the children recruited and those who declined to participate	<b>40</b>
<b>FIGURE 6</b> DUTY recruitment accruals against HTA targets	<b>40</b>
<b>FIGURE 7</b> Urine collection methods	<b>42</b>
<b>FIGURE 8</b> Flow of participants in microbiological diagnosis of UTI in young children	<b>65</b>
<b>FIGURE 9</b> Receiver operating characteristic curves of symptoms, signs and urine dipstick tests with different definitions of urine culture positivity	<b>70</b>
<b>FIGURE 10</b> DUTY flow diagram for diagnostic algorithm	<b>79</b>
<b>FIGURE 11</b> Clean-catch ROC curve for multiple imputation models for symptoms and signs only (solid line) and symptoms, signs and dipstick (dotted line)	<b>84</b>
<b>FIGURE 12</b> Nappy pad ROC curve from multiple imputation for symptoms and signs only (solid line) and symptoms, signs and dipstick (dotted line)	<b>96</b>
<b>FIGURE 13</b> Sampling, testing and treatment strategies	<b>106</b>
<b>FIGURE 14</b> Laboratory-based treatment (LT)	<b>107</b>
<b>FIGURE 15</b> C1, not presumptive not positive	<b>109</b>
<b>FIGURE 16</b> C2, presumptive not positive	<b>110</b>
<b>FIGURE 17</b> C3, not presumptive positive	<b>110</b>
<b>FIGURE 18</b> C4, presumptive positive	<b>111</b>
<b>FIGURE 19</b> Laboratory- and dipstick-based treatment (DT)	<b>112</b>
<b>FIGURE 20</b> Presumptive treatment (PT)	<b>113</b>
<b>FIGURE 21</b> Short-term Markov model	<b>114</b>
<b>FIGURE 22</b> Clean-catch sampling and testing costs vs. percentage of laboratory-confirmed UTI	<b>133</b>

<b>FIGURE 23</b> Clean-catch short-term costs vs. average QALDs, comparison of sampling strategies	133
<b>FIGURE 24</b> Nappy pad sampling and testing costs vs. percentage of laboratory-confirmed UTI	134
<b>FIGURE 25</b> Nappy pad short-term costs vs. average QALDS, comparison of sampling strategies	134
<b>FIGURE 26</b> Clean-catch short-term costs vs. average QALDs, comparison of treatment and testing strategies	137
<b>FIGURE 27</b> Clean-catch sampling and testing costs vs. post-laboratory test sensitivity, comparison of full and points-based models	140
<b>FIGURE 28</b> Clean-catch short-term costs vs. average QALDs, comparison of full and points-based models	140
<b>FIGURE 29</b> Effect of clinical suspicion of UTI on the number of days until symptoms improved among children with UTI	148
<b>FIGURE 30</b> Participant flow in contamination chapter (research laboratory)	148
<b>FIGURE 31</b> Effect of clinical suspicion of UTI on the number of days until symptoms resolved among children with UTI	161
<b>FIGURE 32</b> Effect of clinical suspicion of UTI on the number of days until symptoms improved among children with contaminated urine	161
<b>FIGURE 33</b> Effect of clinical suspicion of UTI on the number of days until symptoms resolved among children with contaminated urine	162
<b>FIGURE 34</b> Flow of studies through the review	202
<b>FIGURE 35</b> Summary ROC plot showing estimates of sensitivity and specificity from studies assessing dipstick positive for both nitrite and LE together with summary ROC curve, summary estimate and 95% confidence and prediction regions	211
<b>FIGURE 36</b> Summary ROC plot showing estimates of sensitivity and specificity from studies assessing dipstick positive for either nitrite and LE together with summary ROC curve, summary estimate and 95% confidence and prediction regions	211
<b>FIGURE 37</b> Medium-term model	281
<b>FIGURE 38</b> Medium-term model, no VUR	282
<b>FIGURE 39</b> Medium-term model, untreated VUR	283
<b>FIGURE 40</b> Medium-term model, treated VUR	284
<b>FIGURE 41</b> Long-term model	285



- FIGURE 42** Comparison of observed and modelled (using a Weibull model: shape = 1.497, scale = 0.091) symptom resolution rates for patients with UTI treated immediately with antibiotics ( $n = 38$ ) 285
- FIGURE 43** Comparison of observed and modelled (using a Weibull model: shape = 1.246, scale = 0.099) symptom resolution rates for patients without UTI ( $n = 733$ ) 285
- FIGURE 44** Symptom resolution in a selection of treatment states 286



# List of boxes

**BOX 1** Final DUTY Clean-Catch Urine Criteria

**92**



# Glossary

**Local laboratory** The NHS laboratory local to, and used by, the primary care site at which the child was recruited.

**Primary care sites** Any primary care site (general practices, walk-in centres, children's emergency departments, out-of-hours general practitioner co-operatives or polyclinics) at which children were recruited.

**Research centre** One of the four partner organisations (at the universities of Bristol, Cardiff, London and Southampton) at which the principal investigators were based.

**Research laboratory** The term used to describe the Specialist Antimicrobial Chemotherapy Unit laboratory, Cardiff.

**Responsible clinician** A medical doctor or nurse practitioner whom the child was consulting for the illness when he/she presented to primary care.

**Urinary tract infection** The term (abbreviated to UTI) used by clinicians to refer to an illness caused by an infection of the urinary tract. In this report we sometimes use the term to refer to 'laboratory positivity', that is the culture of a significant quantity of a uropathogen in the laboratory.



## List of abbreviations

A&E	accident and emergency	LE	leucocyte esterase
AUROC	area under the receiver operating characteristic curve	LHB	local health board
BSAC	British Society for Antimicrobial Chemotherapy	LR	likelihood ratio
CAKUT	congenital abnormalities of the kidney and urinary tract	MCUG	micturating cystourethrogram
CBA	Columbia Blood Agar	MIC	minimum inhibitory concentration
CED	children's emergency department	NICE	National Institute for Health and Care Excellence
CFU	colony-forming units	NIHR	National Institute for Health Research
CI	confidence interval	NISCHR-CRC	National Institute for Social Care and Health Research Clinical Research Centre
CJ	clinical judgement	NMB	net monetary benefit
CLED	cystine-lactose-electrolyte-deficient agar	NPV	negative predictive value
CLRN	comprehensive local research network	NR	not resistant
CRF	case report form	ONS	Office for National Statistics
CSO	clinical studies officer	OR	odds ratio
CSV	comma-separated values	PCRN	primary care research network
DMSA	dimercaptosuccinic acid	PCT	primary care trust
DUTY	Diagnosis of Urinary Tract infection in Young children	PPV	positive predictive value
EQ-5D	European Quality of Life-5 Dimensions	PRS	progressive renal scarring
ESRD	end-stage renal disease	PSA	probabilistic sensitivity analysis
ESRF	end-stage renal failure	QALD	quality-adjusted life-day
EUCAST	European Committee on Antimicrobial Susceptibility Testing	QALY	quality-adjusted life-year
EURICA	The Epidemiology of URinary tract Infection in Children presenting with Acute illness in primary care	R&D	research and development
GP	general practitioner	RBC	red blood cell
HPA	Health Protection Agency	RCT	randomised controlled trial
HTA	Health Technology Assessment	RN	research nurse
HUI	Health Utilities Index	ROC	receiver operating characteristic
ID	identification	SACU	Specialist Antimicrobial Chemotherapy Unit
		SAE	serious adverse event
		SLaM	South London and Maudsley
		SOP	standard operating procedure

## LIST OF ABBREVIATIONS

SPA	suprapubic aspiration	UTI	urinary tract infection
SQL	Structured Query Language	VUR	vesicoureteral reflux
TAPQOL	TNO-AZL (Netherlands Organisation for Applied Scientific Research Academic Medical Centre) Preschool children Quality of Life	WBC	white blood cell
		WIC	walk-in centre
		WTP	willingness to pay



## Plain English summary

The DUTY (Diagnosis of Urinary Tract infection in Young children) study was designed to improve the way in which doctors and nurses find out when a young child has a urine infection so that they can start effective treatment earlier.

A total of 7163 children under 5 years old, who were visiting their general practitioner (GP) or a similar NHS service, joined the study. All were unwell or had urine symptoms. We collected details about the child's illness and their general health, and a urine sample. We preferred the urine sample to be collected by a 'clean catch' (straight from the child into a urine container), but where this was not possible we used a 'nappy pad' (a clean towel put into the child's nappy after cleaning their bottom). Once we had the urine, we did a simple test (urine dipstick) at the GP practice. After that, it was sent to a research laboratory to see if there was an infection.

We looked to see if any parent-reported symptoms or doctor-/nurse-reported examination findings or the dipstick test results made a urine infection more likely. We found that they did, especially when children were able to provide a 'clean-catch' sample. These urine infections were more likely when their parents said that children:

- had pain/crying when they passed urine
- had smelly urine
- had had a previous urine infection and
- did not have a severe cough.

Urine infections were also more likely:

- when the doctor/nurse thought that the child was more unwell, and
- when the dipstick test was positive.



# Scientific summary

## Background

It is not clear which young children with acute illness presenting to primary care should be investigated for urinary tract infection (UTI) and whether or not dipstick testing should be used to inform antibiotic treatment.

## Objectives

The **D**agnosis of **U**rinary **T**ract infection in **Y**oung children (DUTY) study objectives were to (1) develop algorithms, based on symptoms and signs to accurately identify children in whom a urine sample should be obtained; (2) assess whether or not dipstick urinalysis provides additional diagnostic information; (3) model algorithm cost-effectiveness; and (4) compare contamination rates between the clean-catch and 'Newcastle' nappy pad sampling methods.

## Design

The DUTY study was a multicentre, prospective diagnostic cohort study that included: a comparison of reliability and accuracy between NHS laboratories and a single research laboratory to establish the optimum reference standard; the derivation and validation of algorithms to identify children warranting urine collection and establish the added value of urine dipstick testing for antibiotic treatment decisions; a health economic evaluation of validated algorithms; and a comparison of urinary contamination rates from clean-catch and nappy pad samples.

## Setting

NHS 'first-point-of-contact' primary care sites, including 225 general practitioner (GP) surgeries, four children's emergency departments (CEDs) and four walk-in centres across England and Wales.

## Participants

Children < 5 years of age and presenting with any acute (up to 28 days) illness and/or new urinary symptoms. Children were excluded if they were not constitutionally unwell; if they were known to have a neurogenic or surgically reconstructed bladder; if they were using a permanent or intermittent urinary catheter; if the main presenting problem was trauma; or if antibiotics had been taken in the previous 7 days. Clinicians were asked to recruit consecutive eligible children and, where this was not possible, to collect non-recruited children's age and sex.

## Index tests and urine collection methods

Following consent, and blind to the reference standard, index tests (symptoms, signs and dipstick results) were recorded on a case report form. Symptoms included the child's medical history and parent-reported symptoms (graded as absent, mild, moderate or severe, when at their worst during the illness). Clinically qualified NHS staff (GPs, nurse practitioners and emergency department doctors/nurses) performed and

recorded examination findings, which included 'clinicians' global impression of the child's illness severity' and full respiratory and abdominal assessments. In total, 107 symptoms and signs were recorded and, preceding urine dipstick testing, clinicians recorded their opinion of UTI likelihood ('clinical diagnosis'), and their urine sampling and UTI treatment intentions ('clinical judgement'). Urine was collected by 'clean catch' (preferred) or nappy pad.

## Methods to compare culture results from NHS and research laboratories

Microbiology reports from the NHS and research laboratories were classified based on extent and purity of growth and whether or not the species grown was a uropathogen (defined as a member of the *Enterobacteriaceae* group). For NHS laboratories, pure/predominant growths of uropathogens at  $\geq 10^5$  colony-forming units (CFU) per ml were considered positive and research laboratory samples were considered positive if  $\geq 10^5$  CFU/ml of a single uropathogen ('pure growth') or  $\geq 10^5$  CFU/ml of a uropathogen with  $\geq 3 \log_{10}$  difference between the growth of this and the next species ('predominant growth') was present. This analysis included only those children with index test and both NHS and research laboratory results available. Agreement between laboratories was assessed using kappa statistics, with analyses additionally stratified by urine collection method (clean catch or nappy pad) and by age (0 to < 2, 2 to < 3 and 3 to < 5 years). Laboratory accuracy was investigated by comparing the strength of association between NHS and research laboratory UTI positivity and a small number of symptoms, signs and dipstick test results selected because they had previously been reported in the literature to be clearly related to UTI, albeit largely in emergency care settings. Those with the strongest associations and thought suitable for all ages and collection methods were urinary symptoms (pain/crying when passing urine, passing urine more often, changes in urine appearance); temperature  $\geq 39$  °C; and nitrite- or leucocyte-positive results from urine dipstick tests. We used logistic regression models to quantify associations of selected variables with laboratory UTI positivity and we plotted receiver operating characteristic (ROC) curves and used the area under ROC curve (AUROC) to quantify diagnostic accuracy.

### Algorithm development methods

We sequentially evaluated selected index tests in two groups: parent-reported symptoms and clinician-reported signs (from the physical examination); and urine dipstick results. First, we selected those variables with either trend or heterogeneity univariable  $p$ -value  $< 0.01$  for either collection method or when all samples were analysed together. Second, we derived models from among all the selected symptoms and signs, separately for nappy pad and clean-catch samples, using backwards stepwise selection and an exclusion criterion of heterogeneity  $p$ -value  $> 0.1$ . Third, we used backwards stepwise selection with the same exclusion  $p$ -value for models in which dipstick results were added to the previously selected symptoms and signs, to give models including symptoms, signs and dipstick results. For each model, we quantified diagnostic accuracy as the AUROC and compared this with the 'clinician-diagnosis' AUROC. We internally validated the models using the bootstrap procedure. As these coefficient-based models require relatively complex computation to estimate UTI probabilities, we also developed points-based models, the results of which are presented in the main report.

### Health economic methods

We developed decision-analytic models using decision trees and Markov models to identify the optimal urine sampling strategy. We developed a 'clean-catch' model and a 'nappy pad' model to reflect the different symptoms and signs predictive of UTI in older and younger pre-school children and the different diagnostic accuracy of the two urine collection methods. The models synthesised data from the DUTY study and the wider literature to estimate the lifetime costs and health outcomes. We compared six urine sampling risk stratification strategies: three derived from the DUTY risk score reflecting high specificity (DUTY5%), intermediate (DUTY10%) and high sensitivity (DUTY20%) thresholds, one based on 'clinical judgement' and two boundary strategies (sample none, sample all). The model comprised three parts: short term (diagnosis and acute illness; up to 21 days), medium term (recurrent UTI; up to 3 years) and

long term (long-term sequelae; lifetime). Costs were estimated from a NHS perspective and included diagnostic costs and short- and long-term treatment costs. Health outcomes were expressed using quality-adjusted life-years (QALYs) and quality-adjusted life-days (QALDs). Net benefit statistics were used to compare the cost-effectiveness of urine sampling and testing strategies.

### **Contamination comparison methods**

We selected a research laboratory contamination definition for its specificity (and previously termed 'frankly contaminated'), that is > 2 organisms all grown at  $\geq 10^5$  CFU/ml. Univariable associations with contamination were estimated using logistic regression, grouped by clinical and dipstick variables and stratified by urine collection method. We selected variables with  $p$ -values < 0.01 for multivariable modelling.

## **Results: general**

Between April 2010 and April 2012, 516 clinicians from 233 primary care sites screened 14,724 children for eligibility. Of these, 4390 were ineligible; 1276 declined participation; 1684 could not be recruited; 196 were subsequently excluded; and 15 withdrew. This left a recruited sample of 7163 children, of whom 50% were female; 49% were < 2 years old; 82% were white; and 26% had parents who reported education to degree level. A total of 6390 (89%) children provided a urine sample, 6241 (87%) by clean catch or nappy pad. Culture results were available from the NHS, research and both laboratories, respectively, for 5945 (83%), 5017 (70%) and 4828 (67%).

### **Results: comparison of NHS and research laboratory reliability and accuracy**

This was conducted in the 4808 children with culture results from both laboratories originating from clean-catch or nappy pad samples and index test data. NHS laboratories reported UTI positivity in 6.6% (< 3 years) and 3.2% ( $\geq 3$  years). The research laboratory reported positivity in 1.8% and 1.9% for the same age groups. Overall agreement [95% confidence interval (CI)] between the NHS and research laboratories was moderate ( $\kappa = 0.36$ ; 95% CI 0.29 to 0.43). Agreement was better for clean-catch samples (0.54; 95% CI 0.45 to 0.63) than for nappy pads (0.20; 95% CI 0.12 to 0.28). For clean-catch samples, agreement was similar in children aged  $\geq 3$  years (0.55; 95% CI 0.43 to 0.67) and < 3 years (0.52; 95% CI 0.37 to 0.67), which was better than for nappy pad samples in children aged < 3 years (0.20; 95% CI 0.12 to 0.28). Similar patterns were seen when comparisons were further stratified into age groups < 2 and  $\geq 2$  to < 3 years, suggesting that the lower reliability was attributable to nappy pad samples rather than the child's age. The AUROC (95% CI) for the six pre-specified symptoms, signs and dipstick test findings in clean-catch samples for the research and NHS laboratories were 0.86 (95% CI 0.79 to 0.92) and 0.75 (95% CI 0.69 to 0.80), respectively. The corresponding AUROCs for nappy pad samples were 0.79 (95% CI 0.70 to 0.88) and 0.65 (95% CI 0.61 to 0.70). As a result, we used the research laboratory culture result for the algorithm development reference standard and stratified algorithm development and health economic analyses by urine collection method.

### **Results: clean-catch diagnostic algorithm**

Of the 2740 children providing clean-catch urines, 2.2% met the laboratory definition of UTI, 94% were aged  $\geq 2$  years and 54% were female. 'Clinical diagnosis' correctly identified 46.6% of the culture-positive children, with 94.7% specificity and an AUROC 0.77 (95% CI 0.71 to 0.83). Four symptoms (pain/crying while passing urine, smelly urine, history of UTI and absence of severe cough) and three signs (clinician-reported global impression of illness severity, abdominal tenderness and the absence of acute ear abnormality) were independently associated with UTI. The AUROC (95% CI; bootstrap-validated AUROC) for the symptoms and signs model was 0.89 (95% CI 0.85 to 0.95; validated 0.88), increasing to 0.93 (95% CI 0.90 to 0.97; validated 0.90) with leucocytes, nitrites and blood on dipstick testing.

**Results: nappy pad diagnostic algorithm**

Of the 2277 children providing nappy pad samples, 1.3% met the laboratory definition of UTI, of whom 82% were < 2 years old and 48% were female. 'Clinical diagnosis' correctly identified 13.3% of the culture-positive children, with 98.5% specificity and AUROC 0.63 (95% CI 0.53 to 0.72). Four symptoms (parent-reported smelly urine, darker urine, female sex and the absence of a nappy rash) and two dipstick results (leucocytes and nitrites) were independently associated with UTI. The AUROC (95% CI; bootstrap-validated AUROC) for the symptom model was 0.81 (95% CI 0.72 to 0.90, validated 0.78), increasing to 0.87 (95% CI 0.80 to 0.94, validated 0.82) with the dipstick findings.

**Results: health economic analyses (clean-catch samples)**

In clean-catch samples, the 'DUTY5%' (high specificity) threshold urine sampling strategy resulted in fewer urine samples being collected than risk stratification based on clinical judgement (4.8% vs. 9.2%), and slightly higher sensitivity (58.6% vs. 56.7%) and specificity (96.1% vs. 91.4%). The high specificity threshold of the clean-catch model was both cheaper and no less effective than clinical judgement in terms of QALDs in the short term and QALYs in the long term. The absolute difference in short-term net benefits among the three DUTY risk score thresholds evaluated was very small (ranging from £1088 in 'DUTY20%' to £1090 for 'DUTY5%'). The relatively low cost of urine sampling and an antibiotic prescription, the high rate of serendipitous antibiotic prescriptions, and the low prevalence of UTI within the DUTY population all contributed to the narrow range of estimated net benefits. Our results slightly favoured conservative (i.e. high specificity) urine sampling strategies, particularly for GPs concerned about the societal impact of bacterial resistance to antibiotics.

A greater percentage of children would be treated with immediate appropriate antibiotics according to urinary bacterial susceptibility (presumptive treatment 45.2%; dipstick testing 41.8%; laboratory testing 31.2%) and fewer would have delayed antibiotics (presumptive treatment 0.9%; dipstick testing 4.5%; laboratory testing 16.6%) if treatment was presumptive or dipstick-test guided. However, average sampling, testing and treatment costs were higher for the presumptive treatment and dipstick testing strategies (£1.18) than for the laboratory testing strategy (£1.10). Short-term net benefits were similar across all three testing and treatment strategies.

**Results: health economic analyses (nappy pad samples)**

In younger children, if urine was collected using a nappy pad, the distinction in cost-effectiveness between the DUTY risk score and clinical judgement was not clear-cut. This is due to the lower diagnostic value of the DUTY risk score in younger children, the higher contamination rates necessitating repeat urine sampling and the lower accuracy of NHS laboratory results in urine collected using nappy pads raising the possibility that a correct clinical diagnosis is overturned by an incorrect laboratory test result.

Using dipstick tests to determine treatment in children at intermediate risk of UTI slightly increased initial sampling, testing and treatment costs compared with a laboratory test-based treatment strategy (£1.18 vs. £1.10 per patient); however, it increased the proportion of children with UTI treated immediately with antibiotics (41.8% vs. 31.2%). There was no difference in short-term net benefits between dipstick- (£1090) and laboratory test- (£1090) based treatment strategies.

**Results: comparison of contamination rates**

'Frankly contaminated' urine was found in 1.8% and 12.2% of clean-catch and nappy pad samples, respectively, giving a risk ratio (95% CI) of 6.66 (95% CI 4.95 to 8.96;  $p < 0.001$ ). Contamination and UTI was reported more often by the NHS than by research laboratories, especially for nappy pad samples. Probability of contamination was not increased by increasing the time taken for samples to arrive at laboratories or the presence of a nappy rash, but was increased by being female, home sampling, and increased frequency of nappy use.

## Conclusions

Agreement of microbiological UTI diagnosis in routine NHS laboratories and a research laboratory was lower than expected and worse for nappy pads than for clean-catch samples. Accuracy was lower for NHS laboratory than research laboratories and for nappy pad than for clean-catch samples. Algorithms provided better diagnostic accuracy than 'clinical diagnosis' in identifying the children in whom urine collection was warranted and diagnostic accuracy was greater for clean-catch than nappy pad samples. Diagnosis and treatment based on a clean-catch coefficient model was more cost-effective than clinical judgement and although dipstick testing provided additional diagnostic utility, its benefit was offset by dipstick test costs.

### *Implications for health care*

Primary care clinicians should prioritise the use of clean-catch sampling wherever possible. Parent-reported symptoms and clinical examination signs can be efficiently used to identify children who warrant clean-catch urine sampling. Dipstick testing has additional diagnostic value in deciding which children should receive antibiotic treatment, albeit at a higher cost than awaiting the laboratory result. NHS laboratories may wish to adopt the processing and reporting methods used by the research laboratory for paediatric urine samples and update national procedures accordingly.

### *Future research*

Further research is needed to distinguish pathogens from contaminants when multiple urinary bacteria are found in significant concentrations. The impact of using the DUTY clean-catch coefficient algorithm on clinical behaviour and patient outcome in routine clinical practice, and the cost-effectiveness of presumptive versus dipstick versus laboratory-guided antibiotic treatment, should be assessed in randomised trials.

## Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.





# Chapter 1 Introduction

Acute illness in young children is one of the most common reasons for consulting health care worldwide, and urinary tract infection (UTI) is an important cause of serious bacterial illness in children.<sup>1</sup> The National Institute for Health and Care Excellence (NICE) published a guideline on UTI in children in 2007.<sup>2</sup> This guideline emphasised the importance of prompt, microbiologically confirmed diagnosis and treatment of children, particularly in primary care where there is evidence that UTIs are missed. It also recommended a large prospective study to provide the diagnostic evidence needed to help primary care clinicians improve their recognition of children with UTI.<sup>2</sup>

The National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme issued a commissioning brief (see *Appendix 1*) for such a study in 2008. Our four-centre consortium, led by the universities of Bristol and Cardiff, proposed the **D**agnosis of **U**rinary **T**ract infection in **Y**oung children (DUTY) study with the aim of deriving and validating an algorithm for the diagnosis of UTI in children < 5 years old presenting to primary care with any acute (and largely undifferentiated) illness. We proposed a two-step algorithm, reflecting the two-step clinical process in the diagnosis of UTI: first to use symptoms and signs to help clinicians efficiently identify which young children should have their urine tested; and second to determine the added value of dipstick testing for determining which children warrant immediate antibiotic treatment.

We were particularly well placed to understand the challenges of conducting this study, with all the complexities of urine collection from acutely unwell children in the context of busy, UK primary care centres, as the Cardiff group, led by Dr Kathryn O'Brien, had been conducting a smaller study of similar design called 'EURICA' (The Epidemiology of URinary tract Infection in Children presenting with Acute illness in primary care) to establish the prevalence of UTI in pre-school children.

In this introductory chapter, we summarise the background leading to this study and describe the aims and objectives of the DUTY study.

## Prevalence of urinary tract infection

It is important to know how often UTI is the cause of acute illness in children presenting in primary care in the UK as this will inform urine sampling strategies and influence levels of clinical suspicion among general practitioners (GPs).

There are wide variations in the reported rates of UTI in children depending on setting, inclusion criteria and microbiological laboratory test criteria.<sup>3</sup> Most studies report the rate of UTI as determined from laboratory samples which have been requested by clinicians *who suspect UTI to be present*. We cannot rely on urine sampling based on clinician suspicion to determine an accurate prevalence of UTI, as children with non-specific symptoms will be excluded from urine sampling.

Shaikh *et al.* published a review of the prevalence of UTI in children in 2008.<sup>4</sup> They included 14 studies of children < 2 years old and found a pooled prevalence of 7.0% [95% confidence interval (CI) 5.5% to 8.4%]. Urine was collected using suprapubic aspiration (SPA), catheters or clean catch in the majority of studies. None used nappy pads and studies were excluded if > 25% of subjects with UTI had urine collected using bags. The review found a high degree of heterogeneity and a range in prevalence from 3.3% to 13.8%. The review included studies which had not systematically sampled urine from children, and two of the largest studies had included urine samples requested on the basis of clinician suspicion.<sup>5,6</sup>

A more recent review (2012),<sup>7</sup> of 21 studies, which included only studies where urine was systematically sampled (urine sampled from consecutive children from the study population rather than urine sampled according to clinician suspicion of UTI), found similar pooled prevalences to Shaikh *et al.* for UTI [7% for children < 3 months old (5.9% including only studies with more stringent UTI definitions); 8% for children up to 5 years old]. There was significant heterogeneity between studies. Furthermore, almost all of the studies were from the USA and included populations with very different ethnic makeups and levels of circumcision than the UK, both of which factors have been associated with UTI.<sup>3,8</sup> Most studies were from paediatric emergency departments and all excluded children without a fever (where fever was usually defined as > 38 °C). Therefore, these findings may not be generalisable to UK primary care. The EURICA study, conducted in UK general practice ( $n = 1003$ ), with systematic urine sampling, found that 5.9% of children < 5 years presenting with an acute undifferentiated illness had UTI.<sup>9</sup>

## Importance of diagnosis

The accurate and timely diagnosis of UTI is important because appropriate treatment may alleviate short-term suffering and help to prevent longer-term adverse consequences such as renal scarring, impaired renal growth, recurrent pyelonephritis, impaired glomerular filtration, hypertension, end-stage renal disease (ESRD) and pre-eclampsia.<sup>10,11</sup> Some recommendations advise prompt antibiotic treatment for symptoms suggestive of UTI in young children to prevent renal scarring.<sup>2</sup>

That said, some childhood UTI may be self-limiting. Although there are no randomised placebo controlled trials of UTI in children, there is evidence in adults, and some indirect evidence in children, that some UTIs are self-limiting.<sup>1,12,13</sup> However, it is not clear which children have a self-limiting UTI, and, when they recover clinically, if they are still at risk of renal scarring and long-term complications. We do know that in half of adult women, bacteriuria persists following a symptomatic UTI if this is left untreated, even if their clinical symptoms have improved,<sup>14</sup> and an experimental study in pigs found that renal scarring could occur even after symptomatic recovery.<sup>15</sup> In contrast to the guidelines for other self-limiting infections, those for UTI in children advocate its prompt microbiological diagnosis and treatment due to the association with long-term complications.<sup>2</sup>

## Asymptomatic bacteriuria

Asymptomatic bacteriuria is the growth of significant bacteria [ $\geq 10^5$  colony-forming units (CFU) per ml on culture of urine] in a patient with no symptoms. Although the population of interest for the DUTY study are acutely ill, and therefore not asymptomatic, it is possible that some of the children identified as having UTI due to a positive culture result could have asymptomatic bacteriuria with another coincidental illness. It is impossible to distinguish this case from a child with UTI because the presenting symptoms of UTI are often thought to be non-specific.

The significance and treatment of asymptomatic bacteriuria in children remains controversial. Guidelines recommend that asymptomatic bacteriuria should not be treated in children.<sup>2</sup> However, this advice was based on a review of four studies, none of which included children < 4 years old.<sup>2,16-19</sup> A recent Cochrane review concluded that there were insufficient data to form reliable conclusions about the harms and benefits of treating covert bacteriuria in children.<sup>20</sup> A review in 1990 concluded that neonates and preschool children with asymptomatic bacteriuria *should* be treated.<sup>21</sup> We found only one study which followed up infants with asymptomatic bacteriuria for 6 years. None of the nine girls and 27 boys had renal damage on follow-up urography, although some developed pyelonephritis.<sup>22</sup> Numbers were small and some of the infants had received antibiotics for respiratory tract infections.

Several studies (most from the 1970s) have reported the prevalence of asymptomatic bacteriuria in children.<sup>23–31</sup> Rates range from 0% to 1.8% for children < 5 years old, depending on sex and age.<sup>7</sup> NICE points out that children found to have asymptomatic bacteriuria on screening will include ‘those with no discernible history of UTI, some with a previous history of UTI, and some who have had symptomatic UTIs but have not been diagnosed’.<sup>2</sup> The authors of a Cochrane review of interventions for covert bacteriuria in children comment that some children identified with asymptomatic bacteriuria subsequently become symptomatic.<sup>20</sup>

We considered these issues for the design of the DUTY study and concluded that we should only recruit children with constitutional or urinary symptoms associated with their acute illness, such that all children found to have significant bacteriuria with a uropathogenic organism would be considered to have a UTI.

### Longer-term adverse consequences

There is evidence that UTI can lead to renal scarring.<sup>2,32,33</sup> A systematic review of the risk of renal scarring following first childhood UTI included 33 studies with a total of 4891 children.<sup>33</sup> The authors found that 57% had evidence of acute pyelonephritis (defect on early scans; based on 29 studies) and 18% had evidence of renal scarring (persistent defect on follow-up scans; based on 14 studies but this dropped to 15% when only the nine most recent studies were considered). However, there was significant heterogeneity between studies and children with UTI were not necessarily identified systematically. Although this systematic review represents the best available evidence, the finding that renal scarring occurs in 15–18% of children with UTI seems very high and may not be generalisable to a primary care population of children with UTI, if they had been identified through systematic urine sampling.

The risk of renal scarring following UTI seems to be greater in younger children, and renal scarring is uncommon over the age of 4 years.<sup>2,32,34,35</sup> It was previously thought that vesicoureteral reflux (VUR) had to be present for renal scarring to occur, but it is now accepted that renal scarring can occur, without VUR.<sup>2,33,36,37</sup>

It is thought that renal scarring can be prevented if UTI is treated promptly with antibiotics,<sup>2</sup> and there is some evidence that a delay in the treatment of acute UTI is more likely to result in renal scarring.<sup>38–44</sup>

Renal scarring has been associated with long-term complications including renal failure [end-stage renal failure (ESRF)], hypertension and pre-eclampsia.<sup>2,45–47</sup> These are serious, chronic conditions responsible for significant morbidity and costs to the NHS. However, the evidence is weak and bias and/or confounding could be responsible for the observed associations.<sup>47</sup> A recent paper estimated the risk of ESRF following childhood UTI to be 0.1% but the authors suggest that this could be an underestimate.<sup>46</sup> The NICE guideline concludes that ‘there are no appropriate studies that accurately estimate the risks of long-term complications as a result of childhood UTI’, highlighting the need for a long-term cohort study.<sup>2</sup>

### Missed diagnosis

It is thought that many cases of UTI are currently being missed in primary care.<sup>2</sup> One UK-based randomised controlled trial (RCT), in which the intervention was a nurse-led clinic to facilitate urine collection and diagnosis, found a usual-care-arm diagnosis rate half that of the intervention group.<sup>48</sup> This study suggested that even more UTIs are missed in children < 1 year old and in children without specific urinary symptoms (up to 75%).<sup>48</sup>

## Frequency of paediatric urine sampling in primary care

The main barrier to diagnosing UTI is failure to request or obtain a urine sample. GPs may not request a urine sample if they do not suspect UTI, perhaps owing to non-specific symptoms or signs, or perhaps because they do not believe that UTI is sufficiently prevalent or likely in that child. Even if a urine sample is requested, it may not be obtained due to the practical difficulties of obtaining a urine sample from a young child.<sup>49</sup>

An estimate of how often GPs obtain urine samples from acutely ill children consulting can be calculated from several studies published prior to the NICE guidelines.<sup>2</sup> Jadresic found that two urine samples per 100 registered children (< 2 years old) per year were sent to laboratories by GPs.<sup>50</sup> We know that children < 5 years old consult on average six times per year and that 87% of consultations are for acute illness.<sup>51-53</sup> This equates to urine being sampled in 0.4% of illness consultations in children < 2 years old. Another study in Wales gave a similar estimate, of 0.6% of illness consultations involving a urine sample.<sup>54</sup>

The publication of the NICE guideline in 2007 may have raised the level of suspicion of UTI and urine sampling from acutely ill children. We estimated current levels of urine sampling from consultations with acutely ill children in Wales from Public Health Wales data. In 2012, 12,689 urine samples were received by microbiology laboratories from general practices in Wales for children < 5 years old (Dr Robin Howe, Consultant Microbiologist, Public Health Wales, Cardiff, 11 June 2013, personal communication). We do not know how many children were registered with practices in 2012, but from the Office for National Statistics (ONS) 2011 census there were 178,000 children < 5 years old living in Wales in 2011.<sup>55</sup> Assuming that all children were registered with practices; that numbers for 2012 were similar to 2011; that urine sampling reflected normal practice (some practices were participating in the DUTY study); that children consulted six times per year; and that 87% of consultations were for acute illness gives an estimate of urine sampled in 1.4% of acute illness consultations in children < 5 years old in Wales in 2012.

Unless GPs can target urine sampling extremely accurately, it is unlikely that such low levels of urine sampling will allow detection of the majority of UTIs.

## Methods of urine sampling in primary care

There are five main methods of obtaining urine samples from children: SPA, catheter insertion, clean catch, nappy pad and bag collection. Owing to concerns about their invasive nature and the restrictions of time and space in the UK primary care setting, SPA and catheters are not recommended for use in UK primary care.<sup>2</sup> NICE suggest that samples should be collected using a method suitable for the age of the infant or child.<sup>2</sup> 'Clean-catch' samples are preferred, but urine collection pads are suggested if this is not possible.

Our HTA systematic review published in 2006 reviewed methods of urine sampling and included four studies in children < 5 years.<sup>56</sup> We have updated this systematic review for the DUTY study (see *Appendix 2*).

We found a further two primary studies, giving a total of six studies, that assessed urine sampling methods.<sup>57-62</sup> Two studies compared culture of urine bag specimens with culture of SPA samples.<sup>57,59</sup> One reported a sensitivity of 100% and the other of 50%; both studies found specificity to be around 90%. Two studies compared culture results from urine samples obtained by bag specimens with those obtained by catheter.<sup>58,61</sup> The appropriateness of a catheter specimen as the reference standard is questionable, meaning that these results are of limited value. One study compared culture of a nappy pad specimen with culture of SPA samples. This study reported a sensitivity of 100% and specificity of 94%, suggesting excellent agreement between the two sampling methods.<sup>60</sup>

A recently published study assessed a device known as the 'U-test', which is a nappy pad incorporating a urine dipstick.<sup>62</sup> The accuracy results are, therefore, a combination of the nappy pad and the dipstick but show good accuracy with a sensitivity of 100% and specificity of 79%. However, these were compared with a reference standard consisting of a variety of urine collection methods (clean catch, bag, catheter or SPA) and the study had results available for only 25 participants.

The NICE guidelines found 'insufficient data to draw conclusions about urine collection bags and urine collection pads' but advised that either is acceptable for UK primary care.<sup>2</sup> There have been only two studies published since these guidelines and these do not provide sufficiently strong data to change these conclusions, although the limited data suggest that pad specimens may be a more accurate method of urine collection than bag specimens. Furthermore, the pad sampling method has been shown to be more acceptable to parents than the bag method due to the problems of the bag adhesive irritating the child's skin.<sup>63</sup> We prioritised clean-catch urine sampling where this was possible, and the use of nappy pad sampling where it was not, for the DUTY study.

### Which children should have their urine sampled?

When faced with an acutely ill child, the clinician has several decisions to make to diagnose possible UTI:

1. Should a urine sample be obtained from the child?
2. If so, should the urine sample be tested with a dipstick?
3. Should the urine sample be sent to the laboratory for culture?
4. Should antibiotics be prescribed before the culture result is available?
5. Should antibiotics be prescribed once the culture result is available?

Identifying which pre-school children should be sampled in primary care is challenging, particularly in those < 2 years, because most are pre-verbal, symptoms and signs are usually non-specific, and obtaining uncontaminated samples is difficult.<sup>2,49,64</sup> NICE suggests that clinicians should test for UTI in children < 5 years with unexplained fever, vomiting, lethargy, irritability, poor feeding, abdominal pain, offensive urine, haematuria, frequency or dysuria.<sup>2</sup> However, they acknowledge a lack of evidence to support the diagnostic utility of these symptoms and signs, and uncertainty regarding the role of dipstick testing.<sup>64</sup>

### Existing evidence for the diagnostic value of symptoms and signs

We reviewed the evidence for the predictive values of clinical symptoms and signs for UTI in children (see updated systematic review in *Appendix 2*). We identified five primary studies (Craig *et al.*,<sup>1</sup> Gorelick and Shaw,<sup>65</sup> Gorelick *et al.*,<sup>66</sup> Gauthier *et al.*<sup>67</sup> and O'Brien *et al.*<sup>9</sup>) ( $n = 17,793$ ) and one systematic review (eight primary studies)<sup>3</sup> in children aged < 5 years old ( $n = 7892$ ) that assessed clinical symptoms and signs. These were conducted mainly in hospital emergency departments; only one was conducted in general practice.<sup>9</sup> No individual or any combination of symptom(s) or sign(s) were sufficient to rule in a diagnosis of UTI, although some post-test probabilities (e.g. 25% for increased capillary refill time, no fluid intake and suprapubic tenderness) appear high enough to mandate urine testing and empirical treatment while awaiting culture confirmation.

The largest study, which included almost 16,000 children aged < 5 years presenting to the emergency department,<sup>1</sup> derived a clinical prediction rule based on a combination of 27 symptoms and signs. The model was found to have an area under the receiver operating characteristic curve (AUROC) of 0.80 (95% CI 0.78 to 0.82). However, UTI was not identified through systematic urine sampling in this study, with urine cultures obtained in only 21% of children, calling into doubt the representativeness of this model to diagnose UTI.

The most representative primary care study to date (based in UK general practices with systematic urine sampling) also considered the predictive value of presenting symptoms and signs.<sup>9</sup> This study found younger age, urinary frequency and dysuria to be associated with UTI and found no association with fever or the presence of an alternative site of infection. However, the study was powered to determine UTI prevalence, not diagnostic accuracy, and this led to wide CIs around the diagnostic utility estimates.

### Other risk factors

#### Age and sex

Previous studies have suggested that UTI is more common among males up until 3 to 6 months old. For children older than 12 months, UTI is more prevalent in females.<sup>4,43</sup>

In our updated systematic review, we found that age < 3 months was a risk factor for UTI, irrespective of sex [likelihood ratio (LR) 3.9; 95% CI 3.2 to 4.8].<sup>68</sup> The same study found a decreased likelihood of UTI in children aged > 3 years (LR 0.47; 95% CI 0.37 to 0.61). This study also found that being female increased the likelihood of UTI (LR 1.3; 95% CI 1.1 to 1.3) (see *Appendix 2*).

#### Circumcision status

Our review found consistent evidence that circumcision protected against UTI. The systematic review included in our review included six studies in boys aged < 24 months that assessed the association of circumcision and UTI, and reported a pooled LR of 0.33 (95% CI 0.18 to 0.63), suggesting that the likelihood of a UTI was lower in circumcised boys.<sup>56</sup> An additional study in boys aged 0–36 months supported this finding, although the CI was wide (LR 0.07; 95% CI 0.00 to 1.16).<sup>67</sup> It is worth recognising that many of the studies concerning UTI in children have been conducted in countries with much higher levels of circumcision than the UK.<sup>7,69–71</sup> The UK rate of circumcision is approximately 3%, compared with 80% in the USA.<sup>72</sup>

#### Ethnicity

The systematic review included in our review included six studies that evaluated ethnicity in children aged < 24 months and found that non-black race increased the likelihood of UTI (LR 1.4, 95% CI 1.1 to 1.8; systematic review of six studies)<sup>56</sup> (see *Appendix 2*).

#### Past history of urinary tract infection

In our updated systematic review, we found that a prior history of UTI increased the risk of UTI, with a LR of 2.9 (95% CI 1.2 to 7.1) in children < 24 months old and a LR of 2.3 (95% CI 0.3 to 17.4) in children < 12 months old (see *Appendix 2*).

### Summary of symptoms and signs

In our review we found that none of the risk factors, individually or in combination, was sufficient to rule in a diagnosis of UTI. Some combinations of symptoms, signs and proposed clinical prediction rules did reduce the probability of UTI below 2% and may be considered low enough to rule out UTI (see *Appendix 2, Table 84*). However, the studies on which these findings are based did not necessarily systematically sample urine, and were based on populations with a different ethnic mix and rate of circumcision from the UK and, therefore, may not be generalisable to UK general practice.

In the absence of accurate predictive symptoms and signs, a broad urine sampling strategy in children is advocated by NICE.<sup>2</sup> Others have also advocated urine 'screening' in some groups of children, for example in all febrile infants or broader urine sampling strategies in ill children.<sup>70,73,74</sup>

A survey of 200 paediatricians in 1983 concerning the management of febrile infants found that all of the respondents felt that a UTI prevalence of 5% would warrant urine sampling in all; more than 80% felt that a prevalence of more than 3% would warrant urine sampling in all, and about half felt that a prevalence of between 1% and 3% would warrant sampling urine from all febrile children.<sup>74</sup>

A recent retrospective study concluded that urine analysis should be added to the NICE 'traffic light' system for the detection of serious illness in febrile children < 5 years old.<sup>73</sup> This would require a large increase in urine sampling from acutely ill children.

## Dipstick tests

Once a urine sample has been obtained, the clinician has the option of testing the urine sample with a urinary dipstick. The HTA-funded systematic review in 2006<sup>56</sup> found that urinary dipsticks were similar to the overall conclusions based on the updated review, that is that dipstick positive for both leucocyte esterase (LE) and nitrite is useful for ruling in a UTI and negative for both LE and nitrite is useful for ruling out a UTI.<sup>56</sup> NICE recommend only using dipsticks for diagnosis in children  $\geq 3$  years. They recommend that a urine sample should also be sent for culture in most cases (unless both LE and nitrite are negative on dipstick).<sup>2</sup> For children 3 years or older, and occasionally in younger children, NICE advise using a dipstick to help to target empirical antibiotics while waiting for urinary culture results.<sup>2</sup>

We have updated the 2006 systematic review for the DUTY study, which assessed the accuracy of dipstick testing for UTI in children (see *Appendix 2*). We found data from six primary studies<sup>57–62</sup> (including those included in the HTA review) which assessed dipstick testing for LE and nitrite in children aged < 5 years. There was substantial heterogeneity across studies. Negative LRs were too heterogeneous to permit conclusions regarding the utility of LE or nitrite negative for ruling out a diagnosis of UTI, ranging from < 0.01 to 0.88 with a pooled estimate of 0.46 (95% CI 0.18 to 1.13). Positive LRs ranged from 6 to 108 with a pooled estimate of 22.8 (95% CI 11.1 to 46.5) suggesting that a dipstick positive for both LE and nitrite may be useful for ruling in a UTI. Positive LRs for the combination of LE or nitrite positive were also extremely heterogeneous, ranging from 1.8 to 73 with a pooled estimate of 10.5 (95% CI 3.4 to 32.2), making it difficult to draw conclusions regarding the utility of this combination in ruling in a diagnosis of UTI. Negative LRs ranged from 0.16 to 0.32 with a pooled estimate of 0.22 (95% CI 0.16 to 0.30), suggesting that a dipstick negative for both nitrite and LE may be useful in ruling out a diagnosis of UTI.

Overall, the data were too heterogeneous to draw firm conclusions regarding the accuracy of dipstick testing; however, the data suggest that a dipstick positive for both nitrite and LE may be useful for ruling in a diagnosis of UTI, while a dipstick negative for both nitrite and LE may be useful for ruling out a UTI. The NICE guidelines stated that 'further investigation of LE and nitrite dipstick tests alone and in combination, stratified by age and method of urine collection, is required to determine their accuracy in diagnosing UTI.'<sup>2</sup>

## Microbiological diagnosis of urinary tract infection in the laboratory

### *Bacteriuria threshold*

Laboratory diagnosis of UTI is based on colony counts following culture, which reflect the concentration of organisms in urine and hence the likelihood that the bacteria grown arise from a UTI rather than contamination. The standard threshold for diagnosing UTI from culture of  $\geq 10^5$  CFU/ml was established by Kass over 50 years ago.<sup>75</sup> This was based on studies of adult women with acute pyelonephritis and asymptomatic women. Some have questioned if this threshold is appropriate for children,<sup>76,77</sup> with most suggesting a lower threshold<sup>76–79</sup> but one advocating an increase.<sup>80</sup> UTI is typically thought to be caused by a single organism present in a high concentration, usually  $\geq 10^5$  CFU/ml.<sup>81</sup> However, laboratory guidelines differ regarding the urine sampling method, nature and extent of bacterial growth required to confirm UTI.<sup>82,83</sup> For a clean-catch specimen in children,  $> 10^3$  CFU/ml of a single species 'may be diagnostic of UTI'; and a pure growth of between  $10^4$ – $10^5$  CFU/ml is 'indicative of UTI'.<sup>82</sup> Although usually a pure or predominant growth is required for the diagnosis of UTI, the growth of two organisms, each with a growth of  $\geq 10^4$  CFU/ml, would also be considered as positive by these guidelines.<sup>82</sup> NICE guidelines do

not provide a definitive threshold for diagnosing UTI on culture but provide advice about the level of bacterial growth in relation to symptoms and signs.<sup>2</sup> Although NHS laboratories in the UK follow the UK Standards for Microbiological Investigation for the examination of urine, application of the method varies between laboratories.

Some secondary care studies have required two consecutive urine samples with significant bacteriuria to diagnose UTI.<sup>84,85</sup> It has been suggested that obtaining two samples from children in primary care would reduce the risk of false-positive results;<sup>80</sup> however, it is unlikely that this would prove successful in primary care given the challenges and current low levels of urine sampling. It is also unclear if two samples would improve validity as there may be a greater number of false-negative results with this technique. Current guidelines advocate one sample.<sup>2</sup>

Chapter 4 reports the DUTY study analyses used to determine the reference standard and bacteriuria threshold to be used for the development of the DUTY study diagnostic algorithm.

### Culture techniques and uropathogenic organisms

Standard techniques for culturing urine vary but usually rely on either cystine-lactose-electrolyte-deficient (CLED) agar or, more recently, chromogenic agar.<sup>83</sup> No single medium is likely to support the growth of (and detection of) all possibly significant organisms; however, most organisms commonly associated with UTI are supported. Organisms less commonly associated with UTI may not grow at all or may grow at an insufficient rate to be detected, for example *Haemophilus influenzae*, *Streptococcus pneumoniae* and the coagulase-negative staphylococci.<sup>86</sup> There is no definitive list of uropathogenic organisms and the distinction between uropathogenic and non-uropathogenic is not always clear.<sup>82</sup> *Staphylococcus saprophyticus* is commonly associated with UTI in females and growth is supported on the normal media used. Most studies report that most UTI is caused by *Escherichia coli*, both in adult and paediatric populations.<sup>14</sup> Enterobacteriaceae are usually considered to be uropathogens.<sup>30,34–36,82,87</sup> Non-enterobacteriaceae such as enterococci (Lancefield *Streptococcus* group B), coagulase-negative staphylococci, *Staphylococcus aureus*, *S. saprophyticus* and pseudomonads are also considered to be potentially significant isolates.<sup>88</sup> Colonies are counted following culture (for 'between 18–24 hours'). However, reporting varies from laboratory to laboratory and often depends on the clinical information given on the clinical request form.

### Contamination

It is recognised that difficulty in specimen collection and the interpretation of specimens potentially contaminated prior to culture, from skin, faeces and other sources may contribute to the misdiagnosis of UTI.<sup>89–98</sup> Overdiagnosing UTI can lead to unnecessary investigations and treatment, which entail risks of complications and psychological stress to the child and family. Discriminating between contamination with faecal organisms and potential UTI pathogens in the laboratory is difficult as the most common faecal organism and the most common pathogen causing UTI is *E. coli*, and high colony counts may be found in children without UTI.

Contamination rates in collection methods vary, as indeed do the definitions of contamination. Contamination rates range from 0% in clean catch to 48% of bag urines.<sup>98</sup> In a retrospective observational cohort study, contamination in clean catch, catheter specimen of urine and bag was 1%, 12% and 26%, respectively.<sup>98</sup> Definitions of contamination vary from single organism growth < 10<sup>5</sup> CFU/ml OR ≥ 2 organisms to ≥ 2 organisms present at ≥ 10<sup>5</sup> CFU/ml of urine.<sup>89,91</sup>

In a US study, no institutional factors such as access to refrigeration were found to be associated with either low or high contamination rates.<sup>99</sup> Gender and diarrhoeal symptoms also had no association with higher contamination rates in children.<sup>100</sup> Perineal cleansing in female adults had no association with contamination rates, while urine contamination rates were higher in midstream urine collected from toilet-trained children when obtained without perineal/genital cleaning.<sup>89,94</sup> One of the only factors affecting contamination rates which has been published is changing nappy pads every 30 minutes.<sup>91</sup>



Although contamination is generally considered to increase the probability of false-positive results, there is some suggestion that contaminated samples or samples with mixed growth may hide a true UTI, that is lead to false-negative results.<sup>101,102</sup>

## Economic considerations

Urinary tract infection is the fourth most common reason for prescribing antibiotics, accounting for approximately 8% of all antibacterial prescriptions.<sup>103</sup> While the unit costs of laboratory testing and antibiotic prescribing are relatively low,<sup>56</sup> the economic implications of new clinical algorithms for urine sampling and testing may be substantial owing to (1) the large numbers of children who present with non-specific symptoms which might be caused by UTI; (2) the cost of subsequent diagnostic tests used to further evaluate children with recurrent/atypical UTI;<sup>2</sup> (3) the substantial costs and impact on quality of life of a missed diagnosis that leads to rare but serious complications of UTI; and (4) the wider, long-term population impact of diagnostic algorithms on antibiotic prescribing and bacterial resistance.<sup>104</sup>

The few economic evaluations of the diagnosis of UTI in young children have primarily aimed to identify the most cost-effective test or series of tests for diagnosing UTI, rather than address the important issue of exactly which children should be selected for urine sampling and testing in the first place.<sup>2,5,6</sup> There is limited economic evidence on which children should have a urine sample taken, by what sampling method, and which urinalysis tests should be used to guide initial treatment. Guidance is especially needed for children < 3 years of age for whom current NICE clinical guidelines are not based on evidence of cost-effectiveness.<sup>3</sup>

## Summary of DUTY study design

### Summary of the research brief

The background given to the NIHR HTA research brief for improving the recognition of UTI in children in primary care (see *Appendix 1*) was summarised indicating the nature of the problem, that young children with UTI may present with non-specific symptoms such as poor feeding, vomiting, irritability, jaundice (in newborns) or fever alone, suggesting that a broader approach to testing may be appropriate. The research question that the commissioning brief proposed should be answered was: 'Which clinical features of potential infection are useful in making a preliminary diagnosis of UTI in children < 2 years of age and indicate the need for a urine specimen to be taken?'

### Summary of the justification for the DUTY study design

DUTY was a diagnostic cohort study, set up to systematically sample urine from acutely unwell children, with constitutional and/or urinary symptoms, in UK primary care in order to determine the diagnostic value of symptoms and signs for UTI. We proposed to develop and validate an algorithm in two steps: first, to use symptoms and signs to help clinicians to efficiently identify which young children should have their urine tested; and second, to determine the added value of dipstick testing. We included a health economic evaluation to determine the cost-effectiveness of the clinical algorithm compared with existing practice.

We recognised that the age at which children can verbalise their symptoms and achieve bladder control varies between children and that no single age < 5 years would adequately reflect this. We also knew that collecting urine samples from children under the age of 2 years would be challenging and that information from older pre-school children might still be valuable for younger pre-school children. We therefore proposed a study that recruited children up to but not including the age of 5 years.

## Aim of the DUTY study

The aim of the DUTY study was to derive and validate an algorithm to improve the recognition of UTI in children < 5 years of age presenting to primary care with an acute illness. It was envisaged that the algorithm would be constructed to address two questions: (1) which children are at sufficient risk to warrant urine sampling; and (2) to determine the added value of point-of-care urine dipstick testing. The findings from these two stages would be combined to produce the overall algorithm.

### Research objectives

The DUTY study's 'Detailed Project Description' (the final protocol submitted with the final funding application) contained the following five research objectives:

1. To develop an algorithm that accurately identifies children presenting in primary care with an acute illness in whom a urine sample should be obtained, based on socio-demographic factors, medical history, symptoms and signs (see *Chapter 5*).
2. To assess whether dipstick urinalysis for nitrite, LE, protein, blood and glucose gives additional diagnostic information to objective (1) in the identification (ID) of urine samples that should be sent to the laboratory (see *Chapter 5*).
3. To model the cost-effectiveness (cost per correct diagnosis of UTI, cost per symptomatic day avoided and lifetime cost per quality-adjusted life-year (QALY), including potential long-term complications of UTI) from NHS and societal perspectives of one or more diagnostic algorithm guided strategies (see *Chapter 6*).
4. To compare contamination rates for nappy pad versus clean-catch urine sampling methods (see *Chapter 7*).
5. To explore the significance of two categories of positive urine culture (laboratory diagnosed UTI vs. asymptomatic bacteriuria/contamination) by (a) comparing the number of NHS contacts at 3 months in children with positive urine culture stratified at presentation by the responsible clinician's assessment of possible UTI versus plausible alternative diagnosis, and (b) investigating for persisting bacteriuria in a second urine sample.

## Changes to the original funding proposal

### Planned change to age inclusion criterion

In response to reviewers' comments (suggesting eligibility to < 4 years) to our detailed project description, and in communication with the HTA, we decided to increase eligibility to < 5 years. This was a pragmatic decision informed by early findings from the EURICA study (a similar study being conducted by the Cardiff members of the DUTY group) that showed higher recruitment, urine return rates and UTI prevalence in the  $\geq 4$  and < 5 years group and a judgement by the DUTY investigators that children  $\geq 4$  and < 5 years were often similar, with regard to development of language and bladder training skills, to children aged  $\geq 3$  and < 4 years.

### Planned change to research objective 5

The original funding proposal submitted to the HTA on 12 February 2009 contained five research objectives. The outcomes of the first four objectives are reported in this publication but the fifth was changed prior to publication of the final protocol<sup>105</sup> (see *Appendix 3*). This was because the study eligibility criteria meant all children were unwell (or had urinary symptoms) thus making it difficult to determine asymptomatic bacteriuria as no children were asymptomatic. We therefore removed this as an objective and, instead, have undertaken exploratory analysis on the children who had positive or contaminated urine, investigating the impact of their clinician's working diagnosis on symptom duration and number of NHS contacts at 43 months. These latter findings are reported in *Chapter 7*. We also decided that taking a second urine sample in this population would have been too large a logistic undertaking and that persistent bacteriuria was not the focus of the study.

### ***Unanticipated changes to the validation process and reference standard***

Our original intention had been to derive and externally validate the diagnostic algorithm, using an approximate 66/33 data split for derivation/validation and to analyse the clean-catch and nappy pad samples together. However, analyses reported in *Chapter 4* suggested that the urine collection method was having a greater than anticipated effect on laboratory culture results. Therefore, and in consultation with a number of international experts, we decided to stratify algorithm development by the clean-catch and nappy pad collection method (approximately a 50/50 division of data) and to use bootstrapping to validate the algorithms. This decision was made for two reasons. First, the smaller than anticipated numbers of outcome events in the stratified analyses meant that analyses restricted to a development sample would have been underpowered and estimated coefficients imprecise. Second, our statistical advice was that the properties of a random split into development and validation samples could be mimicked by the bootstrap procedure that we adopted, and statistical overoptimism quantified using calibration slopes.

As it would better reflect day-to-day clinical practice, we originally intended to use  $\geq 10^5$  CFU/ml of a known uropathogen from the NHS laboratories as our reference standard. However, as the analyses presented in *Chapter 4* show, there was greater than expected disagreement between local and research laboratories, with evidence that the research laboratory was both more reliable and more accurate than NHS laboratories. We therefore decided to use the research laboratory result as the reference standard.

### ***Other changes from the detailed project description***

#### **Sample size**

The original plan was to recruit 6000 children by end of April 2012. As this target was met early and the prevalence of UTI was lower than expected, we obtained ethical approval to continue recruiting until April 2012 even though this exceeded the original projected study requirement.

#### **Investigation of UTI prevalence**

In our protocol paper (see *Appendix 3*), we said that we would investigate factors influencing UTI prevalence. These results are reported in *Chapter 7*.



## Chapter 2 Overall study methods

### Summary of study design

DUTY was a 3-year multicentre, prospective diagnostic cohort study which aimed to recruit at least 6000 children aged before their fifth birthday, being assessed in primary care for an acute undifferentiated illness. Children were invited to participate when they presented to primary care acutely unwell (with non-traumatic aetiology) of  $\leq 28$  days' duration. Urine samples were obtained from as many eligible, consented children as possible, and data were collected on medical history and presenting symptoms and signs. Urine samples were dipstick tested in primary care and sent for standard microbiological analysis, microscopy and culture at the primary care site's usual local NHS microbiology laboratory, from here on referred to as the 'local laboratory'.

In addition, and where sufficient urine volumes were available, a fraction of the urine sample was decanted and sent for parallel microbiological analysis at the study's designated research laboratory in Cardiff [the Specialist Antimicrobial Chemotherapy Unit (SACU), Public Health Wales Microbiology Cardiff, University Hospital Wales], from here on referred to as the 'research laboratory'. All children with culture positive urines and a random sample of children with urine culture results in other, non-positive categories were followed up by telephone to record symptom duration and health-care resource use at 14 days from recruitment, and through review of the child's primary care medical notes at 3 months.

The primary outcome was a validated diagnostic algorithm using a reference standard derived from research laboratory results (see *Chapter 4*) of  $\geq 10^5$  CFU/ml of a single uropathogen ('pure growth') or  $\geq 10^5$  CFU/ml of a uropathogen with  $\geq 3 \log_{10}$  (1000-fold) difference between the growth of this and the next species ('predominant growth'). We defined uropathogens as members of the Enterobacteriaceae group. We used logistic regression to identify the clinical factors (i.e. demographic, medical history, presenting symptoms and signs and urine dipstick analysis results) most strongly associated with a positive urine culture result to create a rule for use in clinical practice. An economic evaluation was conducted to compare the cost-effectiveness of the candidate prediction rules from the perspectives of families and the NHS.

The study was sponsored by the University of Bristol and run jointly by Bristol and Cardiff University in collaboration with centres at Southampton University and King's College London. The study is summarised in *Figure 1*.

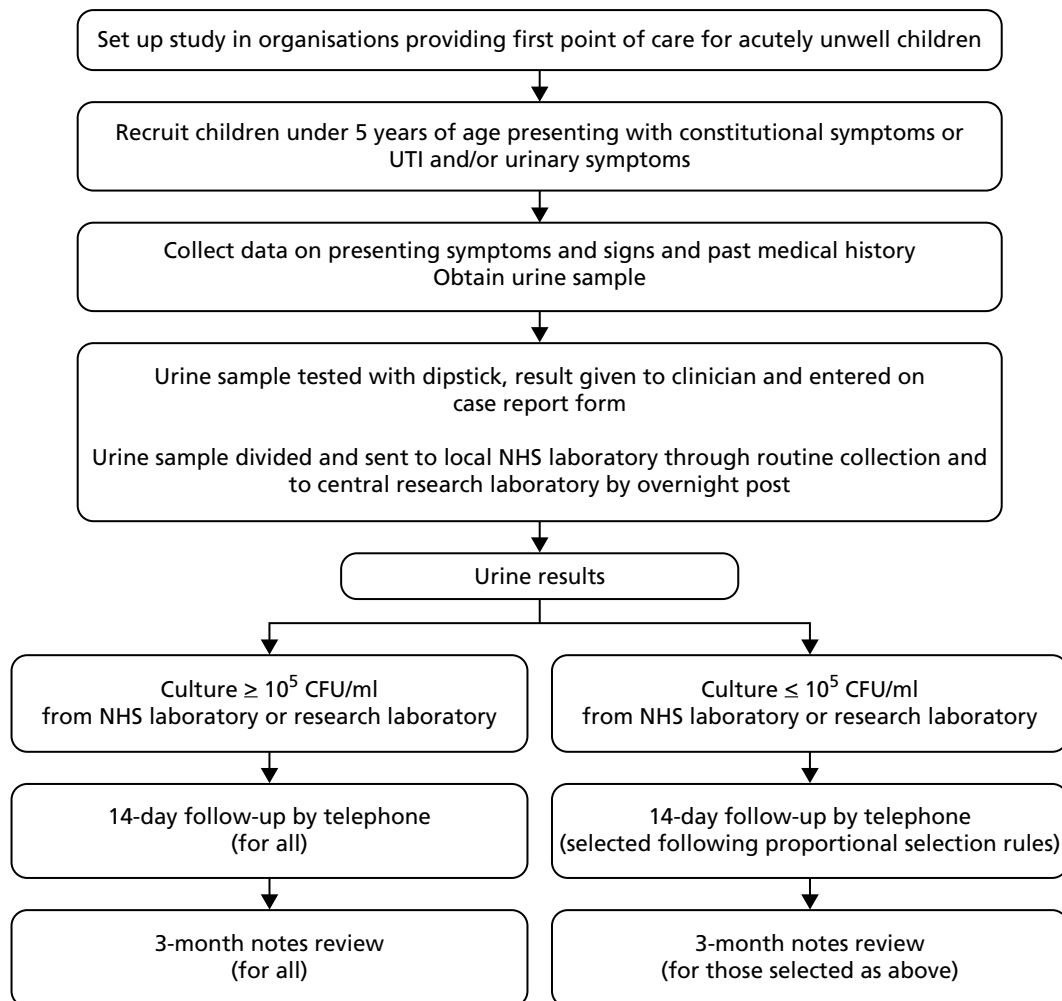


FIGURE 1 DUTY scheme (see Figure 4 for numbers recruited).

## Ethics and research and development approvals

Ethical approval for this multicentre study was given by a National Health Service (NHS) research ethics committee. The initial approval was granted by the National Research Ethics Service Southmead Research Ethics Committee on 9 December 2010, reference #09/H0102/64. NHS research and development (R&D) approval was also gained where the research was undertaken.

## Parental contributions to the study

The study protocol, study documentation and procedures benefited significantly from feedback from parent groups and from continued engagement with parents both as formal advisors and as study participants. During study start-up, the parent members of the NIHR South West Medicines for Children Research Network were supplemented by additional parents, volunteering to serve on the Study Steering Committee and/or review and comment on study documents. This feedback informed, for example, our staffing structure (recruiting staff's competence in working with children was considered absolutely essential) and methods for raising local awareness generally about what was considered a worthwhile and important study.

Once the study was under way, all questionnaires, processes and information (all formats) were reviewed and refined according to ongoing parental feedback about participation in the study.

## Site recruitment

### *Recruitment period and locations*

The study was open to participant recruitment from 7 April 2010 until 30 April 2012. Recruitment was implemented from four research centres at the universities of Bristol, Cardiff, Southampton and King's College London. Each centre recruited children from primary care, defined as any NHS facility providing first-point-of-contact, face-to-face advice for parents of unwell children.

### *Selection of primary care sites*

The majority of UK acute paediatric care is provided by GPs in primary care practices. To ensure we recruited to target, we estimated that we required between 80 and 100 primary care sites per centre, with each centre needing to contribute around 1500 children to achieve the study sample of 6000.

General practitioner practices were supplemented by NHS walk-in centres (WICs) and children's emergency departments (CEDs). WICs manage large numbers of young children with acute, undifferentiated illnesses. For example, the South Bristol WIC estimated 2500 contacts for acute, non-traumatic, illnesses in children < 5 years old annually. Parents also use CEDs as a first point of contact for large numbers of acutely unwell children, particularly outside normal working hours. For example, > 5% of all attendances at UK CEDs are due to fever,<sup>106</sup> which represents approximately 1400 children annually in each of Bristol, Cardiff and Southampton, and more in London. The majority of these are < 5 years old.

### *Site recruitment approach*

In England and Wales, practice recruitment was supported by research staff from the primary care research network (PCRN) and the National Institute for Social Care and Health Research – Clinical Research Centre (NISCHR-CRC) Network.

General practitioner practices were invited to take part in the study either by letter or via a PCRN newsletter. Interested practices were contacted initially by e-mail to provide further information about the study and were followed up by telephone or face-to-face visit with the practice manager or lead research GP to discuss the study in more detail.

Working with the local research networks in both Wales and England provided the study team with the knowledge of which practices had experience of recruiting to studies within primary care and also highlighted practices to approach. The role of the local research networks is described below.

### *The role of the local research networks*

Local research network support, primarily the PCRN and comprehensive local research networks (CLRNs) in England and, in Wales, NISCHR-CRC, was a decisive factor in enabling this study to recruit ahead of target. Collaboration with the local research networks provided the following benefits.

### **Advocacy and advice for general practitioner practice recruitment**

Each study centre worked closely with local research networks to identify and recruit primary care sites to the study. This included attending research meetings, contributing to research network newsletters and bulletins, and working with the networks to approach appropriate sites to elicit expressions of interest.

### **Strategic advice for recruitment across the UK**

Local research networks suggested the extension of recruitment to new areas of England. For example, PCRN South West supported the extension of DUTY study recruitment into the Peninsula area (Devon and Cornwall) to develop local research capacity and increase access to primary care portfolio studies, as well as boosting DUTY recruitment. In addition, the proactive support for the study provided by the Cumbria and Lancashire CLRN was a further significant boost to recruitment figures while enhancing the generalisability of the study findings by including more patients from rural rather than urban areas.

### **Support with gaining research and development approvals**

For some local research network areas (e.g. Western, and the London networks) R&D approvals for recruitment in primary care (with the exception of CEDs in acute trusts) were granted at the level of the primary care trust (PCT). However, for some areas (e.g. Wales, Cornwall, Cumbria and Lancashire) site-specific approvals were required at the level of each GP practice. In these situations the local research network provided administrative time to support the completion of specific approvals for each of the numerous sites participating in the study.

### **Provision of research nurse resource**

The DUTY study benefited from significant additional community research nurse (RN) capacity funded by the CLRNs, PCRNs and NISCHR-CRC. Dedicated sessions of nurse time were provided to support local primary care sites that did not have sufficient recruitment capacity in-house. Clinical staff recruiting to DUTY within the three active CEDs (the Bristol Children's Hospital, the Evelina Children's Hospital and Southampton Paediatric Accident & Emergency) were employed by the acute trusts and funded through the research networks or similar organisations.

### **Provision of training for recruiting staff**

Local research networks not only provided facilities for DUTY researchers to train primary care-based recruiting staff, but offered ongoing training in good clinical practice and paediatric consent to clinicians and researchers involved in all aspects of the study.

### **Recruitment models**

Primary care sites were offered two models of recruitment. The first was known as option 1, in which the majority of the recruitment procedures were undertaken by a RN or clinical studies officer (CSO) external to the primary care site, and funded by the DUTY study or the local PCRN, CLRN or NISCHR-CRC, to work with and recruit for the primary care site's clinical team.

In option 2, recruitment was undertaken entirely 'in-house' by the primary care site's practice team. The acute environment of CED and WIC recruitment meant that, for these settings, the option 1 model was the only viable model.

### **Study staff**

The study grant provided full-time equivalent DUTY RN posts across all four study centres, which were supplemented by additional research staff (RN/CSOs) provided by local PCRNs and CLRNs (in England) and by NISCHR-CRC (in Wales).

Dedicated RNs/CSOs were available to provide external support for option 1 primary care sites, and to support autonomously recruiting option 2 sites through the provision of expert training, mentoring and problem-solving.

### **Site and staff training**

The DUTY study team from each centre would meet with individual sites to explain the study and identify their working requirements and match this to the appropriate model of recruitment in order to meet both the study's needs and those of each individual primary care site. The geographical area of the site was also taken into consideration. Rural practices were encouraged to adopt the option 2 approach, as it was impractical to use DUTY recruiters (RNs/CSOs).



DUTY recruiters were provided with study-specific training including paediatric informed consent, data confidentiality, data collection [using paper-based and electronic case report forms (CRFs)], collection of urine samples and safety reporting. Staff were encouraged to access PCRN or NISCHR-CRC's informed consent and good clinical practice training days.

Sites were also offered practice-based training adjusted to meet their needs in order to tailor recruitment methods to the individual site. Reception staff were advised on how to introduce DUTY to parents of potential participants. Practice staff were trained on the method of processing the urine sample to send to local NHS and research laboratories and clinicians were provided guidance on informed paediatric consent. All staff involved with the study were provided training on data collection and CRF completion.

### **(Local) NHS laboratory recruitment**

The participation of any primary care site in recruitment to the study depended on the support and participation of the local NHS microbiology laboratory to which the site routinely sent urine samples. In each area of recruitment, the local NHS laboratory was approached and service-level agreements put in place prior to involvement in the study. Key staff were provided with a study manual, essential documents and database instructions.

Training of staff was not required as they were only executing their normal routine procedures that were covered within the standard operating procedure (SOP) for their own laboratory. Laboratories were reimbursed with service support costs at a rate locally determined per sample processed.

### **Service support costs**

Service support costs were provided by the local CLRNs for each centre in England and NISCHR-CRC in Wales. They were provided on a 'per-patient' basis for primary care sites and on a 'per-sample' basis for NHS microbiology laboratories.

For option 1 practices, which were supported by DUTY recruiters, the 'per-patient' reimbursement excluded activities undertaken by the DUTY recruiter (e.g. consent, CRF completion, urine sample management and online data entry). For option 2 practices, where all aspects of recruitment were undertaken by site staff, all recruitment activities were reimbursed. Payments were made only for children recruited with valid informed consent and without significant deviation from the study protocol.

To emphasise the importance of obtaining the urine sample in order to achieve the study's primary outcome, service support reimbursements were linked to sites' urine sample retrieval rates. Reimbursements of 100% were dependent on sites having a local laboratory urine sample retrieval rate of at least 90%, and a research sample dispatch rate of at least 85%. Practices that did not meet both of these thresholds were reimbursed only for those children in whom a urine sample was obtained.

For NHS microbiology laboratories, service support costs were reimbursed for all urine samples processed by the laboratory. The laboratory service support costs were based on a Department of Health-agreed cost for standard microscopy and culture, plus a component of administrative time for entering the urine culture results onto the DUTY study web-based database.

## Participant selection

### Eligibility criteria

The study inclusion criteria were designed to be as broad as possible. Children were eligible if they were aged before their fifth birthday and presented to primary care with a new acute illness episode of  $\leq 28$  days' duration.

This illness needed to be associated with (1) at least one 'constitutional' symptom or sign identified by NICE<sup>2</sup> as a potential marker for UTI – that is, fever, vomiting, lethargy/malaise, irritability, poor feeding and failure to thrive; and/or (2) at least one urinary symptom identified by NICE<sup>2</sup> as a potential marker of UTI – that is, abdominal pain, jaundice (children  $< 3$  months only), haematuria, offensive urine, cloudy urine, loin pain, frequency, apparent pain on passing urine and changes to continence.

Therefore, children consulting with other apparently obvious causes for their symptoms such as otitis media or bronchiolitis, as well as those with a history of previous UTI and known abnormalities of the urinary tract, learning difficulties, or reconsulting for an existing illness, were all included, as long as none of the exclusion criteria applied. Where possible, data were included for children recruited and immediately referred from primary care to secondary care. The eligibility criteria are summarised in *Table 1*. The same inclusion and exclusion criteria also made up the CRF section 1: screening form (see *Appendix 4*).

### Widening participation

The study included parents who spoke non-English languages. Parent information sheets and consent forms were translated into other languages as required by participating GP practices (e.g. Welsh, Polish and Brazilian Portuguese). For languages less commonly spoken in the UK, particularly for those in which oral translation was more useful than written translation (e.g. Somali), translational services were accessed.

**TABLE 1** DUTY eligibility criteria

Children only included if ALL criteria were met	Children were excluded if ANY criterion is met
Aged before their fifth birthday	Aged $\geq 5$ years
Presenting at a participating NHS primary care site	Parents are unable or unwilling to assist with the study
Presenting with an acute ( $\leq 28$ days) illness as the main reason for the parent to have requested an appointment	Illness longer than 28 days' duration Presenting with trauma as a predominant concern
Presenting with at least one 'constitutional' symptom or sign identified by NICE <sup>2</sup> as a potential marker for UTI – i.e. fever, vomiting, lethargy/malaise, irritability, poor feeding and failure to thrive <i>and/or</i> at least one urinary symptom identified by NICE <sup>2</sup> as a potential marker of UTI – i.e. abdominal pain, jaundice (children $< 3$ months only), haematuria, offensive urine, cloudy urine, loin tenderness, frequency, apparent pain on passing urine and changes to continence	No urinary or constitutional symptoms as defined by NICE <sup>2</sup> and listed in the left hand column Known neurogenic (e.g. spina bifida) or surgically reconstructed bladder or urinary permanent or intermittent catheterisation (for whom different bacterial concentration cut points are used) Taking any antibiotics in the last 7 days Taking immunosuppressant medication (e.g. antirejection drugs, oral or intramuscular steroids or chemotherapy) Already recruited into the DUTY study Involved in current research or have recently (within 28 days) been involved in any research prior to recruitment There will be no recruitment to the study after the last NHS laboratory transport of the day has departed from that primary care site on Fridays
	<b>For recruitment at A&amp;E settings only:</b> Children will not be eligible if their presentation at A&E is a direct result of GP referral (as they were then not acting as a first point of primary care contact)

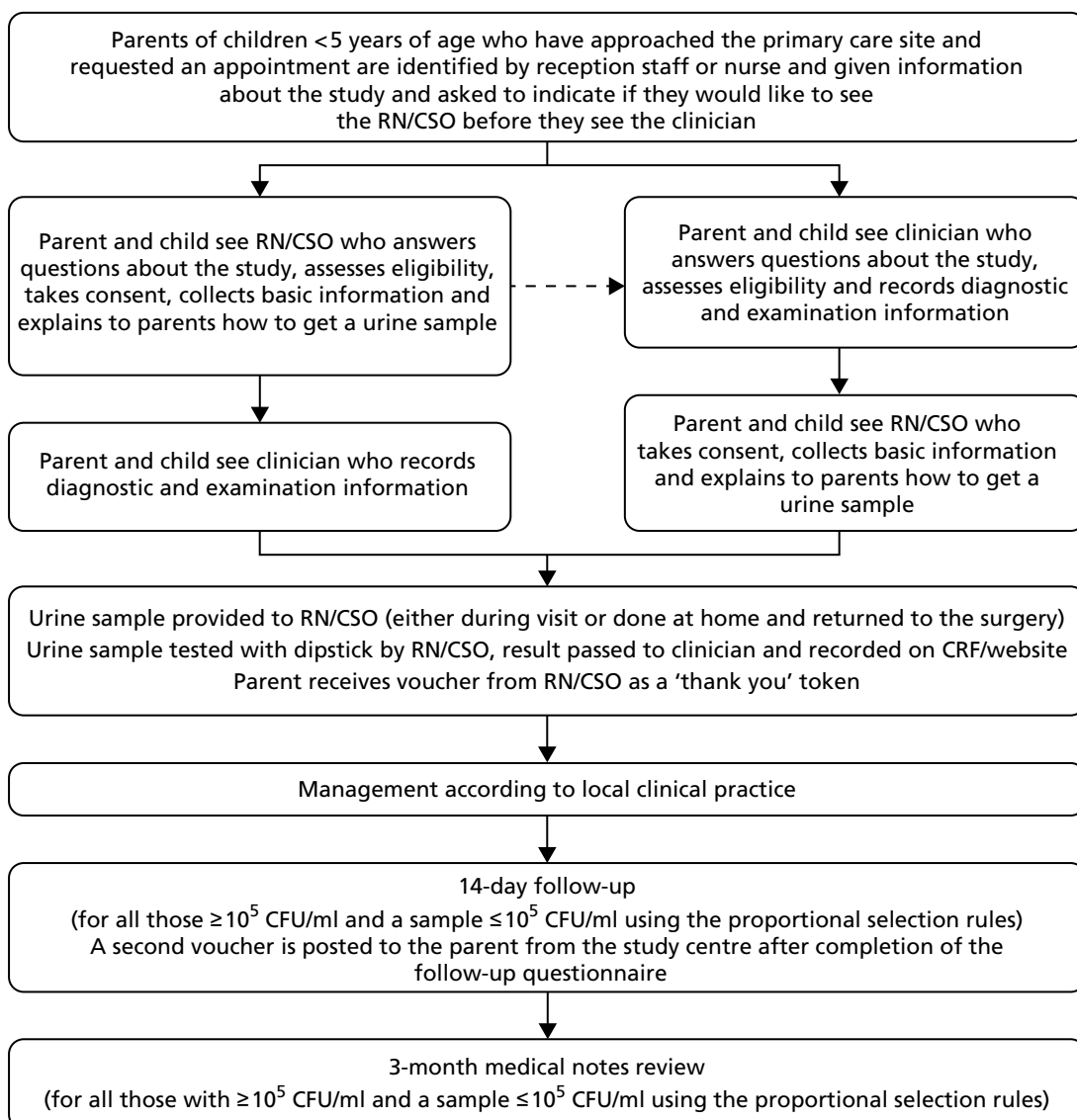
A&E, accident and emergency.

Where possible, interpreters employed by recruiting primary care sites were used to support patient-clinician communications. When these services were not available, translational services were provided via Language Line ([www.language.co.uk](http://www.language.co.uk)). In primary care sites in urban areas, the clinical care team often included a translator whose services were used when available.

## Participant recruitment

The study recruitment process is summarised in *Figure 2*.

The term 'parent' is used to refer to the person with legal responsibility for the child, but it also encompasses other carers (e.g. foster carers and legal guardians).



**FIGURE 2** DUTY study participant flow diagram (the numbers of children recruited are reported in *Figure 4*). Dashed line indicates that the parents can choose to participate either before or after the child sees the doctor/nurse.

### Registration and informed consent

All recruiting primary care sites displayed posters detailing the study in waiting areas. Parents and children were invited into the study in a number of ways:

1. Where possible, the study was mentioned to the parents of children < 5 years old by reception staff when they telephoned for an appointment or during telephone triage. In this instance, the parent was invited to come to the surgery 15 minutes early to receive further information from the RN/CSO.
2. Where the study was not raised at the time of making the appointment, parents of children already booked in were telephoned and advised about the study and invited to attend a little earlier.
3. If contact prior to attendance at the site was not possible, the parent was approached on arrival, given study information sheets and asked if they were happy to discuss participation with the RN/CSO.
4. Once the parent indicated that they were happy to discuss the study, the RN/CSO explained the study, answered any questions, ensuring that they fully understood the implications of participation, and checked the child's eligibility.

Where possible, the RN/CSO recruited the participant while they were waiting to see the GP, in order that the participant was not delayed. However, if the participant's appointment was at risk of being delayed, or it was more convenient for them, the RN/CSO offered to see them after their appointment or visit the parent and child later the same day at their home.

If the parent agreed for their child to participate, written informed consent was obtained from the parent. If the parent was not interested in hearing more about the study, no further approach was made.

In addition to the child being seen by the RN/CSO, they were seen by the child's responsible clinician who managed the child according to their normal practice. The study CRF was then completed for all consented children by the RN/CSO and with clinical examination, diagnosis and management sections completed by the treating clinician (see *Appendix 4*).

### Non-registration

Sites were asked to complete a screening log of all children whose parents were approached by the recruiting clinician and invited to participate in the study. Details were recorded such as their eligibility, whether consent was given or declined, and reason for declining to participate.

## Urine sample collection

The collection of the urine sample was commenced as soon as possible and often while the RN/CSO completed the study CRF, with every effort made to obtain the urine sample while the child was at the recruiting site.

### Rationale for urine collection methods

Suprapubic aspiration and 'in-out' catheterisation carry the lowest risk of contamination<sup>56</sup> but are invasive and unacceptable to parents, and so are uncommon in UK primary care. They were, therefore, not appropriate or feasible for widespread use in the DUTY study. The risk of contamination is low with clean-voided midstream urine or 'clean-catch' samples compared with SPA samples.<sup>2</sup> However, obtaining clean-catch samples can be difficult for children in nappies, and parents find nappy pads or urine bags preferable.<sup>63</sup> Many parents prefer pads to bags as the bags have an adhesive strip that attaches to, and can irritate, the child's skin. These may be suitable alternatives to SPA,<sup>2</sup> but there are few data comparing 'clean-catch' with nappy pad samples. Thus, the preferred primary care method is the 'clean catch', and NICE recommend pad or bag when 'clean catch' is not possible.<sup>2</sup> Therefore, we followed the NICE recommendation for collecting urines, with clean catch being preferred to pads, and pads being preferred to bag.

### Clean catch

The preferred method of 'clean catch' was used for children who were toilet trained or for whom the parent was happy to attempt collection. For the toilet-trained child, a small sterile bowl was used which could fit in a potty or which the parent could hold under the child while sitting on the toilet. For the child still in nappies, the parent first cleaned the nappy area, and then sat with the child on their knee with the bowl placed under the perineal area to collect the urine.

### Nappy pad

As per NICE guidelines,<sup>2</sup> 'Newcastle nappy pads' were used for children still in nappies whose parents did not think clean catch would be successful. First, the parent cleaned the nappy area using water or wipes (the wipes being supplied by the study). A nappy pad was inserted inside a clean nappy, and the nappy refastened. The nappy pad was removed as soon as the child urinated in order to reduce the risk of contamination. The perineum was recleaned and a fresh pad inserted every 30 minutes until micturition or immediately if the pad became contaminated with faeces. Once the child had urinated, and wearing disposable gloves, the RN/CSO removed the pad and urine was extracted into a sterile container as per the manufacturer's instructions.

### Bag urine collection

The child was cleaned in the nappy area by the parent using water or wipes. This was followed by careful drying with sterile towels to ensure that the bag stuck. The RN/CSO applied the bag in such a way to cover the child's genitals. A nappy was then placed over the bag and checked after 30 minutes for urine. The urine was then decanted into a sterile container as per individual manufacturer's instructions.

### Collection of urine at home

If it was not possible to obtain a sample prior to the child leaving the primary care site, the parent was given the necessary equipment and advice on obtaining a urine sample at home. The parent was advised to store the sample in the fridge and return it to their primary care site as soon as possible, ideally within 24 hours. The RN/CSO telephoned parents the next day to remind them to return the sample. Where feasible, the RN/CSO offered to collect the urine sample from the child's home.

### Maximising urine retrieval rates

Obtaining urine samples can be challenging, which is one of the reasons why urine samples are not currently routinely collected in primary care at a rate sufficient to avoid missed diagnoses. As a suboptimal rate of return of samples to the surgery by parents would have diminished power and increased risk of bias, a number of strategies were implemented to maximise retrieval rates. In addition to the role of the RN/CSO in following up parents whose child was not able to provide a sample during the recruitment visit, the data recorded in the study database were used to monitor the retrieval of urine specimens, allowing the research team to identify children for whom urine samples had not been provided, and therefore followed up with the recruiter. Additionally, primary care sites' urine sample return rates were linked to the level of reimbursement via service support costs.

### Urine dipstick methods

The RN/CSO tested the urine sample with a urine dipstick (Siemens Multistix® 8 SG, Siemens Healthcare Diagnostics, Surrey, UK) provided by the study. The dipstick point-of-care test used tested for presence of blood, protein, glucose, ketones, nitrite and LE, and measured pH and specific gravity. A dipstick was placed in the urine as instructed, and results were interpreted according to the manufacturer's instructions.

Urine collection information was collected in section 5 of the CRF (see *Appendix 4*). The time, the urine collection method and the dipstick test results (using the study-supplied Siemens Multistix® 8 SG) were also recorded in this section.

### ***Dispatching urines to the laboratories***

All urine samples, if sufficient quantity of urine was available, were divided into two fractions at the recruiting sites. The priority fraction was sent to the local NHS laboratory for routine diagnostic processing, and the second 'research' fraction to a research laboratory – the SACU, Public Health Wales Microbiology Laboratory, Cardiff – for more in-depth analysis. As only small volumes of urine (minimum 1 ml) were required for each laboratory, it was thought possible that most urine samples would be split into the two fractions.

The NHS 'clinical' fraction was placed in the urine specimen container stipulated by the site's local NHS laboratory and labelled with the child's unique DUTY study ID number on DUTY specific labels as provided in patient recruitment packs. Similar DUTY labels were adhered to specific DUTY microbiology requisition form and the sample sent to the local NHS laboratory using the site's normal method of transport. Any samples returned to the site and not collected within 4 hours were refrigerated on site and processed within 36 hours. Clinicians received and acted on reports from their local laboratory as per usual clinical care. Research laboratory urine results were not routinely fed back to clinicians; this occurred only if a discrepancy was identified between local and research laboratory results.

The remaining portion of urine was decanted into a sterile container (Monovette®, Sarstedt AG & Co., Nümbrecht, Germany) containing boric acid. This was labelled with the child's study ID number and sent by first-class Royal Mail using Post Office-approved Safeboxes™ to the research laboratory.

The urine sampling method, time of sample, time of dipstick test, results of dipstick and quantity of urine were recorded in section 5 of the CRF (see *Appendix 4*).

### **Study thank you vouchers**

All parents received a £5 high street voucher from the RN/CSO as a 'thank you' token for their time in taking part in this first part of the study. A second £5 voucher was sent to those who were proportionally selected for follow-up at 14 days.

### **Data collection**

The DUTY data collection process was complex and the CRF sections were developed with input from the study management group, which included primary care clinicians, consultant nephrologists, health economists, methodologists and statisticians.

To minimise human error, optimise the quality of data entry, and enable effective data collection from multiple sites across England and Wales, we developed a secure, web-based electronic data collection platform. Although this was our preferred data collection mechanism, it was supplemented by a parallel paper-based data collection system for clinicians not wishing to use web-based systems and to cover the event of internet failure. Unique study ID numbers were sequentially generated and used on pre-printed consent forms, paper CRFs, urine sample labels and microbiology requisition forms (for local NHS and research laboratories).

### ***Clinical case report form***

The purpose of the CRF was to capture children's medical histories, examination findings and any known risk factors for UTI. It balanced the need to include as many of the known and potential features associated with UTI with the maximisation of speed and simplicity of completion.

## DUTY index tests

Where available, the selection of index tests (the parent-reported symptoms, clinical signs and dipstick test results) was based on existing evidence. For example, at the time the CRF was developed, ethnicity,<sup>66</sup> circumcision<sup>107</sup> and a history of VUR<sup>108</sup> were known to be associated with UTI in children. Moreover, the NICE clinical guideline<sup>2</sup> on the diagnosis, treatment and long-term management of UTI in children had not long been published, and symptoms and signs were selected from these where evidence of diagnostic utility was available. Furthermore, DUTY was fortunate to be able to build on the experience of the EURICA study,<sup>109</sup> and many of the EURICA symptoms and signs were considered for collection in DUTY.

## Rationale for measuring symptom severity

We decided to measure not only the presence/absence of symptoms, but also their severity. We postulated that the presence of mild symptoms might not be as diagnostically important as severe symptoms. Therefore, parents were asked to rate all 30 symptoms as no problem/slight problem/moderate problem/severe problem, as well as giving a global rating of illness severity from '0' to '10'.

## Case report form summary

The CRF comprised five sections that facilitated data entry by different personnel so as to minimise the burden to busy health-care professionals meeting the demands of day-to-day clinical practice:

1. Eligibility screening and consent (to be completed by recruiting clinician within the recruitment interview with the parent).
2. Registration: background socioeconomic data included date of consultation, name, address, postcode, contact number/s, ethnicity, date of birth and sex. We also asked about the parent's highest educational attainment level and their financial well-being to ascertain the financial burden on families, as postcode mapping may not address this at a household level.
3. Presenting symptoms and medical history included child's presenting symptoms, ongoing health problems (such as asthma or heart disease), antenatal history (gestation at birth and the presence of urinary tract abnormalities on antenatal ultrasound), circumcision, previous UTI, and a sibling or parental history of UTI or other urinary tract diseases.  
Medication history included recent and long-term use of medications (for chronic diseases). We were also interested in medicines that could predispose to UTI and these included the use of laxatives (as a proxy marker for constipation), salbutamol (which could relax bladder smooth muscle) and inhaled steroids (potential immune-suppressant). Toileting and hygiene behaviour was also included, as under- and overwashing and prolonged use of nappies/pull-ups have been postulated as risk factors for UTI.<sup>2</sup>
4. Clinical examination and findings were measured using routine clinical method and included global clinician assessment of illness severity, the child's vital signs and assessments of the child's hydration, consciousness level, throat, ears, chest and abdomen.  
Clinician working diagnosis and management included the clinician's working diagnosis with an accompanying assessment of diagnostic certainty before and after seeing the dipstick urinalysis result. Section 4 of the CRF also asked clinicians to report their subsequent management including the use of antibiotics and referral for secondary care assessment, and whether or not they would have requested a urine sample had the child not been entered into the DUTY study.
5. Urine collection and processing: urine sampling method (clean catch or nappy pad) and urinalysis results with date, time of testing, with a prompt to inform the responsible clinician of the dipstick result and confirmation that the sample had been sent to the local NHS and research laboratories. As one of the DUTY study aims was to assess the added diagnostic value of dipstick urinalysis (over and above the symptoms and signs), dipstick results were included in the CRF as an index test.

## Patient follow-up

DUTY participants were proportionally selected for a follow-up interview at day 14 and medical notes review at 3 months post recruitment. The process by which children were selected and followed up is described below.

### Telephone follow-up at day 14

At 14 days from the recruitment interview, study centre staff contacted parents of all children selected for follow-up according to the proportional selection rules (described in *Table 2*) to record symptom duration and health-care resource use during the 14-day period after recruitment. In older children (> 9 months), the interview also included a parent-completed questionnaire measuring child health-related quality of life (see *Appendix 5*). Owing to the young age of the participants, standard methods [e.g. European Quality of Life-5 Dimensions (EQ-5D)] for measuring health utilities were thought to be invalid. Instead, we used the TNO-AZL (Netherlands Organisation for Applied Scientific Research Academic Medical Centre) Preschool children Quality of Life (TAPQOL) questionnaire, completed at 14 days, to describe the health profiles of children with and without UTI.<sup>110</sup> TAPQOL measures parents' perceptions of health across 12 domains (e.g. sleep, social functioning) and has been shown to be a reliable instrument for both infants and toddlers.<sup>111</sup> No mapping from TAPQOL responses to utility values exists, hence we used a proxy value from a condition (rotavirus) thought to have a similar health-related quality of life impact to UTI. The TAPQOL responses allowed us to validate this choice and make robust comparisons between the health-related quality of life of children with and without UTI.

A booklet version was posted to the parent when telephone contact failed. Details of data collected in the DUTY questionnaire are described below.

1. Symptom duration: we asked if children responded to treatment < 48 hours as NICE has identified failure to respond within 48 hours as a marker of 'atypical UTI' and recommends dimercaptosuccinic acid (DMSA) scan and micturating cystourethrogram (MCUG) in such children.
2. Primary care resource use: practice-based contacts in hours, out-of-hours contacts and community-based contacts (e.g. WIC, NHS Direct or health visitor) and associated expenses for the child's family (fares or mileage, car parking).
3. Hospital resource use [visits to accident and emergency (A&E), attendance at hospital clinics, overnight stays in hospital, use of ambulance services, hospital tests] and associated expenses for the child's family.
4. Consumption of prescribed and over-the-counter medicines.
5. Other out-of-pocket expenses impinging on the child's family, for example time off work, loss of earnings, additional childcare costs.
6. Quality of life (TAPQOL or TNO-AZL Preschool children Quality of Life, after Fekkes *et al.*<sup>110</sup>). The parent was asked to rate the child's quality of life against specific measures of child health including symptoms, sleeping, feeding, behaviour and well-being. The TAPQOL questionnaire is designed for children aged 9 months to 6 years. However, a subset of questions on social, mobility and communications skills are used only in children aged > 18 months.

**TABLE 2** Selection rules used to select children for day-14 and 3-month follow-up

Category	Culture growth category	Definition	Laboratory	Proportion sampled, %
1	$\geq 10^5$ CFU/ml	Pure or predominant species	BOTH NHS laboratory and research laboratory	100 (all)
2	$> 10^3$ and $< 10^5$ CFU/ml	Pure or predominant species	Research laboratory	20
3	$\geq 10^5$ CFU/ml	Two or more species	BOTH NHS laboratory and research laboratory	20
4	$< 10^3$ CFU/ml and 'no growth'		BOTH NHS laboratory and research laboratory	10



### Three-month notes review process

The 3-month notes review (see *Appendix 6*) collected the following data: (1) number and type of consultations (not including routine immunisations, screening checks or NHS Direct contacts); (2) results of any further urine samples; (3) secondary care utilisation [including A&E, hospital clinics, outpatient attendances, admissions and investigations/tests (e.g. ultrasound, MCUG or DMSA scans)]; and (4) the dates, types and doses of any prescribed medications.

The majority of these reviews were completed by the study centre RNs and administrators; however, where recruiting sites were geographically remote from the study centre, the assistance of RNs employed by the PCRN, NISCHR-CRC and practice staff in completing these reviews was sought.

### Day-14 and 3-month selection

Selection of DUTY participants to complete the follow-up interview at day 14 and notes was carried out proportionally according to the growth result of the urine sample provided. All children with a pure predominant positive culture (as defined by the NHS or research laboratory at  $\geq 10^5$  CFU/ml) and a representative number of children with other growth results were selected as described in *Table 2*.

An automatic algorithm within the database was created in order for those children to be selected from growth results posted to the database from the day of recruitment until 28 days post recruitment. Results for any child posted after that time would be deemed too late for them to be selected for a follow-up interview.

### Withdrawal and loss to follow-up

In the majority of cases the only active participation of children was at the initial consultation; therefore, withdrawal from the study was anticipated to be infrequent. For those parents who did wish to withdraw their child from the study, they were asked whether they wished to (1) withdraw completely such that none of the child's data were used; (2) withdraw from all further study activities, but allow all data collected so far to be used; or (3) withdraw from day-14 telephone follow-up but consent to review of patient medical notes at 3 months.

Attrition in those selected for follow-up as a result of the challenges of making contact with busy parents was expected to be a greater threat to follow-up rates than withdrawal by parents. Attrition was minimised by making several attempts to contact parents by telephone. However, if unsuccessful, a postal questionnaire was sent to parents with a stamped addressed envelope. If a telephone interview was not achieved and the postal questionnaire not returned within 2 weeks of sending, the participant was considered lost to the 14-day follow-up, but was still included in the 3-month notes review.

### Discrepant laboratory results and patient safety

All NHS local laboratory results were reported back to clinicians in the same way as for routine care. In order to optimise patient safety the study team informed children's responsible clinicians regarding patients where the NHS local laboratory result was not positive (i.e. contaminated, no growth or not processed) but the research laboratory result indicated clinically significant urine culture positivity. A process was set up by which they were identified and important discrepancies reported to the child's responsible clinician using a purpose-designed letter. While this result was unlikely to drive immediate patient management (as it could be up to a week after the consultation), it was felt that it would be useful for clinicians in cases where the illness was ongoing or recurring.

### Serious adverse events

We did not expect to see any related serious adverse events (SAEs) happening as a result of this research, as the study involved only a non-invasive urine test which is often part of the routine clinical care of patients. However, we did anticipate a relatively high number of unrelated and expected SAEs as a consequence of hospital admissions of acutely unwell children, especially (though not exclusively) for children presenting to recruiting CED sites.

### Processing of urine samples by a 'typical' local NHS laboratory

All local NHS laboratories were 'Clinical Practice Accredited' and NHS laboratory SOPs were used to process DUTY urine samples. All local NHS laboratory SOPs were based on the Public Health England guideline for the investigation of urine.<sup>82</sup> A summary of these processes is given in *Tables 3* and *4*.

### Specimen collection

Specimens were collected in sterile containers (some with boric acid, depending on local NHS laboratory procedures) and, where possible, filled to the line indicated on the tube. All samples were sent to the laboratory as soon as possible after collection, but refrigerated if transport to the laboratory was delayed for more than 4 hours. Specimens in boric acid were refrigerated for no longer than 48 hours.

### Sample processing

On arrival at the laboratory, all DUTY urines and requisition forms were processed as routine specimens using standard laboratory procedures. The patient ID information provided by the requesting primary care site was entered onto the Laboratory Information Management Systems (LIMS). In addition, laboratories were asked to record unique DUTY patient ID, date of birth, date and time of sample collection, sample receipt, and urine processing in laboratory on the DUTY web-based database.

**TABLE 3** Summary of NHS laboratory methods: microscopy

Microscopy	Manual	Automated	Dipstix
Number of laboratories using method	16	8	2 <sup>a</sup>
a One laboratory uses Dipstix as well as manual microscopy. Automated microscopy was performed by the research laboratory.			

**TABLE 4** Summary of NHS laboratory methods: culture

Culture method	Paper foot	Calibrated loop	Multipoint	
Number of laboratories using method	4	19	2	
Culture media	CLED	Chromogenic media	Both	
Number of laboratories using media	7	14	4	
Calibrated loop culture volume	1 µl	2 µl	5 µl	10 µl
Number of laboratories using volume <sup>a</sup>	11	5	1 <sup>b</sup>	1
Number of laboratories using volume <sup>a</sup>	11	5	1 <sup>b</sup>	1
Plate area used	1/4	1/3	1/2	Whole
Number of laboratories using plate area	13	2	3	1
a Two laboratories use a 3-µl volume for culture. b One laboratory uses a 5-µl calibrated loop on certain urines, but normally uses multipoint method. Based on Public Health England's guideline. <sup>82</sup>				

Samples were not processed if (1) there was too little sample volume (< 0.5 ml); (2) a boric acid container was under- or overfilled; (3) the sample was leaking; (4) the data on the form and sample did not match; or (5) if the container was unlabelled.

### Urine microscopy

Microscopy was performed on all specimens by either manual or automated methods. Manual methods vary by laboratory but are based upon the methods set out in Health Protection Agency (HPA) guideline for the 'Investigation of urine'.<sup>83</sup> Validated automated urine analysers were used in some laboratories, in accordance with manufacturers' instructions.

Numbers of white blood cells (WBCs), red blood cells (RBCs), squamous epithelial cells and bacteria were recorded according to local ranges (e.g. WBC: < 10; 10–30; > 30–100; or > 100). The ranges varied by laboratory and were based around the method used.

The presence of other organisms (e.g. *Trichomonas*, *Schistosoma* or *Amoeba*), casts [stating type and quantity (scanty, moderate or numerous)], crystals and yeasts was also recorded.

### Urine culture

Urine culture was performed using local techniques as described in *Table 4*. Standard media, usually UTI chromogenic and CLED media, were inoculated, typically using 1 µl to 2 µl, and plates incubated at 35 °C to 37 °C for 16 to 24 hours or according to local SOPs. Further specialised plates for detection of yeasts, for example, were added if required. Plates were read after incubation by a qualified BioMedical Scientist and results recorded.

### Bacterial identification

The semiquantitative calibrated loop method of culture is only sensitive for screening down to 10<sup>3</sup> CFU/ml if a 5-µl or 10-µl loop is used (e.g. 5 or 10 colonies), or 10<sup>4</sup> CFU/ml if a 1-µl or 2-µl loop is used (e.g. 10 or 20 colonies).

Bacterial ID of significant isolates was performed according to local SOPs but reported according to *Table 5*.

**TABLE 5** Reporting of significant isolates

Bacterial ID	Level of reporting
Anaerobes	'Anaerobes' level
β-Haemolytic streptococci	Lancefield group level
Enterobacteriaceae (except <i>Salmonella</i> species)	'Coliform' level
<i>Enterococcus</i>	Genus level
Pseudomonads	'Pseudomonads' level
<i>S. saprophyticus</i>	Species level
Other coagulase-negative staphylococci	'Coagulase-negative' level
<i>S. aureus</i>	Species level
<i>S. Typhi/Paratyphi</i>	Species level
Yeasts	'Yeasts' level

### **Susceptibility testing**

Susceptibility testing was performed by a variety of standard methods including disc testing and breakpoint testing. Susceptibility testing was performed on all significant isolates using local SOPs based upon susceptibility testing guidelines [for e.g. British Society for Antimicrobial Chemotherapy (BSAC) standardise disc susceptibility testing method (version 11)].<sup>112</sup>

Disc susceptibility testing can be performed either from isolates recovered from urine or direct from urine. Briefly, a standard inoculum was plated onto susceptibility testing media and antimicrobial discs applied. Plates were incubated for 18 to 20 hours in air at 35 °C to 37 °C. Zone sizes were read by a qualified BioMedical Scientist and interpreted using BSAC or other guidelines such as the European Committee on Antimicrobial Susceptibility Testing (EUCAST) and Clinical & Laboratory Standards Institute (CLSI).

The breakpoint method utilises agar plates containing antimicrobials at concentrations around the breakpoint. Significant isolates were emulsified in sterile saline then spot inoculated onto the plates.

Typical antimicrobials tested were firstline antimicrobials: amoxycillin/ampicillin, nitrofurantoin, trimethoprim, ciprofloxacin, co-amoxiclav, cefpodoxime and cephalexin. Cefpodoxime-resistant coliforms were tested for production of extended spectrum beta-lactamase (ESBL) enzymes in some local laboratories.

Growth on the plates by either method indicates susceptibility/resistance to the antimicrobials tested.

### **Reporting procedure and storage of organisms**

NHS laboratories reported results as per local SOP to the requesting clinician and were required to transcribe the results onto the DUTY web-based database within 1 week in order to allow for patient sampling for follow-up. This was also the trigger to activate laboratory payment (see *Appendix 8* for details of laboratory CRF).

Isolates from pure or predominant culture at  $\geq 10^3$  CFU/ml were stored on cryogenic beads at  $-20$  °C or  $-70$  °C.

### **Quality control**

All stock reagents, media, antimicrobials and equipment were monitored for quality assurance at locally specified times. Quality assurance records were completed for lot numbers and expiry dates. All laboratories perform internal quality control and participate in internal quality assurance and external quality assurance schemes.

## **Processing of urine samples by the research laboratory**

The research laboratory had experience in supporting other primary care UTI studies and supported the EURICA study.<sup>109</sup>

### **Data processing**

Urines were sent overnight by Royal Mail SafeBoxes™ by the participating sites. Monovette containers containing boric acid were used to stabilise bacterial counts.

On arrival in the laboratory all DUTY urines and forms were checked for matching identifiers. The following data from each sample were recorded: centre ID, patient ID, date of birth, date and time of sample collection, sample receipt and urine processing in laboratory on the DUTY web-based database. If samples were received out of hours then urines were stored at 4 °C until processed and the date received recorded.

Samples were not processed if (1) there was too little sample volume (< 0.5 ml); (2) the sample was leaking; (3) the sample was received in a non-sterile container; (4) the data on the form and sample did not match; or (5) the container was unlabelled.

Before processing, each Monovette was weighed using a fine-scale balance and the total weight recorded.

### Urine microscopy

Urine microscopy was performed on all urine samples with sufficient volume. If 1 to 2 ml only were received then culture only was performed. Microscopy was performed using the iQ200SPRINT analyser (Beckman Coulter Ltd, High Wycombe, UK). Samples that exceed the pre-set thresholds for a given category were manually reviewed by the BioMedical Scientist.

Microscopy results including WBC counts, RBC counts, bacteria counts, squamous epithelial cell counts, presence of yeasts and presence of casts were recorded.

### Urine culture

Urines were diluted prior to spiral plating in accordance with standard dilution factors and 50 µl of urine dilution was inoculated onto UTI chromogenic agar and Columbia Blood Agar (CBA) and allowed to dry. Spiral plating was repeated for other dilutions. The stylus was washed between urine samples using 5% hypochlorite (× 1) then sterile water (× 2). Agar plates were incubated at 35 °C to 37 °C for 18 to 24 hours in aerobic conditions. If yeasts were seen on microscopy, these were also spiral plated onto two Sabouraud agar plates and incubated at 35 °C to 37 °C and 30 °C for 5 days.

### Colony counting

Total colony counts were performed for all sample CBA plates following the spiral plater manufacturers' instructions. Colonies were counted in sectors of a grid which overlays the agar plate and the number of colonies in each sector translated into CFUs per ml using the manufacturers' tables and counts were recorded. Species-specific counts were performed on all isolates present at > 10<sup>3</sup> CFU/ml in pure or predominant growth cultures on the UTI chromogenic agar.

### Identification of isolates

A preliminary identity of all counted isolates was ascribed using the UTI chromogenic agar. Significant isolates were identified further by basic microbiological tests (e.g. colony morphology, biochemical tests such as indole for *E. coli*). Potentially significant isolates were considered to be Enterobacteriaceae (*E. coli*, *Klebsiella* spp., *Citrobacter* spp., *Enterobacter* spp., *Serratia* spp., *Proteus* spp., *Morganella* spp.) (see Appendix 8).

For samples with mixed ≥ 3 organisms with colony counts of 10<sup>3</sup> CFU/ml and/or no significant isolates, the ID was performed by UTI chromogenic agar alone.

### Susceptibility testing

Antimicrobial susceptibility was determined for all pure or predominant significant organisms according to BSAC disc susceptibility testing guidelines. Isolates were cultured onto CBA plates for single colonies and inspected for purity. A few colonies were suspended in 3 ml sterile water to turbidity 0.5 McFarland and inoculated onto Iso-Sensitest Agar (Oxoid Thermofisher, Basingstoke, UK). The surface was allowed to dry for a few minutes and then the antimicrobial discs applied and plates were aerobically incubated for 18 to 20 hours at 35 °C to 37 °C.

Antimicrobials tested for Gram-positive isolates were trimethoprim, nitrofurantoin, amoxicillin, cefoxitin, vancomycin and novobiocin.

Antimicrobials tested for Gram-negative isolates were trimethoprim, nitrofurantoin, amoxicillin, co-amoxiclav, cephalixin, ciprofloxacin and cefpodoxime.

All antimicrobials except vancomycin were purchased from Sigma-Aldrich Ltd, Poole, Dorset, UK. Disc zone sizes were recorded and interpreted using BSAC guidelines.

### **Detection of antimicrobial substance in urine**

All urines were tested for the presence of any antibacterial substance in the urine. A 0.5 McFarland solution of *Bacillus subtilis* (NCTC 10400) was inoculated onto an Iso-Sensitest Agar plate. After 5–10 minutes' drying time, 10 µl of urine was spotted onto the plate, and the plate incubated for 18 to 24 hours' aerobic incubation at 35 °C to 37 °C. This detects any antimicrobial substances that are in the urine.

### **Storage of urine and cultured organisms**

For samples exhibiting either pure or predominant culture at  $> 10^3$  CFU/ml, each isolate was stored on cryogenic beads at  $-70$  °C. For samples exhibiting mixed  $\geq 3$  organisms at  $10^3$  CFU/ml, a sweep of the growth was saved on cryogenic beads at  $-70$  °C. Any significant isolates were stored on cryogenic beads at  $-70$  °C. Samples with no growth or growth  $< 10^3$  CFU/ml were not stored.

Two aliquots of all urine samples were stored in cryogenic vials at  $-70$  °C, one containing 5% sterile glycerol for bacterial preservation.

### **Data entry to website**

All data recorded were transcribed to the DUTY web-based database on a weekly basis (see *Appendix 8* for details of laboratory CRF). Lists of those patients considered to have a positive UTI result based on research laboratory data were collated and reported to the study centres.

### **Quality control**

All stock reagents, media, antimicrobials and equipment were monitored for quality assurance at locally specified times. Quality assurance records were completed for lot numbers and expiry dates. See *Appendix 8* for research laboratory quality control methods.

### **Returning positive isolates to the research laboratory after the study**

After closure of recruitment for the DUTY study, all isolates retained by the local NHS laboratories were transported to the research laboratory for long-term storage. All isolates were labelled pseudo-anonymously with the study patient ID only. The transfer was made using an Amies charcoal transport swab (Technical Service Consultants Ltd, Heywood, UK) and using Hayes DX couriers ([www.dxdelivery.com](http://www.dxdelivery.com)) (Royal Mail was used as an alternative). Hayes DX ensured that specimen delivery services were fully compliant and International Air Transport Association-approved to meet the current HPA legislation requiring that infectious samples have to be moved in accordance with strict guidelines.

## **Electronic data entry**

The DUTY electronic database and DUTY website were hosted on the electronic PCRn (ePCRn) Citrix servers at the South London and Maudsley (SLaM) NHS Foundation Trust. The Citrix environment provided a secure connection for the users via the client software which was required to be installed on each computer. The secure web-based system was developed by King's College London, using industry best practice and in collaboration with the NHS Research Capability Programme and the PCRn, to provide research tools to enhance recruitment to NIHR portfolio studies in primary care.

The electronic CRF, 14-day follow-up, 3-month notes review, local laboratory and research laboratory input, web-based data collection pages entirely mirrored their paper-based counterparts. They were created in ASP.net (a dynamic web application framework; [www.asp.net](http://www.asp.net)) on top of a dedicated Structured Query Language (SQL) data management server, with data variables forced to comply with entry and validation rules defined in the data element definitions. The SQL data management server incorporated

auditing, back-up and recovery facilities. The study workflow and algorithms were enforced using the same methods, and a visual algorithm on the web pages guided users. The web-based system was piloted for ease of use prior to data entry going live. The data were backed up nightly off-site. The patient identifiable data on the server were encrypted before storage on the database. For added security the encryption/decryption functions were not part of the database but were implemented using linked functions whose code was in a different location from the database on the server. Once logged on to the Citrix environment, users connected to the DUTY website using a separate account. Accounts were linked to roles with different privileges for accessing sections of the website and different privileges for viewing/editing patient data. Data were exported as comma-separated values (CSV) files from the export section of the DUTY website.

### **Data entry in primary care sites**

The web-based database system was presented as the preferred method of data collection, and DUTY recruiting staff were encouraged and supported to enter CRF data onto the database directly or, if using paper-based CRFs during the recruitment interviews, to retrospectively enter the data in a timely manner (consent and registration within 24 hours, and full eCRF data within 5 working days).

### **Data entry in the NHS and research laboratories**

All laboratory data generated from microscopy and culture for each urine sample received from the NHS local laboratory and research laboratory were entered into the DUTY web-based database. Local NHS and research laboratories staff were only able to access the microbiology data collection pages to log the samples on receipt and enter the results within 1 week of data being available.

### **Follow-up data entry in research centres**

Follow-up data collected at day 14 and 3 months from telephone interviews, postal questionnaires and patient records were entered either directly at time of collection, or retrospectively from paper CRFs, onto the web-based follow-up data collection forms by research staff at each centre.

## **Data management**

### **Data quality**

Sites where concerns regarding the possibility of protocol deviations and/or data entry issues not being concordant with electronic data recorded were subject to a data quality audit. The audit questions to be addressed were developed after discussion with both the sponsor and University Hospitals Bristol NHS Foundation Trust. Audits were followed by updated training and agreed reaudits. Following the reaudit of one particular site, by the University Hospital NHS Bristol R&D team on behalf of the sponsor, a study-wide audit was initiated with a random selection of participating primary care sites.

### **Data cleaning**

Where data between paper CRFs and web-based database conflicted, the value on the paper CRF was deemed the true value (making the assumption that the database value was the result of a keying error), unless the paper CRF had already been appropriately annotated with a correction.

All missing data that could not be resolved, for example where no paper record was available, were recorded.

Some 'self-evident correction' rules were developed for areas where it was evident that a data entry mis-key had occurred. Self-evident corrections did not involve any changes to actual data or result values; they were used only for process-orientated variables such as dates.

Other cleaning issues were considered, as shown below in order of importance:

1. duplicate patients/NHS numbers
2. missing urine samples and illogical microbiology data
3. time periods – overlong and negative times (e.g. date of birth and date of recruitment)
4. overall missing data – large numbers of missing data for individual recruits
5. outliers – for example in examination findings (temperature, pulse rate, etc.)
6. relationship of consentor to child – to confirm that consent was given by the parent.

All CRF data queries were provided to each centre for verification either against the paper CRFs or by contacting the recruiting site for clarification.

In order to maximise the urine sample laboratory data, all entries on the CRF/database indicating that a urine sample had been obtained, but the laboratory had not entered the results, were followed up directly with the laboratories.

An inbuilt SQL quality control facility provided a full audit of all amendments made and was available as a CSV download. All amendments made to the database were also manually audited and recorded separately.

### ***Data storage and retention***

All data will be archived until the youngest participant reaches the age of 21 years, in line with each centre's governance framework regulations for clinical research. The archive will consist of the protocol and all of the amendments, patient information sheets, CRFs and the follow-up data, and analysis records. Additionally, all of the legally required documentation relating to ethical approval, sponsorship and indemnity will be archived.

## **Statistical methods**

### ***Research objective algorithm***

The overall objective of the study was to derive and validate a clinical algorithm for the diagnosis of UTI, in children aged before their fifth birthday presenting to primary care with an acute illness. The algorithm was constructed in two stages: ID of children at risk (in whom a urine sample should be obtained) and determination of the added value of point-of-care urine dipstick testing. The determination of when to obtain a urine sample was based on sociodemographic factors, medical history, symptoms and signs. The purpose of the dipstick analysis was to assess whether dipstick urinalysis for nitrite, LE, protein, blood and glucose gives additional diagnostic information to assist in the ID of children who should receive immediate antibiotic treatment and urine samples that should be sent to the laboratory. The findings were combined to produce an overall algorithm.

### ***Sample size calculation***

Our sample size calculation was targeted at our main objective, namely the development of the clinical algorithm. We assumed a candidate predictor with 10% prevalence and UTI prevalence of 2%. With 80% power and a two-sided alpha of 5%, 3000 urine sample results are required to detect an odds ratio (OR) of 2.4, while 3100 results give a 95% CI with width 10% for an algorithm with 80% sensitivity. We originally proposed to recruit 4000 children with a target of recovering urines from at least 77.5% for algorithm derivation and a further 2000 children for external validation. However, we did not originally anticipate the need to stratify analyses by urine collection method. We therefore decided to use all available results to derive the models, with internal bootstrap validation instead of external validation as originally planned.



### Determining prevalence rates

We calculated the prevalence in all samples and also by the collection method using methods appropriate for small proportions.<sup>113</sup> We also looked at the degree of variation in prevalence stratified by collection method. First, we considered the degree of variation in prevalence between practices using a two-level random-effect logistic regression model with practice/site as a random effect and no fixed effects. Second, we used a LR test between this multilevel model and logistic regression to show if there was any evidence of clustering, using a conservative  $p$ -value threshold of 0.1. Finally, using the chosen model type, multilevel logistic or logistic regression, we then added a number of variables to the model: area, type of recruitment site, age, sex and an Index of Multiple Deprivation (IMD) deprivation score for each child. We then assessed the effect these variables had on prevalence by listing the crude and adjusted ORs (with 95% CIs) for each of these variables.

### Economic evaluation

We developed decision-analytic models using decision trees and Markov models to identify the optimal urine sampling strategy for acutely unwell children under the age of 5 presenting to primary care. As diagnostic accuracy depended on the urine collection method, we developed a 'clean-catch' model and a 'nappy pad' model. The models, which synthesise data from the DUTY study and the wider literature, estimate the lifetime costs and health outcomes for six urine sampling strategies including three derived from the DUTY risk score. A lifetime horizon is important to capture the rare but serious complications of UTI that may occur later in life.

In secondary analyses, we extended the models to compare three testing and treatment strategies to explore the role of dipstick testing in guiding laboratory testing and antibiotic prescriptions. Costs were estimated from a NHS perspective and included diagnostic, short-term treatment and long-term complication costs. Health outcomes were expressed using QALYs and quality-adjusted life-days (QALDs).

The model was made up of three parts: short term (diagnosis and acute illness; 21 days), medium term (recurrent UTI; 3 years) and long term (long-term sequelae; lifetime).

A full description of the methods used in this analysis is given in *Chapter 6*.

### Other analyses

We planned to use methods appropriate for small proportions<sup>113</sup> to estimate the prevalence (with 95% CI) of culture positive urines in acutely unwell children under 5 years. This was undertaken on all children with cultured urines. The degree of variation in prevalence between practices and geographical areas was explored. This analysis was also explored by looking at difference by recruitment site type (general practice, WICs, out-of-hours providers and CEDs).

Recruited children for whom urine samples were obtained were compared with those for whom no urine sample was obtained in terms of clinical presentation and demographics.

We compared the probability of contamination in samples that were retrieved via a 'clean-catch' method with those using nappy pads and investigated the factors associated with contamination. We examined the impact of the time between obtaining the urine, transportation (including day of the week) and laboratory analysis on the rates of positive and contaminated urine samples (e.g. exploring if delayed samples such as those taken after daily laboratory collection have an impact on contamination rates) using recommended methods.<sup>114</sup>

## Preventing and minimising bias

The following design and analytic strategies were employed to minimise bias:

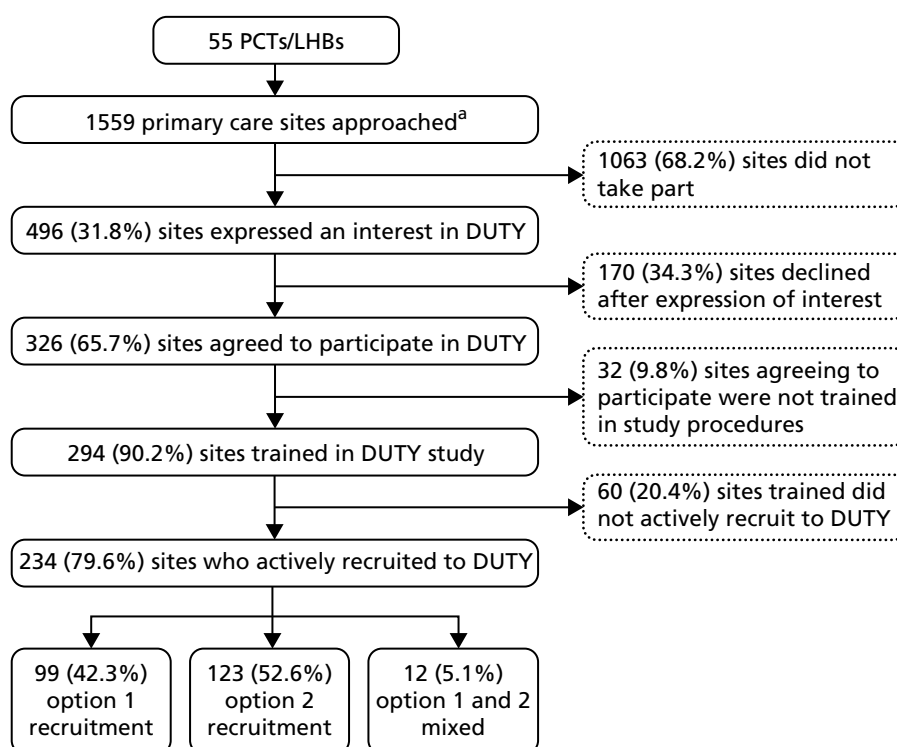
1. Selection bias: where possible we recruited consecutive children; we asked sites to keep a screening log of patients who were approached but did not take part in the study and the reasons for this.
2. Index test technology: all tests (symptoms, signs, nappy pads, dipstick tests) were carried out blind to the reference standard using standardised equipment and protocols (the blinding was easily achieved as the report on the urine sample took a minimum of 2 days to be processed).
3. Incorporation bias: the reference standard consisted of culture alone and did not incorporate any of the index tests.
4. Review bias: observers assessing the index tests differed from and were blind to the reference standard (and vice versa).
5. Verification bias: as many children as possible had a urine sample to assess the reference standard. Children in whom it was not possible to obtain a sample were excluded from the analysis. We compared the characteristics of children with and without urine culture results.
6. Disease progression bias: we measured the time between clinical assessment and obtaining the urine samples.
7. Treatment paradox: for most children, it was anticipated that antibiotic treatment would be started after the urine sample had been obtained, but to assess the degree to which this was the case, we tested for the presence of urinary antimicrobial substances.
8. Handling of missing values and withdrawals: these parameters were measured and modal and multiple imputations used in the analyses.
9. Appropriateness of the reference standard: research laboratory samples were processed by two staff members using a single, standardised procedure. We used a microbiological definition of UTI of  $\geq 10^5$  CFU/ml of a single uropathogen ('pure growth') or  $\geq 10^5$  CFU/ml of a uropathogen with  $\geq 3 \log_{10}$  (1000-fold) difference between the growth of this and the next species ('predominant growth'). We defined uropathogens as members of the Enterobacteriaceae group.

## Chapter 3 General results

This chapter presents the general results of the DUTY study, providing descriptive statistics regarding recruiting sites and participating laboratories, CRF data and the proportion of participants who were followed up. Recruitment accruals across the four study centres and a comparison of the recruited participants with those who declined to participate are also presented.

### Site recruitment

A breakdown of the number of sites approached, those expressing an interest, those that agreed to participate and those actively recruiting for each of the four study centres are presented in *Figure 3* and *Table 6*.



**FIGURE 3** Recruitment of primary care sites. a, Incomplete data set: owing to involvement of PCRN it is not possible to determine the total number of primary care sites approached.

**TABLE 6** Recruitment of primary care sites involved in DUTY across study centres

Recruitment of primary care sites	Bristol, <i>n</i>	Cardiff, <i>n</i>	London, <i>n</i>	Southampton, <i>n</i>	Total, <i>n</i> (%) <sup>b</sup>
PCTs/LHBs involved	20	6	11	18	55 (N/A)
Primary care sites approached <sup>a</sup>	139	465	90	865	1559 (N/A)
Sites expressing an interest in DUTY	139	113	47	197	496 (31.8)
Sites agreeing to participate in DUTY	112	51	44	119	326 (20.9)
Sites trained in study procedures	112	49	42	91	294 (18.9)
Sites that actively recruited to DUTY	96	44	36	58	234 (15.0)

LHB, local health board; N/A, not applicable.

a Incomplete data set as full records of primary care sites approached not kept by PCRN.

b Percentage of primary care sites approached.

A total of 55 PCTs in England and local health boards (LHBs) in Wales were involved in setting up recruitment sites for the DUTY study. This figure was correct at the commencement of the study; however, some of the PCTs have since reconfigured. We are aware of 1467 sites that were approached about participating in the DUTY study; however, this is an underestimation as it was not possible to determine the actual number of primary care sites that were approached about the DUTY study in the areas covered by the some of the PCRNPs involved.

A total of 496 sites expressed an interest in the study and 326 (65.7%) agreed to participate. Of these, 294 (90.2%) sites were trained in the DUTY study processes. However, there were 32 sites (9.8%) that agreed to participate but did not progress any further; this was due to reasons such as difficulties faced by DUTY RNs in setting up meetings at the practice to go through the DUTY training.

Of the 294 sites that were trained, 79.6% of these actively recruited at least one participant to the study. Sites were monitored by their local centre and efforts were made to support practices to increase the recruitment potential of slow recruiting sites.

Table 6 shows that 96 (41.0%) of the total 234 sites actively recruiting were co-ordinated by the Bristol centre, and 58 (24.8%) were co-ordinated from the Southampton centre. The Cardiff centre co-ordinated 44 sites (18.9%) and the London centre co-ordinated 35 sites (15.0%).

### Types of primary care site

The majority of primary care sites were GP surgeries ( $n = 226$ , 96.6%), with four WICs and four CEDs. The Cardiff centre was the only centre to recruit exclusively from GP surgeries ( $n = 44$ ). The final recruitment figures were 6797 (94.9%) in GP surgeries, 284 (4.0%) in CEDs and 82 (1.1%) in WICs. Three WICs were co-ordinated by the Bristol centre and one by Southampton. The Bristol centre also co-ordinated two CEDs, and London and Southampton each co-ordinated one CED site.

### Recruitment models used by primary care sites

The two models of recruitment (option 1 and option 2) offered to primary care sites are described in Chapter 2. Owing to the practicalities of recruiting in busy clinical environments, the WICs and CEDs employed only the option 2 recruitment model. Table 7 shows the distribution of recruitment methods by centre.

Over half of the primary care sites ( $n = 124$ ) employed an option 2 recruitment model, while 99 (42.5%) employed an option 1 model. There were 12 (5.1%) sites that used a mixed recruitment model, where both the site staff and the DUTY RN/CSOs would recruit at the site.

**TABLE 7** Recruitment models used by primary care sites involved in DUTY across study centres

Recruitment model used by primary care sites	Bristol, <i>n</i>	Cardiff, <i>n</i>	London, <i>n</i>	Southampton, <i>n</i>	Total, <i>N</i> (%)
Option 1 recruitment model	41	27	17	14	99 (42.3)
Option 2 recruitment model	48	14	17	44	124 (52.6)
Mixed-options recruitment models	7	3	2	0	12 (5.1)
Total primary care sites	96	44	36	58	234

## NHS laboratories

All the local NHS laboratories related to the sites in the Bristol centre agreed to participate, process the urine samples and then enter the results on the DUTY study online database. In the Cardiff centre, all local NHS laboratories approached agreed to take part, but one laboratory did not enter results on the DUTY database. Two NHS laboratories in the London centre (South West London area) were approached but declined to participate. In the Southampton centre, two laboratories declined to participate in the study; therefore, related primary care sites were not included. Out of the 28 laboratories that had agreed to participate, 24 entered results of the DUTY samples processed onto the DUTY database. The urine sample results from the local NHS laboratories that did not use the DUTY database were retrieved via the reports sent to the GP surgery and were entered onto the database retrospectively by the centre study managers.

## Participant recruitment

The flow of participants through the DUTY study is shown in *Figure 4*. This includes the number of children and their families approached and assessed for eligibility, the number recruited to the study, urine sample retrieval rates, urinalysis and the number of participants followed up at 14 days via interview and at 3 months via note reviews. Reporting is in accordance with the STROBE checklist for reporting observational studies.<sup>115</sup>

### Eligibility assessment

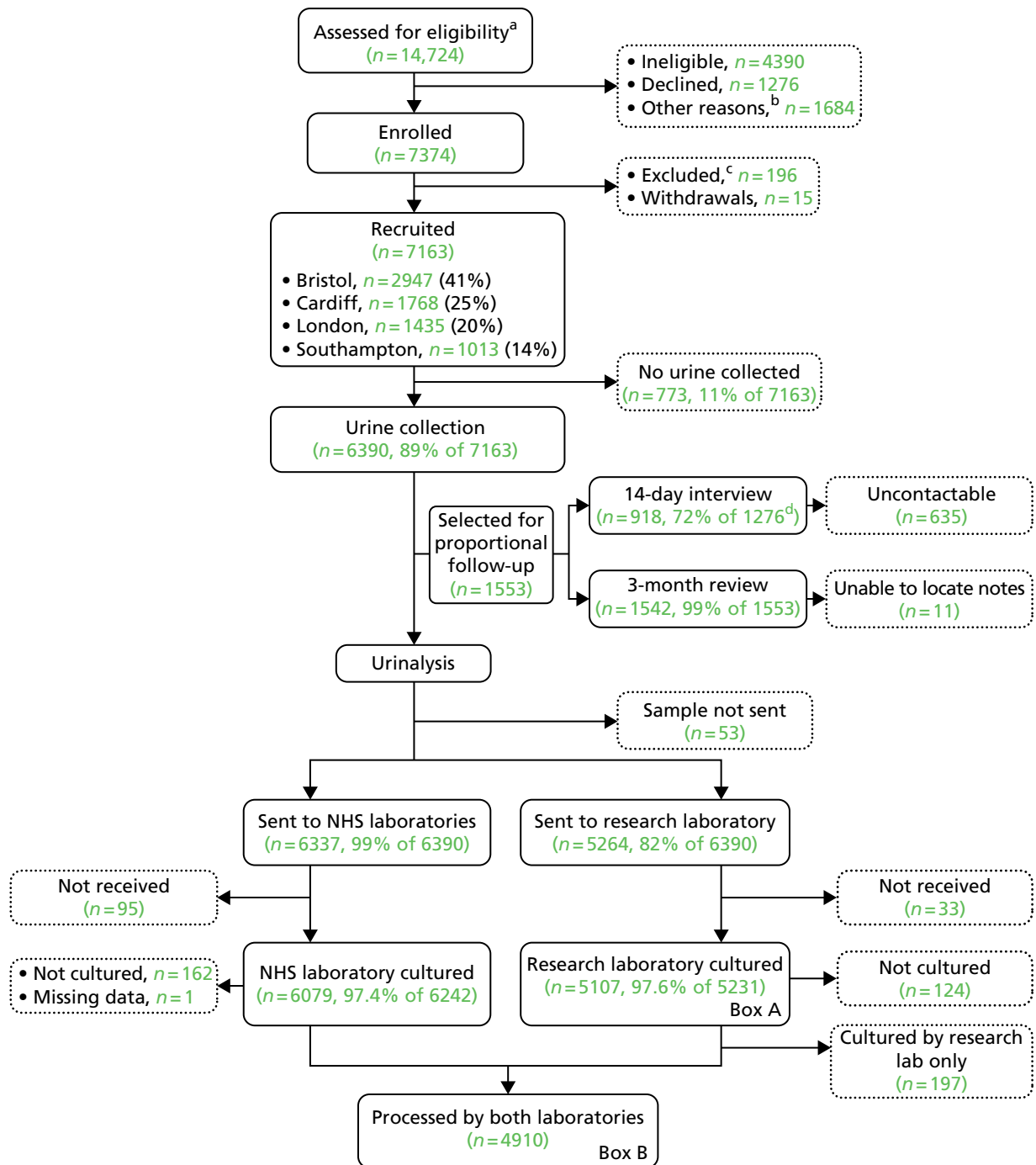
A total of 14,724 children were assessed for eligibility across the four study centres according to completed screening logs. However, this figure is an underestimation of the total number of children assessed, as not all sites completed screening logs and some were incomplete.

Of the 234 primary care sites taking part, 198 (85%) completed and returned at least one screening log to the study centres. These show that 7350 children were screened but not recruited (see *Figure 4*), as they declined (1276) or were not eligible (4390), or for other reasons (1684) including that they left the primary care site prior to invitation (811), they did not give consent (214) or there was a language barrier (112) and an appropriate translator was not available at the time of recruitment. *Table 8* shows the distribution of these children by centre.

### Comparison of children recruited to DUTY with those not recruited

From the screening log information, which recorded the sex and date of birth of the children visiting the primary care sites, we were able to compare the age and sex of the children who were recruited to DUTY ( $n = 7163$ ) with that of the children whose parents declined to participate ( $n = 1276$ ) (*Table 9*).

Comparing the proportion of males in the two samples with a two-sample proportion test shows strong evidence that the proportion of males is higher in the declined sample, with a mean difference of 5.2% (95% CI 2.2% to 8.2%;  $p < 0.001$ ). The mean age in the declined sample was 24.06 months and in the recruited sample it was 26.88 months. Comparing the mean age in the two samples with an independent sample *t*-test also shows strong evidence that the mean age is higher in the recruited sample, with a mean difference of 2.04 months (95% CI 1.08 to 3 months;  $p < 0.001$ ).



**FIGURE 4** DUTY study participant flow diagram. a, Incomplete data set: number approached represents data received from screening logs; b, other reasons include left prior to invitation, no consent or there was a language barrier; c, includes  $n=44$  retrospectively ineligible due to GP referral (protocol amendment 6),  $n=55$  data quality issues and  $n=97$  cases removed during data cleaning; and d, omits 277 cases not selected by DUTY database (see *Participant follow-up, Proportional selection for follow-up*). Box A used in Chapters 3, 6 and 8. Box B used in Chapters 4, 5 and 7.

**TABLE 8** Children screened but not recruited by the four study centres

Reasons for non-recruitment	Bristol, <i>n</i> (%)	Cardiff, <i>n</i> (%)	London, <i>n</i> (%)	Southampton, <i>n</i> (%)	Total, <i>N</i> (%)
Declined	265 (20.8)	420 (32.9)	379 (29.7)	212 (16.6)	1276 (17.4)
Not eligible	1236 (28.2)	1294 (29.5)	1276 (29.1)	584 (13.3)	4390 (59.7)
Other (including missing)	455 (27.0)	706 (41.9)	339 (20.1)	184 (10.9)	1684 (22.9)
Total	1956 (26.6)	2420 (32.9)	1994 (27.1)	980 (13.3)	7350 (100)

**TABLE 9** Comparison of ages and sex of children recruited with those declined to participate

Variable	Category	Declined, <i>n</i> (%)	Recruited, <i>n</i> (%)
Age of child	< 6 months	166 (13.2)	651 (9.1)
	6 to < 12 months	178 (14.2)	1141 (15.9)
	1 to < 2 years	353 (28.1)	1681 (23.5)
	2 to < 3 years	216 (17.2)	1347 (18.8)
	3 to < 4 years	194 (15.4)	1333 (18.6)
	4 years plus	151 (12.0)	1010 (14.1)
	Missing	18 <sup>a</sup>	0
	Total	1258	7163
Sex	Male	691 (54.4)	3526 (49.2)
	Female	579 (45.6)	3637 (50.8)
	Missing	6	0

a Twelve missing dates of birth and six ages of zero or lower.

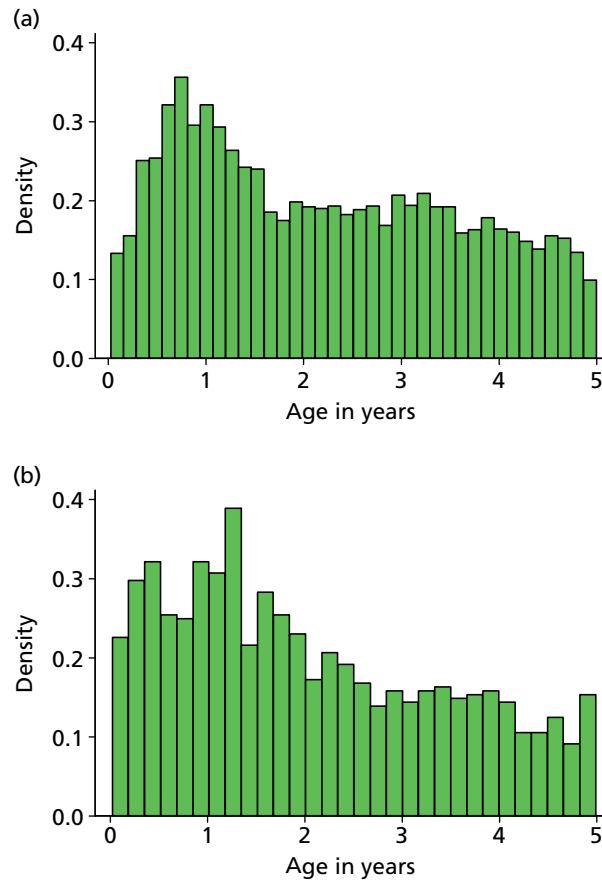
Differences in sex and age are not particularly large and the small *p*-value and narrow 95% CIs are due to the large sample size. *Figure 5* shows that the distribution of age in the two samples is similar, with both skewed towards the younger age.

### Recruitment accruals

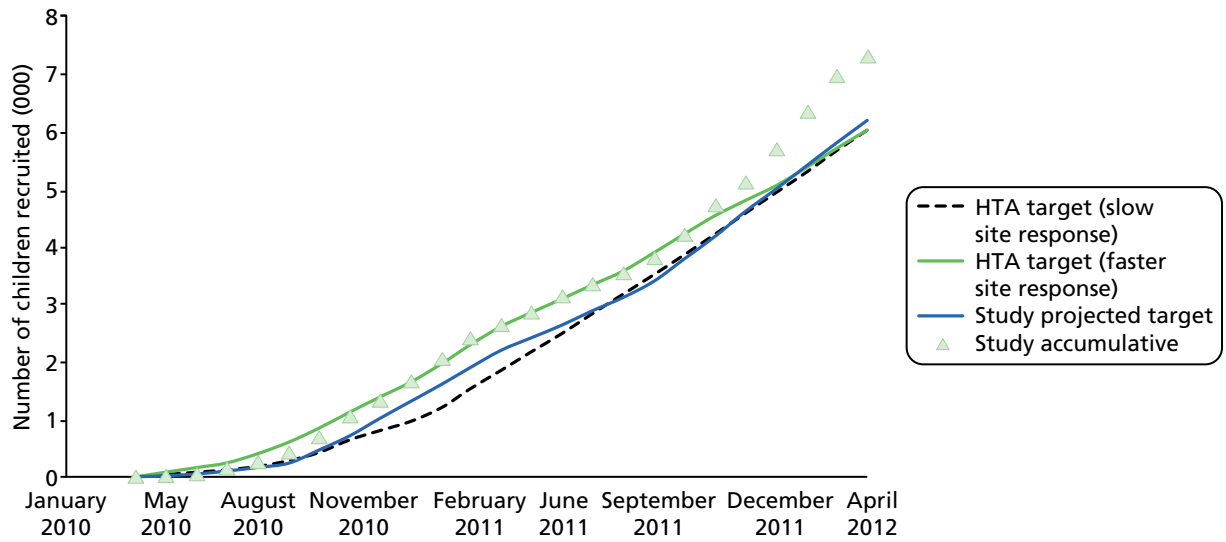
A total of 7374 children were enrolled into the study (50.1% of those assessed for eligibility); of these 196 were excluded and 15 were withdrawals (see *Exclusions* and *Withdrawals*, below).

Of the 7163 total recruited children, 6797 (94.9%) were recruited in GP surgeries, 284 (4.0%) in CEDs and 82 (1.1%) in WICs. The Bristol centre recruited 2947 (41.1%), the Cardiff centre recruited 1768 (24.7%), the London centre recruited 1435 (20.0%) and the Southampton centre recruited 1013 (14.1%) participants to the study during the recruitment period.

When comparing the total number of children recruited to the DUTY study with the original recruitment targets set by the HTA for study viability assessment, it can be seen that that the study consistently exceeded its projected study target (*Figure 6*). The sample size requirement (6000) was reached in January 2012, when we sought approval (protocol amendment 9) to continue recruiting until the end of the scheduled period (30 April 2012) in order to maximise the statistical power of the sample and provide more precise estimates of association of index tests with UTI in the final algorithm, and to take account of missing data.



**FIGURE 5** Histograms of the ages of the children recruited and those who declined to participate. Histogram of age for (a) recruits and (b) decliners.



**FIGURE 6** DUTY recruitment accruals against HTA targets.



### Exclusions

A total of 196 children were retrospectively excluded, including 44 children who were retrospectively ineligible because they had been recruited in a CED after being referred by their GP. These participants had not been recruited on their first point of contact in primary care, and were likely to be selected as more unwell. The study protocol was subsequently amended (protocol amendment 6) to ensure that only children consulting the CED for the first time, without prior GP or other primary care contact, were included. An additional 55 children were excluded from a site co-ordinated by the London centre due to poor data quality issues. This site received intensive training and recruited a further 10 participants; however, the data quality was still poor, and therefore the site was closed to further recruitment. During data cleaning a further 97 children were also found retrospectively to be ineligible owing to absence of urinary and/or constitutional symptoms, previous (duplicate) recruitment or invalid consent, or because someone other than the parent had (inappropriately) consented.

### Withdrawals

There were 20 withdrawals during the course of the study, of which nine consented to use the data already collected. However, in four cases data were inadvertently removed from the database instead of retaining data collected up to the point of withdrawal. Therefore, all data were withdrawn for 15 children, whereas five withdrawals were made with the retention of data already collected. The most common reason for withdrawal was the parent changing his or her mind about the child participating in the study.

### Serious adverse events

Serious adverse events were categorised as 'related' if additional information was available for the responsible clinician as a result of the child being in the DUTY study, or if the admission to hospital was delayed or altered as a result of the child being recruited to the study. For example, if a urine sample would not have routinely been collected had the child not been recruited, but the urine sample results had provided information that changed the way the GP would have normally diagnosed or treated the child, the SAE would have been categorised as related. Conversely, SAEs were categorised as 'unrelated' if a hospital admission was not affected by study participation.

Serious adverse events were also reported as 'expected' or 'unexpected', with classification of 'expected' used if hospitalisation was consistent with the study population of acutely unwell children presenting to primary care, and 'unexpected' if the SAE deviated from what would be expected in that population. The co-chief investigators reviewed all SAEs to reach consensus on these classifications.

A total of 79 of 7163 (1.1%) children from the entire study were reported as being hospitalised, that is having a SAE, three of whom were admitted for UTI. In Bristol there were 36 (1.2%) of the 2947 children recruited (45.6% of all SAEs), from the Cardiff centre 19 of 1768 (1.1%, 24.1% of all), from London centre 17 of 1435 (1.2%, 21.5% of all), and from Southampton centre 4 of 1013 (0.4%, 5.1% of all). The majority of these were expected.

Three (3.7%) SAEs were classified as being related to taking part in DUTY as a result of dipstick test results. Two were diagnosed by the responsible clinician as having UTI. One child was referred for an ultrasound scan that showed bilateral nephrocalcification and was subsequently referred to a paediatrician. The second was admitted to hospital for intravenous antibiotics for treatment of the UTI and the third child was admitted for a respiratory infection and a possible diagnosis of diabetes.

One adverse event was reported, where the parent of a child recruited to DUTY concluded that the nappy pad inserted into the child's nappy had caused a nappy rash. The nappy pad was removed, and though a urine sample was retrieved it was dipstick tested only and not sent to the laboratories.

### Positive unintended consequences of the DUTY study

One child was diagnosed with type I diabetes as a direct consequence of taking part in the DUTY study. This was as a result of the dipstick showing high glucose levels in combination with their medical history. Type I diabetes can be missed as the symptoms may often be missed in primary care.

### Urine collection

Figure 4 shows that urine samples were collected from 6390 (89%) children, sent to the NHS (priority) laboratory in 6337 (99%) and the research laboratory in 5264 (82%) of cases. A small proportion of samples were obtained ( $n = 53$ , 0.8%) but not sent to either laboratory; these were dipstick tested at the primary care site only.

Of the urine samples recorded as being sent to NHS laboratories, 95 (2%) were not received, 162 (2.5%) were not cultured and the results from one urine sample were missing. Of the samples sent to the research laboratory, 33 (1%) were not received and 124 (2%) were not cultured, mainly due to too little volume of urine for the laboratory to process, or because samples had leaked in transit. Therefore, of the 6390 urine samples collected from DUTY participants, results were available for 6079 (95%) from NHS laboratories and 5107 (80%) results from the research laboratory. Of these, 4910 (76.8%) were available from both laboratories.

### Urine collection methods

Of the 6390 urine samples collected, 3721 (58.2%) were collected in the surgery and 2632 (41.2%) were collected in the child's home and returned to the surgery or collected by the RN/CSOs. In 37 cases, no data were recorded on where the sample was collected.

Figure 7 shows the method of urine collection, which included the DUTY-preferred method of clean catch which accounts for just under half (47.5%) of the samples collected ( $n = 3036$ ), while 50.2% were collected using our alternative method of Newcastle nappy pads ( $n = 3205$ ). A small proportion of samples were collected by urine bag ( $n = 100$ , 1.6%), with a further 49 cases where the method of collection was not recorded (0.8%).

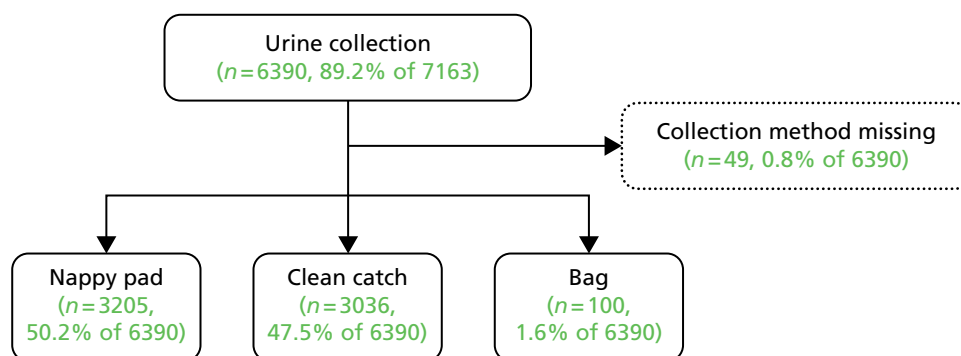


FIGURE 7 Urine collection methods.

## Microbiological culture results

The results presented in this section represent those data entered by the local NHS laboratories and the research laboratory onto the DUTY online database (see *Appendix 7*).

*Table 10* shows the urine culture results as reported by the research laboratory. The prevalence for the pure/predominant growth at  $\geq 10^5$  CFU/ml was 3.2%, lower than the local NHS laboratories, which reported 7.8% of samples with  $\geq 10^5$  CFU/ml pure/predominant growth.

Other differences between the research and local NHS laboratory categories of bacterial growth include higher rates of growth in the local NHS than research laboratories of  $\geq 10^5$  CFU/ml – mixed growth (> 2 species) at 15.6% versus 6.4%; much lower rates of growth in the local NHS than research laboratories of  $10^3$  to  $10^5$  CFU/ml – mixed growth (two species) at 1.5% versus 14.5%; lower rates of growth in the local NHS than research laboratories of  $10^3$  to  $10^5$  CFU/ml – mixed growth (> 2 species) at 7.4% versus 18.8%; and much higher rates of reported 'no growth' in the local NHS than research laboratories at 31.7% versus 8.6%.

**TABLE 10** Frequency of urine sample growth in each of the research and NHS laboratory result categories

Result categories	Research laboratory, <i>n</i>	Research laboratory, %	NHS laboratory, <i>n</i>	NHS laboratory, %
$\geq 10^5$ to CFU/ml: pure/predominant	169	3.2	488	7.8
$\geq 10^5$ CFU/ml: mixed growth (two species)	185	3.5	72	1.2
$\geq 10^5$ CFU/ml: mixed growth (> 2 species)	337	6.4	976	15.6
$10^3$ to $10^5$ CFU/ml: pure/predominant	361	6.9	199	3.2
$10^3$ to $10^5$ CFU/ml: mixed growth (two species)	757	14.5	91	1.5
$10^3$ to $10^5$ CFU/ml: mixed growth (> 2 species)	982	18.8	460	7.4
$< 10^3$ CFU/ml: pure/predominant	56	1.1	15	0.2
$< 10^3$ CFU/ml: mixed growth (two species)	78	1.5	15	0.2
$< 10^3$ CFU/ml: mixed growth (> 2 species)	48	0.9	69	1.1
No growth	452	8.6	1978	31.7
No significant growth	1682	32.2	1716	27.5
Unable to process urine	124	2.4	73	1.2
Culture not performed	N/A		89	1.4
Missing data	N/A		1	0.0
Total	5231	100	6242	100

N/A, not applicable.

## Participant follow-up

This section presents data on the number of children who were selected for follow-up at 14 days via interview and 3 months via a primary care medical notes review at their GP surgery. The proportional representation of follow-up cases, based on the urine culture result categories used by the database algorithm to select participants for follow-up, is also described.

### Proportional selection for follow-up

The algorithm for proportional selection of participants for follow-up was set up to ensure that children with a positive culture as defined by the NHS laboratory or research laboratory at  $\geq 10^5$  CFU/ml were followed up (see *Chapter 2, Patient follow-up, Day-14 and 3-month selection*). For a short period, the algorithm was not working as intended, in that if the results from the laboratories had different levels of growth, the lower category was chosen, meaning that the higher ( $\geq 10^5$  CFU/ml) result category was under-represented. This error was detected too late to follow up 277 children at 14 days, but 3-month notes review data were collected for these children.

Table 11 shows the proportion of participants that were selected and the number of 14-day interviews that were completed for each of the result categories (omitting the 277 urine samples that failed to be selected by the algorithm). Of the 13 urine sample result categories, nine were near the target proportion set, or on, or above it. There were 6390 children for whom a urine sample had been taken; 6314 had results entered into the DUTY database during the course of the study and 1276 (20%) children were successfully selected for follow-up.

**TABLE 11** Proportions of cases selected for follow-up in each of the urine sample result categories and numbers completed

Highest result category from either laboratory	Data in result category, <i>n</i>	Target for selection, %	Selected for interview, <i>n</i> (%)	Interview completed, <i>n</i> (%)
$10^5$ CFU/ml: pure/predominant	855	100	579 (67.7)	413 (71.3)
$\geq 10^5$ CFU/ml: mixed growth (two species)	132	20	42 (31.8)	30 (71.4)
$\geq 10^5$ CFU/ml: mixed growth (> 2 species)	866	20	132 (15.2)	89 (67.4)
$10^3$ to $10^5$ CFU/ml: pure/predominant	1040	20	192 (18.5)	149 (77.6)
$10^3$ to $10^5$ CFU/ml: mixed growth (two species)	283	10	28 (9.9)	19 (67.9)
$10^3$ to $10^5$ CFU/ml: mixed growth (> 2 species)	508	10	41 (8.1)	26 (63.4)
$< 10^3$ CFU/ml: pure/predominant	59	10	10 (16.9)	4 (40.0)
$< 10^3$ CFU/ml: mixed growth (two species)	34	10	5 (14.7)	2 (40.0)
$< 10^3$ CFU/ml: mixed growth (> 2 species)	37	10	4 (10.8)	3 (75.0)
No growth	1626	10	176 (10.8)	125 (71.0)
Culture not performed	40	10	1 (2.5)	0 (0.0)
Unable to process urine	74	10	2 (2.7)	2 (100.0)
No significant growth	760	10	64 (8.4)	56 (87.5)
Total with urine sample	6314		1276 (20.2)	918 (71.9)

Omits 277 samples missed for selection at 14-day follow-up by the algorithm.

All of the pure/predominant samples of  $\geq 10^5$  CFU/ml should have been selected by the algorithm for follow-up; however, only 579 were actually selected, giving a selection rate of 67.7%. There were 413 interviews completed, giving a completion rate of 71.3% for this category.

### Fourteen-day follow-up

There were 1276 participants who were selected for 14-day interviews, for whom 918 (71.9%) were successfully completed. *Table 12* shows the completion rates across the four study centres.

Of the 1276 participants successfully selected for follow-up, interviews could not be completed for 358 (28.1%). The majority of cases ( $n = 164$ , 45.8%) were missed because the parent could not be contacted, while in 129 cases (30.6%) no reason was recorded, and a further 13 interviews (3.6%) were not completed for other reasons (e.g. parent refused to complete the interview or the participant was withdrawn).

The majority of the 14-day follow-ups were completed by telephone interview. However, postal questionnaires were sent if, after three or four attempts at contacting the parents by telephone were made, no successful contact resulted. Postal questionnaires were implemented from September 2011 onwards. In total, 69 14-day follow-ups (7.5%) were completed by postal questionnaire. We can only estimate the response rate of the postal questionnaires, as only three of the centres recorded the number of postal questionnaires sent out. With the information we do have from those centres, the collective response rate for the whole of the study was 45.9%. The Cardiff centre had a response rate of 44.2% (19 out of 43 questionnaires returned), London had a response rate of 31.3% (10 out of 32 questionnaires returned) and Southampton had a 100% response rate (10 out of 10 questionnaires returned). The Bristol centre had 32 questionnaires returned. Fifty-two postal questionnaires were not returned (14.5%).

### Three-month notes review

Of the 1553 children selected for follow-up, 3-month note reviews were completed for 1542 (99.3%) (*Table 13*). Only 11 3-month note reviews were not completed, the main reason being that the child had left the surgery. In the London centre, reviews could not be completed in instances where the child was recruited outside their usual GP practice (e.g. WIC or CED) and no PCT R&D approval was in place for the recruited GP practice.

**TABLE 12** Day 14 follow-up interviews completed for the four study centres

14-day follow-up	Bristol, <i>n</i> (%)	Cardiff, <i>n</i> (%)	London, <i>n</i> (%)	Southampton, <i>n</i> (%)	Total, <i>N</i> (%)
Selected	541 (42.4)	353 (27.6)	199 (15.6)	184 (14.4)	1276
Completed	419 (77.4)	243 (68.8)	125 (63.8)	131 (71.2)	918 (71.9)
Not completed	122 (22.6)	110 (31.2)	74 (36.2)	53 (28.8)	358 (28.1)

**TABLE 13** Three-month note reviews completion by study centres

Three-month note reviews	Bristol, <i>n</i>	Cardiff, <i>n</i>	London, <i>n</i>	Southampton, <i>n</i>	Total, <i>N</i> (%)
Selected	673	388	255	237	1553 (100.0)
Completed	672	384	249	237	1542 (99.3)
Not completed	1	4	6	0	11 (0.7)

## Participant characteristics: descriptive statistics

This section describes the characteristics of participants as recorded on sections 2 to 5 of the CRF (see *Appendix 4*). Descriptive statistics have been presented for the overall sample.

### Description of study population

Table 14 shows the demographic characteristics of the 7163 recruited children. Approximately half of the children were aged < 2 years with equal numbers of boys and girls (49.2% and 50.8%, respectively).

**TABLE 14** Demographic data (from section 2 of the CRF) for the study population overall

Characteristics	n (%)
Age	
Less than 6 months	650 (9.1)
6 months to less than 12 months	1140 (15.9)
1 year to less than 2 years	1682 (23.5)
2 years to less than 3 years	1348 (18.8)
3 years to less than 4 years	1333 (18.6)
4 years and over	1010 (14.1)
Gender	
Male	3526 (49.2)
Female	3637 (50.8)
Ethnicity groupings	
White	5895 (82.3)
Mixed	371 (5.2)
Asian	301 (4.2)
Black	471 (6.6)
Other	36 (0.5)
Missing	89 (1.2)
Highest parental level of qualification	
Degree (or equivalent)	1829 (25.5)
Diploma (or equivalent)	971 (13.6)
'A' level	742 (10.4)
GCSE/'O' level	1710 (23.9)
Other	0 (0)
None	400 (5.6)
Missing	1511 (21.1)
Cost of living	
Find it a strain to get by week to week	381 (5.3)
Have to be careful about money	2894 (40.4)
Able to manage without much difficulty	1806 (25.2)
Quite comfortably off	614 (8.6)
Missing	1468 (20.5)

GCSE, General Certificate of Secondary Education.

The sample was predominantly white (82.3%), with one-quarter of the parents educated to degree level or equivalent, and nearly another quarter educated to GCSE level or equivalent. The financial status of parents was also assessed by asking a question regarding the cost of living, where 40.4% reported they had to be careful about money, and almost one-quarter said that they were able to manage without much difficulty.

The comparison of DUTY recruits with the national census data for age, sex and ethnicity can be seen in *Table 15*. This shows that the children recruited to DUTY were broadly similar in age to the national profile, bearing in mind that children recruited to DUTY were presenting unwell to primary care.

### Description of presenting symptoms and signs

*Table 16* demonstrates the frequencies of the symptoms and signs categories (see *Appendix 4*, section 3 of the case report form) for all 7163 children who were recruited to the DUTY study. The children had been unwell for a median of 4 days (including the day of recruitment) before consulting at the primary care site with their acute illness. In 5261 (73%) children, it was the first time parents had consulted their doctor or nurse with the illness episode. The most prevalent sign and symptom was that the 'child was not themselves', with this being reported as a moderate/severe problem in 4682 (65.4%) participants. The second most prevalent sign/symptom was confusion/disorientation, which was reported as a moderate/severe problem in 4628 (64.6%) participants. Fever at any time during the illness was a moderate/severe problem in 3803 (53.1%) participants. Other most prevalent presenting symptoms and signs, reported as a moderate/severe problem, included cough ( $n = 3771$ , 52.6%), blocked or runny nose ( $n = 3695$ , 51.6%), refused feeds/eating less than normal ( $n = 3670$ , 51.2%), and fever at any time during the past 24 hours ( $n = 3057$ , 42.7%).

**TABLE 15** Comparison of DUTY recruits with national census data

Characteristics	From Census 2011, <i>n</i> (%)	From DUTY, <i>n</i> (%)
Age <sup>a</sup>		
Less than 1 year	711,529 (20.3)	1790 (25.0)
1 year to less than 2 years	704,155 (20.1)	1682 (23.5)
2 years to less than 3 years	698,777 (20.0)	1348 (18.8)
3 years to less than 4 years	699,399 (20.0)	1333 (18.6)
4 years to less than 5 years	682,890 (19.5)	1010 (14.1)
Gender <sup>b</sup>		
Male	27,573,376 (49.2)	3526 (49.2)
Female	28,502,536 (50.8)	3637 (50.8)
Ethnicity groupings <sup>b</sup>		
White	48,209,395 (86.0)	5895 (82.3)
Mixed	1,224,400 (2.2)	371 (5.2)
Asian	4,213,531 (7.5)	301 (4.2)
Black	1,864,890 (3.3)	471 (6.6)
Other	563,696 (1.0)	36 (0.5)
Missing	0 (0.0)	89 (1.2)

a Census 2011: those aged under 5 years old in England and Wales.

b Census 2011: all England and Wales.

TABLE 16 Symptoms and signs

Symptoms and signs	<i>n</i>	Median	IQR
How many days (including today) has your child been unwell)?	7163	4.0	2.0–7.0
		<i>n</i> (%)	
Compared with yesterday, is your child the same, better or worse?			
Same		2799 (39.1)	
Better		1445 (20.2)	
Worse		2902 (40.5)	
Missing		17 (0.2)	
Please rate your overall impression of your child's current illness when it is at its worst			
0		47 (0.7)	
1		187 (2.6)	
2		339 (4.7)	
3		684 (9.5)	
4		941 (13.1)	
5		1326 (18.5)	
6		1250 (17.5)	
7		1277 (17.8)	
8		760 (10.6)	
9		216 (3.0)	
10		121 (1.7)	
Missing		15 (0.2)	
Child not themselves			
No problem		580 (8.1)	
Slight problem		1885 (26.3)	
Moderate problem		3250 (45.4)	
Severe problem		1432 (20.0)	
Missing		16 (0.2)	
Confused or disorientated			
No problem		1138 (15.9)	
Slight problem		1381 (19.3)	
Moderate problem		2458 (34.3)	
Severe problem		2170 (30.3)	
Missing		16 (0.2)	
Disturbed sleep			
No problem		5952 (83.1)	
Slight problem		709 (9.9)	
Moderate problem		388 (5.4)	
Severe problem		98 (1.4)	
Missing		16 (0.2)	



TABLE 16 Symptoms and signs (continued)

Symptoms and signs	<i>n</i>	Median	IQR
Fever at any time during this illness			
No problem		1802 (25.2)	
Slight problem		1542 (21.5)	
Moderate problem		2346 (32.8)	
Severe problem		1457 (20.3)	
Missing		16 (0.2)	
Fever now or in the past 24 hours			
No problem		2796 (39.0)	
Slight problem		1294 (18.1)	
Moderate problem		1925 (26.9)	
Severe problem		1132 (15.8)	
Missing		16 (0.2)	
Chills or shivering			
No problem		5149 (71.9)	
Slight problem		909 (12.7)	
Moderate problem		811 (11.3)	
Severe problem		278 (3.9)	
Missing		16 (0.2)	
New generalised rash with this illness			
No problem		5777 (80.7)	
Slight problem		678 (9.5)	
Moderate problem		486 (6.8)	
Severe problem		206 (2.9)	
Missing		16 (0.2)	
Nappy rash or similar			
No problem		5912 (82.5)	
Slight problem		707 (9.9)	
Moderate problem		372 (5.2)	
Severe problem		156 (2.2)	
Missing		0 (0.0)	
Muscle aches or pain all over			
No problem		2389 (33.4)	
Slight problem		540 (7.5)	
Moderate problem		429 (6.0)	
Severe problem		91 (1.3)	
Missing		3714 (51.8)	

continued

TABLE 16 Symptoms and signs (continued)

Symptoms and signs	<i>n</i>	Median	IQR
Headaches			
No problem		2351 (32.8)	
Slight problem		494 (6.9)	
Moderate problem		419 (5.8)	
Severe problem		89 (1.2)	
Missing		3810 (53.2)	
Refused feeds/eating less than normal			
No problem		1584 (22.1)	
Slight problem		1893 (26.4)	
Moderate problem		2429 (33.9)	
Severe problem		1241 (17.3)	
Missing		16 (0.2)	
Poor weight gain or weight loss			
No problem		4242 (59.2)	
Slight problem		514 (7.2)	
Moderate problem		247 (3.4)	
Severe problem		71 (1.0)	
Missing		2089 (29.2)	
Vomiting			
No problem		4467 (62.4)	
Slight problem		1300 (18.1)	
Moderate problem		903 (12.6)	
Severe problem		477 (6.7)	
Missing		16 (0.2)	
Diarrhoea (at any time)			
No problem		5074 (70.8)	
Slight problem		1058 (14.8)	
Moderate problem		663 (9.3)	
Severe problem		352 (4.9)	
Missing		16 (0.2)	
Diarrhoea (in the past 24 hours)			
No problem		5700 (79.6)	
Slight problem		697 (9.7)	
Moderate problem		474 (6.6)	
Severe problem		276 (3.9)	
Missing		16 (0.2)	

TABLE 16 Symptoms and signs (continued)

Symptoms and signs	<i>n</i>	Median	IQR
Constipation in the last week			
No problem		5824 (81.3)	
Slight problem		764 (10.7)	
Moderate problem		419 (5.8)	
Severe problem		140 (2.0)	
Missing		16 (0.2)	
Abdominal pain/tummy ache/pulling up legs			
No problem		4703 (65.7)	
Slight problem		1165 (16.3)	
Moderate problem		964 (13.5)	
Severe problem		315 (4.4)	
Missing		16 (0.2)	
Passing urine more often			
No problem		4149 (57.9)	
Slight problem		487 (6.8)	
Moderate problem		464 (6.5)	
Severe problem		158 (2.2)	
Missing		1905 (26.6)	
Any changes in urine appearance			
No problem		4502 (62.9)	
Slight problem		1027 (14.3)	
Moderate problem		609 (8.5)	
Severe problem		140 (2.0)	
Missing		885 (12.4)	
Pain/crying when passing urine			
No problem		4780 (66.7)	
Slight problem		294 (4.1)	
Moderate problem		190 (2.7)	
Severe problem		94 (1.3)	
Missing		1805 (25.2)	
Day or bed wetting when previously dry			
No problem		2054 (28.7)	
Slight problem		238 (3.3)	
Moderate problem		161 (2.2)	
Severe problem		82 (1.1)	
Missing		320 (4.5)	
Wears nappies day and night		4308 (60.1)	

continued

TABLE 16 Symptoms and signs (continued)

Symptoms and signs	<i>n</i>	Median	IQR
Blocked or runny nose			
No problem		1621 (22.6)	
Slight problem		1831 (25.6)	
Moderate problem		2434 (34.0)	
Severe problem		1261 (17.6)	
Missing		16 (0.2)	
Cough			
No problem		1846 (25.8)	
Slight problem		1530 (21.4)	
Moderate problem		2222 (31.0)	
Severe problem		1549 (21.6)	
Missing		16 (0.2)	
Wheeze			
No problem		4406 (61.5)	
Slight problem		1250 (17.5)	
Moderate problem		1058 (14.8)	
Severe problem		433 (6.0)	
Missing		16 (0.2)	
Short of breath, difficulty breathing or grunting			
No problem		5061 (70.7)	
Slight problem		1082 (15.1)	
Moderate problem		710 (9.9)	
Severe problem		294 (4.1)	
Missing		16 (0.2)	
Chest pains			
No problem		3386 (47.3)	
Slight problem		142 (2.0)	
Moderate problem		77 (1.1)	
Severe problem		24 (0.3)	
Missing		3534 (49.3)	
Earache or holding ear(s)			
No problem		3717 (51.9)	
Slight problem		1003 (14.0)	
Moderate problem		760 (10.6)	
Severe problem		436 (6.1)	
Missing		1247 (17.4)	

TABLE 16 Symptoms and signs (continued)

Symptoms and signs	<i>n</i>	Median	IQR
Sore throat			
No problem		2557 (35.7)	
Slight problem		973 (13.6)	
Moderate problem		896 (12.5)	
Severe problem		355 (5.0)	
Missing		2382 (33.3)	
More unwell than with similar previous illnesses			
No problem		4246 (59.3)	
Slight problem		1104 (15.4)	
Moderate problem		1357 (18.9)	
Severe problem		440 (6.1)	
Missing		16 (0.2)	
Darker urine?			
No		6299 (87.9)	
Yes		764 (10.7)	
Missing		100 (1.4)	
Cloudy urine?			
No		6890 (96.2)	
Yes		173 (2.4)	
Missing		100 (1.4)	
Smelly urine?			
No		5952 (83.1)	
Yes		1111 (15.5)	
Missing		100 (1.4)	
Bloody urine?			
No		7039 (98.3)	
Yes		24 (0.3)	
Missing		100 (1.4)	
Other urine appearance problems?			
No		6980 (97.4)	
Yes		83 (1.2)	
Missing		100 (1.4)	
Any other symptoms?			
No		6117 (85.4)	
Yes		1030 (14.4)	
Missing		16 (0.2)	

continued

TABLE 16 Symptoms and signs (continued)

Symptoms and signs	<i>n</i>	Median	IQR
Does the child have an eye problem?			
No		6894 (96.2)	
Yes		253 (3.5)	
Missing		16 (0.2)	
Does the child have decreased urinary frequency or lower urine volume?			
No		7082 (98.9)	
Yes		65 (0.9)	
Missing		16 (0.2)	
Not counting today, approximately how many times has your child previously consulted a doctor or nurse for this episode of illness?			
0		5261 (73.4)	
1		1157 (16.2)	
2		361 (5.0)	
3		169 (2.4)	
4		78 (1.1)	
5		48 (0.7)	
6		15 (0.2)	
7		14 (0.2)	
8		16 (0.2)	
9		2 (0.0)	
10		13 (0.2)	
10+		9 (0.1)	
Missing		20 (0.3)	
Does your child have any ongoing health problems?			
No		5883 (82.1)	
Yes		1261 (17.6)	
Missing		19 (0.3)	

IQR, interquartile range.

### Description of clinical examination findings

Clinical examinations were performed by either a doctor or a nurse, and the reported findings recorded on the CRF (see *Appendix 4*, section 4 of the case report form) for all recruited patients are presented in *Table 17*.

For the 7163 children recruited, the responsible clinician's global impression score was recorded as between 0 and 4 in over 90% of cases. Hydration and consciousness level were reported as normal in 94.6% and 96.8% of cases, respectively. From a general examination, 68.0% of cases were recorded as normal. Normal findings were recorded by the responsible clinician on examination of the throat, ear, chest and abdomen in 60.5%, 63.0%, 79.6% and 81.2% for recruited children, respectively (see *Table 17*).

**TABLE 17** Clinical examination findings for all 7163 recruited participants

Clinical examination	n (%)
Please give your global impression of the child on a scale of 0–10 <sup>a</sup>	
0	604 (8.4)
1	1825 (25.5)
2	1966 (27.4)
3	1417 (19.8)
4	675 (9.4)
5	284 (4.0)
6	192 (2.7)
7	101 (1.4)
8	35 (0.5)
9	8 (0.1)
10	2 (0.0)
Hydration	
Not examined	21 (0.3)
Normal	6776 (94.6)
Some dehydration	308 (4.3)
Severe dehydration	1 (0.0)
Missing	57 (0.8)
Consciousness level	
Not examined	11 (0.2)
Normal	6933 (96.8)
Drowsy	87 (1.2)
Irritable	74 (1.0)
Missing	58 (0.8)
General examination	
Not examined	22 (0.3)
Normal	4871 (68.0)
Abnormal	2213 (30.9)
Missing	57 (0.8)
Throat examination	
Not examined	1065 (14.9)
Normal	4334 (60.5)
Abnormal	1707 (23.8)
Missing	57 (0.8)
continued	

**TABLE 17** Clinical examination findings for all 7163 recruited participants (*continued*)

Clinical examination	n (%)
Ear examination	
Not examined	843 (11.8)
Normal	4512 (63.0)
Abnormal	1751 (24.4)
Missing	57 (0.8)
Chest examination	
Not examined	483 (6.7)
Normal	5704 (79.6)
Abnormal	919 (12.8)
Missing	57 (0.8)
Abdomen examination	
Not examined	1108 (15.5)
Normal	5819 (81.2)
Abnormal	179 (2.5)
Missing	57 (0.8)

a 0 = completely well, 10 = extremely unwell.

For the 2213 (30.9%) cases with an abnormal general examination the following findings were recorded: pallor (27.8%); flushed (35.1%); jaundice (0.4%); lymphadenopathy (15.6%); distressed (11.5%); rash (7.6%); and other (17.3%). For the 1707 cases with an abnormal throat examination the following findings were recorded: red or inflamed (84.7%); swollen (17.3%); quinsy (10.0%); lymph nodes (1.1%); tonsillitis (1.55); mouth ulcer (1.2%); oral thrush (1.1%); and other (4.2%). For the 1751 with an abnormal ear examination an acute abnormality was recorded in 1576 cases (90.0%) and a chronic abnormality in 4.7%. The main findings recorded were pink (43.9%); red or bulging (42.6%); fluid (1.8%); acute perforation (0.1%); chronic perforation (2.2%); glue ear (0.6%); grommets (0.3%); otitis externa (1.3%); otorrhoea (2.2%); otitis media (3.2%); wax (6.45); and other (3.0%). For the 919 cases with an abnormal chest examination the following findings were recorded: bronchial breathing with distribution (7.95); wheeze (44.5%); crackles (57.8%); recession (8.1%); grunting (1.5%); nasal flaring (0.5%); transmitted sounds with distribution (2.6%); and other (4.4%). For the 179 cases with an abnormal abdomen examination the following findings were recorded: mass or organomegaly (3.4%); loin tenderness (5.6%); suprapubic tenderness (22.9%); other abdominal tenderness (25.7%); genital inflammation or nappy rash (10.1%); and other (33.0%). These data are not included in the report, but are available from the authors on request. *Chapter 5* provides a description of the inclusion of symptoms and signs in the prediction rule statistical analysis (see *Methods, Statistical analysis*) and the results of the rule (see *Results, Clean-catch models* and *Nappy pad models*).

### Medical history

For almost 75% of children, the pregnancy was reported as full term, with over 65% being breastfed for a period of time. In 89 cases, parents reported that they had been told that the child's kidney, bladder or urinary system was abnormal after a pregnancy ultrasound test. A small proportion of boys were reported as having been circumcised (3.5%) (*Table 18*).



**TABLE 18** Parent-reported medical history for the 7163 recruited participants

Medical history	n (%)		
What length was your pregnancy?			
Early (37 weeks or less)	733 (10.2)		
Full term	5260 (73.4)		
Late	1121 (15.6)		
Missing	49 (0.7)		
Was your child breastfed?			
No	2191 (30.6)		
Less than 3 months	1896 (26.5)		
More than 3 months	2995 (41.8)		
Missing	81 (1.1)		
Has your child been circumcised?			
No	3199 (44.7)		
Yes	253 (3.5)		
Missing	74 (1.0)		
Female – N/A	3637 (50.8)		
Were you ever told that your child's kidney, bladder or urinary system was abnormal in any way after a pregnancy ultrasound test?			
No	7053 (98.5)		
Yes	89 (1.2)		
Missing	21 (0.3)		
Is your child taking any medications?			
No	3349 (46.8)		
Yes	3799 (53.0)		
Missing	15 (0.2)		
Does your child use nappies?			
No	1783 (24.9)		
Day	23 (0.3)		
Night	1028 (14.4)		
Day and night	4308 (60.1)		
Missing	21 (0.3)		
	Age of child	N	Median (IQR)
How many nappies has your child used in the last 24 hours?	0	1787	6.0 (5.0–8.0)
	1	1678	6.0 (5.0–7.0)
	2	1340	3.0 (1.0–5.0)
	3	1331	0.0 (0.0–1.0)
	4	1006	0.0 (0.0–0.0)
Number of baths/showers in a week?			
		7114	6.0 (4.0–7.0)

IQR, interquartile range.

For the 3799 cases reported as currently taking medication, paracetamol and ibuprofen were the most commonly reported (76.5% and 31.0%, respectively), and the following were reported for the remainder of cases: laxatives (4.2%); steroid inhaler (4.4%); beta2 agonist inhaler (8.6%); antihistamines (1.9%); cough medicine (4.7%); and other (11.4%).

Parents reported that 4308 (60.1%) children wore nappies both day and night, with a further 1028 (14.4%) wearing nappies in the night only. One-quarter of the DUTY recruits were reported as not using nappies at all.

*Table 19* shows the reported number of cases with a history of any congenital abnormalities of the kidney and urinary tract (CAKUT) diseases, UTI infections, bladder or renal problems in the child or direct family members. The most prevalent reported item was history of UTI infection, both in the child (4.3%) and in the family (14.1%), with maternal history of UTI infection representing the highest proportion of mother-, father- or sibling-reported infection.

For the children in whom clinical observations could be recorded, the mean or median recordings are presented in *Table 20*. 13.6% (or 14.3% of those with a measurement) had an extreme temperature of  $\geq 38^{\circ}\text{C}$  and 3.6% (or 5.4% of those with a measurement) had an extreme oxygen saturation of  $< 94\%$ .

### Clinical diagnoses and antibiotic treatment

Most clinicians did not have a working diagnosis of UTI prior to knowledge of the dipstick result. *Table 21* indicates the number and proportions of different diagnoses made.

*Table 22* shows that a total of 2246 patients were treated with antibiotics (31.3%), and, of these, 86.7% were issued with an immediate prescription. Amoxicillin was the most common antibiotic prescribed by clinicians, accounting for 49.4% of scripts issued ( $n = 1109$ ), followed by trimethoprim at 11.4% ( $n = 256$ ) and penicillin at 9.9% ( $n = 222$ ) (see *Table 22*). When these antibiotics are split up by the diagnoses for which they were prescribed, amoxicillin is generally the most frequent antibiotic. However, when tonsillitis was diagnosed, penicillin was the most commonly prescribed antibiotic, and when UTI, gastroenteritis or a viral illness were diagnosed, trimethoprim was the most frequently prescribed antibiotic.

**TABLE 19** Parent-reported medical history of previous urinary, bladder or renal problems

Medical history	<i>n</i> yes (%)
Does the child have a history of any CAKUT diseases?	45 (0.6)
Does the mother, father or sibling have a history of any CAKUT diseases?	120 (1.7)
Does the child have a history of any urinary, bladder or renal infections?	310 (4.3)
Does the mother, father or sibling have a history of any urinary, bladder or renal infections?	1008 (14.1)
Does the child have a history of non-CAKUT bladder problems including enuresis/dysfunctional voiding problems?	2 (0.0)
Does the mother, father or sibling have a history of non-CAKUT bladder problems including enuresis/dysfunctional voiding problems?	58 (0.8)
Does the child have a history of any non-CAKUT renal problem?	13 (0.2)
Does the mother, father or sibling have a history of any non-CAKUT renal problem?	173 (2.4)
Does the child have a history of any other renal/urinary problem?	11 (0.2)
Does the mother, father or sibling have a history of any other renal/urinary problem?	25 (0.4)

**TABLE 20** Clinical observations recorded for recruited participants

Clinical observations (continuous)	<i>n</i>	Mean	SD
Temperature (°C)	6833	37.05	0.869
Pulse rate (beats per minute)	5610	119.6	20.51
Oxygen saturation (%)	4778	98.0	96.0, 98.0
Respiratory rate (rpm)	5732	28.0	24.0, 32.0
Clinical observations (categorised)	<i>n</i> (%)		
Capillary refill time (seconds)			
< 2	5320 (74.3)		
2 to 5	913 (12.7)		
> 5	13 (0.2)		
Missing	917 (12.8)		
Temperature (°C)			
< 38	5859 (81.8)		
≥ 38	974 (13.6)		
Missing	330 (4.6)		
Oxygen saturation (%)			
< 94	260 (3.6)		
≥ 94	4518 (63.1)		
Missing	2385 (33.3)		

**TABLE 21** Treating clinicians' working diagnosis

What is your working diagnosis?	<i>n</i> (%)
No UTI	6673 (93.2)
UTI	361 (5.0)
UTI + another diagnosis	67 (0.9)
Missing	62 (0.9)

TABLE 22 Antibiotic treatment prescribed by clinical diagnosis

Antibiotic treatment	URTI, n (%)	Chest infection, pneumonia, bronchitis and n (%)	Exacerbation of asthma (infective or non-infective), n (%)	Tonsillitis, n (%)	Otitis media, n (%)	Pharyngitis, n (%)	Gastroenteritis, n (%)	Viral illness, n (%)	Other, n (%)	UTI + another diagnosis, n (%)	Total, N (%)		
Amoxicillin (liquid suspension)	123 (45.4)	307 (72.6)	11 (73.3)	16 (76.2)	48 (17.4)	453 (79.9)	6 (31.6)	36 (14.1)	3 (20.0)	45 (33.1)	55 (24.8)	6 (24.0)	1109 (49.4)
Cefalexin	5 (1.8)	3 (0.7)	0 (0.0)	0 (0.0)	2 (0.7)	2 (0.4)	1 (5.3)	12 (4.7)	1 (6.7)	1 (0.7)	3 (1.4)	1 (4.0)	31 (1.4)
Clarithromycin (liquid suspension)	4 (1.5)	6 (1.4)	0 (0.0)	1 (4.8)	8 (2.9)	4 (0.7)	1 (5.3)	1 (0.4)	0 (0.0)	0 (0.0)	3 (1.4)	0 (0.0)	28 (1.2)
Co-amoxiclav (liquid suspension)	0 (0.0)	5 (1.2)	0 (0.0)	1 (4.8)	5 (1.8)	4 (0.7)	0 (0.0)	8 (3.1)	0 (0.0)	1 (0.7)	16 (7.2)	1 (4.0)	41 (1.8)
Erythromycin (liquid suspension)	8 (3.0)	14 (3.3)	1 (6.7)	0 (0.0)	13 (4.7)	22 (3.9)	2 (10.5)	1 (0.4)	0 (0.0)	2 (1.5)	8 (3.6)	0 (0.0)	71 (3.2)
Flucloxacillin (liquid suspension)	2 (0.7)	0 (0.0)	0 (0.0)	1 (4.8)	0 (0.0)	1 (0.2)	0 (0.0)	4 (1.6)	1 (6.7)	1 (0.7)	25 (11.3)	1 (4.0)	36 (1.6)
Ceftriaxone	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.7)	1 (0.5)	0 (0.0)	2 (0.1)
Nitrofurantoin	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.4)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.0)
Penicillin	17 (6.3)	0 (0.0)	0 (0.0)	0 (0.0)	166 (60.1)	2 (0.4)	5 (26.3)	0 (0.0)	0 (0.0)	10 (7.4)	23 (10.4)	1 (4.0)	222 (9.9)
Trimethoprim	36 (13.3)	1 (0.2)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.2)	0 (0.0)	138 (53.9)	7 (46.7)	51 (37.5)	14 (6.3)	8 (32.0)	256 (11.4)
Other (including topical creams, eye or ear drops)	14 (5.2)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	3 (0.5)	0 (0.0)	4 (2.9)	0 (0.0)	42 (18.9)	0 (0.0)	0 (0.0)	63 (2.8)
Missing	56 (20.7)	82 (19.4)	3 (20.0)	2 (9.5)	34 (12.3)	75 (13.2)	4 (21.1)	51 (19.9)	2 (13.3)	14 (10.3)	22 (9.9)	6 (24.0)	351 (15.6)

URTI, urinary respiratory tract infection.

## Study monitoring

This section describes the monitoring procedures employed in the DUTY study to identify protocol deviations.

Data quality issues were identified in one primary care site, in relation to incomplete CRFs, source data, paper CRF and database data entries. As a result of an internal audit of the site, all data for recruits from that site were removed, and the site was retrained. A further audit was completed after further recruitment had taken place. Similar issues were found in terms of poor data quality, and as such this site was closed to recruitment.

As a result, a whole study-wide audit was undertaken. This involved a sample of 16 study sites selected from all four centres that were audited to (1) ensure that the study was being conducted to protocol standards; (2) identify any protocol deviations; and (3) assess data quality. The sites were selected where at least 20 children had been recruited into the study. For each site, 10 records were randomly selected and audited by either a research manager or an experienced RN.

The audit points included evidence of good clinical practice training; consent being taken accurately (as evidenced by the correct completion of the consent form); the completeness of the CRF (either electronic or paper and the accurate transcription from paper to electronic where both had been completed); evidence of medical attendance in the child's medical notes; and any other areas of concern that had the potential to significantly affect either the data quality or the safety of the child.

Errors were divided into those regarded as (1) minor, such as not changing the day/date, when entering information electronically, by 1 day after recruitment (the database defaulted to the day it was being used), and (2) significant, such as where one form had the urine collection recorded as 'clean catch' on paper and 'nappy pad' on the database or the database recorded the presence of a symptom, such as diarrhoea, where this did not appear on the paper CRF. Significant errors had the potential to adversely affect data quality or the recognition of contamination in a urine sample.

Overall, we found good concordance between medical notes and entry onto the CRF, and also between paper CRFs and the online database. Minor errors accounted for 5.2% and major errors for 0.2% of the data examined. In conclusion, minor errors were typically in the form of miskeying dates, as described above, and under-reporting eligibility, and only two significant errors were found in the review. In the data cleaning these typographical errors were able to be corrected.



# Chapter 4 Microbiological diagnosis of urinary tract infection by NHS and research laboratories

## Introduction

Laboratory diagnosis is based on colony counts following culture, which reflect the concentration of organisms in urine and hence the likelihood that the bacteria grown arise from a UTI rather than contamination. UTI is typically caused by a single organism that is present in a high concentration, usually  $\geq 10^5$  CFU/ml.<sup>81</sup> However, laboratory guidelines differ regarding the nature and extent of bacterial growth required to confirm UTI.<sup>82,83</sup> NICE guidelines do not provide a definitive threshold for diagnosing UTI on culture but provide advice about the level of bacterial growth in relation to symptoms and signs.<sup>2</sup> Based on paediatric data, others have since proposed higher thresholds, for example  $\geq 10^6$  CFU/ml.<sup>80</sup> It is not surprising, then, that laboratory guidelines differ regarding the nature and extent of bacterial growth required to confirm UTI according to the patient's age, symptoms and urine collection method.<sup>82,83</sup> Although NHS laboratories in the UK follow the UK Standards for Microbiological Investigation for examination of urine, application of the method varies between laboratories.

For this reason, and when sufficient urine was available, we sent a fraction of the urine to a single research laboratory where it was to be processed by a small number of technicians and according to a SOP. Although we wanted the algorithm results to be generalisable to current NHS practice, we did not know whether the NHS or research laboratory was providing the more reliable and accurate results.

Therefore, the aim of this chapter is to report the comparison of reliability and accuracy of UTI laboratory diagnosis between routine NHS laboratories and a single research laboratory. We investigated associations of pre-specified symptoms and signs related to UTI, and urine dipstick test results identified from the literature, with different laboratory definitions of urine culture positivity.

## Methods

### *Urine collection and sample processing*

Urine samples were obtained by clean catch, where possible, for children who were toilet trained or for whom the parent was happy to attempt such collection. A full description of the urine collection methods is given in *Chapter 2*.

NHS laboratories processed urine samples using their local SOPs and recorded data according to local reporting procedures, including, where possible, quantifying bacterial growth (as  $< 10^3$ ;  $10^3$  to  $< 10^5$ ; or  $\geq 10^5$  CFU/ml), purity of growth (pure/predominant; mixed growth two species; mixed growth  $> 2$  species), organism speciation for up to two species, and microscopy for white and red cells. The research laboratory quantified absolute colony counts (range  $10^1$ – $10^{10}$  CFU/ml) for all organisms present and established species ID for organisms present at  $\geq 10^3$  CFU/ml.

### *Statistical analysis*

Results reported by the NHS and research laboratories were initially classified based on extent and purity of growth and whether or not the species grown was a uropathogen, defined as a member of the Enterobacteriaceae group. For NHS laboratory results, samples for which the laboratory reported pure/predominant growth of a uropathogen at  $\geq 10^5$  CFU/ml were considered UTI positive. For research laboratory results, samples were considered UTI positive if there was growth of  $\geq 10^5$  CFU/ml of a single uropathogen ('pure growth') or growth of  $\geq 10^5$  CFU/ml of a uropathogen with  $\geq 3 \log_{10}$  difference

between the growth of this and the next species ('predominant growth'). Agreement between laboratories was assessed using kappa statistics, with analyses additionally stratified by urine collection method (clean catch or nappy pad) and by age (0 to < 2, 2 to < 3 and 3 to < 5 years).

Analyses were restricted to samples with a result from both the NHS and research laboratories. Samples classified as UTI positive by both NHS and research laboratories results were denoted 'Agree UTI positive (step 1)'. Where there was disagreement between the NHS and research laboratories, we considered whether or not the combined evidence was consistent with a UTI. When the result classified as negative in one of the laboratories met either of the two conditions (1) growth of  $\geq 10^5$  CFU/ml of a uropathogen with lesser growth of at most one other species, or (2) pure/predominant growth of  $10^3$  to  $10^5$  CFU/ml of a uropathogen, samples were denoted 'Agree UTI positive (step 2)'.

A priori (before examining their associations with different definitions of microbiological positivity from different laboratories), we selected a small number of variables (symptoms, signs and dipstick test results) reported in the literature to be clearly related to presence of a UTI.<sup>116</sup> Those thought suitable for all ages and collection methods were urinary symptoms (pain/crying when passing urine, passing urine more often, changes in urine appearance); temperature  $\geq 39$  °C, and nitrite- or leucocyte-positive results from urine dipstick tests. Additional symptoms and signs thought to be relevant mainly for older children were daytime or bed wetting when previously dry, and a history of UTI. Most parent-reported symptoms were recorded using categories 'no', 'slight', 'moderate' or 'severe problem'; we decided a priori (based only on inspection of symptom frequencies) to dichotomise these into 'no or slight problem' and 'moderate or severe problem'. Responses of 'not known' were coded with 'no or slight problem'. A small number of samples for which there were missing data on most or all urine symptoms (five samples), or for which urine dipstick tests were not available (12 samples), or there was missing information on prior infection (three samples) were excluded. Missing data on temperature (204 children) were coded as < 39 °C. The remaining, sporadic missing values (on five children) were coded as 'no or slight problem'.

We used separate logistic regression models to quantify associations of the selected variables with UTI positivity in the NHS and research laboratories, and for the different outcomes 'agree UTI', 'disagree (NHS positive, research negative)' and 'disagree (NHS negative, research positive)', all compared with 'agree UTI negative'. We plotted receiver operating characteristic (ROC) curves and AUROCs to quantify diagnostic utility. The maximum value of the AUROC is 1 (perfect prediction) while a value of 0.5 corresponds to no association with any predictor. The symptoms 'day or bed wetting when previously dry' and 'history of UTI' were recorded in too few of the children who provided nappy pad samples to permit their associations with microbiology results to be examined. For these children, therefore, we examined associations of the remaining six signs, symptoms and dipstick results with microbiology results. For children who provided a clean-catch sample, we fitted two sets of logistic regression models, one including all eight symptoms and dipstick results (the 'eight variable model') and the other, for comparability, including the same six symptoms and results as for children providing nappy pad samples (the 'six variable model').

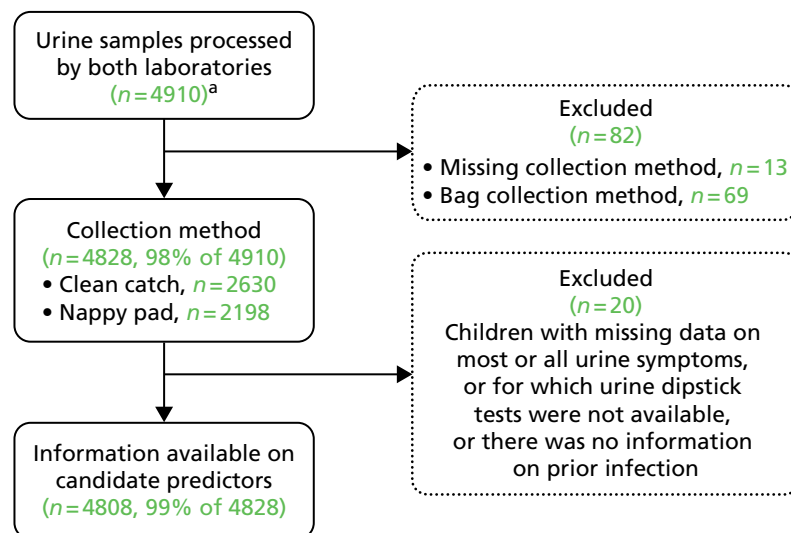
In sensitivity analyses we (1) stratified by age (< 3 and  $\geq 3$  years); (2) allowed for 'not known' categories in the questionnaire responses for variables for which such responses occurred sufficiently frequently; (3) stratified according to whether samples were coded as 'agree UTI positive' at step 1 or step 2; (4) stratified by whether the sample was collected at the surgery or at home; (5) stratified by time between recruitment and laboratory sample receipt (< 24 hours and  $\geq 24$  hours); (6) stratified NHS laboratory results according to extent of pure/predominant growth ( $\geq 10^5$ ,  $\geq 10^3$  to <  $10^5$  CFU/ml); (7) stratified research laboratory results according to extent of pure/predominant growth ( $\geq 10^7$ ;  $\geq 10^6$  to <  $10^7$ ;  $\geq 10^5$  to <  $10^6$ ;  $\geq 10^4$  to <  $10^5$ ; and  $\geq 10^3$  to <  $10^4$  CFU/ml); (8) stratified NHS and research laboratory results according to whether the WBC count was < 30 or  $\geq 30$ /mm<sup>3</sup>; (9) stratified research laboratory results according to whether growth was pure or predominant, and (10) as the research laboratory received only the urine available after the priority fraction was placed in the NHS collection sample, we stratified by research laboratory urine volume (using the cut-point of median urine Monovette weight for clean-catch and nappy pad samples). All analyses were carried out using Stata version 12 (StataCorp LP, College Station, TX, USA).



## Results

Of the 7163 children included in the study, 6241 provided a urine sample using the clean-catch or nappy pad collection methods, for whom 5945 and 5071 culture results were available from the NHS and research laboratories, respectively. A total of 4828 children had results from both laboratories, of whom 4808 had information available on candidate predictors. [The starting point for the number of urine results in this chapter was determined by the use of the individuals with cultured specimens in both laboratories (see box B in *Figure 4*).] The detail of the urine samples in this chapter is shown in *Figure 8*. Most children (4543, 94.5%) were recruited from GP surgeries (*Table 23*).

There were approximately equal numbers of boys and girls. Of 2884 children aged < 3 years, urine samples for 758 (26.3%) were collected using clean catch. By contrast, 1861 (96.7%) of 1924 children aged 3–5 years provided a clean-catch sample. Among children aged < 3 years, samples were obtained in the surgery in 1470 cases (51.0%), compared with 1477 (76.8%) among children aged 3–5 years. Parents reported the following symptoms as a moderate or severe problem: pain or crying when passing urine in 217 (4.5%) children; passing urine more often in 484 (10.1%); day or bed wetting when previously dry in 209 (4.3%); and a change in urine appearance in 523 (10.9%). The symptoms that were most commonly reported as 'not known' among children aged < 3 years were pain or crying when passing urine (980, 34.0%) and passing urine more often (1038, 40.0%). A history of UTI was reported in 221 (4.6%) children, 140 of whom were aged  $\geq 3$  years. Only 185 (3.8%) children had a temperature  $\geq 39$  °C. Dipstick urine results were positive for nitrites in 416 (8.7%) and for leucocytes in 670 (13.9%) children. Both nitrite (12.9% compared with 2.2%) and leucocyte (16.0% compared with 10.8%) positivity were more common in children aged < 3 years than in children aged  $\geq 3$  years. Monovette weights (grams) were available for all but 22 of the research laboratory urine samples. Median weight was 16.23 g (IQR 13.67–17.03 g) for clean-catch samples, and 11.08 g (IQR 9.53–13.26 g) for nappy pad samples.



**FIGURE 8** Flow of participants in microbiological diagnosis of UTI in young children. a, See *Figure 4* for how this number was reached.

**TABLE 23** Characteristics of children and urine samples collected via clean catch or nappy pads, for the 4808 children with both NHS laboratory and research laboratory results

Variable	Category	Age, <i>n</i> (%)	
		< 3 years	3–5 years
Gender	Male	1439 (49.9)	919 (47.8)
	Female	1445 (50.1)	1005 (52.2)
Age (years)	0 to < 1	1016 (35.2)	0
	1 to < 2	942 (32.7)	0
	2 to < 3	926 (32.1)	0
	3 to < 4	0	1099 (57.1)
	4 to < 5	0	825 (42.9)
Ethnicity	White	2429 (84.2)	1575 (81.9)
	Non-white	431 (14.9)	328 (17.1)
	Not known	24 (0.8)	21 (1.1)
Recruitment site	GP surgery	2716 (94.2)	1827 (95.0)
	Emergency department	128 (4.4)	66 (3.4)
	WIC	40 (1.4)	31 (1.6)
Sample method	Clean catch	758 (26.3)	1861 (96.7)
	Nappy pad	2126 (73.7)	63 (3.3)
Location of sample collection	Surgery	1470 (51.0)	1477 (76.8)
	Home	1414 (49.0)	447 (23.2)
Time between recruitment and laboratory sample receipt	NHS < 24 hours	2045 (70.9)	1432 (74.4)
	NHS ≥ 24 hours	839 (29.1)	492 (25.6)
	RL < 24 hours	816 (28.3)	608 (31.6)
	RL ≥ 24 hours	2068 (71.7)	1316 (68.4)
Pain/crying when passing urine	No or slight problem	1812 (62.8)	1734 (90.1)
	Moderate or severe problem	92 (3.2)	125 (6.5)
	Not known	980 (34.0)	65 (3.4)
Passing urine more often	No or slight problem	1618 (56.1)	1604 (83.4)
	Moderate or severe problem	228 (7.9)	256 (13.3)
	Not known	1038 (40.0)	64 (3.3)
Changes in urine appearance	No or slight problem	2206 (76.5)	1539 (80.0)
	Moderate or severe problem	297 (10.3)	226 (11.8)
	Not known	381 (13.2)	159 (8.3)
Day or bed wetting when previously dry	No or slight problem	364 (12.6)	1551 (80.6)
	Moderate or severe problem	45 (1.6)	164 (8.5)
	Wears nappies day and night	2377 (82.4)	70 (3.6)
	Not known	98 (3.4)	139 (7.2)

**TABLE 23** Characteristics of children and urine samples collected via clean catch or nappy pads, for the 4808 children with both NHS laboratory and research laboratory results (*continued*)

Variable	Category	Age, n (%)	
		< 3 years	3–5 years
History of UTI	No	2699 (93.6)	1708 (88.8)
	Yes	81 (2.8)	140 (7.3)
	Not known	104 (3.6)	76 (4.0)
Temperature	< 39 °C	2780 (96.4)	1843 (95.8)
	≥ 39 °C	104 (3.6)	81 (4.2)
Urine dipstick nitrite	Negative	2511 (87.1)	1881 (97.8)
	Positive	373 (12.9)	43 (2.2)
Urine dipstick leucocytes	Negative/trace	2423 (84.0)	1715 (89.1)
	Positive	461 (16.0)	209 (10.8)
NHS laboratory result	Negative	2695 (93.5)	1862 (96.8)
	Positive	189 (6.6)	62 (3.2)
RL result	Negative	2833 (98.2)	1887 (98.1)
	Positive	51 (1.8)	37 (1.9)

RL, research laboratory.

A total of 251 (5.2%) and 88 (1.8%) samples were classified as positive according to the NHS and research laboratory result, respectively. The causative organism distributions were similar between laboratories: in the NHS, *E. coli* 71.7%, unidentified coliforms 19.5%, other coliforms 2.8%, *Proteus* spp. 6.0%; and in the research laboratory, *E. coli* 84.1%, other coliform (*Klebsiella* spp., *Enterobacter* spp., *Serratia* spp., *Citrobacter* spp., *Morganella* spp.) 10.2%, *Proteus* spp. 5.7%. NHS laboratory positivity was more common (6.6%) in children aged < 3 years than in those aged 3–5 years (3.2%). By contrast, rates of research laboratory positivity were similar in these age groups (1.8% and 1.9%, respectively). Only 64 (1.3%) samples were positive in both laboratories and coded as 'agree UTI positive (step 1)'. In 187 (3.9%) the NHS laboratory result was positive but the research laboratory result negative, while in 24 (0.5%) the research laboratory result was positive but the NHS laboratory result negative (*Table 24*).

Overall agreement between the NHS and research laboratories was moderate ( $\kappa = 0.36$ ; 95% CI 0.29 to 0.43; see *Table 24*). Agreement was better for clean-catch samples (0.54; 95% CI 0.45 to 0.63) than for nappy pads (0.20; 95% CI 0.12 to 0.28). For children aged ≥ 3 years, too few nappy pad samples were available to allow assessment of reliability. For clean-catch samples, reliability was similar in children aged ≥ 3 years (0.55; 95% CI 0.43 to 0.67) and < 3 years (0.52; 95% CI 0.37 to 0.67), which was better than for nappy pad samples in children aged < 3 years (0.20; 95% CI 0.12 to 0.28). Similar patterns were seen when comparisons were further stratified into age groups < 2 and ≥ 2 to < 3 years. It thus appeared that lower reliability was attributable to nappy pad samples rather than to child's age.

There was little evidence that passing urine more often or day/bed wetting when previously dry were associated with UTI positivity (*Table 25*). Associations of pain or crying when passing urine, and dipstick nitrite and leucocyte positivity, were markedly stronger in clean-catch than nappy pad samples and with research laboratory than with NHS laboratory positivity. The association with temperature ≥ 39 °C appeared stronger in clean-catch than nappy pad samples, though CIs were wide. Associations with change in urine appearance and history of prior UTI did not differ markedly between NHS and research laboratory results.

**TABLE 24** Extent of agreement between NHS and research laboratory result

Age group and sample collection method	<i>n</i>	NHS -ve, RL -ve	NHS -ve, RL +ve	NHS +ve, RL -ve	NHS +ve, RL +ve (agree UTI step 1)	Kappa (95% CI)
Both collection methods	4808	4533	24	187	64	0.36 (0.29 to 0.43)
Clean catch	2619	2501	14	59	45	0.54 (0.45 to 0.63)
Nappy pad	2189	2032	10	128	19	0.20 (0.12 to 0.28)
≥ 3 years	1924	1852	10	35	27	0.53 (0.41 to 0.65)
Clean catch	1861	1792	10	32	27	0.55 (0.43 to 0.67)
Nappy pad	63	60	0	3	0	N/A (N/A)
< 3 years	2884	2681	14	152	37	0.29 (0.21 to 0.36)
Clean catch	758	709	4	27	18	0.52 (0.37 to 0.67)
Nappy pad	2126	1972	10	125	19	0.20 (0.12 to 0.28)
< 2 years	1958	1809	7	121	21	0.23 (0.15 to 0.31)
Clean catch	173	155	0	12	6	0.47 (0.23 to 0.72)
Nappy pad	1785	1654	7	109	15	0.19 (0.10 to 0.27)
≥ 2 and < 3 years	926	872	7	31	16	0.44 (0.29 to 0.59)
Clean catch	585	554	4	15	12	0.54 (0.36 to 0.72)
Nappy pad	341	318	3	16	4	0.27 (0.05 to 0.50)

-ve, negative; +ve, positive; N/A, not applicable; RL, research laboratory.  
 N/A: cannot compute kappa statistic because no samples were classified as positive by the research laboratory.

**TABLE 25** Associations of signs, symptoms and urine dipstick tests with separate NHS and research laboratory results, and 'agree UTI positive'

Signs, symptoms and dipstick results by laboratory	Clean catch (eight-variable model)		Clean catch (six-variable model)		Nappy pad (six-variable model)	
	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value
NHS laboratory						
Pain/crying passing urine	2.8 (1.6 to 5.1)	<0.001	2.9 (1.6 to 5.1)	<0.001	1.1 (0.4 to 3.1)	0.838
Passing urine more often	0.6 (0.3 to 1.1)	0.089	0.6 (0.3 to 1.1)	0.073	0.7 (0.3 to 1.5)	0.370
Change in urine appearance	2.9 (1.8 to 4.9)	<0.001	3.0 (1.8 to 4.9)	<0.001	2.1 (1.3 to 3.5)	0.005
Day/bed wetting, previously dry	1.0 (0.5 to 2.1)	0.961	N/A		N/A	
History of UTI	1.7 (0.9 to 3.3)	0.129	N/A		N/A	
Temperature ≥ 39 °C	1.8 (0.8 to 3.9)	0.129	1.7 (0.8 to 3.8)	0.157	0.7 (0.2 to 2.2)	0.526
Dipstick: nitrite +ve	7.3 (3.9 to 13.6)	<0.001	7.6 (4.1 to 14.1)	<0.001	2.0 (1.4 to 2.9)	0.001
Dipstick: leucocyte +ve	3.1 (1.9 to 5.0)	<0.001	3.1 (1.9 to 5.1)	<0.001	3.1 (2.1 to 4.4)	<0.001
<i>n</i> observations ( <i>n</i> +ve)	2619 (104)		2619 (104)		2189 (147)	
AUROC (95% CI)	0.75 (0.70 to 0.81)		0.75 (0.69 to 0.80)		0.65 (0.61 to 0.70)	

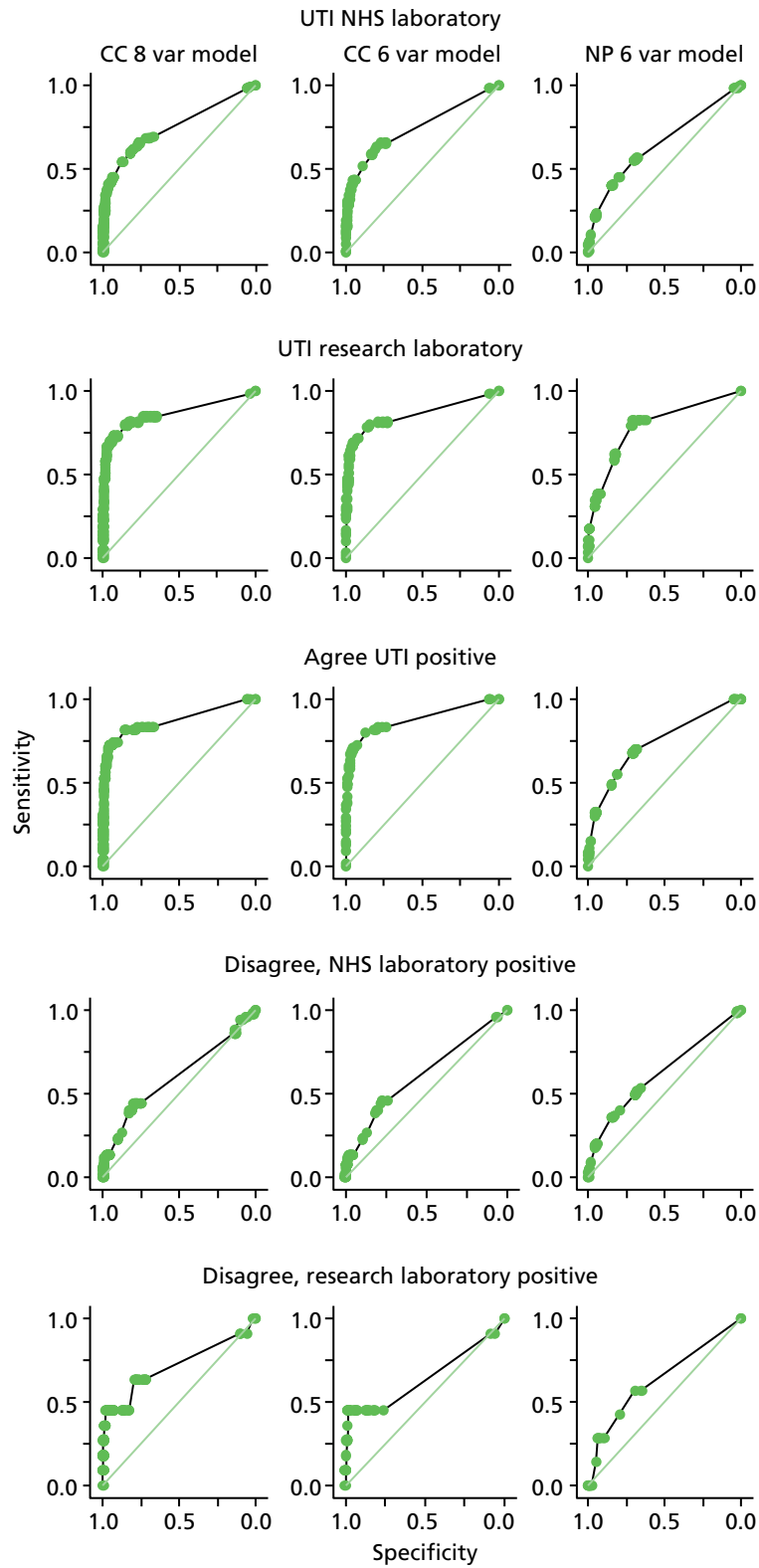
**TABLE 25** Associations of signs, symptoms and urine dipstick tests with separate NHS and research laboratory results, and 'agree UTI positive' (*continued*)

Signs, symptoms and dipstick results by laboratory	Clean catch (eight-variable model)		Clean catch (six-variable model)		Nappy pad (six-variable model)	
	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value
Research laboratory						
Pain/crying passing urine	5.9 (3.0 to 11.7)	<0.001	6.0 (3.0 to 11.8)	<0.001	1.4 (0.3 to 7.0)	0.716
Passing urine more often	0.7 (0.3 to 1.6)	0.457	0.8 (0.4 to 1.7)	0.543	1.2 (0.3 to 4.4)	0.839
Change in urine appearance	2.8 (1.4 to 5.6)	0.004	3.1 (1.6 to 6.1)	0.001	3.1 (1.2 to 7.9)	0.019
Day/bed wetting, previously dry	1.6 (0.7 to 3.7)	0.299	N/A		N/A	
History of UTI	2.7 (1.2 to 6.3)	0.017	N/A		N/A	
Temperature $\geq 39^{\circ}\text{C}$	2.0 (0.7 to 5.9)	0.223	1.7 (0.6 to 5.1)	0.333	1.1 (0.1 to 8.8)	0.930
Dipstick: nitrite +ve	10.7 (5.1 to 22.3)	<0.001	11.2 (5.4 to 23.1)	<0.001	5.2 (2.4 to 11.3)	<0.001
Dipstick: leucocyte +ve	5.2 (2.8 to 9.8)	<0.001	5.3 (2.8 to 10.0)	<0.001	4.1 (1.9 to 8.9)	<0.001
n observations (n +ve)	2619 (59)		2619 (59)		2189 (29)	
AUROC (95% CI)	0.87 (0.80 to 0.93)		0.86 (0.79 to 0.92)		0.79 (0.70, 0.88)	
'Agree UTI positive' (steps 1 and 2 combined)						
Pain/crying passing urine	6.0 (2.9 to 12.5)	<0.001	6.1 (3.0 to 12.7)	<0.001	1.0 (0.2 to 4.9)	0.993
Passing urine more often	0.5 (0.2 to 1.2)	0.121	0.6 (0.2 to 1.3)	0.160	0.4 (0.1 to 2.0)	0.279
Change in urine appearance	3.7 (1.8 to 7.6)	<0.001	4.1 (2.0 to 8.3)	<0.001	3.5 (1.6 to 7.6)	0.001
Day/bed wetting, previously dry	1.7 (0.7 to 4.3)	0.258	N/A		N/A	
History of UTI	2.1 (0.9 to 5.2)	0.103	N/A		N/A	
Temperature $\geq 39^{\circ}\text{C}$	2.0 (0.6 to 6.2)	0.245	1.8 (0.6 to 5.5)	0.339	0.6 (0.1 to 5.1)	0.674
Dipstick: nitrite +ve	17.1 (8.0 to 36.6)	<0.001	18.0 (8.5 to 38.4)	<0.001	3.2 (1.7 to 5.9)	<0.001
Dipstick: leucocyte +ve	4.6 (2.3 to 8.9)	<0.001	4.7 (2.4 to 9.1)	<0.001	4.1 (2.2 to 7.5)	<0.001
n observations (n +ve)	2556 (55)		2556 (55)		2079 (47)	
AUROC (95% CI)	0.87 (0.81 to 0.94)		0.88 (0.82 to 0.94)		0.74 (0.67 to 0.81)	

-ve, negative; +ve, positive; N/A, not applicable.  
N/A: variable was not considered for the six-variable model.

For both clean-catch and nappy pad samples, values of the AUROC were consistently lower for NHS than research laboratory positivity (see *Table 25* and *Figure 9*). For clean-catch samples (eight-variable model) the AUROC was 0.75 (95% CI 0.70 to 0.81) for NHS laboratory positivity and 0.87 (0.80 to 0.93) for research laboratory positivity. Values of the AUROC were similar in the six-variable models. Values of the AUROC were lower in nappy pad than clean-catch samples: 0.65 (95% CI 0.61 to 0.70) for NHS laboratory positivity and 0.79 (95% CI 0.70 to 0.88) for research laboratory positivity.

Of the 211 samples that were coded as positive in one laboratory only, 38 (10 from clean-catch samples and 28 from nappy pad samples) were coded as 'agree UTI positive (step 2)' when the NHS and research laboratory results were considered together. Thus, a total of 102 (2.1%) samples were considered as 'agree UTI' on the basis of both the NHS and research laboratory results. Strong associations of pain/crying



**FIGURE 9** Receiver operating characteristic curves of symptoms, signs and urine dipstick tests with different definitions of urine culture positivity. CC, clean catch; NP, nappy pad.

passing urine and nitrite and leucocyte dipstick positivity with an overall 'agree UTI' microbiology result in clean-catch samples were lower or absent in nappy-pad samples (see *Table 25*). For the six-variable model, the AUROC was higher in clean-catch (0.88; 95% CI 0.82 to 0.94) than nappy pad samples (0.74; 95% CI 0.67 to 0.81) (see *Table 25* and *Figure 9*).

Values of the AUROC were much lower for samples classified as 'disagree (NHS positive, research negative)' (six-variable model 0.62; 95% CI 0.54 to 0.69 for clean-catch and 0.62, 95% CI 0.57 to 0.68 for nappy pad samples). Only a small number (11 clean catch, 7 nappy pad) of samples were classified as 'disagree (NHS negative, research positive)' and values of the AUROC were also lower for these than for samples classified as 'agree UTI'. Values of the AUROC in sensitivity analyses conducted using the six-variable model are shown in *Table 26*.

Results for clean-catch samples using the eight-variable model and six-variable model were similar, and are available from the authors on request. For clean-catch samples, the values of the AUROC were similar for children aged < 3 and 3 to < 5 years, for both NHS laboratory and research laboratory positivity. Allowing for 'not known' responses for children aged < 3 years made little difference to the AUROC, for either NHS or research laboratory positivity. The AUROC was markedly higher for samples that were positive in both NHS and research laboratories than in those coded as 'agree UTI' after considering both results, for both clean-catch and nappy pad samples. For the research laboratory, but not for NHS laboratories, values of the AUROC were higher for samples collected in surgery than those collected at home. Values of the AUROC were similar in samples received by both laboratories within 24 hours and samples received after 24 hours except for nappy pad samples in the research laboratory, where the diagnostic accuracy appeared higher for samples received within 24 hours. For both NHS and research laboratory positivity, the AUROC increased with increasing threshold of pure/predominant growth count. For research laboratory positivity, values of the AUROC were markedly lower for pure/predominant growth < 10<sup>5</sup> CFU/ml. Values of the AUROC were markedly higher in samples with WBC count  $\geq 30/\text{mm}^3$ , except for research laboratory positivity in nappy pad samples. There was little evidence that values of the AUROC were higher for research laboratory positivity with pure than with predominant growth. There was little difference in AUROCs when stratifying by research laboratory urine volume (Monovette weight) in clean-catch and nappy pad samples.

**TABLE 26** Areas under the ROC curve (95% CI) from sensitivity analyses using six-variable logistic regression models

Characteristics of stratified analysis	Clean catch		Nappy pad	
	<i>n</i> observed ( <i>n</i> positive)	AUROC (95% CI)	<i>n</i> observed ( <i>n</i> positive)	AUROC (95% CI)
NHS positive age < 3 years	758 (45)	0.75 (0.67 to 0.84)	2126 (144)	0.66 (0.61 to 0.71)
NHS positive age 3 to < 5 years	1861 (59)	0.74 (0.67 to 0.81)	63 (3)	Too few
NHS positive age < 3 years, with 'not known' <sup>a</sup>	758 (45)	0.76 (0.68 to 0.85)	2126 (144)	0.66 (0.61 to 0.71)
RL positive age < 3 years	758 (22)	0.88 (0.80 to 0.97)	2126 (29)	0.79 (0.70 to 0.88)
RL positive age 3 to < 5 years	1861 (37)	0.85 (0.77 to 0.94)	63 (0)	Too few
RL positive age < 3 years, with 'not known' <sup>a</sup>	758 (22)	0.89 (0.81 to 0.98)	2126 (29)	0.81 (0.72 to 0.89)
'Agree UTI' age < 3 years	730 (21)	0.91 (0.83 to 0.98)	2017 (45)	0.75 (0.68 to 0.83)
'Agree UTI' age 3 to < 5 years	1826 (34)	0.87 (0.79 to 0.95)	62 (2)	Too few
'Agree UTI' age < 3 years, with 'not known' <sup>a</sup>	730 (21)	0.93 (0.89 to 0.98)	2017 (45)	0.76 (0.68 to 0.84)
'Agree UTI positive (step 1)'	2546 (45)	0.92 (0.87 to 0.97)	2051 (19)	0.85 (0.76 to 0.94)
'Agree UTI positive (step 2)'	2511 (10)	0.76 (0.60 to 0.91) <sup>b</sup>	2060 (28)	0.66 (0.55 to 0.76)
NHS surgery sample	2012 (84)	0.73 (0.67 to 0.80)	935 (64)	0.66 (0.58 to 0.73)
NHS home sample	607 (20)	0.81 (0.71 to 0.92)	1254 (83)	0.66 (0.59 to 0.72)
RL surgery sample	2012 (47)	0.90 (0.84 to 0.96)	935 (12)	0.84 (0.73 to 0.95) <sup>b</sup>
RL home sample	607 (12)	0.73 (0.55 to 0.90) <sup>b</sup>	1254 (17)	0.76 (0.64 to 0.89) <sup>b</sup>
NHS sample receipt < 24 hours	1959 (76)	0.76 (0.69 to 0.82)	1518 (108)	0.67 (0.31 to 0.72)
NHS sample receipt ≥ 24 hours	660 (28)	0.73 (0.64 to 0.83)	671 (39)	0.62 (0.53 to 0.70) <sup>b</sup>
RL sample receipt < 24 hours	794 (15)	0.89 (0.77 to 1.00)	630 (7)	0.94 (0.87 to 1.00) <sup>b</sup>
RL sample receipt ≥ 24 hours	1825 (44)	0.84 (0.76 to 0.92)	1559 (22)	0.75 (0.64 to 0.86) <sup>b</sup>
NHS <i>p/p</i> <sup>c</sup> ≥ 10 <sup>5</sup>	2619 (104)	0.75 (0.69 to 0.80)	2189 (147)	0.65 (0.61 to 0.70)
NHS <i>p/p</i> <sup>c</sup> ≥ 10 <sup>3</sup> to < 10 <sup>5</sup>	2515 (47)	0.58 (0.51 to 0.66) <sup>b</sup>	2042 (40)	0.57 (0.48 to 0.65) <sup>b</sup>
RL <i>p/p</i> <sup>c</sup> ≥ 10 <sup>7</sup>	2593 (33)	0.89 (0.81 to 0.97)	2166 (6)	0.74 (0.60 to 0.89) <sup>b</sup>
RL <i>p/p</i> <sup>c</sup> ≥ 10 <sup>6</sup> to < 10 <sup>7</sup>	2573 (13)	0.84 (0.70 to 0.98)	2166 (6)	0.96 (0.92 to 1.00) <sup>b</sup>
RL <i>p/p</i> <sup>c</sup> ≥ 10 <sup>5</sup> to < 10 <sup>6</sup>	2573 (13)	0.79 (0.64 to 0.94)	2169 (9)	0.81 (0.68 to 0.94) <sup>b</sup>
RL <i>p/p</i> <sup>c</sup> ≥ 10 <sup>4</sup> to < 10 <sup>5</sup>	2560 (24)	0.59 (0.51 to 0.68) <sup>b</sup>	2160 (61)	0.59 (0.54 to 0.64) <sup>b</sup>
RL <i>p/p</i> <sup>c</sup> ≥ 10 <sup>3</sup> to < 10 <sup>4</sup>	2560 (110)	0.57 (0.52 to 0.62)	2160 (93)	0.61 (0.56 to 0.66)
NHS positive and WBC ≥ 30/mm <sup>3</sup>	2572 (57)	0.85 (0.79 to 0.91)	2068 (26)	0.74 (0.62 to 0.86)
NHS positive and WBC < 30/mm <sup>3</sup>	2562 (47)	0.63 (0.55 to 0.71)	2163 (121)	0.64 (0.59 to 0.69)
RL positive and WBC ≥ 30/mm <sup>3</sup>	2599 (39)	0.97 (0.93 to 1.00)	2164 (4)	0.79 (0.48 to 1.00) <sup>b</sup>
RL positive and WBC < 30/mm <sup>3</sup>	2580 (20)	0.67 (0.54 to 0.81) <sup>b</sup>	2185 (25)	0.80 (0.71 to 0.89) <sup>b</sup>
RL pure growth ≥ 10 <sup>5</sup>	2604 (44)	0.84 (0.76 to 0.92)	2172 (12)	0.83 (0.72 to 0.94) <sup>b</sup>
RL predominant growth ≥ 10 <sup>5</sup>	2575 (15)	0.92 (0.84 to 1.00)	2177 (17)	0.76 (0.64 to 0.89) <sup>b</sup>
RL urine weight < median <sup>d</sup>	1306 (30)	0.90 (0.82 to 0.97) <sup>b</sup>	1092 (9)	0.77 (0.59 to 0.94) <sup>b</sup>
RL urine weight ≥ median <sup>d</sup>	1303 (27)	0.83 (0.72 to 0.93)	1085 (19)	0.81 (0.70 to 0.91)

RL, research laboratory.

a Including 'not known' responses for 'pain/crying when passing urine' and 'passing urine more often'.

b Not all variables included because of perfect prediction.

c Pure/predominant growth.

d The median Monovette weights cut-point was 16.23 g for clean-catch samples and 11.08 g for nappy pad samples.



## Discussion

### Summary of findings

Based on a large, unselected cohort of children presenting to primary care in England and Wales with acute illness, the reliability of microbiological diagnosis of UTI in routine NHS laboratories and a research laboratory was lower than expected and worse for urine samples collected using nappy pads than for clean-catch samples. Associations of microbiological positivity with pre-specified symptoms, signs and urine dipstick test results were lower for NHS laboratories than the research laboratory, and for nappy pad samples than clean-catch samples. Urines giving a 'positive' result in a NHS laboratory but not in the research laboratory had only modest associations with the preselected symptoms, signs and dipstick test results. These findings did not appear to be attributable to the younger age of the children providing nappy pad samples. Discrimination improved with increasing bacteriuria concentration and with the presence of WBCs in the urine sample (pyuria). Results of urine microbiology should, therefore, not be regarded as dichotomous result, but rather as a continuous phenomenon to be interpreted in the clinical context, with UTI possible even when bacterial concentrations are between  $10^3$  and  $10^4$  CFU/ml, and becoming increasingly probable with higher concentrations of pure or predominant bacterial growth in the presence of pyuria.

### Results in context with other studies

To our knowledge, this is the largest and most generalisable primary care-based study comparing the diagnostic performance of NHS laboratories with a research laboratory, and using nappy pad and clean-catch collection methods. However, the number of UTI positive samples was relatively small, especially for clean-catch samples in younger children and for the research laboratory. We do not know which samples were sent to NHS laboratories in containers with boric acid: all samples were sent via the routine mechanisms for that laboratory. All research laboratory samples were sent in containers with boric acid by Royal Mail post, introducing longer delays (which were not associated with UTI status) before processing than with NHS laboratories. In current UK primary care, nappy pad and bag samples are often the only feasible methods for obtaining urine samples from young children: there is usually insufficient space or staffing for parents and children to wait to provide clean-catch specimens. Moreover, most primary care clinicians (other than those with specialist paediatric or emergency department training) are not trained in SPA or catheter techniques, and neither of these are acceptable to parents. Nappy pads have been shown to be acceptable to parents<sup>63</sup> and endorsed by NICE.<sup>2</sup>

In the early development of microbiology, laboratory diagnosis of UTIs required isolation of the same organism from repeated urine samples. It was recognised in the 1940s that high urine bacterial counts were related to UTI, and subsequently proposed that a single sample with a high count could support a laboratory diagnosis of UTI and allow early treatment.<sup>117</sup> Early proposed thresholds to define positivity were derived from detailed investigation of fresh urine samples and ranged from  $\geq 10^3$  CFU/ml<sup>118</sup> to  $\geq 3 \times 10^3$  CFU/ml<sup>119</sup> and  $\geq 10^5$  CFU/ml.<sup>120</sup> Recently, it was suggested that a threshold of  $\geq 10^6$  CFU/ml would be more appropriate.<sup>80</sup> Current laboratory guidelines differ with respect to recommended thresholds. The UK Standards for Microbiological Investigations do not have specific paediatric guidance and suggest a threshold of a 'single organism  $\geq 10^4$  CFU/ml indicating UTI', but other thresholds are also discussed.<sup>82</sup> European paediatric guidelines suggest a threshold of  $\geq 10^4$  CFU/ml if symptoms are present and  $\geq 10^5$  CFU/ml if no symptoms are present for midstream specimens, and lower thresholds for specimens collected by bladder catheterisation or SPA.<sup>83</sup> US guidance suggests that clinicians require both urinalysis evidence of infection and a threshold of  $\geq 5 \times 10^4$  CFU/ml.<sup>121</sup>

Microbiological examination of urine requires quantification of bacteria and the ability to differentiate mixed from pure growths. The pour-plate method proved too labour-intensive given the large numbers of urine samples submitted to routine microbiology laboratories in the UK:<sup>82</sup> in 2012, 663,355 samples (12,689 from children aged < 5 years) were submitted in Wales alone, equating to some 12.1 million samples annually (250,000 from children aged < 5 years) in England and Wales. More rapid methods using calibrated loops, filter paper strips or multipoint methods to deliver a standard inoculum were developed

in response to the need for rapid throughput.<sup>122–124</sup> All were validated against viable counts performed by pour plates or the method of Miles and Misra.<sup>125</sup> The Standards for Microbiological Investigation followed by most UK laboratories provide options for urine culture using these methods to inoculate CLED or chromogenic agar, but they have not previously been calibrated against clinical symptoms. Difficulties in defining mixed growths and achieving accurate bacterial counts may be due to the small volumes of urine inoculated onto small areas of agar.<sup>80,122</sup> Spiral plating, which was used by the research laboratory in this study and involves a much larger inoculum (50 µl) over an entire 9-cm agar plate, is a more accurate method of quantifying bacterial counts and allows easy differentiation of mixed cultures.<sup>126</sup> Enhanced diagnostic performance might also be achieved through improvements in procedures for sample collection and transport.

NHS laboratory UTI positivity was consistently less strongly associated with urinary symptoms and dipstick results than research laboratory UTI positivity. A substantial number of UTI positive NHS laboratory samples [128 (5.8%) nappy pad and 59 (2.3%) clean catch] were negative in the research laboratory, and associations of the NHS laboratory positive results with symptoms, signs and dipstick results were modest. These findings suggest that the diagnostic performance of current routine NHS laboratory testing is suboptimal and may lead to overtreatment and unnecessary investigations.

### **Clinical and microbiological implications**

Even for samples processed in the research laboratory, the diagnostic utility of microbiology based on nappy pad samples was less than for clean-catch samples. The prevalence of UTI positivity diagnosed in the research laboratory was lower for nappy pad (1.3%) than for clean-catch (2.3%) samples, suggesting that UTIs are missed in nappy pad samples because of contamination. Therefore, primary care clinicians should try to obtain clean-catch samples in even very young children in whom they suspect a UTI,<sup>127</sup> for example by providing time and space to support urine collection. If an algorithm based on parent-reported symptoms can provide earlier ID of the children at greatest risk of UTI, then parents could be advised to obtain a urine sample prior to attending primary care.

In adult medicine, results from urine microbiology can be interpreted in the clinical context of the patient's presentation. However, in young children the significant difficulties in obtaining uncontaminated samples, together with the non-specific nature of the presenting symptoms, mean that there is greater reliance on the laboratory result. More detailed routine microbiological examination of paediatric urine samples would have resource implications that could better be justified if urines were selected for testing through an algorithm that increased the prior probability of positivity. Our results suggest that NHS laboratories should distinguish primary care paediatric (age < 5 years) samples from adult samples and consider reporting these in more detail, and that national procedures should be correspondingly updated.

Adoption of the more accurate but labour-intensive research laboratory methods would not be appropriate to use for all urines processed in NHS laboratories but, in many laboratories, extra or augmented methods are used for 'special urines', usually received in small numbers, to enhance reporting. It seems reasonable that research laboratory methods be implemented in NHS laboratories for the processing of paediatric urines if associated with enhanced diagnostic performance.

### **Implications for the DUTY study**

For the purposes of the DUTY study, samples will be considered UTI positive (called 'the reference standard' in *Chapter 5*) if there is growth in the research laboratory of  $\geq 10^5$  CFU/ml of a single uropathogen ('pure growth') or growth of  $\geq 10^5$  CFU/ml of a uropathogen with  $\geq 3 \log_{10}$  difference between the growth of this and the next species ('predominant growth').

# Chapter 5 Derivation and validation of a clinical algorithm to guide diagnosis and treatment of urinary tract infection in pre-school children in primary care

## Introduction

The main aim of the DUTY study was to derive and validate a two-step algorithm for the diagnosis of UTI in children under 5 years presenting to primary care with acute illnesses. The first step was to identify the children in whom urine sampling is warranted and the second to establish if dipstick testing provides additional value in identifying children warranting antibiotic treatment. There are three stages in the development of such algorithms: derivation, validation, and assessment of impact on clinical behaviour and patient outcome.<sup>128</sup> In this chapter, we report the first two of these stages against STARD criteria.<sup>129</sup>

## Methods

### Participants

The methods of recruitment are described in detail in *Chapter 2* and our study protocol (see *Appendix 3*).<sup>106</sup>

### Index tests and urine collection

All participating NHS staff received training in study procedures and only clinically qualified staff (GPs, nurse practitioners or emergency department doctors/nurses) performed and recorded clinical examination findings. Following consent, index tests (symptoms, signs or dipstick results) were recorded on a CRF (see *Appendix 4*) using a paper- or a web-based data collection system, with the latter requiring a unique log-in account for each staff member. Symptoms included the child's medical history and parent-reported symptoms (graded as absent, mild, moderate or severe when at their worst during the illness). Signs, from a full clinical examination, included 'clinicians' global impression of the child's illness severity', rated on a scale of 0 (child displaying no constitutional upset) to 10 (child displaying life-threatening signs requiring immediate hospitalisation) and 'abdominal tenderness' (any of suprapubic, loin or other abdominal tenderness). In total, 107 symptoms and signs were recorded and, preceding urine dipstick testing, clinicians recorded their working diagnosis, including a rating of the likelihood of UTI ('clinical diagnosis') as 'UTI fairly to very certain', 'no UTI fairly certain or uncertain' or 'no UTI certain to very certain'.

Our preferred urine collection method was 'clean catch'. NICE-recommended 'Newcastle nappy pads'<sup>2</sup> were used for children still using nappies and those for whom the parent/guardian did not think clean catch would be successful. Urine collection bags were used in a few instances, but data from these samples were excluded from the present study. At the primary care site, urine samples were dipstick tested (using Siemens Multistix 8 SG) for blood, protein, glucose, ketones, nitrite, LE, pH and specific gravity (eight dipstick index tests). For full details of urine collection methods, see *Chapter 2, Urine sample collection*, and our protocol paper (see *Appendix 3*).<sup>106</sup> All 107 index tests and the clinicians' working diagnosis were measured blind to the reference standard.

### Reference standard

Research laboratory samples were processed by two staff members using a single, standardised procedure. We used a microbiological definition of UTI of  $\geq 10^5$  CFU/ml of a single uropathogen ('pure growth') or  $\geq 10^5$  CFU/ml of a uropathogen with  $\geq 3 \log_{10}$  (1000-fold) difference between the growth of this and the next species ('predominant growth') (see *Chapter 4*). We defined uropathogens as members of the

Enterobacteriaceae family. Aside from knowing children's dates of birth (used as part of the urine ID process), research laboratory staff were blind to children's index test characteristics. We compared the characteristics of children for whom research laboratory culture results were and were not available.

### Statistical analysis

We examined the frequency of symptom and sign categories, blind to their associations with urine culture results, and merged the least frequent categories prior to analyses. We used logistic regression to estimate associations of index tests with urine culture positivity. We fitted separate models for clean-catch and nappy pad samples because of the markedly different age distributions of children providing samples using these methods and concerns about contamination reducing the sensitivity of urine culture in nappy pad samples. The  $p$ -values were derived using LR tests. For ordinal variables, both heterogeneity and trend  $p$ -values were derived. 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 the response 'don't know' to questions about the presence of symptoms such as pain or crying when passing urine. In both univariable and multivariable analyses, missing data were coded as the modal value, usually absence of the symptom. We repeated multivariable analyses using the chained equations approach to multiple imputation: estimates and Wald  $p$ -values<sup>130</sup> based on 50 imputed data sets derived using Rubin's rules.<sup>131</sup>

We sequentially evaluated selected index tests in two groups: parent-reported symptoms and clinician-reported signs, and urine dipstick results. This reflects the clinical process, where symptoms and signs are used to identify children from whom urine should be collected and dipstick testing may additionally help clinicians to decide which children should be treated. First, we selected those variables with either trend or heterogeneity univariable  $p$ -value  $< 0.01$  for either collection method or when all samples were analysed together. Second, we derived models from among all of the selected symptoms and signs, separately for nappy pad and clean-catch samples, using backwards stepwise selection and an exclusion criterion of heterogeneity  $p$ -value  $> 0.1$ . Third, we used backwards stepwise selection with the same exclusion  $p$ -value for models in which dipstick results were added to the previously selected symptoms and signs, to give models including symptoms, signs and dipstick results. We investigated the effect of using more liberal  $p$ -value thresholds of 0.1 and 0.2 at each stage, and found no important differences in the final models (results available from the authors on request).

For each model, we quantified diagnostic accuracy as the AUROC (also known as the  $c$ -statistic). We also estimated AUROC values for 'clinician diagnoses' of UTI. To formally test the additional value of dipsticks we calculated the difference in AUROC, along with 95% CI and  $p$ -value, between the symptoms and signs model and the symptom, sign and dipstick model. We internally validated the models, calculating the AUROC and calibration slope, using the bootstrap procedure described by Steyerberg.<sup>132</sup> The calibration slope is a shrinkage factor that compares predicted with observed probabilities, where 1 is perfect calibration and 0 reflects no agreement between predicted and observed probabilities. Multiplying the estimated coefficients (log-odds ratios) in a predictive model by the calibration slope would correct the estimated positive predictive values for statistical overoptimism due to the model fitting. For each model, we selected cut-points corresponding to a range of values for sensitivity, and then calculated the corresponding specificity, negative and positive predictive values, and the proportion of children classified positive. This was completed using the linear predictors that were combined using Rubin's rules.<sup>131</sup> CIs were calculated for these parameters by transforming to the log-odds scale, and then using the standard error formula for log-odds, and then transforming back to a proportion. As log-odds are undefined for a proportion of 0 or 1, we were unable to calculate CIs when these parameters were equal to either 0% or 100%.

We also investigated the robustness of our clean-catch models in younger children by fitting the clean-catch models to those with clean-catch samples who were under 3 years of age. Finally, to investigate the robustness of our nappy pad models, we fitted the nappy pad models to those with clean-catch samples who were under 3 years of age.

### **Fever of unknown origin**

In order to investigate if 'fever of unknown origin' was associated with UTI, we identified children with fever and no symptoms or signs suggestive of an alternative source. We did three analyses using the three fever variables – two parent-reported ('fever now or in the past 24 hours' and 'fever at any time during this illness') and one from the physical examination (temperature of  $\geq 38^\circ\text{C}$ ) – for children without any of the following symptoms and signs: new generalised rash; vomiting; diarrhoea (any time and last 24 hours); blocked/runny nose; cough; wheeze; shortness of breath; chest pains; earache; sore throat; oxygen saturations  $< 94\%$ ; any throat abnormality; any ear abnormality; and any chest abnormality. The categories 'slight', 'moderate' and 'severe problem' for the parent-reported fever variables were merged owing to the small number of UTI positives within these categories. We conducted two analyses for each of the clean-catch and nappy pad samples: one used a complete case approach and the other assumed 'modal/normal' imputation for each of the 14 variables above.

### **Points-based models**

As the above models are relatively complex and would need a computer or calculator to estimate risk scores, we also generated points-based models where the risk of UTI could be more easily calculated. To do this, we first dichotomised all variables in our original models, including dipsticks. To make the resulting rule as user-friendly as possible, we dichotomised all variables at the 'present/absent' threshold. We removed clinical examination variables, as these variables contributed the least to the models and would take a disproportionate amount of time to collect in relation to their value. After deriving these dichotomised variable models using the multiple imputed data, we calculated a corresponding points-based model. Using methods similar to previous studies,<sup>133,134</sup> points were calculated for each variable by dividing that coefficient's variable by the smallest coefficient in the model and then rounding to the nearest integer. We quantified the diagnostic accuracy of the resulting points-based models using the AUROC, and compared these with the 'full models'. To formally test the additional value of dipsticks, we calculated the difference in AUROC, along with its 95% CI and *p*-value, between the symptoms and signs model and the symptoms, signs and dipstick model. This comparison was estimated for both the coefficient and the integer points-based models. The points-based models cannot be validated using bootstrapping as this method relies on the coefficients being estimable in different (bootstrap) samples and, in the 'points-model', these are fixed. However, to assess possible optimism of the points-based models, we internally validated the dichotomised models without the points' scores. For the points-based models, we selected cut-points corresponding to a range of integer risk scores, and then calculated the corresponding sensitivity, specificity, negative and positive predictive values, and the proportion of children classified positive. This was completed with the points risk scores that were combined using Rubin's rules.<sup>131</sup>

Additionally, for the points-based models we selected cut-points corresponding to the complete range of risk scores for each possible combination of symptoms and signs, and then calculated the corresponding sensitivity, specificity, negative and positive predictive values, and the proportion of children classified positive.

### **Added value of dipstick testing**

In addition to looking at the difference in the AUROC between the models with and without the dipstick test results, we fitted multinomial logit models for the association of the selected symptoms and signs with the dipstick test results included in the symptoms, signs and dipstick model. We then quantified the additional value of dipstick testing in diagnosis of UTI using a three-step simulation approach based on the points-based models. The steps were (1) sample coefficient values from the multivariate normal distribution of the multinomial logit parameter estimates; (2) based on the sampled coefficients, randomly

generate a set of dipstick results; and (3) compute the probability of UTI based on 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 as it was observed in only 3 out of 2740 individuals and led to numerical instabilities. We grouped the distribution of these probabilities and changes in probabilities according to whether or not primary care clinicians considered it reasonable to recommend urine sampling based on symptoms and signs and the cost-effectiveness analyses; results were weighted according to the relative frequencies of different combinations of symptoms and signs within the groups.

## Results

### Participants

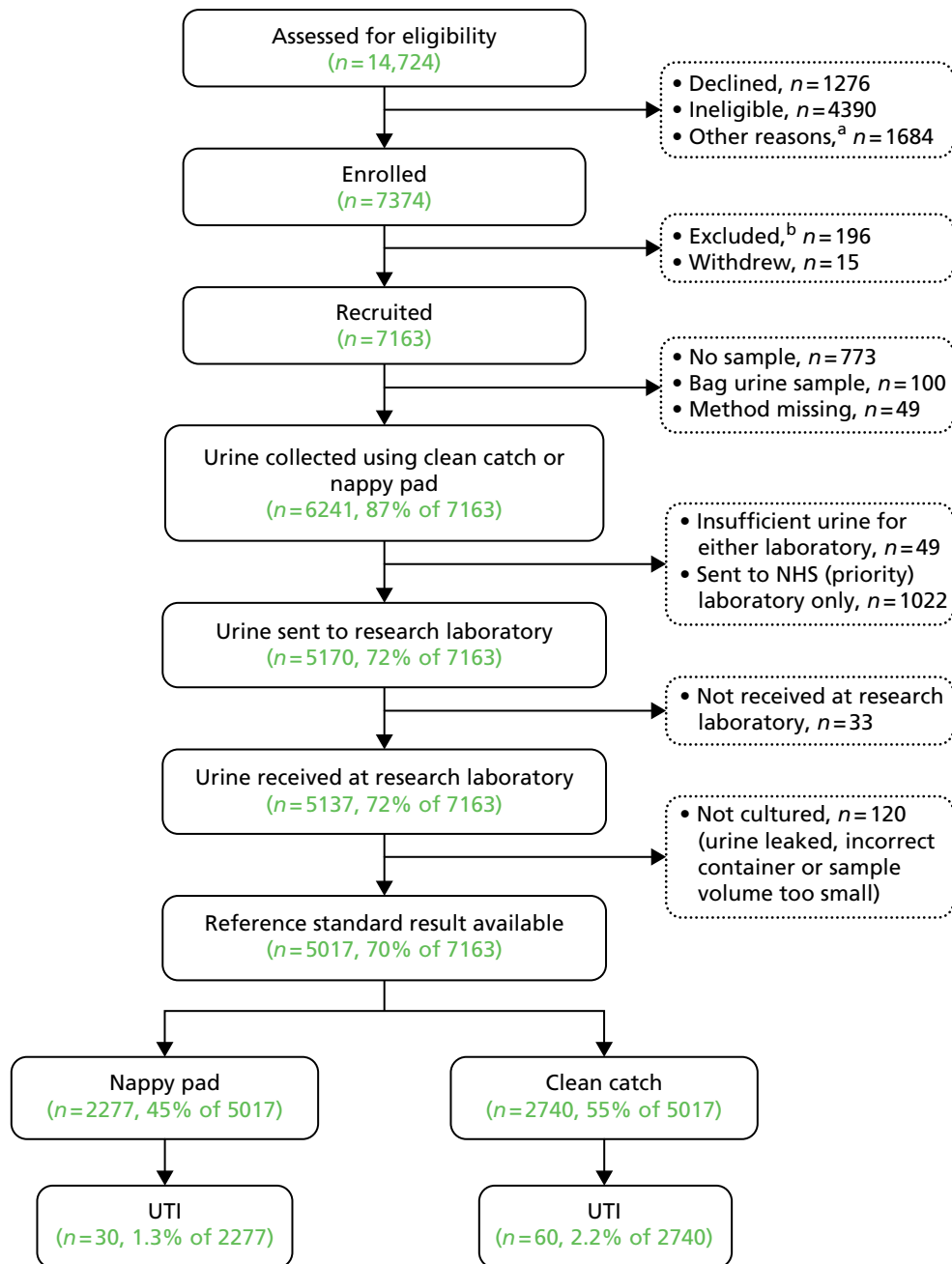
The recruitment of participants is described in *Chapter 3*. In summary, 14,724 children were screened for eligibility, of whom 4390 were immediately ineligible, 1276 declined, 1684 could not be recruited for other reasons, 196 were subsequently excluded and 15 withdrew (see *Figure 4*). Urine was collected using clean-catch or nappy pad methods from 6241 (87%) children and sent to the NHS (priority) laboratory in 6192 (86%) and the research laboratory in 5170 (72%); these figures and where they flow from can be seen in *Figure 10*. Reference standard results were available from the research laboratory (our final analytic sample) in 5017 (70%), of which 2740 (55%) were collected via clean catch and 2277 (45%) via nappy pad (see *Figures 4* and *10*). The number in *Figure 10* in the box 'urine sent to research laboratory' does not include those urines collected by bag or those for which we did not know the collection method.

Comparing recruited children with ( $n = 5017$ ) and without ( $n = 2146$ ) a research laboratory culture result, we found that children with a culture result were older (59% vs. 33% were  $\geq 2$  years); more likely to have a mild illness (83% vs. 80% clinician global impression  $< 3$ ); more likely to report pain/crying on passing urine (12% vs. 8%); more likely to report darker urine (14% vs. 7%); more likely to have UTI suspected clinically (7% vs. 4%); less likely to report nappy rash (16% vs. 20%); and less likely to be nitrite positive (9% vs. 14%). However, they were similar in terms of sex (female 51% vs. 50%); ethnicity (non-white 17% vs. 18%); parental highest qualification (diploma or degree 49% vs. 50%); deprivation (21% vs. 21% for most deprived quintile score); parental impression of overall illness severity (19% vs. 15%  $< 3$ ); days unwell prior to recruitment (both median 4 days); and the prevalence of all the other symptoms, signs and dipstick index tests found to be independently associated with UTI. Further details of these comparisons are given in *Chapter 3*.

The most common non-UTI clinical diagnoses in the clean-catch and nappy pad groups, respectively, were 'upper respiratory tract infection' (28% and 35%), 'viral illness' (15% and 18%) and otitis media (10% and 9%). Gastroenteritis was thought to be present in 3.6% and 5.7%, respectively. Antimicrobial substances, which can arise from the use of both systemic antibiotics and locally applied cleaning agents, were found in 4.5% and 6.6% of clean-catch and nappy pad samples, respectively, and for both collection methods samples were more likely to be present in children with than without UTI (*Table 27*, and see *Table 38*). A total of 79 children were reported hospitalised, none as a result of study participation. We are not aware of any adverse events resulting from the measurement of index or reference tests.

### Clean-catch models

*Table 27* shows that 2.2% of the 2740 children providing clean-catch samples met the criteria for a microbiological diagnosis of UTI, and that 94% were aged  $\geq 2$  years, 54% were female, and 96% of samples were provided within 24 hours of index test measurement. Clinicians correctly identified 47% of the 60 UTIs. Missing data or 'not known' responses were generally infrequent other than in the youngest children (nappy pad samples), in whom symptoms, such as pain/crying on passing urine, cannot be determined (see *Tables 27* and *38*).



**FIGURE 10** DUTY flow diagram for diagnostic algorithm. a, Includes left prior to invitation ( $n=811$ ), no consent ( $n=214$ ), language barrier ( $n=112$ ); and b, excluded as retrospectively ineligible/duplicate recruitment/invalid consent ( $n=141$ ) and poor data quality ( $n=55$ ).

**TABLE 27** Clean-catch samples: children’s characteristics and crude ORs for index tests associated with UTI

Demographics and index tests <sup>a</sup>	n (%) <sup>b</sup>	UTI positive (%) <sup>c</sup>	Crude OR (95% CI)
<b>Gender</b>			
Male	1267 (46.2)	13 (1.0)	1 (ref.)
Female	1473 (53.8)	47 (3.2)	3.18 (1.71 to 5.90)
<b>Age of child</b>			
< 6 months	34 (1.2)	1 (2.9)	1.13 (0.15 to 8.77)
6 to < 12 months	54 (2.0)	3 (5.6)	2.19 (0.62 to 7.77)
1 to < 2 years	91 (3.3)	2 (2.2)	0.84 (0.19 to 3.70)
2 to < 3 years	612 (22.3)	16 (2.6)	1 (ref.)
3 to < 4 years	1073 (39.2)	21 (2.0)	0.74 (0.39 to 1.44)
4 years plus	876 (32.0)	17 (1.9)	0.74 (0.37 to 1.47)
<b>Clinician diagnosis prior to dipstick</b>			
Not UTI certain/very certain	1149 (41.9)	6 (0.5)	0.28 (0.12 to 0.69)
Not UTI fairly certain/uncertain	1417 (51.7)	26 (1.8)	1 (ref.)
UTI fairly to very certain	168 (6.1)	28 (16.7)	10.75 (6.13 to 18.8)
Missing	6 (0.2)	0 (0.0)	
<b>Pain/crying when passing urine<sup>a</sup></b>			
No problem	2234 (81.5)	22 (1.0)	1 (ref.)
Slight problem	182 (6.6)	6 (3.3)	2.97 (1.21 to 7.29)
Moderate problem	128 (4.7)	12 (9.4)	9.01 (4.45 to 18.2)
Severe problem	51 (1.9)	15 (29.4)	36.30 (17.81 to 74.0)
Missing/not known	145 (5.3)	5 (3.4)	
<b>Smelly urine<sup>a</sup></b>			
No problem	2108 (76.9)	20 (0.9)	1 (ref.)
Slight problem	174 (6.4)	10 (5.7)	5.87 (2.76 to 12.5)
Moderate problem	179 (6.5)	16 (8.9)	9.46 (4.93 to 18.2)
Severe problem	51 (1.9)	10 (19.6)	23.5 (10.6 to 52.3)
Missing/not known	228 (8.3)	4 (1.8)	
<b>Cough</b>			
No problem	773 (28.2)	24 (3.1)	1 (ref.)
Slight problem	556 (20.3)	16 (2.9)	0.93 (0.48 to 1.76)
Moderate problem	829 (30.3)	17 (2.1)	0.66 (0.35 to 1.23)
Severe problem	579 (21.1)	3 (0.5)	0.16 (0.05 to 0.54)
Missing/not known	3 (0.1)	0 (0.0)	
<b>History of UTI<sup>a</sup></b>			
No	2449 (89.4)	43 (1.8)	1 (ref.)
Yes	177 (6.5)	12 (6.8)	3.81 (1.99 to 7.31)
Missing/not known	114 (4.2)	5 (4.4)	
<b>Clinician global impression of illness severity (0–10)<sup>a</sup></b>			
0–1	989 (36.1)	14 (1.4)	1 (ref.)
2	739 (27.0)	14 (1.9)	1.35 (0.64 to 2.85)
3	531 (19.4)	14 (2.6)	1.89 (0.89 to 4.00)



**TABLE 27** Clean-catch samples: children's characteristics and crude ORs for index tests associated with UTI (*continued*)

Demographics and index tests <sup>a</sup>	n (%) <sup>b</sup>	UTI positive (%) <sup>c</sup>	Crude OR (95% CI)
4–5	363 (13.2)	12 (3.3)	2.39 (1.09 to 5.21)
≥ 6	115 (4.2)	6 (5.2)	3.85 (1.45 to 10.21)
Missing	3 (0.1)	0 (0.0)	
Abdominal exam: any abdominal pain <sup>a</sup>			
No	2237 (81.6)	46 (2.1)	1 (ref.)
Yes	63 (2.3)	8 (12.7)	7.34 (3.33 to 16.19)
Missing	440 (16.1)	6 (1.4)	
Ear exam: any acute abnormality <sup>a</sup>			
No	1783 (65.1)	50 (2.8)	1 (ref.)
Yes	635 (23.2)	4 (0.6)	0.23 (0.08 to 0.64)
Missing	322 (11.8)	6 (1.9)	
Dipstick: leucocytes <sup>a</sup>			
Negative	2272 (82.9)	17 (0.7)	1 (ref.)
Trace	154 (5.6)	6 (3.9)	5.40 (2.10 to 13.9)
+	110 (4.0)	2 (1.8)	2.47 (0.56 to 10.8)
++	148 (5.4)	19 (12.8)	19.61 (9.95 to 38.6)
+++	48 (1.8)	16 (33.3)	66.6 (30.9 to 143.3)
Missing	8 (0.3)	0 (0.0)	
Dipstick: nitrites <sup>a</sup>			
Negative	2658 (97.0)	35 (1.3)	1 (ref.)
Positive	74 (2.7)	25 (33.8)	38.4 (21.4 to 68.9)
Missing	8 (0.3)	0 (0.0)	
Dipstick: blood <sup>a</sup>			
Negative	2297 (83.8)	29 (1.3)	1 (ref.)
Non-haem	246 (9.0)	8 (3.3)	2.64 (1.19 to 5.84)
Haem trace	50 (1.8)	6 (12.0)	10.70 (4.23 to 27.08)
Haem +	67 (2.4)	4 (6.0)	4.98 (1.70 to 14.60)
Haem ++ or +++	72 (2.6)	13 (18.1)	17.29 (8.56 to 34.94)
Missing	8 (0.3)	0 (0.0)	

Ref., reference.

a Index tests independently associated with UTI in multivariable models. Missing values were assigned to the modal category.

b All children column gives the percentage of observations within that category.

c 'UTI positive' column gives the percentage of positives relative to the number of observations within that category.

### 'Full' coefficient-based models

Tables 27 and 28, respectively, show the crude and adjusted ORs for the index tests (marked with footnote marker d in Table 28) selected for the clean-catch model, for children predominantly ≥ 2 and < 5 years. Table 28 shows that parent-reported pain/crying while passing urine, smelly urine, history of UTI, and absence of severe cough were independently associated with UTI and, for the first two of these, evidence of increasing strength of association for increasing severity of symptoms with UTI. Clinician-reported global impression of illness severity, abdominal tenderness on examination and the absence of acute ear abnormality on examination were independently associated with UTI, again with the former showing a

**TABLE 28** Clean-catch samples: ‘full’ coefficient models based on symptoms and signs and on symptoms, signs and dipstick results, including results based on MI

Index tests	Symptoms and signs model		Symptoms, signs and dipstick model	
	Adjusted OR (95% CI) <sup>a</sup>	MI adjusted OR (95% CI)	Adjusted OR (95% CI) <sup>a</sup>	MI adjusted OR (95% CI)
Pain/crying when passing urine				
No problem	1 (ref.)	1 (ref.)	1 (ref.)	1 (ref.)
Slight problem	1.68 (0.64 to 4.41)	1.87 (0.70 to 4.99)	1.01 (0.30 to 3.45)	1.19 (0.34 to 4.11)
Moderate problem	5.81 (2.58 to 13.10)	6.06 (2.60 to 14.12)	3.26 (1.19 to 8.92)	3.55 (1.25 to 10.04)
Severe problem	21.69 (9.19 to 51.23)	23.79 (9.91 to 57.10)	15.23 (5.17 to 44.89)	16.55 (5.46 to 50.11)
Smelly urine				
No problem	1 (ref.)	1 (ref.)	1 (ref.)	1 (ref.)
Slight problem	5.08 (2.23 to 11.60)	5.30 (2.24 to 12.55)	3.70 (1.35 to 10.11)	3.99 (1.39 to 11.42)
Moderate problem	6.43 (3.07 to 13.46)	6.56 (2.99 to 14.37)	5.55 (2.33 to 13.23)	5.83 (2.30 to 14.77)
Severe problem	11.85 (4.60 to 30.57)	12.09 (4.58 to 31.90)	5.51 (1.64 to 18.46)	6.00 (1.72 to 20.90)
History of UTI				
No	1 (ref.)	1 (ref.)	1 (ref.)	1 (ref.)
Yes	3.16 (1.46 to 6.83)	3.07 (1.40 to 6.72)	2.86 (1.15 to 7.11)	2.80 (1.13 to 6.96)
Cough				
No problem	1 (ref.)	1 (ref.)	1 (ref.)	1 (ref.)
Slight problem	1.33 (0.63 to 2.82)	1.37 (0.64 to 2.93)	1.33 (0.53 to 3.32)	1.37 (0.54 to 3.46)
Moderate problem	1.36 (0.65 to 2.86)	1.45 (0.68 to 3.10)	2.23 (0.94 to 5.29)	2.35 (0.98 to 5.64)
Severe problem	0.23 (0.06 to 0.92)	0.24 (0.06 to 0.97)	0.29 (0.05 to 1.60)	0.29 (0.05 to 1.64)
Clinician global impression of illness severity (0–10)				
0–1	1 (ref.)	1 (ref.)	1 (ref.)	1 (ref.)
2	2.19 (0.94 to 5.14)	2.19 (0.92 to 5.17)	2.50 (0.92 to 6.77)	2.49 (0.90 to 6.87)
3	3.09 (1.32 to 7.23)	3.16 (1.32 to 7.55)	3.22 (1.19 to 8.74)	3.19 (1.15 to 8.85)
4–5	4.35 (1.73 to 10.92)	4.73 (1.86 to 12.04)	3.68 (1.22 to 11.12)	4.12 (1.34 to 12.62)
≥ 6	9.23 (2.91 to 29.28)	9.71 (2.98 to 31.61)	8.27 (2.04 to 33.57)	9.11 (2.18 to 37.97)
Abdominal exam: any abdominal pain				
No	1 (ref.)	1 (ref.)	1 (ref.)	1 (ref.)
Yes	2.76 (1.04 to 7.34)	2.52 (0.95 to 6.72)	1.42 (0.34 to 6.01)	1.22 (0.29 to 5.19)
Ear exam: any acute abnormality				
No	1 (ref.)	1 (ref.)	1 (ref.)	1 (ref.)
Yes	0.25 (0.08 to 0.81)	0.22 (0.07 to 0.71)	0.40 (0.12 to 1.27)	0.34 (0.10 to 1.10)
Dipstick: leucocytes				
Negative			1 (ref.)	1 (ref.)
Trace			2.04 (0.63 to 6.60)	1.99 (0.61 to 6.56)
+			0.66 (0.11 to 3.94)	0.61 (0.10 to 3.92)
++			7.38 (3.04 to 17.92)	7.24 (2.94 to 17.81)
+++			16.78 (5.47 to 51.44)	16.62 (5.20 to 53.13)

**TABLE 28** Clean-catch samples: 'full' coefficient models based on symptoms and signs and on symptoms, signs and dipstick results, including results based on MI (*continued*)

Index tests	Symptoms and signs model		Symptoms, signs and dipstick model	
	Adjusted OR (95% CI) <sup>a</sup>	MI adjusted OR (95% CI)	Adjusted OR (95% CI) <sup>a</sup>	MI adjusted OR (95% CI)
Dipstick: nitrites				
Negative			1 (ref.)	1 (ref.)
Positive			7.33 (3.09 to 17.40)	7.54 (3.12 to 18.23)
Dipstick: blood				
Negative			1 (ref.)	1 (ref.)
Non-haem			0.86 (0.29 to 2.54)	0.86 (0.29 to 2.60)
Haem trace			5.54 (1.43 to 21.53)	5.42 (1.34 to 21.93)
Haem +			3.22 (0.85 to 12.23)	3.51 (0.90 to 13.63)
Haem ++ or +++			1.90 (0.59 to 6.11)	1.94 (0.58 to 6.47)
AUROC (95% CI)	0.892 (0.84 to 0.94)	0.899 (0.85 to 0.95)	0.926 (0.89 to 0.96)	0.933 (0.90 to 0.97)
Validated AUROC <sup>b</sup>	0.871	0.876	0.904	0.903
Δ ROC <sup>c</sup> (95% CI)			0.034 (0.01 to 0.06)	0.034 (0.01 to 0.06)
Δ ROC <sup>c</sup> p-value			0.007	0.009

MI, multiple imputation; ref., reference.

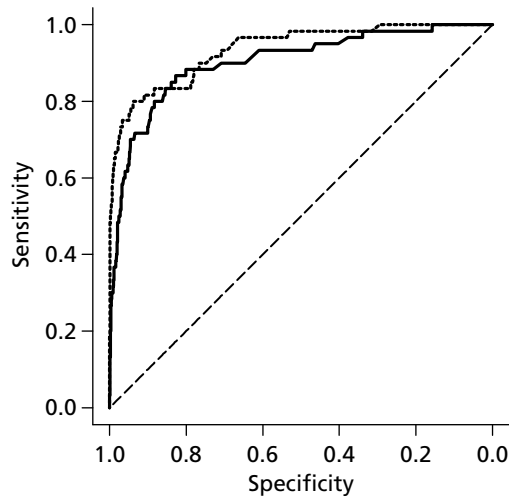
a Missing values coded to modal category.

b Internal validation using the bootstrap procedure.

c The difference in ROC between symptoms and signs model and symptoms, signs and dipstick model.

gradation of association with increasing illness severity. Leucocytes, nitrites and blood on dipstick testing were strongly and independently associated with UTI. The increase in the AUROC while using the symptoms, signs and dipstick model was around 0.06 for both the modal and the multiple imputation, and the low *p*-value means that there is very strong evidence that the dipsticks increase the AUROC when added to the symptoms and signs model. The validated multiple imputation AUROC for the symptoms and signs model (0.876) demonstrated very good diagnostic accuracy (see *Table 28*). Adding dipstick findings increased the validated AUROC to 0.903. For comparison, the AUROC for 'clinician diagnosis' was 0.774 (0.714 to 0.833). *Figure 11* shows the multiple imputed ROC curves for the models with and without dipstick urinalysis. Model calibration slopes were 0.832 (> 0.8 indicates a model with good calibration).

The upper section of *Table 29* shows the diagnostic test characteristics corresponding to various choices of cut-point for positivity using the coefficient-based symptoms and signs model, which could be used to identify the children in whom urine collection is warranted. The risk scores for these cut-points can be calculated using the parameters in *Table 30*. For a sensitivity of 60%, 70% or 80%, respective urine sampling of approximately 5%, 7% or 13% of unwell children aged ≥ 2 years presenting to primary care would be required. At these cut-points, urine samples would be positive for UTI in 27%, 23% and 13% of children, with corresponding specificities of 96.3%, 94.6% and 88.3%. The lower section of *Table 29* shows the diagnostic test characteristics for the model based on the symptoms, signs and dipstick test results. If a decision to treat presumptively with antibiotics was made using the 80% sensitivity cut-point, this model would improve the specificity from 88% to 94% and would reduce the overall number of children treated from 13% to 8%, compared with the symptoms and signs only model.



**FIGURE 11** Clean-catch ROC curve for multiple imputation models for symptoms and signs only (solid line) and symptoms, signs and dipstick (dotted line).

**TABLE 29** Clean-catch diagnostic test characteristics (95% CI) for a range of cut-points, using both the symptoms and signs model (upper part of table) and the symptoms, signs and dipstick model (lower part of table)

Risk score ( $\geq$ )	Sensitivity, %	Specificity, %	PPV, %	NPV, %	Children positive, %
Symptoms and signs model					
-0.195	20.0 (11.7 to 32.0)	99.8 (99.5 to 99.9)	66.7 (42.9 to 84.2)	98.2 (97.7 to 98.7)	0.7 (0.4 to 1.0)
-0.87	30.0 (19.8 to 42.7)	99.5 (99.1 to 99.7)	56.3 (39.0 to 72.1)	98.4 (97.9 to 98.9)	1.2 (0.8 to 1.6)
-1.698	40.0 (28.5 to 52.8)	98.2 (97.6 to 98.6)	32.9 (23.1 to 44.4)	98.7 (98.1 to 99.0)	2.7 (2.1 to 3.3)
-1.98	50.0 (37.6 to 62.4)	97.5 (96.9 to 98.1)	31.3 (22.8 to 41.2)	98.9 (98.4 to 99.2)	3.5 (2.9 to 4.3)
-2.34	60.0 (47.2 to 71.5)	96.3 (95.5 to 97.0)	26.7 (19.9 to 34.7)	99.1 (98.6 to 99.4)	4.9 (4.2 to 5.8)
-2.75	70.0 (57.3 to 80.2)	94.6 (93.7 to 95.4)	22.6 (17.1 to 29.1)	99.3 (98.9 to 99.6)	6.8 (5.9 to 7.8)
-3.515	80.0 (68.0 to 88.3)	88.3 (87.0 to 89.4)	13.3 (10.1 to 17.2)	99.5 (99.1 to 99.7)	13.2 (12.0 to 14.5)
-3.884	85.0 (73.6 to 92.0)	83.9 (82.4 to 85.2)	10.6 (8.1 to 13.6)	99.6 (99.2 to 99.8)	17.6 (16.2 to 19.1)
-4.86	93.3 (83.5 to 97.5)	61.0 (59.1 to 62.8)	5.1 (3.9 to 6.6)	99.8 (99.4 to 99.9)	40.2 (38.4 to 42.0)
-5.7	96.7 (87.6 to 99.2)	37.8 (35.9 to 39.6)	3.4 (2.6 to 4.3)	99.8 (99.2 to 100.0)	63.0 (61.2 to 64.8)
-6.664	100	15.7 (14.4 to 17.1)	2.6 (2.0 to 3.3)	100	84.6 (83.2 to 85.9)
Symptoms, signs and dipstick model					
1.43	20.0 (11.7 to 32.0)	99.9 (99.7 to 100.0)	85.7 (57.3 to 96.4)	98.2 (97.7 to 98.7)	0.5 (0.3 to 0.9)
0.321	40.0 (28.5 to 52.8)	99.9 (99.7 to 100.0)	88.9 (70.7 to 96.4)	98.7 (98.2 to 99.0)	1.0 (0.7 to 1.4)
-1.15	60.0 (47.2 to 71.5)	99.3 (98.8 to 99.5)	64.3 (51.0 to 75.7)	99.1 (98.7 to 99.4)	2.0 (1.6 to 2.6)
-3.275	80.0 (68.0 to 88.3)	93.8 (92.9 to 94.7)	22.5 (17.4 to 28.6)	99.5 (99.2 to 99.7)	7.8 (6.8 to 8.8)
-3.98	83.3 (71.7 to 90.8)	88.3 (87.0 to 89.5)	13.8 (10.6 to 17.7)	99.6 (99.2 to 99.8)	13.2 (12.0 to 14.6)
-5.237	96.7 (87.6 to 99.2)	66.3 (64.5 to 68.1)	6.0 (4.7 to 7.7)	99.9 (99.6 to 100.0)	35.0 (33.3 to 36.8)
-5.825	98.3 (89.1 to 99.8)	53.1 (51.2 to 54.9)	4.5 (3.5 to 5.7)	99.9 (99.5 to 100.0)	48.1 (46.2 to 49.9)
-6.69	100	29.5 (27.8 to 31.2)	3.1 (2.4 to 3.9)	100	71.2 (69.4 to 72.8)

NPV, negative predictive value; PPV, positive predictive value.  
Results based on models using multiple imputation to deal with missing values.

**TABLE 30** Parameters to calculate risk score for the clean-catch 'full' coefficient-based models

Variables	Coefficients	
	Symptoms and signs model	Symptoms, signs and dipstick model
Intercept	-6.0152	-7.0011
Pain/crying when passing urine		
No problem	0	0
Slight problem	0.6263	0.1734
Moderate problem	1.8009	1.2658
Severe problem	3.1693	2.8062
Smelly urine		
No problem	0	0
Slight problem	1.6683	1.3838
Moderate problem	1.8805	1.7631
Severe problem	2.4923	1.7924
History of UTI		
No	0	0
Yes	1.1227	1.0299
Cough		
No problem	0	0
Slight problem	0.3156	0.3118
Moderate problem	0.3736	0.8541
Severe problem	-1.4324	-1.2293
Clinician global impression of illness severity (0–10) <sup>a</sup>		
0–1	0	0
2	0.7818	0.9103
3	1.1509	1.1613
4–5	1.5543	1.4148
≥ 6	2.2732	2.2090
Abdominal exam: any abdominal pain		
No	0	0
Yes	0.9244	0.1987
Ear exam: any acute abnormality		
No	0	0
Yes	-1.5122	-1.0919
Dipstick: leucocytes		
Negative		0
Trace		0.6904
+		-0.4912
++		1.9791
+++		2.8103

continued

**TABLE 30** Parameters to calculate risk score for the clean-catch ‘full’ coefficient-based models (*continued*)

Variables	Coefficients	
	Symptoms and signs model	Symptoms, signs and dipstick model
Dipstick: nitrites		
Negative		0
Positive		2.0204
Dipstick: blood		
Negative		0
Non-haem		-0.1465
Haem trace		1.6902
Haem +		1.2556
Haem ++ or +++		0.6638

a 0 = completely well, 10 = extremely unwell.

Clean-catch samples were available for 88, 91 and 612 children aged < 12, 12–23 and 24–35 months, with UTI diagnosed in 4, 2 and 16 of these children, respectively. We acknowledge that the small number of UTIs may have led to results from multiple imputation analyses being unstable and estimated AUROC values being overoptimistic.

**Fever and fever of unknown origin (clean-catch samples)**

Table 31 shows that neither of the parent-reported fever variables met our clean-catch multivariable model inclusion criteria on univariable analysis, that is they were not associated with UTI. None of the children with a temperature of ≥ 38 °C and no other localising symptoms or signs providing a clean-catch sample had a UTI.

**Points-based models**

Table 32 shows the ORs (95% CI) and AUROC (95% CI) for the dichotomised points-based models with and without the dipstick tests and, for comparison, the AUROC for the respective full models. It shows the relative (integer) contribution of symptoms, signs and dipstick to UTI and that the assigned points do change with the addition of the dipstick tests. The table also shows that for symptoms and signs model, the AUROC marginally falls from the full validated symptoms and signs model 0.867 to the dichotomised coefficient-based points-based model-validated AUROC of 0.849 to the (non-bootstrapped) AUROC of 0.856 for the points-based model. The AUROC for the dichotomised symptoms, signs and dipstick points models remain highly discriminatory at 0.900 and little different from the full symptoms, signs and dipstick

**TABLE 31** Fever and fever of unknown origin

Symptom	Category	Complete cases		Modal imputation	
		OR (95% CI)	p-value	OR (95% CI)	p-value
Fever now or in past 24 hours: unknown origin	No	1 (ref.)	0.464	1 (ref.)	0.876
	Yes	2.36 (0.31 to 18.24)		1.18 (0.16 to 8.72)	
Fever at any time during this illness: unknown origin	No	1 (ref.)	0.163	1 (ref.)	0.484
	Yes	3.42 (0.77 to 15.12)		1.74 (0.41 to 7.32)	

Ref., reference.

**TABLE 32** Clean-catch points-based models using multiple imputation

Dichotomised variables	Symptoms and signs model		Symptoms, signs and dipstick model	
	Adjusted OR (95% CI)	Points	Adjusted OR (95% CI)	Points
Pain/crying passing urine presence	5.46 (3.06 to 9.75)	2	3.27 (1.70 to 6.30)	2
Smelly urine presence	6.83 (3.75 to 12.44)	2	5.20 (2.70 to 9.99)	2
UTI history	2.72 (1.33 to 5.55)	1	2.17 (0.94 to 5.02)	1
Absence of severe cough	4.66 (1.37 to 15.92)	2	3.13 (0.88 to 11.08)	2
Severe illness presence <sup>a</sup>	4.57 (1.77 to 11.79)	2	4.51 (1.61 to 12.66)	2
Dipstick: leucocytes positive			4.36 (2.22 to 8.56)	2
Dipstick: nitrites positive			8.07 (3.90 to 16.70)	3
Dipstick: blood positive			2.13 (1.12 to 4.05)	1
AUROC	0.860 (0.810 to 0.910)	0.856 (0.808 to 0.903)	0.902 (0.8 to 0.950)	0.900 (0.853 to 0.948)
Validated AUROC	0.849		0.892	
$\Delta$ ROC curve <sup>b</sup> (95% CI)			0.042 (0.01 to 0.07)	0.045 (0.02 to 0.07)
$\Delta$ ROC curve <sup>b</sup> <i>p</i> -value			0.004	0.003
Calibration slope	0.947		0.942	

<sup>a</sup> A score of  $\geq 6$  on the global impression scale.

<sup>b</sup> The difference in ROC curve between the symptoms and signs model and the symptoms, signs and dipstick model.

validated model AUROC of 0.903. The difference in AUROC for both the points- and coefficient-based models was just over 0.04 and the low *p*-values show that there is very strong evidence that the dipsticks do increase the AUROC when added to the symptoms and signs model. The calibration slopes of 0.95 indicate a very well-calibrated clean-catch points-based model.

The clean-catch model in under-3-year-olds (*Table 33*) is parametrically similar to the clean-catch model in all children (see *Table 32*), as the ORs have overlapping CIs. In general the ORs are slightly higher within the under-3-year-olds, except for pain/crying passing urine presence and absence of severe cough, which are slightly lower. Given this evidence, we conclude that our model is applicable to all children under 5 with clean-catch samples.

*Table 34* shows the diagnostic test characteristics [sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV)] and the proportion rule positive for different cut-points for the points-based models for symptoms and signs, and symptoms, signs and dipstick results (based on models using multiple imputation to deal with missing values). The cost-effectiveness analyses presented in *Chapter 6* suggest that the optimal threshold to sample urine would be the 51.7% sensitivity (high specificity) threshold (shaded in green in *Table 34*). The cost-effectiveness analyses could not discern the optimal antibiotic treatment strategy. *Table 35* gives an illustrative example for 1000 children. It shows the number of samples that would have to be collected and how many would have UTI present if a clinician used the points-based model.

**TABLE 33** Clean-catch rule of points-based models using multiple imputation for under-3-year-olds

Dichotomised variables	Symptoms and signs model, adjusted OR (95% CI)	Symptoms, signs and dipstick model, adjusted OR (95% CI)
Pain/crying passing urine presence	4.27 (1.54 to 11.85)	2.32 (0.58 to 9.38)
Smelly urine presence	13.36 (3.92 to 45.48)	23.13 (4.60 to 116.46)
UTI history	3.76 (1.05 to 13.41)	2.61 (0.35 to 19.72)
Absence of severe cough	2.23 (0.44 to 11.31)	2.00 (0.28 to 14.42)
Severe illness presence <sup>a</sup>	8.42 (1.87 to 37.82)	18.17 (2.90 to 113.76)
Dipstick: leucocytes positive		13.63 (3.43 to 54.18)
Dipstick: nitrites positive		8.20 (2.15 to 31.28)
Dipstick: blood positive		3.32 (1.03 to 10.72)

a A score of  $\geq 6$  on the global impression scale.

**TABLE 34** Clean-catch diagnostic test characteristics (95% CI) for a range of cut-points for the points-based model, using both symptoms and signs model (upper part of table) and symptoms, signs and dipstick model (lower part of table)

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

Shaded cells are the thresholds at which urine could be cost-effectively sampled (upper table).



**TABLE 35** Outcome for illustrative 1000 children  $\geq 2$  years presenting unwell to primary care using  $\geq 5$  points (sensitivity = 51.7%, specificity = 94.6%)

DUTY model	UTI present	UTI absent	Total
Points based: collect urine	11	53	64
Points based: do not collect urine	11	926	936
Total	22	978	1000

Number of contaminated samples (data from *Chapter 7*; see *Table 68*).  
 $\geq 10^5$  CFU/ml  $\geq 3$  organisms (1.8% of 64) = 1.  
 $\geq 10^5$  CFU/ml  $\geq 2$  organisms (2.9% of 64) = 2.  
 $\geq 10^4$  CFU/ml  $\geq 3$  organisms (6.4% of 64) = 4.

### Added value of dipstick testing

*Table 36* shows the diagnostic test characteristics (sensitivity, specificity, PPV and NPV) and the proportion rule positive for different cut-points for the coefficient-based points-based models for any combination of symptoms and signs.

*Table 37* shows, for each symptom and sign combination, the probability of UTI (in this case, the modelled PPV), the points total, the number of UTIs and total observations, and the absolute change in probability using the dipstick model from our simulations. It shows that there is a clear trend towards an increased change (both increasing and decreasing) in the UTI probability with added dipstick as the probability of UTI post symptoms and signs increases. We then grouped the data into those symptoms and signs combinations where we have or have not thought that urine sampling would be reasonable. The change in probability for those in whom we would not suggest sampling (group 2) is very slight and clustered around zero so dipstick testing would add very little diagnostic value. However, for those in whom we would suggest urine sampling (group 1), there is a clinically important change in probability and dipstick testing is worthwhile in the ID of those with and without a UTI.

*Box 1* shows how the points-based model could be presented to practising clinicians and includes practical guidance regarding when to obtain a urine sample. Data from *Chapter 6* show it is not clear which of the following antibiotic treatment strategies is most cost-effective: immediate presumptive treatment of all sampled children, immediate dipstick guided treatment or treatment delayed until receipt of laboratory result.

### Nappy pad models

*Table 38* shows that 1.3% of children providing nappy pad samples met the criteria for a microbiological diagnosis of UTI, of whom 82% were  $< 2$  years and 48% were female, and that 92% of samples were provided within 24 hours of index test measurement. Clinicians correctly identified 13% of the 30 UTIs.

### 'Full' coefficient-based model

*Table 39* shows adjusted ORs for the index tests (footnote marker d in *Table 38*) selected for the nappy pad model, for children predominantly  $< 2$  years. Parent-reported smelly urine, darker urine, female sex and the absence of a nappy rash were independently associated with UTI; for the first two of these there was evidence of graded associations. No clinical examination findings were independently associated with UTI in this age group. The presence of leucocytes and nitrites from dipstick urine testing were independently associated with UTI, though more modestly than in clean-catch samples. The increase in AUROC between the symptoms and signs model and the symptoms, signs and dipstick model was about 0.065 in the multiple imputation analysis. However, with a  $p$ -value of 0.035 there was only slight evidence to suggest an increase in the AUROC when adding the dipsticks to the symptoms and signs model. The symptoms and signs model had reasonable diagnostic accuracy (validated AUROC for the multiple imputed model was 0.778) and diagnostic accuracy increased (validated AUROC 0.821) with addition of dipstick findings. For comparison, the AUROC for 'clinician diagnosis' was 0.626 (95% CI 0.532 to 0.719).

**TABLE 36** Clean-catch diagnostic test characteristics (95% CI) for all possible cut-points of the coefficient-based points-based model, using the symptoms and signs model

Combination	Sensitivity, %	Specificity, %	PPV, <sup>a</sup> %	NPV, <sup>b</sup> %	Children positive, %
+,+,+,+,+	0	100	100	97.8 (97.2 to 98.3)	0
+,+,+,+,-	0	100	100	97.8 (97.2 to 98.3)	0
+,+,+,-,+	8.3 (3.5 to 18.5)	99.7 (99.5 to 99.9)	41.7 (18.5 to 69.2)	98.0 (97.4 to 98.4)	0.4 (0.2 to 0.8)
+,+,-,+,+	8.3 (3.5 to 18.5)	99.7 (99.4 to 99.9)	38.5 (17.0 to 65.6)	98.0 (97.4 to 98.4)	0.5 (0.3 to 0.8)
+,-,+,+,+	8.3 (3.5 to 18.5)	99.7 (99.4 to 99.9)	38.5 (17.0 to 65.6)	98.0 (97.4 to 98.4)	0.5 (0.3 to 0.8)
-+,+,+,+	11.7 (5.7 to 22.5)	99.7 (99.4 to 99.8)	43.8 (22.5 to 67.6)	98.1 (97.5 to 98.5)	0.6 (0.4 to 1.0)
+,+,+,-,-	41.7 (29.9 to 54.4)	96.8 (96.1 to 97.4)	22.5 (15.7 to 31.2)	98.7 (98.2 to 99.0)	4.1 (3.4 to 4.9)
+,+,-,+,-	41.7 (29.9 to 54.4)	96.8 (96.1 to 97.4)	22.5 (15.7 to 31.2)	98.7 (98.2 to 99.0)	4.1 (3.4 to 4.9)
+,-,+,+,-	41.7 (29.9 to 54.4)	96.6 (95.8 to 97.2)	21.4 (14.9 to 29.7)	98.7 (98.1 to 99.0)	4.3 (3.6 to 5.1)
-+,+,+,-	43.3 (31.5 to 56.0)	96.3 (95.6 to 97.0)	21.0 (14.7 to 29.0)	98.7 (98.2 to 99.1)	4.5 (3.8 to 5.4)
+,+,-,-,+	43.3 (31.5 to 56.0)	96.2 (95.4 to 96.8)	20.2 (14.1 to 28.0)	98.7 (98.2 to 99.1)	4.7 (4.0 to 5.6)
+,-,+,-,+	48.3 (36.1 to 60.8)	95.5 (94.6 to 96.2)	19.3 (13.8 to 26.4)	98.8 (98.3 to 99.2)	5.5 (4.7 to 6.4)
+,-,-,+,+	48.3 (36.1 to 60.8)	95.4 (94.6 to 96.2)	19.2 (13.7 to 26.3)	98.8 (98.3 to 99.2)	5.5 (4.7 to 6.4)
-+,+,-,+	51.7 (39.2 to 63.9)	94.8 (93.9 to 95.6)	18.2 (13.1 to 24.8)	98.9 (98.4 to 99.2)	6.2 (5.4 to 7.2)
-+,-,+,+	51.7 (39.2 to 63.9)	94.8 (93.9 to 95.6)	18.2 (13.1 to 24.8)	98.9 (98.4 to 99.2)	6.2 (5.4 to 7.2)
-,-,+,+,+	53.3 (40.8 to 65.5)	94.4 (93.5 to 95.2)	17.7 (12.8 to 23.9)	98.9 (98.4 to 99.2)	6.6 (5.7 to 7.6)
+,+,-,-,-	53.3 (40.8 to 65.5)	93.5 (92.5 to 94.4)	15.6 (11.3 to 21.2)	98.9 (98.4 to 99.2)	7.5 (6.6 to 8.5)
+,-,+,-,-	65.0 (52.2 to 75.9)	87.0 (85.6 to 88.2)	10.1 (7.4 to 13.5)	99.1 (98.6 to 99.4)	14.2 (12.9 to 15.5)
+,-,-,+,-	65.0 (52.2 to 75.9)	86.8 (85.5 to 88.0)	9.9 (7.3 to 13.3)	99.1 (98.6 to 99.4)	14.3 (13.1 to 15.7)
-+,+,-,-	76.7 (64.4 to 85.7)	80.2 (78.7 to 81.7)	8.0 (6.0 to 10.5)	99.4 (98.9 to 99.6)	21.0 (19.5 to 22.6)
-+,-,+,-	76.7 (64.4 to 85.7)	80.1 (78.6 to 81.6)	8.0 (6.0 to 10.5)	99.4 (98.9 to 99.6)	21.1 (19.6 to 22.7)
-,-,+,+,-	81.7 (69.8 to 89.5)	77.8 (76.2 to 79.3)	7.6 (5.8 to 9.9)	99.5 (99.1 to 99.7)	23.5 (22.0 to 25.1)
+,-,-,-,+	81.7 (69.8 to 89.5)	77.6 (76.0 to 79.1)	7.5 (5.7 to 9.8)	99.5 (99.1 to 99.7)	23.7 (22.2 to 25.4)
-+,+,-,-,+	81.7 (69.8 to 89.5)	77.2 (75.6 to 78.8)	7.4 (5.7 to 9.7)	99.5 (99.0 to 99.7)	24.1 (22.5 to 25.7)
-,-,+,-,+	85.0 (73.6 to 92.0)	74.5 (72.8 to 76.1)	6.9 (5.3 to 9.0)	99.6 (99.1 to 99.8)	26.8 (25.2 to 28.5)
-,-,-,+,-	85.0 (73.6 to 92.0)	74.4 (72.7 to 76.0)	6.9 (5.3 to 9.0)	99.6 (99.1 to 99.8)	26.9 (25.3 to 28.6)
+,-,-,-,-	88.3 (77.5 to 94.3)	70.0 (68.3 to 71.7)	6.2 (4.8 to 8.0)	99.6 (99.2 to 99.8)	31.2 (29.5 to 33.0)
-+,-,-,-,-	88.3 (77.5 to 94.3)	67.3 (65.5 to 69.1)	5.7 (4.4 to 7.4)	99.6 (99.2 to 99.8)	33.9 (32.2 to 35.7)
-,-,+,-,-	96.7 (87.6 to 99.2)	17.3 (15.9 to 18.8)	2.6 (2.0 to 3.3)	99.6 (98.3 to 99.9)	83.0 (81.5 to 84.4)
-,-,-,+,-	98.3 (89.1 to 99.8)	16.9 (15.5 to 18.3)	2.6 (2.0 to 3.3)	99.8 (98.5 to 100.0)	83.5 (82.0 to 84.8)
-,-,-,-,+	100	15.9 (14.6 to 17.3)	2.6 (2.0 to 3.3)	100	84.5 (83.0 to 85.8)
-,-,-,-,-	100	0	2.2 (1.7 to 2.8)	100	100

a Positive predictive value.

b Negative predictive value.

If a test was present it is marked as +; if it was negative it is marked -. The first descriptor is for smelly urine, the second for if pain passing urine is present, the third if severe cough is absent, the fourth if there is severe illness, and the fifth if there is history of UTI.

**TABLE 37** Additional diagnostic value of dipstick testing and also the prevalence of each combination of symptoms and signs (clean-catch samples)

Signs and symptoms <sup>a</sup>	Modelled PPV <sup>b</sup>	Points total	UTI/total	Absolute change in probability; median (2.5th and 97.5 percentiles)	Grouping <sup>c</sup>
+,+,+,+,+	0.731	9	0/0	0.239 (0.043, 0.433)	1
+,+,+,+,-	0.500	8	0/1	0.336 (0.039, 0.436)	1
+,+,+,-,+	0.372	7	7/17	0.286 (0.081, 0.503)	1
+,+,-,+,+	0.368	7	0/1	0.248 (0.004, 0.542)	1
+,-,+,+,+	0.332	7	0/0	0.217 (0.030, 0.575)	1
-+,+,+,+	0.284	7	0/1	0.209 (0.022, 0.575)	1
+,+,+,-,-	0.179	6	19/101	0.138 (0.020, 0.585)	1
+,+,-,+,-	0.176	6	0/0	0.118 (0.038, 0.647)	1
+,-,+,+,-	0.155	6	0/7	0.098 (0.042, 0.523)	1
-+,+,+,-	0.128	6	1/9	0.091 (0.013, 0.610)	1
+,+,-,-,+	0.113	5	0/4	0.084 (0.003, 0.579)	1
+,-,+,-,+	0.098	5	2/23	0.070 (0.014, 0.585)	1
+,-,-,+,+	0.096	5	0/1	0.056 (0.015, 0.497)	1
-+,+,-,+	0.080	5	2/22	0.062 (0.007, 0.495)	1
-+,+,-,+	0.078	5	0/0	0.053 (0.024, 0.583)	1
-,-,+,+,+	0.068	5	0/6	0.043 (0.017, 0.400)	1
+,+,-,-,-	0.045	4	0/6	0.031 (0.012, 0.464)	2
+,-,+,-,-	0.038	4	8/202	0.025 (0.011, 0.280)	2
+,-,-,+,-	0.038	4	0/4	0.019 (0.001, 0.364)	2
-+,+,-,-	0.031	4	6/198	0.023 (0.004, 0.353)	2
-+,+,-,-	0.030	4	0/0	0.019 (0.005, 0.267)	2
-,-,+,+,-	0.026	4	3/69	0.015 (0.002, 0.262)	2
+,-,-,-,+	0.023	3	0/5	0.014 (0.004, 0.056)	2
-+,+,-,-,+	0.018	3	0/4	0.013 (0.006, 0.151)	2
-,-,+,-,+	0.016	3	3/81	0.010 (0.004, 0.033)	2
-,-,-,+,-	0.015	3	0/1	0.007 (0.001, 0.054)	2
+,-,-,-,-	0.008	2	2/75	0.004 (0.000, 0.029)	2
-+,+,-,-,-	0.007	2	0/21	0.004 (0.001, 0.017)	2
-,-,+,-,-	0.006	2	6/1424	0.003 (0.000, 0.017)	2
-,-,-,+,-	0.006	2	1/14	0.002 (0.002, 0.027)	2
-,-,-,-,+	0.003	1	0/25	0.002 (0.000, 0.013)	2
-,-,-,-,-	0.001	0	0/418	0.000 (0.000, 0.006)	2
Grouping	Weighted; median		UTI/total	Weighted; median (2.5th, 97.5th percentiles)	
1	0.179		31/193	0.108 (0.013, 0.585)	
2	0.006		29/2547	0.003 (0.000, 0.041)	

a 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 history of UTI.

b Positive predictive value.

c 1 = points-based model suggests urine collection cost-effective ( $\geq 5$  points); 2 = do not collect urine ( $< 5$  points).

**BOX 1** Final DUTY Clean-Catch Urine Criteria

1. The *DUTY Clean-Catch Urine Criteria* are for children in whom a clean-catch sample is possible (usually children  $\geq 2$  years).
2. Use the symptoms and signs in *Table A* to decide if a clean-catch urine should be collected ( $\geq 5$  points or can be operationalised as 'any three of the five' symptoms and signs has been shown to be cost-effective). Clinicians concerned about overdiagnosis and treatment can select a higher specificity (at least six points) threshold. Higher sensitivity thresholds (e.g.  $\geq 3$  points or  $\geq 4$  points) would reduce underdiagnosis, but these thresholds have not been shown to be cost-effective.
3. It is not clear which of the following possible antibiotic treatment strategies is most cost effective: (1) immediate presumptive treatment of all sampled children; (2) immediate dipstick-guided treatment; or (3) laboratory guided (delayed) treatment.
4. For children urine sampled at the  $\geq 5$ -point threshold, the probability of UTI will be 18%. Although not demonstrably cost-effective, dipstick testing can raise or lower this probability (*Table B*).
5. Consider advising all (urine and non-urine sampled) children's parents to seek medical advice if their child gets worse.
6. The *DUTY Clean-Catch Urine Criteria* are designed to supplement and not replace clinical judgement.

**TABLE A** Points to select children for urine sampling

Clinical characteristic (present/absent)	Points <sup>a</sup>
<i>Symptoms and signs</i>	<i>To guide urine collection</i>
Pain/crying passing urine presence	2
Smelly urine presence	2
UTI history	1
Absence of severe cough	2
Severe illness presence <sup>b</sup>	2
<i>Collect clean-catch urine if symptoms and signs points total <math>\geq 5</math> or 'any three of the five'</i>	

<sup>a</sup> Refer to *Table 36* (upper) for probability of UTI with any total score.

<sup>b</sup> Score of  $\geq 6$  on the clinician global illness severity scale (range 0–10).

**TABLE B** Points to select children for antibiotic treatment

Clinical characteristic (present/absent)	Points <sup>a</sup>
<i>Symptoms, signs and dipstick</i>	<i>To guide antibiotic treatment</i>
Pain/crying passing urine presence	2
Smelly urine presence	2
UTI history	1
Absence of severe cough	2
Severe illness presence <sup>b</sup>	2
Dipstick: leucocytes positive	2
Dipstick: nitrites positive	3
Dipstick: blood positive	1

<sup>a</sup> Refer to *Table 36* for probability of UTI with total score.

<sup>b</sup> Score of  $\geq 6$  on the clinician global illness severity scale (range 0–10).

**TABLE 38** Nappy pad samples: children's characteristics and crude ORs for index tests associated with UTI

Demographics/index tests <sup>a</sup>	<i>n</i> (%) <sup>b</sup>	UTI (%) <sup>c</sup>	Crude OR (95% CI) <sup>d</sup>
Total	2277	30 (1.3)	
Age of child			
< 6 months	369 (16.2)	5 (1.4)	1.72 (0.54 to 5.46)
6 to < 12 months	603 (26.5)	11 (1.8)	2.33 (0.90 to 6.04)
1 to < 2 years	884 (38.8)	7 (0.8)	1 (ref.)
2 to < 3 years	353 (15.5)	7 (2.0)	2.53 (0.88 to 7.28)
3 to < 4 years	58 (2.5)	0 (0.0)	N/A
≥ 4 years	10 (0.4)	0 (0.0)	N/A
Time from index tests to taking urine sample			
Sample before recruitment	120 (5.3)	2 (1.7)	1.33 (0.31 to 5.67)
< 24 hours	1982 (87.0)	25 (1.3)	1 (ref.)
24 hours to < 48 hours	109 (4.8)	3 (2.8)	2.22 (0.66 to 7.45)
48 hours to < 72 hours	18 (0.8)	0 (0.0)	N/A
≥ 72 hours	48 (2.1)	0 (0.0)	N/A
Clinician diagnosis prior to dipstick			
No UTI certain to very certain	1033 (45.4)	8 (0.8)	0.52 (0.22 to 1.19)
Uncertain or no UTI fairly certain	1201 (52.7)	18 (1.5)	1 (ref.)
UTI fairly to very certain	38 (1.7)	4 (10.5)	7.76 (2.49 to 24.18)
Missing	5 (0.2)	0 (0.0)	
Gender <sup>a</sup>			
Male	1183 (52.0)	9 (0.8)	1 (ref.)
Female	1094 (48.0)	21 (1.9)	2.55 (1.16 to 5.60)
Smelly urine <sup>a</sup>			
No problem	1518 (66.7)	12 (0.8)	1 (ref.)
Slight problem	206 (9.0)	4 (1.9)	2.20 (0.73 to 6.61)
Moderate problem	138 (6.1)	5 (3.6)	4.18 (1.52 to 11.50)
Severe problem	26 (1.1)	4 (15.4)	20.21 (6.29 to 64.97)
Missing/not known	389 (17.1)	5 (1.3)	
Darker urine <sup>a</sup>			
No problem	1764 (77.5)	19 (1.1)	1 (ref.)
Slight problem	83 (3.6)	2 (2.4)	2.19 (0.51 to 9.43)
Moderate or severe problem	41 (1.8)	4 (9.8)	9.59 (3.17 to 29.02)
Missing/not known	389 (17.1)	5 (1.3)	

continued

**TABLE 38** Nappy pad samples: children’s characteristics and crude ORs for index tests associated with UTI (*continued*)

Demographics/index tests <sup>a</sup>	<i>n</i> (%) <sup>b</sup>	UTI (%) <sup>c</sup>	Crude OR (95% CI) <sup>d</sup>
Nappy rash <sup>a</sup>			
No problem	1715 (75.3)	29 (1.7)	1 (ref.)
Slight to severe problem	560 (24.6)	1 (0.2)	0.10 (0.01 to 0.77)
Missing	2 (0.1)	0 (0.0)	
Dipstick: leucocytes <sup>a</sup>			
Negative	1759 (77.3)	13 (0.7)	1 (ref.)
Trace	125 (5.5)	1 (0.8)	1.09 (0.14 to 8.38)
+	119 (5.2)	4 (3.4)	4.69 (1.50 to 14.61)
++	177 (7.8)	4 (2.3)	3.12 (1.01 to 9.66)
+++	91 (4.0)	8 (8.8)	12.99 (5.24 to 32.20)
Missing	6 (0.3)	0 (0.0)	
Dipstick: nitrites <sup>a</sup>			
Negative	1916 (84.1)	13 (0.7)	1 (ref.)
Positive	355 (15.6)	17 (4.8)	7.39 (3.55 to 15.35)
Missing	6 (0.3)	0 (0.0)	

N/A, not applicable; ref., reference.

a Index tests independently associated with UTI in multivariable models.

b Total column gives the percentage of observations within that category.

c UTI column gives the percentage of positives relative to the number of observations within that category.

d Index tests associated with UTI.

**TABLE 39** Nappy pad samples: models based on symptoms and signs, and on symptoms, signs and dipstick results, including results based on multiple imputation

Index tests	Symptoms and signs model		Symptoms, signs and dipstick model	
	Adjusted OR (95% CI) <sup>a</sup>	MI adjusted OR (95% CI)	Adjusted OR (95% CI) <sup>a</sup>	MI adjusted OR (95% CI)
Gender				
Male	1 (ref.)	1 (ref.)	1 (ref.)	1 (ref.)
Female	2.61 (1.16 to 5.83)	2.45 (1.08 to 5.55)	1.71 (0.71 to 4.08)	1.70 (0.70 to 4.11)
Smelly urine				
No problem	1 (ref.)	1 (ref.)	1 (ref.)	1 (ref.)
Slight problem	1.99 (0.64 to 6.17)	2.47 (0.77 to 7.91)	1.76 (0.54 to 5.78)	2.28 (0.68 to 7.68)
Moderate problem	3.75 (1.21 to 11.65)	5.11 (1.66 to 15.76)	3.26 (0.98 to 10.87)	4.64 (1.39 to 15.53)
Severe problem	17.81 (4.82 to 65.86)	22.04 (6.05 to 80.25)	8.42 (2.03 to 34.88)	12.94 (3.20 to 52.30)
Darker urine				
No problem	1 (ref.)	1 (ref.)	1 (ref.)	1 (ref.)
Slight problem	2.51 (0.55 to 11.49)	2.50 (0.56 to 11.22)	2.50 (0.51 to 12.30)	2.52 (0.52 to 12.21)
Moderate or severe problem	3.66 (0.97 to 13.82)	2.98 (0.81 to 10.96)	3.59 (0.90 to 14.42)	3.19 (0.82 to 12.37)
Nappy rash				
No problem	1 (ref.)	1 (ref.)	1 (ref.)	1 (ref.)
Slight to severe problem	0.07 (0.01 to 0.55)	0.07 (0.01 to 0.52)	0.08 (0.01 to 0.59)	0.07 (0.01 to 0.56)
Dipstick: leucocytes				
Negative			1 (ref.)	1 (ref.)
Trace			0.80 (0.10 to 6.37)	0.74 (0.09 to 6.13)
+			3.06 (0.88 to 10.68)	3.01 (0.83 to 10.91)
++			2.13 (0.62 to 7.30)	2.25 (0.65 to 7.84)
+++			6.23 (2.20 to 17.67)	5.53 (1.90 to 16.10)
Dipstick: nitrites				
Negative			1 (ref.)	1 (ref.)
Positive			5.93 (2.72 to 12.93)	6.35 (2.86 to 14.12)
ROC (95% CI)	0.769 (0.68 to 0.85)	0.805 (0.72 to 0.89)	0.858 (0.79 to 0.93)	0.870 (0.80 to 0.94)
Validated ROC <sup>b</sup>	0.744	0.778	0.799	0.821
Δ ROC <sup>c</sup> (95% CI)			0.089 (0.02 to 0.16)	0.065 (0.00 to 0.13)
Δ ROC <sup>c</sup> p-value			0.012	0.036
Calibration slope <sup>c</sup>	0.695	0.749	0.647	0.708

MI, multiple imputation; ref., reference.

a Missing values coded to modal category.

b Internal validation using the bootstrap procedure.

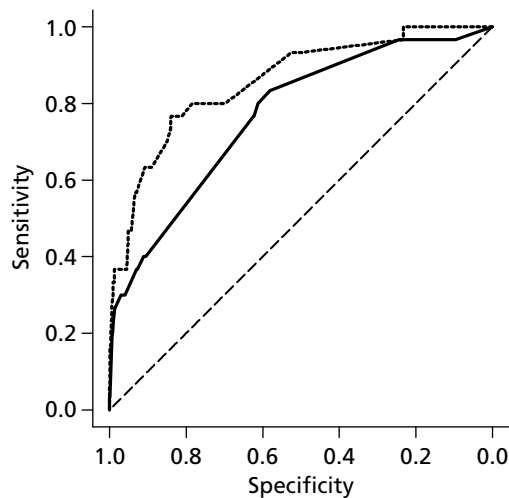
c The difference in ROC between the symptoms and signs model and the symptoms, signs and dipstick model.

Figure 12 shows the multiple imputed ROC curves for the models with and without dipstick urinalysis. The calibration slopes being around 0.7 shows that we have a large degree of overoptimism in our nappy pad models.

The upper section of Table 36 shows the diagnostic test characteristics corresponding to various choices of cut-point for positivity using the symptoms and signs model, which could be used to identify the children in whom urine collection is warranted. The risk scores for these cut-points can be calculated using the parameters in Table 40. For a sensitivity of 50%, 63% or 90%, respective urine sampling of approximately 9%, 20% or 43% of unwell children aged < 2 years presenting to primary care would be required. At these cut-points, urine samples would be positive for UTI in 8%, 4% and 3% of children, with corresponding specificities of 92%, 81% and 58%. The lower section of Table 36 shows the diagnostic test characteristics for the model based on the symptoms, signs and dipstick test results. If a decision to treat presumptively with antibiotics was made using the (only comparable) 63% sensitivity cut-point, this model would improve the specificity from 81% to 91% and would reduce the overall number of children treated from 20% to 10%, compared with the symptoms and signs only model. However, as discussed in Chapter 6, the cost-effectiveness of this strategy is not clear.

### Fever and fever of unknown origin (nappy pad samples)

Table 41 shows that neither of the parent-reported fever variables met our nappy pad multivariable model inclusion criteria on univariable analysis, that is they were not associated with UTI. None of the children with a temperature of  $\geq 38^\circ\text{C}$  and no other localising symptoms or signs providing a nappy pad sample had a UTI.



**FIGURE 12** Nappy pad ROC curve from multiple imputation for symptoms and signs only (solid line) and symptoms, signs and dipstick (dotted line).



**TABLE 40** Parameters to calculate risk score for the nappy pad 'full' coefficient-based models

Index tests	Coefficients	
	Symptoms and signs model	Symptoms, signs and dipstick model
Intercept	-5.2709	-6.0041
Gender		
Male	0	0
Female	0.8972	0.5293
Smelly urine		
No problem	0	0
Slight problem	0.9042	0.8251
Moderate problem	1.6310	1.5343
Severe problem	3.0927	2.5601
Darker urine		
No problem	0	0
Slight problem	0.9166	0.9243
Moderate or severe problem	1.0913	1.1598
Nappy rash		
No problem	0	0
Slight to severe problem	-2.6792	-2.6098
Dipstick: leucocytes		
Negative		0
Trace		-0.2988
+		1.1029
++		0.8105
+++		1.7095
Dipstick: nitrites		
Negative		0
Positive		1.8488

**TABLE 41** Relationship between fever of unknown origin and UTI in nappy pad samples

Symptom	Category	Complete cases		Modal imputation	
		OR (95% CI)	p-value	OR (95% CI)	p-value
Fever now in or in past 24 hours: unknown origin	No	1 (ref.)	0.064	1 (ref.)	0.310
	Yes	17.6 (1.55 to 199.9)		3.48 (0.45 to 26.72)	
Fever at any time during this illness: unknown origin	No	1 (ref.)	0.092	1 (ref.)	0.339
	Yes	12.5 (1.15 to 135.4)		3.19 (0.42 to 24.39)	

Ref., reference.

### Points-based models

Table 42 shows the points-based multiple imputation model ORs (95% CIs) using the dichotomised symptoms, with and without dichotomised dipstick test results, along with their corresponding integer-based points and the AUROCs for the dichotomised coefficient and points models. These show that, while not as discriminatory as the clean-catch models, the points-based model AUROCs, remain similar to the dichotomised coefficient AUROCs and to the full validated model AUROCs with and without the dipstick tests. It is also noteworthy that the ‘darker urine’ and ‘absence of nappy rash’ points change on the addition of the dipstick tests to the model, such that clinicians would need to retotal the points in the nappy pad points-based model in the light of the dipstick results. The difference in AUROC for both the points- and coefficient-based points-based models was around 0.06, and the low *p*-values show that there was only slight evidence that the dipsticks do increase the AUROC when added to the symptoms and signs model. The calibration slopes being over 0.85 in these points-based models shows that we have good calibration in the points-based nappy pad models.

Table 43 shows the diagnostic test characteristics (sensitivity, specificity, PPV and NPV) and the proportion rule positive for different cut points for the coefficient-based models. Similarly, Table 44 shows the diagnostic test characteristics (sensitivity, specificity, PPV and NPV) and the proportion rule positive for different cut-points for the points-based models for symptoms, and symptoms and dipstick results (based on models using multiple imputation to deal with missing values). This shows that if clinicians were willing to collect nappy pad samples on 13.4% of younger (< 2 years) preschool children, they would need to use a cut-point of  $\geq 5$  with an achieved sensitivity of 53.3%, that is a sensitivity considerably higher than ‘clinician diagnosis’ alone in this age group. However, as we discuss in Chapter 6, the cost-effectiveness of this strategy is not clear. Table 45 shows the diagnostic test characteristics (sensitivity, specificity, PPV and NPV) and the proportion rule positive for different cut points for the coefficient-based points-based models for the symptoms and signs model for any combination of symptoms and signs.

Table 45 shows the diagnostic characteristics of all possible combinations of symptoms and signs for the dichotomised nappy pad model.

**TABLE 42** Nappy pad points-based dichotomised models with and without dipstick test

Dichotomised variables	Symptoms and signs model		Symptoms, signs and dipstick model	
	Adjusted OR (95% CI)	Points	Adjusted OR (95% CI)	Points
Female	2.48 (1.12 to 5.51)	1	1.88 (0.81 to 4.37)	1
Smelly urine presence	4.77 (2.14 to 10.61)	2	4.62 (2.02 to 10.53)	2
Darker urine presence	3.19 (1.20 to 8.49)	1	2.80 (0.99 to 7.95)	2
Absence of nappy rash	12.74 (1.71 to 94.62)	3	13.96 (1.84 to 105.60)	4
Dipstick: leucocytes positive			2.89 (1.31 to 6.40)	2
Dipstick: nitrites positive			6.09 (2.82 to 13.15)	3
AUROC	0.801 (0.716 to 0.885)	0.799 (0.714 to 0.884)	0.860 (0.793 to 0.927)	0.862 (0.795 to 0.927)
Validated AUROC	0.795		0.842	
$\Delta$ ROC <sup>a</sup> (95% CI)			0.059 (0.00 to 0.12)	0.062 (0.00 to 0.12)
$\Delta$ ROC <sup>a</sup> <i>p</i> -value			0.019	0.046
Calibration slope	0.887		0.866	

a The difference in ROC between the symptoms and signs model and the symptoms, signs and dipstick model.

**TABLE 43** Nappy pad diagnostic test characteristics (95% CI) for a range of cut-points using both the symptoms and signs model (upper part of table) and the symptoms, signs and dipstick model (lower part of table)

Risk score ( $\geq$ )	Sensitivity, %	Specificity, %	PPV, %	NPV, %	Children positive, %
<b><i>Symptoms and signs model</i></b>					
-1.64	20.0 (9.3 to 37.9)	99.6 (99.2 to 99.8)	40.0 (19.2 to 65.2)	98.9 (98.4 to 99.3)	0.7 (0.4 to 1.1)
-2.55	30.0 (16.4 to 48.3)	99.0 (98.5 to 99.3)	28.1 (15.3 to 45.8)	99.1 (98.6 to 99.4)	1.4 (1.0 to 2.0)
-2.76	40.0 (24.3 to 58.1)	97.0 (96.2 to 97.6)	15.2 (8.8 to 24.9)	99.2 (98.7 to 99.5)	3.5 (2.8 to 4.3)
-3.46	40.0 (24.3 to 58.1)	95.1 (94.1 to 95.9)	9.8 (5.6 to 16.4)	99.2 (98.7 to 99.5)	5.4 (4.5 to 6.4)
-3.54	50.0 (32.8 to 67.2)	91.9 (90.7 to 93.0)	7.6 (4.6 to 12.2)	99.3 (98.8 to 99.6)	8.7 (7.6 to 9.9)
-3.64	53.3 (35.8 to 70.1)	90.0 (88.7 to 91.2)	6.6 (4.1 to 10.6)	99.3 (98.8 to 99.6)	10.6 (9.4 to 11.9)
-4.36	60.0 (41.9 to 75.7)	83.7 (82.1 to 85.2)	4.7 (3.0 to 7.3)	99.4 (98.9 to 99.6)	16.9 (15.4 to 18.5)
-4.37	63.3 (45.1 to 78.4)	80.6 (79.0 to 82.2)	4.2 (2.7 to 6.5)	99.4 (98.9 to 99.7)	19.9 (18.3 to 21.6)
-4.4	90.0 (73.2 to 96.7)	57.9 (55.8 to 59.9)	2.8 (1.9 to 4.0)	99.8 (99.3 to 99.9)	42.8 (40.8 to 44.8)
-5.3	96.7 (79.8 to 99.5)	24.1 (22.4 to 25.9)	1.7 (1.2 to 2.4)	99.8 (98.7 to 100.0)	76.2 (74.4 to 77.9)
<b><i>Symptoms, signs and dipstick model</i></b>					
-1	20.0 (9.3 to 37.9)	99.8 (99.5 to 99.9)	54.5 (26.8 to 79.7)	98.9 (98.4 to 99.3)	0.5 (0.3 to 0.9)
-1.6	40.0 (24.3 to 58.1)	99.5 (99.1 to 99.7)	50.0 (31.0 to 69.0)	99.2 (98.7 to 99.5)	1.1 (0.7 to 1.6)
-3.626	63.3 (45.1 to 78.4)	91.0 (89.8 to 92.1)	8.6 (5.6 to 13.1)	99.5 (99.0 to 99.7)	9.7 (8.6 to 11.0)
-4.16	80.0 (62.1 to 90.7)	83.6 (82.0 to 85.1)	6.1 (4.1 to 9.0)	99.7 (99.3 to 99.9)	17.2 (15.7 to 18.8)
-5.48	93.3 (76.9 to 98.3)	50.8 (48.7 to 52.8)	2.5 (1.7 to 3.6)	99.8 (99.3 to 100.0)	49.8 (47.8 to 51.9)
-5.955	96.7 (79.8 to 99.5)	44.5 (42.5 to 46.6)	2.3 (1.6 to 3.3)	99.9 (99.3 to 100.0)	56.0 (53.9 to 58.0)
-6.02	100	22.7 (21.0 to 24.4)	1.7 (1.2 to 2.4)	100	77.6 (75.9 to 79.3)

Results based on models using multiple imputation to deal with missing values.

**TABLE 44** Nappy pad diagnostic test characteristics (95% CI) for a range of cut-points of the points-based model using both the symptoms and signs model (upper part of the table) and the symptoms, signs and dipstick model (lower part of the table)

Cut-point ( $\geq$ )	Sensitivity, %	Specificity, %	PPV, %	NPV, %	Children positive, %
<b><i>Symptoms and signs model</i></b>					
7	13.3 (5.1 to 30.6)	99.4 (99.0 to 99.6)	22.2 (8.6 to 46.5)	98.8 (98.3 to 99.2)	0.8 (0.5 to 1.3)
6	43.3 (27.1 to 61.2)	94.0 (92.9 to 94.9)	8.8 (5.2 to 14.5)	99.2 (98.7 to 99.5)	6.5 (5.6 to 7.6)
5	53.3 (35.8 to 70.1)	87.1 (85.7 to 88.5)	5.2 (3.2 to 8.4)	99.3 (98.8 to 99.6)	13.4 (12.1 to 14.9)
4	90.0 (73.2 to 96.7)	57.8 (55.7 to 59.8)	2.8 (1.9 to 4.0)	99.8 (99.3 to 99.9)	42.9 (40.8 to 44.9)
3	96.7 (79.8 to 99.5)	21.2 (19.6 to 23.0)	1.6 (1.1 to 2.3)	99.8 (98.5 to 100.0)	79.0 (77.3 to 80.6)
2	96.7 (79.8 to 99.5)	18.7 (17.2 to 20.4)	1.6 (1.1 to 2.2)	99.8 (98.3 to 100.0)	81.5 (79.8 to 83.0)
1	96.7 (79.8 to 99.5)	9.3 (8.1 to 10.5)	1.4 (1.0 to 2.0)	99.5 (96.7 to 99.9)	90.8 (89.6 to 91.9)
<b><i>Symptoms, signs and dipstick model</i></b>					
14	6.7 (1.7 to 23.1)	100 (99.7 to 100.0)	66.7 (15.4 to 95.7)	98.8 (98.2 to 99.1)	0.1 (0.0 to 0.4)
13	6.7 (1.7 to 23.1)	100 (99.7 to 100.0)	66.7 (15.4 to 95.7)	98.8 (98.2 to 99.1)	0.1 (0.0 to 0.4)
12	23.3 (11.6 to 41.5)	99.5 (99.1 to 99.7)	38.9 (19.8 to 62.1)	99.0 (98.5 to 99.3)	0.8 (0.5 to 1.3)
11	33.3 (19.0 to 51.6)	98.8 (98.3 to 99.2)	27.0 (15.2 to 43.3)	99.1 (98.6 to 99.4)	1.6 (1.2 to 2.2)
10	40.0 (24.3 to 58.1)	96.0 (95.1 to 96.7)	11.8 (6.8 to 19.6)	99.2 (98.7 to 99.5)	4.5 (3.7 to 5.4)
9	53.3 (35.8 to 70.1)	92.5 (91.3 to 93.5)	8.6 (5.4 to 13.7)	99.3 (98.9 to 99.6)	8.1 (7.1 to 9.3)
8	66.7 (48.4 to 81.0)	88.9 (87.6 to 90.2)	7.4 (4.8 to 11.2)	99.5 (99.1 to 99.7)	11.8 (10.6 to 13.2)
7	80.0 (62.1 to 90.7)	74.9 (73.1 to 76.6)	4.1 (2.8 to 6.0)	99.6 (99.2 to 99.8)	25.8 (24.1 to 27.7)
6	83.3 (65.7 to 92.9)	66.0 (64.1 to 68.0)	3.2 (2.2 to 4.7)	99.7 (99.2 to 99.9)	34.6 (32.7 to 36.6)
5	96.7 (79.8 to 99.5)	47.6 (45.5 to 49.6)	2.4 (1.7 to 3.4)	99.9 (99.3 to 100.0)	53.0 (51.0 to 55.1)
4	100	20.0 (18.4 to 21.7)	1.6 (1.1 to 2.3)	100	80.3 (78.6 to 81.9)
3	100	14.2 (12.8 to 15.7)	1.5 (1.1 to 2.2)	100	86.0 (84.5 to 87.4)
2	100	11.3 (10.1 to 12.7)	1.5 (1.0 to 2.1)	100	88.8 (87.4 to 90.0)
1	100	6.9 (6.0 to 8.1)	1.4 (1.0 to 2.0)	100	93.1 (92.0 to 94.1)

Results based on models using multiple imputation to deal with missing values.

**TABLE 45** Nappy pad diagnostic test characteristics (95% CI) for a range of cut-points of the coefficient-based points-based model, using the symptoms and signs model (upper part of table)

Combination <sup>a</sup>	Sensitivity, %	Specificity, %	PPV, %	NPV, %	Children positive, %
+,+,+,+	13.3 (5.1 to 30.6)	99.4 (99.0 to 99.6)	22.2 (8.6 to 46.5)	98.8 (98.3 to 99.2)	0.8 (0.5 to 1.3)
+,+,+,-	23.3 (11.6 to 41.5)	98.8 (98.2 to 99.1)	20.0 (9.8 to 36.4)	99.0 (98.5 to 99.3)	1.5 (1.1 to 2.1)
+,+,-,+	43.3 (27.1 to 61.2)	94.0 (92.9 to 94.9)	8.8 (5.2 to 14.5)	99.2 (98.7 to 99.5)	6.5 (5.6 to 7.6)
+,-,+,+	46.7 (29.9 to 64.2)	92.5 (91.4 to 93.5)	7.7 (4.6 to 12.6)	99.2 (98.8 to 99.5)	8.0 (6.9 to 9.2)
+,+,-,-	53.3 (35.8 to 70.1)	86.5 (85.0 to 87.8)	5.0 (3.1 to 8.0)	99.3 (98.8 to 99.6)	14.1 (12.7 to 15.5)
+,-,+,-	63.3 (45.1 to 78.4)	82.4 (80.8 to 83.9)	4.6 (2.9 to 7.1)	99.4 (98.9 to 99.7)	18.2 (16.7 to 19.8)
-,+,,+	63.3 (45.1 to 78.4)	81.6 (80.0 to 83.2)	4.4 (2.8 to 6.8)	99.4 (98.9 to 99.7)	19.0 (17.4 to 20.6)
+,-,-,+	90.0 (73.2 to 96.7)	57.9 (55.8 to 59.9)	2.8 (1.9 to 4.0)	99.8 (99.3 to 99.9)	42.8 (40.8 to 44.8)
-,+,-,-	90.0 (73.2 to 96.7)	52.8 (50.7 to 54.8)	2.5 (1.7 to 3.6)	99.7 (99.2 to 99.9)	47.8 (45.7 to 49.8)
+,-,-,-	96.7 (79.8 to 99.5)	24.2 (22.4 to 26.0)	1.7 (1.2 to 2.4)	99.8 (98.7 to 100.0)	76.1 (74.3 to 77.8)
-,+,-,+	96.7 (79.8 to 99.5)	21.2 (19.5 to 22.9)	1.6 (1.1 to 2.3)	99.8 (98.5 to 100.0)	79.1 (77.3 to 80.7)
-,-,+,-	96.7 (79.8 to 99.5)	20.4 (18.8 to 22.1)	1.6 (1.1 to 2.3)	99.8 (98.5 to 100.0)	79.8 (78.1 to 81.4)
-,+,-,-	96.7 (79.8 to 99.5)	18.4 (16.9 to 20.1)	1.6 (1.1 to 2.2)	99.8 (98.3 to 100.0)	81.8 (80.1 to 83.3)
-,-,+,-	96.7 (79.8 to 99.5)	16.7 (15.2 to 18.3)	1.5 (1.1 to 2.2)	99.7 (98.1 to 100.0)	83.4 (81.9 to 84.9)
-,-,-,+	96.7 (79.8 to 99.5)	9.4 (8.3 to 10.7%)	1.4 (1.0 to 2.0)	99.5 (96.7 to 99.9)	90.7 (89.4 to 91.8)
-,-,-,-	100	0	1.3 (0.9 to 1.9)	100	100

a The first plus represents absence of nappy rash, the second plus presence of smelly urine, the third plus presence of darker urine and the fourth plus being female.

## Discussion

### Summary of findings

Based on a large cohort of children in the first 4 years of life presenting with acute illness to primary care, 2.2% and 1.3% of urine samples obtained by clean-catch and nappy pad methods met criteria for a microbiological diagnosis of UTI in the research laboratory. In children providing clean-catch samples (predominantly aged  $\geq 2$  years), four symptoms and three signs were independently associated with a microbiological diagnosis of UTI, with a combined validated AUROC of 0.876. Diagnostic accuracy increased to 0.903 with the addition of three dipstick results and the added value of dipstick testing increased with increasing UTI probability post symptoms and signs. In children providing nappy pad samples (predominantly aged  $< 2$  years), four symptoms (one in common with clean catch, 'smelly urine') and no signs were associated with a microbiological diagnosis of UTI. Validated AUROCs were more modest: 0.778 for symptoms and signs, increasing to 0.821 with the addition of two dipstick test results. Both of the symptom/signs models provide better diagnostic accuracy than 'clinical diagnosis' in identifying the children warranting urine sampling, with the clean-catch model outperforming the nappy pad model [clean-catch AUROC 0.744 (95% CI 0.714 to 0.833), nappy pad AUROC 0.626 (95% CI 0.532 to 0.719)]. Dipstick testing provides additional diagnostic utility for treatment decisions. Simplification of full coefficient-based models into dichotomised points-based models resulted in modest reductions in diagnostic utility (as discussed in *Chapter 6, Comparison of testing and treatment strategies for the clean-catch DUTY5% strategy*), but greater immediate clinical accessibility. Also of note is that the calibration slope and difference between the AUROC and validated AUROC are better in the points-based models. Neither 'fever' nor 'fever of unknown origin' was found to be useful for diagnosing UTI.

### Strengths and limitations

To our knowledge, this is the largest primary care diagnostic accuracy study of clinical symptoms, signs and dipstick tests for diagnosing UTI in young children. Participating children should be representative of those presenting to primary care: key characteristics were similar for participating children and those who were invited but declined participation. We achieved high rates of data completeness across a large number of primary care sites and maintained the blinding of staff to the reference standard. All index tests were measured according to routine clinical practice using standardised reporting forms and equipment, and nearly all were completed within 24 hours of urine sample retrieval, minimising the impact of disease progression bias. Only a small proportion of samples had antimicrobial substances present, and this was even smaller in those with UTI, suggesting a minimal treatment paradox. Our reference standard was specific to the commonest uropathogens and did not include any of the index tests. Two members of staff, blind to all of the index tests except age, performed the microbiological cultures and interpreted results, using a standardised process in a single laboratory.

The main limitations of our study are the relatively small number of UTI diagnoses and the unanticipated differences between clean-catch and nappy pad samples in the reliability and accuracy of the laboratory results. Together, these factors meant that we had insufficient numbers of children with UTI to both derive and externally validate the algorithms, and limited precision of the diagnostic accuracy estimates, especially in the nappy pad models. Although bootstrap validation is an accepted technique to use, leading to reduced AUROC for our models, estimates may be optimistic in relation to values that we would have achieved with external validation. Despite index tests selection being based on extensive discussion and literature review, we omitted to investigate infectious contacts as an index test.<sup>116,135</sup>

### Results in context with other studies

Microbiological definitions of UTI, in terms of both the bacterial species and their concentration, vary between UK,<sup>2</sup> European<sup>136</sup> and other guidelines, and also according to patients' age, sex and urine collection method. Most UTIs are caused by organisms in the Enterobacteriaceae group, with over 90% being caused by *E. coli*, *Proteus mirabilis*, *Klebsiella pneumoniae* and *Enterobacter cloacae*.<sup>82</sup> Our study reference standard, therefore, defined uropathogens as members of the Enterobacteriaceae group at the UK guidelines<sup>2,82</sup> threshold of a pure/predominant growth of  $\geq 10^5$  CFU/ml. We used a rigorous criterion (minimum 3-log difference between the predominant and next most concentrated organism) for defining predominance. This definition could have reduced estimated prevalence if some UTIs were incorrectly classified as contamination: the only other UK primary care-based study of which we are aware estimated prevalence to be 5.9%.<sup>9</sup> Collecting an uncontaminated urine specimen is most difficult in the youngest children, and no study has yet reliably distinguished pathogen from contaminant, especially when they coexist. Our definition of UTI excluded atypical bacteria causing UTIs, which are also thought to be more common in younger children.<sup>137,138</sup>

We are aware of four primary studies<sup>1,65-67</sup> of 18,796 children and one systematic review of eight primary studies<sup>3</sup> in children aged < 5 years (7892 children) that assessed the diagnostic value of clinical symptoms and signs. The only other primary care study found younger age and urinary frequency and dysuria to be independently associated with UTI.<sup>9</sup> Among the remaining studies, largely conducted in emergency departments, abdominal pain, back pain, dysuria, frequency, and new-onset urinary incontinence increased the likelihood of a UTI.<sup>3</sup> Stridor, audible wheeze, circumcision, temperature < 39 °C with a source, abnormal chest sounds, chest crackles, age  $\leq 3$  years, not feeling hot, and breathing difficulty decreased the likelihood of UTI. The largest study, which included almost 16,000 children aged < 5 years presenting to the emergency department,<sup>1</sup> derived a diagnostic model based on a combination of 27 symptoms and signs. This model was found to have an AUROC of 0.80 (95% CI 0.78 to 0.82). Data from six primary studies of dipstick testing in children aged < 5 years<sup>70,138-142</sup> suggest that a dipstick positive for both nitrite and leucocyte may be useful for ruling in a diagnosis of UTI, while dipstick negative for both nitrite and leucocyte may be useful for ruling it out. However, these data were heterogeneous and should be interpreted with caution.

Our clean-catch model includes clinically intuitive items. Previous investigation of malodorous urine has shown conflicting results,<sup>67,139</sup> but our study strongly supports its diagnostic value. We investigated, but did not find evidence for, a number of non-specific symptoms (including fever, vomiting, lethargy, irritability and poor feeding) previously found to be diagnostic of UTI<sup>116</sup> and recommended for clinical use by NICE.<sup>2</sup> It remains possible that such symptoms are of use in the secondary care settings in which studies reporting their utility were conducted, or in children with a different illness spectrum. We found that the presence of severe cough and abnormal ear findings on examination (suggestive of alternative diagnoses to UTI) were associated with a reduced risk of UTI. Such inverse associations are unlikely to reflect biological mechanisms but to arise because both they and UTI are causes of children attending primary health care. Such 'conditioning on a common effect' induces inverse associations between factors that are independent in the source population of well and unwell children.<sup>140</sup> These associations are, nonetheless, of diagnostic utility.

Our nappy pad model, and to some extent the clean-catch model for children aged 2–3 years, is the first primary care study to identify parent-reported symptoms that can be used to select preverbal children warranting urine sampling and presumptive antibiotic treatment. Female sex and parent-reported smelly or darker urine all appear biologically plausible as contributing to the diagnosis of UTI. However, the reasons for the apparently substantial reduction in the risk of UTI associated with presence of a nappy rash are not clear. The inverse association could arise through conditioning on the common effect of primary care attendance,<sup>140</sup> but this is unlikely to produce such a substantial association. Alternative explanations are that rash may be a risk factor for contamination of urine (though we do not find any evidence of this, as reported in *Chapter 7*) and hence mask the presence of a UTI, or that skin products used to treat nappy rash could render the urine sterile. An increased likelihood of contamination of nappy pad samples might also explain the more modest associations of symptoms and dipstick test results with UTI than were found in clean-catch samples. These differences could also arise from differences in illness profiles between older and younger children.

Given the concerns regarding the vulnerability of children's kidneys to damage from upper renal tract infection, it may surprise some clinicians that fever and fever of unknown origin (as a widely perceived marker of pyelonephritis) were not found to be diagnostic in either of the clean-catch nappy pad algorithms. NICE reviewed whether or not symptoms and signs could be used to localise UTI (upper or lower) and concluded that they 'cannot be used to predict acute pyelonephritic changes on DMSA scanning' (p. 64).<sup>2</sup> The NICE guidance also states that children with bacteriuria and a temperature of 38 °C or no fever but loin pain or tenderness should be considered to have pyelonephritis (p. 76). However, studies have shown that both pyelonephritis and renal scarring can occur in the absence of fever<sup>33,141</sup> and NICE's own review found that symptoms and signs were not predictive. The DUTY study was planned to address the HTA commissioning brief as closely as possible, which was to identify the clinical features of paediatric UTI in primary care, without placing any a priori emphasis on a particular symptom (such as fever). We have clearly demonstrated that, among unwell children presenting to primary care (these are the ones in whom we do not wish to miss UTI), fever is *not* diagnostic. This indicates that many UTIs in general practice do not present with fever, which is a new and important finding. It should not be assumed that, because UTI has been traditionally linked with fever, especially in hospital settings and in advanced illness, our finding that fever is not diagnostic of UTI in primary care is wrong.

### **Clinical and research implications**

Parent-reported symptoms and clinical signs can be used to identify preschool children presenting to primary care in whom urine should be collected, and along with urine dipstick results, who should receive immediate antibiotic treatment. Dipsticks have been considered unhelpful in young children<sup>2</sup> until now. Diagnostic utility is better for urine collected using clean catch than for urine collected using nappy pads, but nappy pad samples still provide better diagnostic accuracy than current clinical practice for children < 2 years, in whom the diagnosis of UTI is most challenging. That said, clinicians should be very cautious about using the nappy pad collection method in children with a nappy rash, for whom they should try to collect urine via clean catch. We believe that our results can be applied to other resource-rich nations with similar 'first point of contact' health-service provision, but may not be applicable to the spectra of illness in

preschool children presenting to primary care in resource-poor settings or those referred to secondary care following an initial primary care assessment.

In terms of clinical operationalisation, the dichotomised points-based models provide a simplified tool for clinical use but, as discussed in *Chapter 6*, at reduced cost-effectiveness compared with the preferred, coefficient-based model. That said, in the absence of a computer to aid coefficient-based algorithm use, for older ( $\geq 2$  years) children likely to provide clean-catch samples, we think that it is reasonable to select a cut-point of  $\geq 5$  (sensitivity of 51.7%) which would be clinically operationalised as 'collect urine if any three of the five symptoms and signs (history or UTI, smelly urine or pain/crying passing urine, absence of severe cough and severe illness) are present'. At this threshold, urine dipstick testing would be unlikely to increase the probability of UTI high enough to warrant treatment in children for whom the points-based model did not recommend urine collection. This would result in urine collection in 6.4% of 'DUTY eligible' children aged  $\geq 2$  years. The cost-effectiveness analyses presented in *Chapter 6* suggest that the use of dipstick testing to guide antibiotic treatment has additional diagnostic value, but at a higher cost than awaiting the laboratory result.

For younger ( $< 2$  years) children providing nappy pad samples, while selecting a threshold of  $\geq 5$  points appears attractive to improve sensitivity, the analyses presented in *Chapter 6* suggest that this may not be cost-effective.

Further research is needed to distinguish pathogens from contaminants when bacteria are found in significant concentrations in urine and to establish the cost-effectiveness of different sensitivity/specificity cut-points for the clean-catch and nappy pad models for use in routine clinical practice, using routine health service laboratories. A RCT is needed to assess impact on clinical behaviour and patient outcome, the third of the three steps in the development of a clinical algorithm.<sup>128</sup>

### Conclusions

We have found novel symptoms and signs that are useful for identifying preschool children presenting to primary care who should have urine sampled. Diagnostic utility is better for clean-catch than nappy pad samples, but both perform better than 'clinician diagnosis' alone. Dipstick testing provides additional diagnostic value for antibiotic treatment. The full model may be operationalised using online/desktop technology, but a points-based model is available for immediate use. We present a cost-effectiveness analysis of these strategies in *Chapter 6*.



# Chapter 6 Health economic analysis and modelling of diagnostic strategies

## Introduction

Urinary tract infection is the fourth most common reason for prescribing antibiotics, accounting for approximately 8% of all antibacterial prescriptions.<sup>103</sup> While the unit costs of laboratory testing and antibiotic prescribing are relatively low,<sup>56</sup> the economic implications of new clinical algorithms for urine sampling and testing may be substantial owing to (1) the large numbers of children who present with non-specific symptoms which might be caused by UTI; (2) the cost of subsequent diagnostic tests used to further evaluate children with recurrent/atypical UTI;<sup>2</sup> (3) the substantial costs and impact on quality of life of a missed diagnosis that leads to rare but serious complications of UTI; and (4) the wider, long-term population impact of diagnostic algorithms on antibiotic prescribing and bacterial resistance.<sup>104</sup>

The few economic evaluations of the diagnosis of UTI in young children<sup>56,142,143</sup> have primarily aimed to identify the most cost-effective test or series of tests for diagnosing UTI rather than address the important issue of which children should be selected for testing in the first place. There is limited evidence on which children should have a urine sample taken and by what sampling method, and which urinalysis tests should be used to guide initial treatment. Guidance is especially needed for children < 3 years of age where current NICE clinical guidelines<sup>2</sup> are not based on evidence of cost-effectiveness. In this chapter we develop decision-analytic models to estimate the costs and outcomes of various diagnostic strategies to identify children with elevated risk of UTI in whom a urine sample should be collected. The models also explore the role of dipstick testing in the diagnosis and treatment of UTI in children < 5 years old.

## Methods

### Overview

We developed decision-analytic models using decision trees and Markov models to identify the optimal urine sampling strategy for acutely unwell children < 5 years old presenting to primary care. As *Chapter 5* showed that the diagnostic accuracy depends on the urine collection method and that NHS laboratory results do not have perfect diagnostic accuracy, we developed a 'clean-catch' model and a 'nappy pad' model that account for imperfect NHS laboratory results. The models, which synthesise data from the DUTY study and the wider literature, estimate the lifetime costs and health outcomes for six urine sampling strategies, including three derived from the DUTY risk score.

In secondary analyses, we extend the models to compare three testing and treatment strategies to explore the role of dipstick testing in guiding laboratory testing and antibiotic prescriptions. We also compared the performance of a simpler points-based DUTY algorithm against the full DUTY algorithm. The DUTY points-based algorithm attributes an integer score to each sign and symptom predictive of UTI and sums these to produce a DUTY risk score between 0 and 7.

Costs were estimated from a NHS perspective and included diagnostic costs and short- and long-term treatment costs. Health outcomes were expressed using QALYs and QALDs. The model is made up of three parts: short term (diagnosis and acute illness; up to 21 days), medium term (recurrent UTI; up to 3 years) and long term (long-term sequelae; lifetime).

## Model structure

### Short-term decision tree

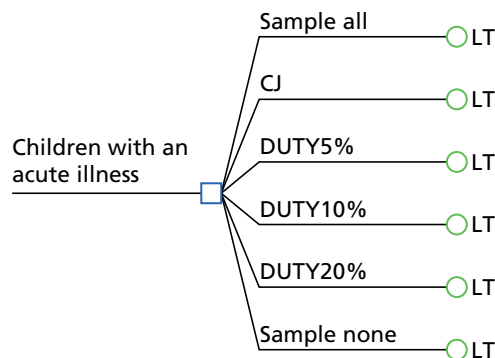
The short-term decision tree deals with the events before and during the index consultation. At presentation, a child may have UTI which could develop into a pyelonephritic attack. Children with pyelonephritis suffer a reduced quality of life during the infection, and the number of pyelonephritic attacks in childhood is predictive of the probability of progressive renal scarring (PRS), which may lead to life-limiting long-term complications. Some children may also present with VUR, which increases their susceptibility to UTI infections, and hence PRS, in future years if it is not diagnosed and treated.

### Comparison of sampling strategies

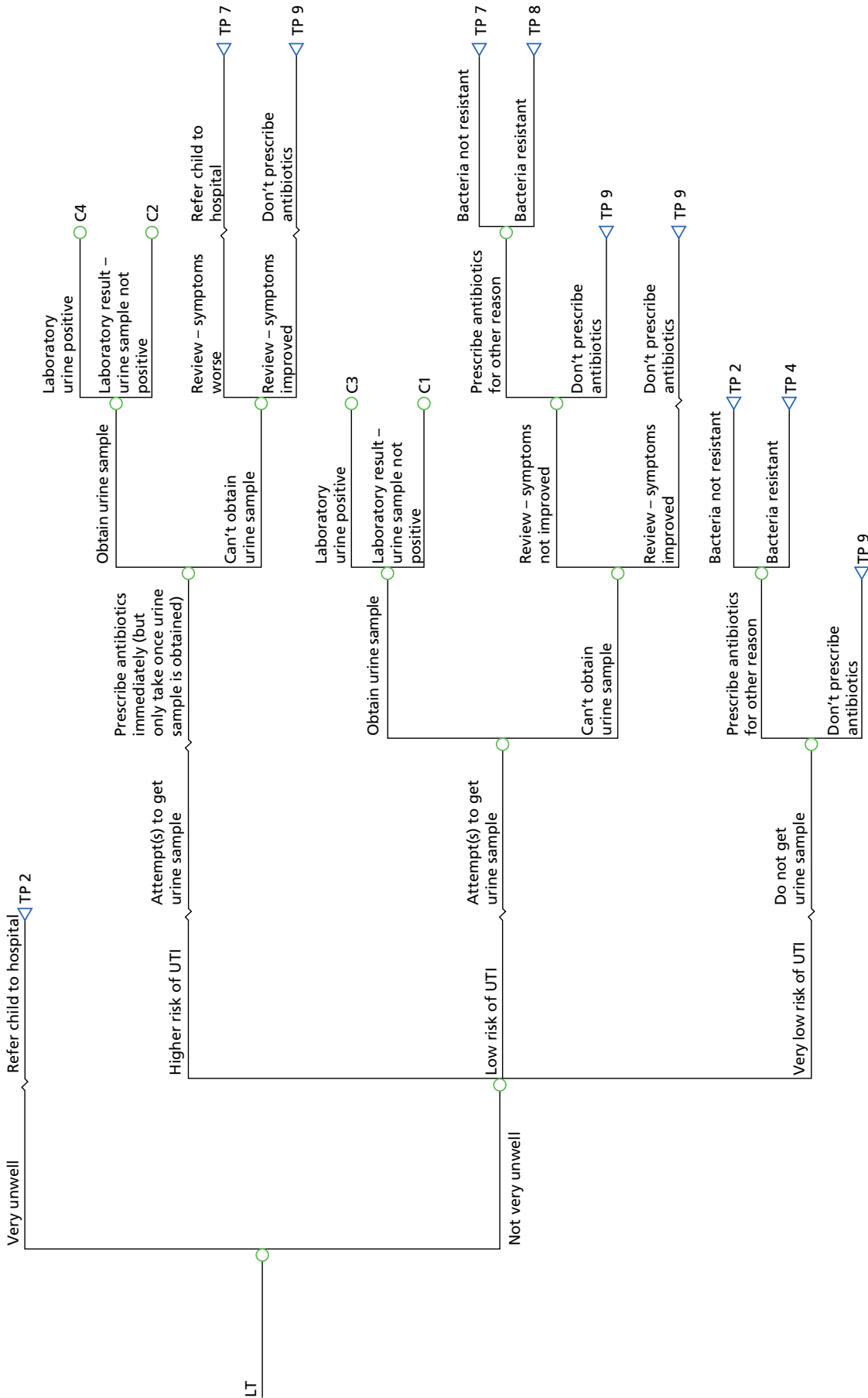
There are six urine sampling strategies, designed to reflect the full variety of clinical behaviour, compared in the primary analyses (*Figure 13*). In all strategies, a proportion of children who are very unwell are referred directly to hospital for testing and treatment (*Figure 14*). For each strategy, the decision to get a urine sample is based on whether the child is deemed at 'very low', 'low' or 'higher' risk of UTI. The model includes two 'boundary strategies' for urine sampling. In the first of these (sample all), all children meeting the DUTY eligibility criteria are judged to be at 'higher' risk of UTI and GPs request a urine sample in all. In the second strategy (sample none), all children meeting the DUTY eligibility criteria are judged to be at 'very low risk' of UTI and GPs do not obtain a urine sample. These boundary strategies are not intended to reflect clinical reality, but provide a reference point against which other intermediate urine sampling strategies can be compared. The risk stratification for the intermediate sampling strategies is based either on GPs' clinical judgement or on the DUTY risk score. We considered three possible DUTY risk score cut-points reflecting a range from high specificity to high sensitivity and hence there are six diagnostic strategies compared in the decision model:

1. sample none ('sample none')
2. sample based on clinical judgement ('CJ')
3. sample based on a high specificity cut-off of the DUTY risk score ('DUTY5%')
4. sample based on an intermediate specificity and sensitivity cut-off of the DUTY risk score ('DUTY10%')
5. sample based on a high sensitivity cut-off of the DUTY risk score ('DUTY20%')
6. sample all ('sample all').

No urine sample is requested in children deemed at 'very low' risk of UTI (see *Figure 14*). Urine sampling is attempted on children deemed at 'low' risk of UTI with antibiotic treatment delayed until a laboratory result is returned (see *Figure 14*). With a positive laboratory result, a suitable antibiotic is prescribed depending on the sensitivities of the bacteria. No antibiotics are prescribed if the laboratory test result is negative. Children at low risk who cannot provide a sample are reviewed in 2 days, and if symptoms have improved no antibiotic is prescribed; if symptoms do not improve, antibiotics may be given if indicated based on the working diagnosis. If the sample is contaminated, a repeat is sought. A urine sample is requested in children considered at 'higher' risk of UTI and antibiotics prescribed immediately with the



**FIGURE 13** Sampling, testing and treatment strategies.



**FIGURE 14** Laboratory-based treatment (LT). The labels at the terminal nodes represent the treatment pathway of children with UTI; all children without UTI enter the treatment pathway (TP) TP1 regardless of their diagnostic pathway.

advice that they should not be taken until the urine sample is obtained (see *Figure 14*). If no sample is obtained, symptoms are reviewed within 2 days, and if symptoms have improved the antibiotic is not started; if symptoms have not improved, the child is referred to hospital. When a sample is provided, the antibiotics are started immediately and the sample is sent to the laboratory. If the bacteria are found to be resistant to the prescribed antibiotic the GP will change the antibiotic prescription, and if no infection is found the GP may ask for the antibiotic to be stopped. If the initial sample is contaminated, a repeat sample is sought. In all urine sampling strategies a proportion of children with a missed diagnosis of UTI are serendipitously treated with antibiotics prescribed for another working diagnosis (see *Figure 14*).

### Comparison of dipstick testing and treatment strategies

In secondary analyses we explored the impact of alternative testing and treatment strategies on our findings. These testing and treatment strategies were developed through discussions with the GPs on the project team (AH, CB, PL, BD and KO'B) and aim to reflect a range of current practice in primary care (*Figures 15–20*). Specifically, we considered two alternatives to the testing and treatment strategy where treatment of 'low' risk children is delayed until the laboratory test result is known (see *Figures 15–19*). In the dipstick test and treatment strategy (see *Figures 16–20*), urine samples obtained in children judged at 'low' risk of UTI are tested with a dipstick at the primary care site. Those with a negative dipstick result are sent to the laboratory with treatment delayed until a result is returned. Those with a positive dipstick result are prescribed immediate antibiotics and the urine sample sent to the laboratory; if the bacteria are found to be resistant, the GP will change antibiotic, and if no infection is found the GP may ask for the antibiotic to be stopped. In the presumptive treatment strategy (see *Figures 18, 19 and 21*) the GP prescribes antibiotics immediately for children judged at 'low' or 'higher' risk of UTI with the advice that they should not be taken until the urine sample is obtained.

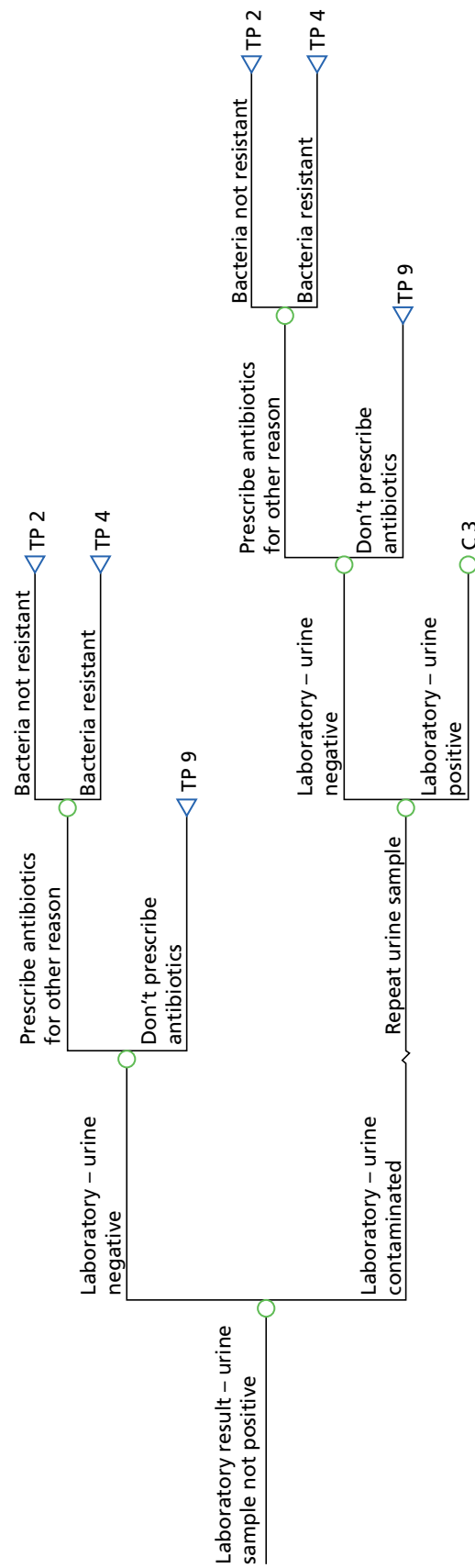
### Short-term Markov model

There are four possible Markov health states for children in the 21 days following the initial consultation (*Figure 21*). A child can be symptomatic but not have UTI ('symptomatic not UTI'); a child can be symptomatic due to UTI and pyelonephritic attack ('symptomatic UTI + PA'); a child can be symptomatic due to UTI alone ('symptomatic UTI'); and a child can become 'asymptomatic'.

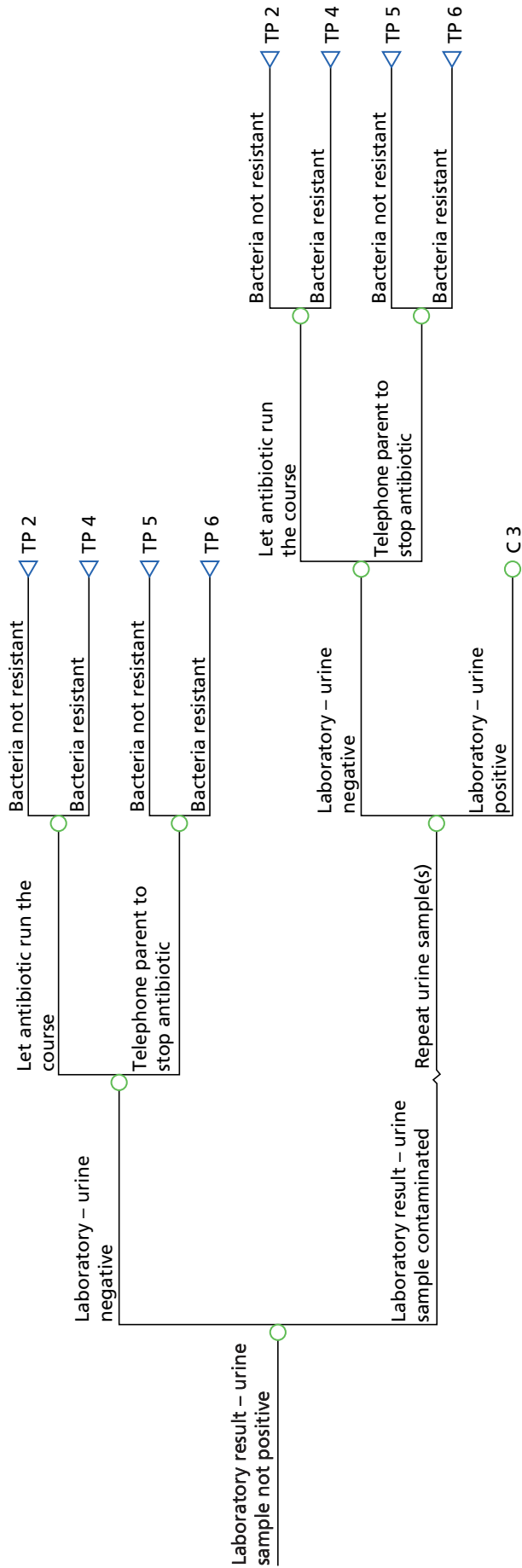
Each of the four Markov health states has a utility or preference score. Preference scores (anchored at 1 for perfect health and 0 for a health state as bad as death) are used to value health outcomes (morbidity and mortality) in one summary measure, the QALY or QALD. The less time a child spends in a symptomatic health state, the higher their QALDs.

The daily transition probabilities from any symptomatic state to the asymptomatic state depend on the natural progression of the disease, the accuracy of diagnosis, the timeliness of treatment and the sensitivity of the bacteria to the antibacterial prescribed. For example, a child correctly diagnosed with UTI and treated immediately with an appropriate antibiotic will, on average, become asymptomatic more rapidly than a child with undiagnosed UTI who is not prescribed antibiotics. In total there are nine treatment pathways:

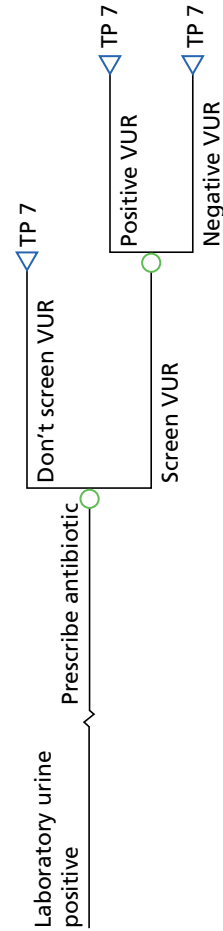
- not UTI
- immediate treatment [not resistant (NR)]
- immediate treatment [resistant (R)], change antibiotic based on laboratory test result (NR)
- immediate treatment (R)
- immediate treatment (NR), stop treatment based on false-negative laboratory test result
- immediate treatment (R), stop treatment based on false-negative laboratory test result
- delayed treatment (NR)
- delayed treatment (R)
- not treated.



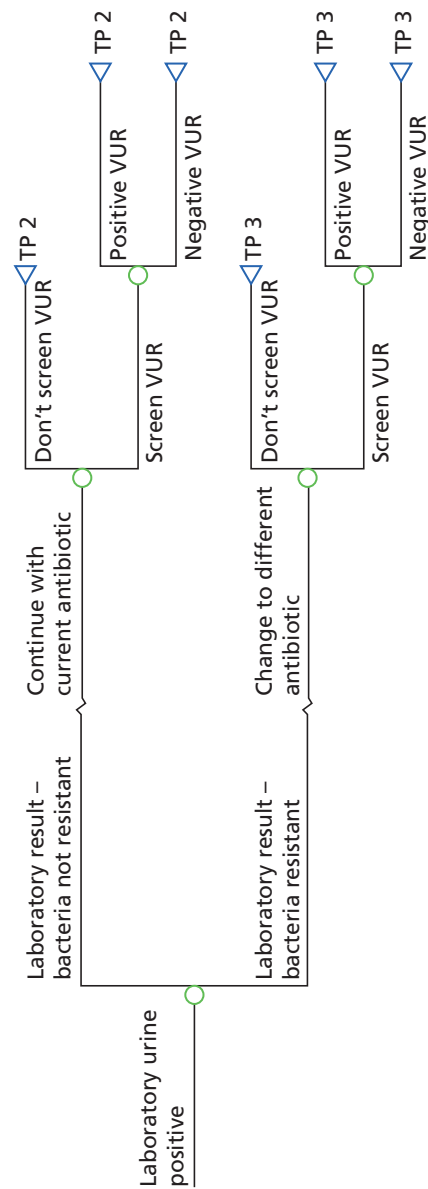
**FIGURE 15** C1, not presumptive not positive. The labels at the terminal nodes represent the treatment pathway of children with UTI; all children without UTI enter the treatment pathway (TP) TP1 regardless of their diagnostic pathway.



**FIGURE 16** C2, presumptive not positive. The labels at the terminal nodes represent the treatment pathway of children with UTI; all children without UTI enter the treatment pathway (TP) TP1 regardless of their diagnostic pathway.



**FIGURE 17** C3, not presumptive positive. The labels at the terminal nodes represent the treatment pathway of children with UTI; all children without UTI enter the treatment pathway (TP) TP1 regardless of their diagnostic pathway.



**FIGURE 18** C4, presumptive positive. The labels at the terminal nodes represent the treatment pathway of children with UTI; all children without UTI enter the treatment pathway (TP) TP1 regardless of their diagnostic pathway.

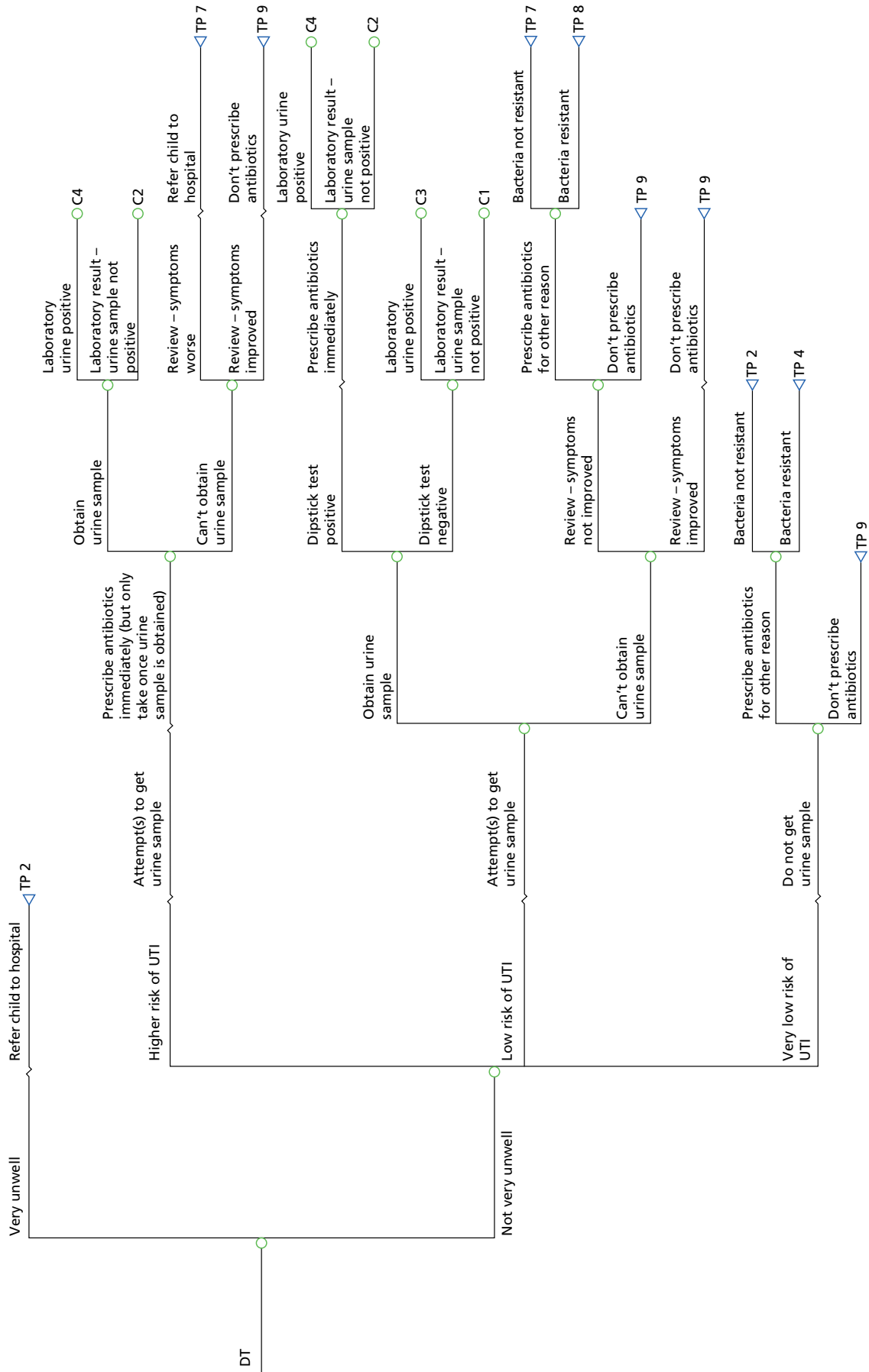
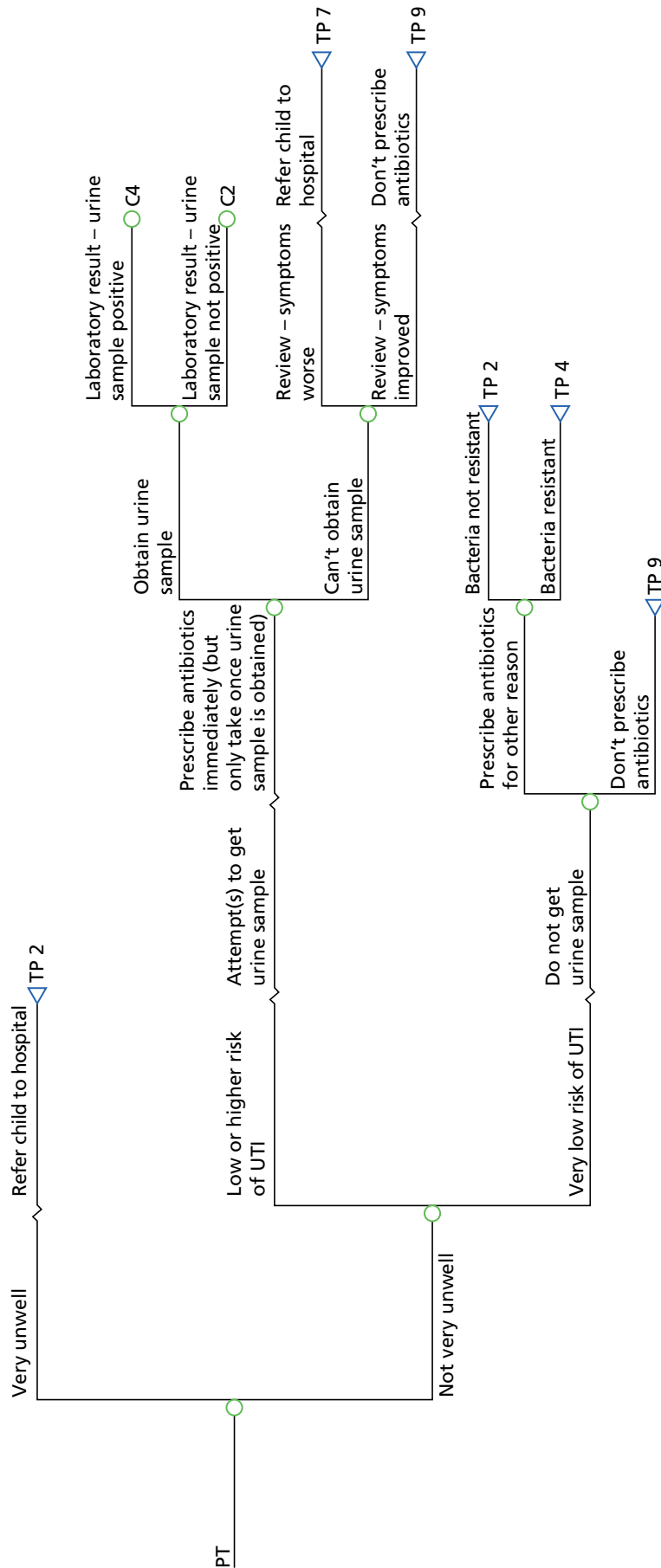
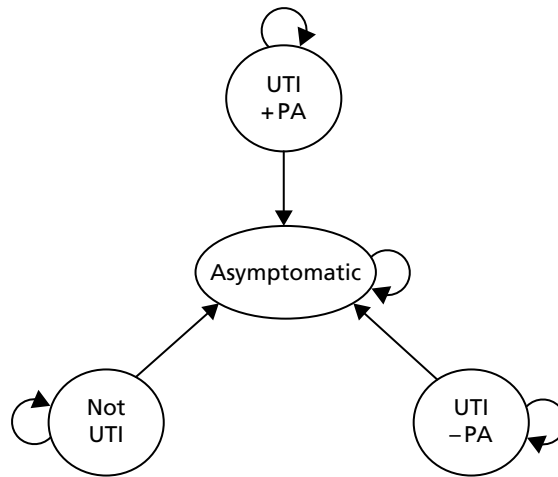


FIGURE 19 Laboratory- and dipstick-based treatment (DT). The labels at the terminal nodes represent the treatment pathway of children with UTI; all children without UTI enter the treatment pathway (TP) TP1 regardless of their diagnostic pathway.





**FIGURE 20** Presumptive treatment (PT). The labels at the terminal nodes represent the treatment pathway of children with UTI; all children without UTI enter the treatment pathway (TP) TP1 regardless of their diagnostic pathway.



**FIGURE 21** Short-term Markov model.

The post-diagnosis daily costs of NHS treatment depend on both the underlying disease (i.e. UTI or not) and the treatment (e.g. no antibiotics, delayed antibiotics or immediate antibiotics). Some costs are fixed and occur at the start of the period, for example the cost of the initial antibiotics prescription. Other costs occur throughout the 21-day period, for example repeat visits to the GP for ongoing symptoms.

### Medium-term model

The medium-term model (see *Appendix 9, Figures 37–40*) simulates the number of recurrent UTIs and pyelonephritic attacks a child experiences in the 3 years following the index consultation. All children enter the medium-term model with no UTIs or one UTI and either zero or one pyelonephritic attacks (see *Appendix 9, Figure 37*). The medium-term model then estimates the number of recurrent UTI infections and GP consultations over a 3-year period following the index consultation. We assume that only one recurrent UTI or GP consultation may take place in each year. Children who had a UTI as the index consultation experience an increased probability of recurrent UTI.

We assume that only one recurrent UTI or GP consultation may take place in each year. Children who had a UTI as the index consultation experience an increased probability of recurrent UTI. Children without UTI at the index consultation experience a lower probability of a first infection until their first UTI. Children who do not have a UTI episode may still consult their GP for another reason and be incorrectly tested or treated for UTI. A higher number of false-positive diagnoses will occur in strategies with lower specificity. In the medium-term model the diagnosis and acute illness costs and utilities of subsequent GP visits are identical to the initial GP visit and, therefore, dependent on the urine sampling strategy adopted.

Children with untreated VUR have a higher probability of recurrent UTI. At each consultation a child with undetected VUR may be referred for imaging, and receive treatment, providing UTI is diagnosed and the VUR is detected. All patients with a positive ultrasound VUR diagnosis are referred onto a confirmatory, more invasive and costly MCUG scan before prophylactic antibiotic treatment is commenced. Children correctly diagnosed with VUR are assumed to be treated with prophylactic antibiotics, according to NICE guidelines.<sup>2</sup>

### Long-term model

The structure of the long-term model is based on previous work.<sup>56</sup> The model provides a link between the number of pyelonephritic attacks by the end of the medium-term model and infection-related renal scarring, which could potentially lead to ESRD requiring treatment by dialysis or renal transplant (see *Appendix 9, Figure 41*). The long-term model calculates the lifetime cost, quality of life and mortality consequences of these severe complications.

## Model inputs

We used a variety of sources to estimate model parameters and assigned probability distributions to represent the uncertainty around these point estimates. This enables probabilistic sensitivity analysis (PSA) to estimate the probability that the DUTY risk scores are more cost-effective than clinical judgement.

## Prevalence

A total of 5107 DUTY study samples were cultured in the research laboratory. Of these, 73 collected with the bag method and 17 collected with an unknown collection method were removed because of the unsuitability of the collection method. Of the remaining urines, 327 were found to be contaminated by our primary definition in the research laboratory, leaving 4690 samples with known UTI status. The prevalence of UTI in the population of children presenting to primary care was based on the research laboratory test results of the DUTY study (Table 46). Ninety (1.9%) of these 4690 children had a research laboratory-confirmed UTI, including 60 (2.2%) of 2690 in children with clean-catch collection and 30 (1.5%) of 2000 in children with nappy pad collection. We estimated the prevalence of pyelonephritis among those with UTI as the proportion of children with laboratory-confirmed UTI and fever (> 38 °C) recorded.<sup>2</sup> Seven children with UTI did not have temperature recorded and so were excluded from this analysis. Thirteen of the remaining 87 children with UTI had pyelonephritis. In the sensitivity analysis we increased the prevalence of pyelonephritis among children with UTI up to 57% as has been reported in a systematic review including secondary care cases.<sup>33</sup> The prevalence of VUR among children with UTI is estimated from a previous meta-analysis of 27 studies.<sup>33</sup> Children who do not have UTI at the index consultation are assumed not to have VUR.

## Risk stratification using clinical judgement and the DUTY risk scores

In order to model sampling and treatment based on clinical judgement, we used GP responses to questions on the DUTY CRF about working diagnosis and planned management. Children are stratified into three risk categories ('very low', 'low' and 'higher') in the model, and, therefore, the definition of 'clinical judgement' is more complex than the definition of 'clinician diagnosis' used in Chapter 5, where diagnostic accuracy is based on a binary categorisation. The proportion of very low-risk children was identified as those for whom the GP answered 'no' to the question 'If this child was NOT in the DUTY study would you have requested a urine sample?' or indicated a working diagnosis of 'not UTI'. Higher-risk children were those for whom the GP had a working diagnosis of UTI and answered 'yes' to the question 'Before seeing the dipstick results, are you planning on treating this child with antibiotics for suspected UTI?'. Finally, low-risk children were those for whom the GP had a working diagnosis of UTI and the GP answered 'yes' to the question 'If this child was NOT in the DUTY study would you have requested a urine sample?'. Categorising patients in this way requires complete data on a number of variables from the CRF; we excluded patients from the analysis if this information was not available or they did not have research laboratory-confirmed UTI status.

Exploratory analysis demonstrated that the prevalence of research laboratory-confirmed UTI increased with increasing 'clinical judgement' risk strata; 1.1% of patients categorised as very low risk had laboratory-confirmed UTI compared with 9.4% and 13.2% of those classified as low and higher risk,

**TABLE 46** Prevalence of UTI in children presenting to primary care

Parameter	Estimate	Distribution	Source
UTI prevalence	0.019	Beta(91,4601)	DUTY
PA (among those with UTI)	0.157	Beta(14,75)	DUTY record of fever
VUR (among those with UTI)	0.240	Odds~LN <sup>a</sup> (-1.153,0.113 <sup>2</sup> )	Shaikh <i>et al.</i> <sup>33</sup>

PA, pyelonephritic attack.  
a Log-normal.

respectively. Using these methods indicated that around 5% of children meeting the DUTY inclusion criteria would have had a urine sample requested based on clinical judgement.

We compared the outcomes of the boundary and clinical judgement strategies with those derived from the DUTY risk score. The diagnostic accuracy of the DUTY risk score depends on the cut-points used to identify the children who should not be sampled (very low risk), the children who should be sampled and treated based on the laboratory test result (low risk) and the children who should be sampled and treated presumptively (higher risk). For each urine collection method we choose three cut-points corresponding to the selection for urine testing of approximately 5% (DUTY5%), 10% (DUTY10%) or 20% (DUTY20%) of children. Hence the DUTY5% strategy represents a relatively high specificity strategy where the proportion of children selected for urine sampling is similar to the clinical judgement sampling strategy. In contrast, the DUTY20% strategy represents a relatively high sensitivity strategy where the proportion of children selected for urine sampling is about four times higher than the clinical judgement strategy. For each DUTY risk score strategy we assumed that the half of those children sampled who had the highest risk scores would be deemed at higher risk of UTI. For example, in the DUTY5% strategy, children with risk scores in the highest 2.5 percentiles were classed at higher risk of UTI and treated presumptively, and children with risk scores between the 25th and 5th percentiles were classed as low risk and treatment was based on the result of the laboratory test.

The accuracy of the clinical judgement and DUTY risk score stratifications used in the clean-catch and nappy pad models are shown in *Table 47*. In clean-catch samples, the DUTY5% urine sampling strategy outperformed clinical judgement. For example, in the clean-catch subgroup, DUTY5% identified 58.2% of children who had UTI as being at low to higher risk of UTI and identified 96.2% of children who did not have UTI as being at very low risk of UTI. In comparison, clinical judgement identified 56.4% of children who had UTI as being at low to higher risk of UTI and identified 91.4% of children who did not have UTI as being at very low risk of UTI. In nappy pad samples, clinical judgement only classified 16.7% of children who had UTI as being at low to higher risk of UTI and identified 97.6% of children who did not have UTI as being at very low risk of UTI. Sensitivity was higher based on the DUTY5% (40.0%), DUTY10% (53.3%) and DUTY20% (63.3%) algorithms, but specificity was lower (95.2%, 90.1% and 80.6%, respectively).

### Risk stratification using clinical judgement and the DUTY points-based risk scores

The accuracy of the simpler DUTY points-based algorithm stratifications used in the clean-catch and nappy pad models is shown in *Table 48*. We chose cut-points at 3, 4, 5 and 6 on the DUTY points score. A secondary cut-point was chosen so that approximately half would be deemed at 'higher' risk of UTI to allow direct comparison with the full DUTY risk scores. However, as the points-based risk score had only eight possible values, a close match was not always possible. The diagnostic accuracy of the points-based algorithm was slightly lower than the full DUTY risk scores. For example, in clean-catch samples the full DUTY20% risk scores identified a similar proportion of patients with UTI as at low or higher risk of UTI (85%) and more people without UTI as at very low risk (80.6% vs. 73.6%) than the points-based DUTY  $\geq 3$  algorithm.

### Short-term model probabilities

The short-term model probabilities are presented in *Table 49*. A minority of children presenting to primary care will be referred immediately to hospital before the decision of whether or not to attempt a urine sample is considered. We identified these 'very unwell' children as those for whom the GP answered 'yes' to the question 'Before seeing the dipstick results, would you have referred this child to a paediatrician or admitted this child to hospital?'. The probabilities of obtaining a sample, having a contaminated sample and having a uropathogen resistant to antibiotics were based on observations from the DUTY study. Just over 91% of children in the DUTY study provided a urine sample. The probability of contamination, stratified by collection method, was based on a NHS laboratory report of heavy mixed growth with a count  $> 10^5$  CFU/ml of more than two organisms. In DUTY children with UTI, resistance to both non-UTI antibiotics (assumed to be amoxicillin) and UTI antibiotics (assumed to be trimethoprim) was based on the

TABLE 47 Risk stratification for CJ and full DUTY risk scores<sup>a</sup>

Risk stratification	Not UTI				UTI				
	% sampled	Very low risk	Low risk	Higher risk	Distribution	Very low risk	Low risk	Higher risk	Distribution
Clean catch									
CJ	9.60	2282 (91.43)	84 (3.37)	130 (5.21)	Dir(2283,85,131)	24 (43.64)	9 (16.36)	22 (40.00)	Dir(25,10,23)
DUTY5%	4.97	2407 (96.20)	52 (2.08)	43 (1.72)	Dir(2408,53,44)	23 (41.82)	12 (21.82)	20 (36.36)	Dir(24,13,21)
DUTY10%	10.05	2284 (91.29)	123 (4.92)	95 (3.80)	Dir(2285,124,96)	16 (29.09)	7 (12.73)	32 (58.18)	Dir(17,8,33)
DUTY20%	20.81	2017 (80.62)	267 (10.67)	218 (8.71)	Dir(2018,268,219)	8 (14.55)	8 (14.55)	39 (70.91)	Dir(9,9,40)
Nappy pad									
CJ	2.64	1891 (97.57)	24 (1.24)	23 (1.19)	Dir(1892,25,24)	25 (83.33)	3 (10.00)	2 (6.67)	Dir(26,4,3)
DUTY5%	5.38	1848 (95.16)	39 (2.01)	55 (2.83)	Dir(1849,40,56)	18 (60.00)	1 (3.33)	11 (36.67)	Dir(19,2,12)
DUTY10%	10.55	1750 (90.11)	98 (5.05)	94 (4.84)	Dir(1751,99,95)	14 (46.67)	4 (13.33)	12 (40.00)	Dir(15,5,13)
DUTY20%	20.08	1565 (80.59)	185 (9.53)	192 (9.89)	Dir(1566,186,193)	11 (36.67)	3 (10.00)	16 (53.33)	Dir(12,4,17)

CJ, clinical judgement; Dir, Dirichlet distribution.

<sup>a</sup> Excludes patients who did not have UTI results, those referred immediately to secondary care or where missing data did not allow calculation of clinical judgement decision.

TABLE 48 Risk stratification for DUTY points-based risk scores<sup>a</sup>

Risk stratification	% sampled	Not UTI				UTI			
		Very low risk	Low risk	Higher risk	Distribution	Very low risk	Low risk	Higher risk	Distribution
Clean catch									
DUTY $\geq 6$ (6,7) <sup>b</sup>	4.58	2409 (96.28)	82 (3.28)	11 (0.44)	Dir(2410,83,12)	31 (56.36)	17 (30.91)	7 (12.73)	Dir(32,18,8)
DUTY $\geq 5$ (5,6)	6.92	2354 (94.08)	55 (2.20)	93 (3.72)	Dir(2355,56,94)	26 (47.27)	5 (9.09)	24 (43.64)	Dir(27,6,25)
DUTY $\geq 4$ (4,5)	22.96	1959 (78.30)	395 (15.79)	148 (5.92)	Dir(1960,396,149)	11 (20.00)	15 (27.27)	29 (52.73)	Dir(12,16,30)
DUTY $\geq 3$ (3,5)	27.65	1842 (73.62)	512 (20.46)	148 (5.92)	Dir(1843,513,149)	8 (14.55)	18 (32.73)	29 (52.73)	Dir(9,19,30)
Nappy pad									
DUTY $\geq 6$ (6,7)	6.95	1819 (93.67)	110 (5.66)	13 (0.67)	Dir(1820,111,14)	16 (53.33)	10 (33.33)	4 (13.33)	Dir(17,11,5)
DUTY $\geq 5$ (5,6)	15.42	1655 (85.22)	164 (8.44)	123 (6.33)	Dir(1656,165,124)	13 (43.33)	3 (10.00)	14 (46.67)	Dir(14,4,15)
DUTY $\geq 4$ (4,5)	45.13	1079 (55.56)	576 (29.66)	287 (14.78)	Dir(1080,577,288)	3 (10.00)	10 (33.33)	17 (56.67)	Dir(4,11,18)
DUTY $\geq 3$ (3,4)	79.77	398 (20.49)	681 (35.07)	863 (44.44)	Dir(399,682,864)	1 (3.33)	2 (6.67)	27 (90.00)	Dir(2,3,28)

Dir, Dirichlet distribution.

a Excludes patients who did not have UTI results, those referred immediately to secondary care or where missing data did not allow calculation of clinical judgement decision.

b Numbers in brackets represent the cut-points used in the points-based model. For example, in the clean-catch DUTY  $\geq 6$  algorithm all children with a score of  $< 6$  were considered at very low risk, children with a score of 6 were considered at low risk and children with a score of  $\geq 7$  were considered at higher risk.

**TABLE 49** Short-term model probabilities

Parameter	Estimate	Distribution	Source
Very unwell – referred directly to hospital	0.037	Beta(263,6840)	DUTY
Sample obtained	0.915	Beta(6426,596)	DUTY
Contamination			DUTY
Clean catch	0.056	Beta(163,2740)	
Nappy pad	0.258	Beta(785,2261)	
Antibiotic resistance (not UTI Abx)	0.531	Beta(51,45)	DUTY
Antibiotic resistance (UTI Abx)	0.281	Beta(27,69)	DUTY
Reconsultation before symptom resolution	0.192	Beta(43,181)	Hay <i>et al.</i> <sup>144</sup>
Antibiotics for non-UTI reason	0.267	Beta(209,573)	DUTY 14-day
Stop antibiotic given no UTI	0.075	Uniform(0.05,0.10)	Expert opinion
Referred for VUR scan	0.077	Beta(26,312)	DUTY
Abx, antibiotics.			

research laboratory reports of resistance to these antibiotics. We estimated the proportion of children who would return to primary care before symptom resolution from a cohort study,<sup>144</sup> based in UK primary care, of 222 children presenting with acute respiratory tract infection and cough. We estimated the proportion of children where an antibiotic was prescribed for another disease (not UTI) by calculating the proportion of children without UTI in DUTY whose parents reported antibiotic use within 2 days of the initial consultation in the 14-day follow-up questionnaire.

The probability of a GP stopping antibiotic treatment midway through a prescription following a negative laboratory test in a presumptively treated child was based on expert opinion from three GPs (AH, CB and BD). We estimated the proportion referred for further investigation by calculating the proportion of children in the DUTY cohort with a positive NHS laboratory UTI diagnosis who had a record of an ultrasound scan within 3 months of the index consultation.

### Diagnostic accuracy of dipstick and NHS laboratory test results

The diagnostic accuracy of dipstick tests and NHS laboratory results were both defined against the reference standard of the research laboratory result in DUTY (*Table 50*). A positive dipstick test was defined as a positive result (+, ++ or +++) for either leucocytes or nitrates. Estimates of the diagnostic accuracy of ultrasound scans are taken from a previous meta-analysis.<sup>56</sup> MCUG is the gold-standard test for VUR diagnosis and hence it is assumed to be 100% sensitive and specific.

### Symptom resolution

The daily symptom resolution transition probabilities are based on data from the DUTY 14-day interviews. We used responses to the question ‘How many days since [name of child] joined the DUTY study (day 0) was it until [his/her] symptoms improved?’ in the 14-day questionnaire data to calculate symptom duration. We used the responses from children treated immediately, defined as receiving antibiotics within 2 days of the index consultation, and children without UTI to parameterise two Weibull survival models. The Weibull model is a parametric survival model which is used to model time to event data; in this case the event is symptom resolution. A small proportion of parents of children both with and without UTI reported symptom recovery times > 14 days and hence we extrapolated our estimates to 21 days, by which time the vast majority of children were predicted to have become asymptomatic. The estimated symptom resolution rates from our model were a good fit for the observed rates for both children with treated UTI and those without UTI (see *Appendix 9, Figures 42 and 43*).

**TABLE 50** Diagnostic test performance

Parameter	Sensitivity		Specificity		Source
	Estimate	Distribution	Estimate	Distribution	
Dipstick					DUTY
Clean catch	0.758	Beta(47,15)	0.840	Beta(2215,423)	
Nappy pad	0.781	Beta(25,7)	0.688	Beta(1438,652)	
NHS laboratory test <sup>a</sup>					DUTY
Clean catch	0.780	Beta(46,13)	0.976	Beta(2352,58)	
Nappy pad	0.769	Beta(20,6)	0.922	Beta(1363,116)	
Ultrasound for VUR	0.440	Odds~LN <sup>b</sup> (-0.243,0.2352)	0.775	Odds~LN (1.238,0.2862)	Whiting <i>et al.</i> <sup>56</sup>
MCUG for VUR	1.000	Fixed	1.000	Fixed	Assumption

a After removing samples which were found to be contaminated in the NHS laboratory.

b Log-normal.

As a urine sample was requested in all children participating in DUTY and the vast majority of children with laboratory-confirmed UTI were treated with antibiotics, we could not directly observe the symptom resolution rate where antibiotics had not been prescribed. Therefore, we estimated the untreated UTI symptom resolution rates based on treatment effect sizes in the literature. To our knowledge, no RCTs of antibiotics for UTI in children exist. We identified a systematic review,<sup>145</sup> which included five RCTs estimating the treatment effect of antibiotics for women with cystitis. In four of these studies the treatment regimens were thought not to be representative of current UK general practice for treating UTI (e.g. single-dose amoxicillin 750 mg) and hence we used data from a RCT which evaluated the effectiveness of nitrofurantoin in 78 women aged between 15 and 54 years.<sup>146</sup> The study found that in women with bacteriologically proven UTI, nitrofurantoin was significantly more effective than placebo in achieving symptomatic relief at 3 days [risk ratio 0.55 (95% CI 0.34 to 0.89)] (*Table 51*).

We assumed that for children in whom the uropathogens were resistant to the prescribed antibiotic, the rate of symptom resolution was reduced by 30% based on expert opinion from a panel of GPs and microbiologists (see *Table 51*). We assumed that children with UTI who receive antibiotics only once the NHS laboratory test result is known or who remain untreated because a urine sample cannot be obtained or is contaminated have treatment delayed for a period of 2 days before antibiotics are taken, at which point their transition probabilities to the asymptomatic state become identical to those receiving treatment. We assumed that the treatment benefit from antibiotics persisted for only 7 days and hence all children

**TABLE 51** Short-term Markov model probabilities

Parameter	Estimate	Distribution	Source
Antibiotic treatment effect <sup>a</sup>	0.550	RR <sup>b</sup> ~LN <sup>c</sup> (-0.599, 0.247)	Christiaens <i>et al.</i> <sup>146</sup>
Effect of resistance on symptom resolution rates <sup>d</sup>	0.700	Uniform (0.5, 0.9)	Expert opinion

a Risk ratio comparing symptom resolution rates for children not treated with antibiotics with those in children treated with antibiotics.

b Risk ratio.

c Log-normal.

d Risk ratio comparing symptom resolution rates for uropathogens resistant to antibiotics vs. uropathogens sensitive to antibiotics.



with UTI had identical symptom resolution probabilities between days 8 and 21. The average symptom duration is displayed in *Table 52*. Symptom resolution rates for a selection of treatment pathways are displayed in *Appendix 9* (see *Figure 44*).

### Medium-term Markov model probabilities

After the initial episode, the medium-term model tracks patients as they consult their GP for new symptoms, potentially symptoms of UTI, during the following 3 years. We estimated the proportion of patients who would consult their GP during a 12-month period for symptoms not caused by UTI from a cohort study of 7727 children (aged  $\leq 56$  months) attending primary care for acute symptoms.<sup>147</sup>

The probability of recurrent UTI in patients with a previous history of UTI was estimated from a previous meta-analysis based on six primary studies.<sup>33</sup> Wide variations in incidence rates for UTI in children without a previous UTI have been reported due to differences in study setting and diagnostic criteria for UTI.<sup>148–150</sup> We elected to use data from a study in English general practice,<sup>149</sup> which calculated the incidence of UTI across 18 months in 2789 children up to 14 years old, owing to its use of strict diagnostic criteria and its UK primary care setting (*Table 53*).

We estimated the effect of treatment with prophylactic antibiotics on UTI recurrence rates for patients with VUR using results from a previous meta-analysis.<sup>15</sup> This estimate is moderate and open to dispute as it is largely based on studies with inadequate blinding.<sup>15</sup> We assumed that the prevalence of pyelonephritis in children with recurrent UTI and the probability of being referred for VUR were the same as the index infection and that the sensitivity of GP diagnosis during recurrent UTI infections was identical to index infection and, therefore, varied by urine sampling strategy.

### Long-term probabilities

The long-term health consequences of UTIs, and specifically the number of pyelonephritic attacks, are based on natural history evidence.<sup>16–19</sup> We used data from a prospective study including 1777 children to estimate the strength of association between the number of pyelonephritic attacks and the probability of

**TABLE 52** Symptom duration, costs and utilities in each treatment pathway during acute illness

TP	Average symptom duration, days (95% CI)	Average acute illness cost, £ (95% CI)	Average acute illness QALDs (95% CI)
TP1: not UTI	6.03 (5.69 to 6.40)	42.61 (32.03 to 53.99)	20.38 (19.79 to 20.97)
TP2: immediate treatment (NR)	4.72 (3.80 to 5.90)	50.24 (32.58 to 69.74)	20.54 (20.02 to 20.98)
TP3: immediate treatment (R), change antibiotic based on laboratory test result (NR)	5.06 (4.10 to 6.34)	53.39 (35.31 to 73.40)	20.50 (19.93 to 20.98)
TP4: immediate treatment (R)	5.72 (4.40 to 7.36)	62.51 (40.67 to 88.67)	20.42 (19.73 to 20.97)
TP5: immediate treatment (NR), stop treatment based on false-negative laboratory test result	5.63 (4.40 to 7.11)	62.83 (41.03 to 86.76)	20.43 (19.78 to 20.97)
TP6: immediate treatment (R), stop treatment based on false-negative laboratory test result	6.07 (4.72 to 7.67)	67.34 (44.26 to 93.17)	20.38 (19.66 to 20.97)
TP7: delayed treatment (NR)	5.22 (4.26 to 6.47)	54.87 (36.24 to 75.44)	20.48 (19.89 to 20.98)
TP8: delayed treatment (R)	5.90 (4.68 to 7.43)	64.24 (42.44 to 89.26)	20.40 (19.71 to 20.97)
TP9: not treated	6.29 (4.66 to 8.08)	69.72 (43.79 to 97.92)	20.35 (19.59 to 20.97)

TP, treatment pathway.

**TABLE 53** Medium-term Markov model probabilities

Parameter	Estimate	Distribution	Source
Consult with no UTI	0.69	Beta(21194,9396)	Hay <i>et al.</i> <sup>147</sup>
Consult with UTI given			
No UTI history	0.0034	<sup>a</sup> Beta(10.33,2780.67)	Dickinson <i>et al.</i> <sup>149</sup>
Previous UTI			
No VUR/treated VUR	0.08	Odds~LN <sup>b</sup> (-2.442,0.2182)	Shaikh <i>et al.</i> <sup>33</sup>
Treatment effect for treated VUR	RR <sup>c</sup> = 0.68	RR~LN (-0.385,0.2802)	Nagler <i>et al.</i> <sup>151</sup>

a Numerator adjusted to account for 18-month follow-up period.

b Log-normal.

c Relative risk comparing UTI recurrence rates in children with VUR with children without VUR.

renal scarring.<sup>20</sup> The link between renal scarring and ESRD is based on registry data (*Table 54*).<sup>21</sup> We assumed that the mean age of ESRD onset following infection-related scarring was 13.67 years,<sup>152</sup> with a range of 7 to 24 years,<sup>153</sup> based on the results of two observational studies. We also assumed that half of children with ESRD would be treated with dialysis, with the remainder treated with renal transplant. Mean survival times for patients without ESRD, treated with dialysis and treated with transplant are taken from national statistics, a systematic review of home versus hospital haemodialysis, and renal registry data, respectively.<sup>154–156</sup>

### Costs

Where appropriate, costs were inflated to 2011 prices using the *Unit Costs of Health and Social Care*.<sup>158</sup> We used a uniform distribution, with the lower bound 50% lower and an upper bound 50% greater than the estimated mean, to describe parameter uncertainty.

**TABLE 54** Long-term probabilities

Parameter	Estimate	Distribution	Source
PRS			Jodal <sup>40</sup>
0 PA	0.05	Beta(8,135)	
1 PA	0.09	Beta(33,335)	
2 PA	0.16	Beta(16,79)	
3 PA	0.34	Beta(13,24)	
4 PA	0.58	Beta(15,11)	
ESRD given PRS	0.05		Alexander <i>et al.</i> <sup>157</sup>
Mean age of ESRD onset	13.67	Triangle(7,24)	Arant, <sup>152</sup> Jacobson <i>et al.</i> <sup>153</sup>
Future treatment			
Transplant	0.500		
Dialysis	0.500		
Mean survival			
No ESRD	73.00	Uniform(69.4, 76.7)	ONS <sup>154</sup>
Dialysis	12.25	Uniform(11.6, 12.9)	Mowatt <i>et al.</i> <sup>155</sup>
Transplant	21.60	Uniform(20.5, 22.7)	Kaufman <sup>156</sup>

PA, pyelonephritic attack.

### Cost of the initial urine sample and test

We used time and motion methods, as described in *Appendix 9*, to evaluate the time taken and resources used for urine collection. We used results from a report in pathology testing in England<sup>159</sup> to estimate the cost of a laboratory test for UTI (*Table 55*). We assumed that laboratory urine tests were included in the least expensive and highest volume basket of pathology tests. Within this basket, prices ranged from £4.04 to £9.43 between the pilot sites in the report; we used the median cost reported across all sites of £6.13. An additional 45 seconds (£2.33) of GP time is assumed for interpretation of each urine result and 5 minutes (£15.50) for contacting patients in the event of a positive laboratory result after a delayed prescription, or when a negative laboratory result was returned after a presumptive prescription and the GP decided to stop the antibiotic treatment.

**TABLE 55** Costs (£)

Cost	Estimate <sup>a</sup>	Distribution	Source
Short-term decision tree			
Sample			DUTY
No dipstick	6.78	Uniform(3.39,10.17)	
Dipstick	7.81	Uniform(3.91,11.72)	
Sample attempted but not obtained	1.41	Uniform(0.71,2.12)	
Nappy pad collection kit	1.83	Uniform(0.92,2.75)	DUTY
Laboratory test	6.13	Uniform(3.07,9.2)	Carter Report <sup>159</sup>
Dipstick	0.38	Uniform(0.19,0.57)	DUTY
GP result interpretation cost	2.33	Uniform(1.17,3.5)	Expert opinion
GP stop antibiotic	15.50	Uniform(7.75,23.25)	Expert opinion
GP start/change antibiotic	15.50	Uniform(7.75,23.25)	Expert opinion
UTI antibiotics	2.80	Uniform(1.4,4.2)	Prescription cost analysis <sup>160</sup>
Non-UTI antibiotics	1.24	Uniform(0.62,1.86)	Prescription cost analysis <sup>160</sup>
VUR scan (ultrasound)	50.00	Uniform(25,75)	Reference costs <sup>161</sup>
MCUG	137.06	Uniform(68.53,205.59)	Whiting <i>et al.</i> <sup>56</sup>
Short-term Markov model			
UTI			DUTY
Fixed cost (2 days)	20.46	Uniform(10.30,69)	
Daily cost	13.91	Uniform(6.96,20.87)	
Non-UTI			DUTY
Fixed cost (2 days)	19.58	Uniform(9.79,29.37)	
Daily cost	6.82	Uniform(3.41,10.23)	
Medium-term model			
Diagnosis			DUTY
Acute illness	Table		DUTY
Antibiotic prophylaxis	23.80	Uniform(11.9,35.7)	Nagler, <sup>151</sup> prescription cost analysis <sup>160</sup>
Long-term model			
Dialysis per year	21,655	Uniform(10827.5,32482.5)	Baboolal <i>et al.</i> <sup>162</sup>
Transplant <sup>b</sup>	19,456	Uniform(9728,29184)	Reference costs <sup>161</sup>

a All costs were included in the model as uniform distributions, with a lower bound 50% lower than the estimated mean and an upper bound 50% greater than the estimated mean.

b Assuming HRG LA02A: kidney transplant 19 years and over from cadaver heart-beating donor.

We assumed that children treated with antibiotics for a non-UTI diagnosis received an oral solution of amoxicillin with dose 125 mg/5 ml and that children with a UTI diagnosis were treated with an oral solution of trimethoprim with dose 50 mg/5 ml. These were, respectively, the most commonly prescribed antibiotics for non-UTI and UTI diagnoses in the DUTY study. Costs for these antibiotics were estimated using prescription cost-analysis data.<sup>160</sup> Consumable costs (e.g. nappy pads) were estimated using the price paid in the DUTY study. Imaging costs for ultrasound and MCUG scans were estimated from national reference costs<sup>161</sup> and a previous economic model,<sup>56</sup> respectively.

### **Acute illness and medium-term cost**

We used data from a carer-reported resource-use questionnaire, completed by a subsample of 918 parents at 14 days, to estimate the acute illness cost for children with UTI and other diagnoses. Details of this analysis are provided in *Appendix 9*. We estimated a fixed cost due to the initial GP visit and prescription, if provided, and a variable daily cost reflecting the ongoing costs of further care for children whose symptoms do not resolve quickly.

The medium-term diagnosis and acute illness costs for repeat GP visits by children with or without UTI were the same as the index consultation (see *Table 52*). We used the most common regimen reported in a systematic review of antibiotic prophylaxis,<sup>151</sup> 1–2 mg per kilo daily, and assumed a child weight of 15 kg to calculate the annual amount of antibiotics needed during treatment. We assumed that prophylactic treatment was with an oral suspension of trimethoprim with dose 50 mg/5 ml (see *Table 55*).<sup>160</sup>

### **Long-term costs**

The long-term model requires data on the average number of years spent in a given state (e.g. number of years spent on dialysis before death) and the costs in each health state. We assumed that patients experience no increased cost until the onset of ESRD. Some patients will be treated by dialysis with an ongoing annual cost until death estimated from a UK study based on semistructured interviews with experts in nephrology management (see *Table 55*).<sup>162</sup> Other patients are treated with a renal transplant and experience a one-off treatment cost, estimated using the national reference costs, at the time of procedure and no further costs until death.

### **Utilities**

All utilities used in the model are reported in *Table 56*. We used a uniform distribution, with a lower bound 20% lower and an upper bound 20% greater than the estimated mean, to describe parameter uncertainty.

**TABLE 56** Utilities

Utility	Estimate	Distribution	Source
Short- and medium-term models			
UTI/no PA	0.943	Uniform(0.75,1.00)	Brisson <i>et al.</i> <sup>164</sup>
UTI/PA	0.711	Uniform(0.57,0.85)	Whiting <i>et al.</i> <sup>56</sup>
Non-UTI	0.943	Uniform(0.75,1.00)	Brisson <i>et al.</i> <sup>164</sup>
Long-term model			
No ESRD	1.000	Fixed	N/A
Dialysis	0.430	Uniform(0.34,0.52)	Churchill <i>et al.</i> <sup>163</sup>
Transplant	0.840	Uniform(0.67,1.00)	Churchill <i>et al.</i> <sup>163</sup>

N/A, not applicable; PA, pyelonephritic attack.

a All utilities were included in the model as uniform distributions, with a lower bound 20% lower than the estimated mean and an upper bound 20% greater than the estimated mean.

### **Short- and medium-term utilities**

We measured health-related quality of life in DUTY using the TAPQOL questionnaire.<sup>110</sup> This showed broad similarity in the quality of life between children in the DUTY study who had working diagnoses of UTI and other diseases; full details are provided in *Appendix 9*. We therefore conducted a search of the Cost-Effectiveness Analysis Registry<sup>165</sup> for all studies reporting utility scores for acutely unwell infants. This search identified studies in a number of diseases, but no estimates for infants with UTI. We elected to use utility estimates from two studies on rotavirus as it was thought that the symptoms most closely matched those of UTI. One study, which used the Health Utilities Index (HUI2) questionnaire to elicit utility values from caregivers of children under the age of 3 presenting with gastroenteritis in Canada with a confirmed rotavirus infection, reported a utility value of 0.943. Another used response from GP-completed EQ-5D questionnaires and reported utility values of 0.781 and 0.688 in babies and children, respectively. We used the caregiver-reported HUI2-derived utility scores in our baseline analysis and, additionally, used the GP EQ-5D-derived utility scores as a sensitivity analysis to explore the impact of more severe quality of life decrements on our conclusions. We used utility values for pyelonephritis in adults reported in the literature.<sup>166</sup> The average medium-term acute illness utilities for children with UTI were the same as the index infection (see *Table 52*).

### **Long-term utilities**

We assumed that patients who do not develop ESRD have a perfect quality of life until death. Patients with PRS experience no quality of life decrement until the onset of ESRD. Half of patients are treated by dialysis until death, while the remainder are treated with a renal transplant and experience an increased quality of life, when compared with those on dialysis, until death. Utility estimates for patients on dialysis and following renal transplant were estimated from a published time-trade-off exercise of 103 transplant and 60 hospital haemodialysis patients (see *Table 56*).<sup>163</sup>

### **Assumptions**

The structure of the short-, medium- and long-term models and the links between them require a number of structural assumptions to be made.

#### **Short-term decision tree**

- The proportion of children in whom it is not possible to obtain a urine sample and the proportion where the sample is contaminated are equal across those with and without UTI.
- Following a contaminated sample, a repeat uncontaminated sample will always be obtained on the second attempt.
- NHS laboratory tests are perfectly accurate at detecting the sensitivity of uropathogens to antibiotics.

#### **Short-term Markov model**

- Children for whom antibiotics are administered only when the laboratory test result is known have delayed treatment. They receive no benefit of treatment for 2 days, until the laboratory results are known and treatment commences. They receive the full benefit of antibiotic treatment thereafter.
- The small minority of children who are very unwell and referred to hospital are diagnosed correctly and treated immediately.

#### **Medium-term model**

- The probability of detection, cost and utility decrement of a recurrent UTI is the same as the index infection.
- Only one recurrent UTI infection may occur in each year.
- A child consults their GP at most once per year.
- Patients with a positive ultrasound scan for VUR are sent for a confirmatory MCUG, which is 100% sensitive and specific.
- Once patients have a first UTI, they experience an increased risk of further UTIs in the subsequent years of the medium-term model.

## Long-term model

- A positive association exists between the number of pyelonephritic attacks and the probability for renal scarring.
- We did not attempt to model negative externalities of an antibiotic prescription, such as the long-term impact on bacterial resistance.

### *Analytical methods*

We converted 95% CIs to standard errors using methods detailed in the Cochrane handbook when parameter standard errors were not reported directly.<sup>167</sup> The model was implemented in WinBUGS 1.4.3 (MRC Biostatistic Unit, Cambridge, UK)<sup>168</sup> using Markov chain Monte Carlo methods. We took a Bayesian approach to the analysis, which requires the specification of a prior distribution for each parameter reflecting uncertainty in the value of the parameter before the data are known. We used diffuse prior distributions for all parameters, which should not have any significant effect on the parameter estimates, allowing the observed data to dominate the estimation. We calculated the expected costs and benefits, in terms of QALYs, for each strategy. The expected costs include the cost of the initial diagnosis, treatment and long-term complications. Benefits include those in the short term, due to the treatment of initial infection, and those due to long-term prevention of complications. We ranked all of the strategies according to their net monetary benefit (NMB), which translates the expected costs and benefits of a given strategy onto a monetary scale for some value of willingness to pay (WTP) for a QALY. The NMB for strategy  $i$  and a WTP  $\lambda$  is:

$$\text{NMB}_i = [\text{Benefit}_i \times \lambda] - \text{Cost}_i \quad (1)$$

We calculated NMB at a WTP value of £20,000. On this basis, each urine sampling strategy is ranked in terms of expected NMB and the optimal strategy is the one with the largest NMB. Outcomes and costs beyond the first year after presentation were discounted at a rate of 3.5%.<sup>169</sup>

We conducted a PSA to account for parameter uncertainty. We calculated the probability of each intervention being cost-effective at a selection of WTP values. We undertook deterministic sensitivity analysis to evaluate the robustness of conclusions to structural assumptions made within the model.

### *Sensitivity analysis*

Several types of uncertainty exist within our model; we conducted further analysis to assess the impact of changes in the methodology or parameter estimates on our findings. Parameter uncertainty was explored through specification of a probability distribution for each of the model parameters. When available, we used the sampling distribution of the parameter; otherwise, expert opinion was used to identify a range of plausible values. We assessed the impact of using a discount rate of 1.5% for both costs and outcomes. We undertook one-way deterministic analysis to assess the generalisability of our results and to identify how robust our estimates were to large changes in the point estimate of key parameters. We conducted sensitivity analysis on parameters when they were thought to be particularly influential on our outcomes, for example the prevalence of UTI, or when no good-quality evidence existed to inform the base-case parameter estimate, for example the utility of children with UTI. The upper and lower bounds for each parameter included in the sensitivity analysis were generated through discussion with members of the project team.

## Results

### Comparison of sampling strategies

#### Diagnostic pathway

In all urine sampling strategies, just over 3% of children who are very unwell are referred directly to hospital for testing and treatment. In the 'sample all' strategy, GPs attempt to obtain a urine sample in the remaining 96.3%, while in the 'sample none' strategy no children are sampled. Using the 'clinical judgement' urine sampling strategy, GPs would attempt to get a sample in 9.2% of older children sampled with a clean catch, compared with only 2.8% of younger children sampled using a nappy pad (*Table 57*). The DUTY5%, DUTY10% and DUTY20% strategies thresholds were selected to sample approximately 5%, 10% and 20% of children, respectively, across both collection methods. When sought, samples were successfully obtained in 92% of cases in all strategies. Presumptive treatment levels were highest in the 'sample all' strategy, where all children in whom a sample was obtained were prescribed immediate antibiotics. Sixty-two per cent of sampled children in the clean-catch 'clinical judgement' strategy were prescribed immediate antibiotics, while around half of the children in whom samples were obtained were treated presumptively across all the DUTY strategies and the nappy pad 'clinical judgement' strategy.

#### Diagnostic accuracy

The boundary strategies 'sample none' and 'sample all' are 100% specific or 100% sensitive, respectively, in selecting children for urine sampling across both collection methods (*Table 58*). In the clean-catch samples, clinical judgement has lower diagnostic accuracy than the DUTY strategies; for example, DUTY5% has higher sensitivity (58.6% vs. 56.7%) and higher specificity (96.1% vs. 91.4%) than clinical judgement. DUTY10%, which samples approximately the same proportion of children as clinical judgement, has similar specificity (91.2% vs. 91.4%) and substantially higher sensitivity (70.7% vs. 56.7%). The DUTY algorithm also outperformed clinical judgement in nappy pad samples; for example, the DUTY5% strategy offered substantially improved sensitivity (42.7% vs. 21.1%) despite having relatively similar specificity (95.1% vs. 97.5%) when compared with clinical judgement. The diagnostic accuracy of both clinical judgement and the DUTY algorithm is substantially lower in nappy pad samples than clean-catch samples. For example, the clean-catch DUTY5% strategy had much higher sensitivity (58.6% vs. 42.7%) than the nappy pad DUTY5% strategy despite sampling a similar proportion of children. PPVs are relatively low, and NPVs are high, across all strategies and both collection methods, primarily due to the low prevalence of UTI. The sensitivity of each strategy is lower at the post-NHS laboratory result stage as initially correct decisions to collect a urine sample may be over-ridden by false-negative NHS laboratory results; however, the ordering of urine sampling strategies, in terms of their sensitivity and specificity, remains unchanged. The high specificity of the laboratory test results in a significantly increased PPV at the post-laboratory test stage, although the imperfect sensitivity results in a slightly lower NPV in all strategies.

#### Treatment pathway

A higher percentage of children with UTI receive appropriate antibiotic treatment in the urine sampling strategies with higher sensitivity. For example, in the clean-catch DUTY5% strategy, 47.8% of children receive immediate or delayed treatment with an appropriate antibiotic; the comparable percentage in the DUTY20% strategy is 60.6% (*Table 59*). The level of serendipitous treatment was high; despite no urine samples being conducted in the 'sample none' strategy, 15.8% of children with UTI received immediate non-resistant treatment, and a total of 29.5% received some form of antibiotic treatment for working diagnoses other than UTI. Conversely, the imperfect sensitivity of the urine sampling strategies, the inability to obtain samples from some children and the imperfect sensitivity of the laboratory test resulted in 19.0% of children with UTI not being treated with antibiotics even in the clean-catch DUTY20% strategy. Similar trends were found in nappy pad samples, although levels of appropriate treatment were lower, across both the clinical judgement and the DUTY algorithm strategies, owing to a lower accuracy of both the initial sampling decision and the laboratory test in nappy pad samples.

TABLE 57 Diagnostic pathway

Outcome	Clean catch					Nappy pad						
	'Sample none'	CJ	DUTY5%	DUTY10%	DUTY20%	'Sample all'	'Sample none'	CJ	DUTY5%	DUTY10%	DUTY20%	'Sample all'
Percentage attempt sample	0.000	9.218	4.755	9.636	19.909	96.297	0.000	2.787	5.461	10.434	19.594	96.297
Percentage obtain sample	0.000	8.437	4.352	8.819	18.221	88.132	0.000	2.551	4.998	9.550	17.933	88.132
Percentage prescribe antibiotics immediately	0.000	5.202	2.135	4.294	8.718	88.132	0.000	1.227	3.115	4.896	9.456	88.132
Percentage wait for laboratory test before antibiotic prescription	0.000	3.234	2.217	4.525	9.504	0.000	0.000	1.324	1.883	4.654	8.478	0.000

CJ, clinical judgement.  
DUTY5%: highly specific DUTY algorithm.  
DUTY10%: moderately sensitive and specific DUTY algorithm.  
DUTY20%: highly sensitive DUTY algorithm.



TABLE 58 Diagnostic accuracy

Outcome	Clean catch					Nappy pad						
	'Sample none'	CJ	DUTY5%	DUTY10%	DUTY20%	'Sample all'	'Sample none'	CJ	DUTY5%	DUTY10%	DUTY20%	'Sample all'
Sensitivity (GP)	0.000	0.567	0.586	0.707	0.843	1.000	0.000	0.211	0.427	0.547	0.636	1.000
Specificity (GP)	1.000	0.914	0.961	0.912	0.806	0.000	1.000	0.975	0.951	0.900	0.805	0.000
PPV (GP)	0.000	0.115	0.231	0.137	0.079	0.019	0.000	0.142	0.146	0.098	0.061	0.019
NPV (GP)	0.981	0.991	0.992	0.994	0.996	0.000	0.981	0.984	0.988	0.990	0.991	0.000
Sensitivity (NHS laboratory)	0.000	0.405	0.419	0.505	0.603	0.715	0.000	0.148	0.300	0.384	0.447	0.703
Specificity (NHS laboratory)	1.000	0.998	0.999	0.998	0.996	0.978	1.000	0.998	0.996	0.993	0.986	0.928
PPV (NHS laboratory)	0.000	0.805	0.904	0.835	0.734	0.392	0.000	0.601	0.619	0.511	0.385	0.163
NPV (NHS laboratory)	0.981	0.988	0.989	0.990	0.992	0.994	0.981	0.983	0.986	0.988	0.989	0.994

CJ, clinical judgement.  
DUTY5%: highly specific DUTY algorithm.  
DUTY10%: moderately sensitive and specific DUTY algorithm.  
DUTY20%: highly sensitive DUTY algorithm.

TABLE 59 Treatment pathway

Outcome	Clean catch				Nappy pad							
	'Sample none'	CJ	DUTY5%	DUTY10%	DUTY20%	'Sample all'	'Sample none'	CJ	DUTY5%	DUTY10%	DUTY20%	'Sample all'
UTI												
Immediate (NR)	15.793	33.583	31.192	42.753	48.351	65.927	15.793	18.928	33.411	33.816	40.173	65.868
Delayed treatment (NR)	0.000	12.824	16.617	10.629	12.250	1.571	0.000	8.571	4.878	11.217	9.286	1.571
Other treatment (R)	13.687	16.707	15.701	19.180	20.401	25.908	13.687	13.547	17.532	16.962	18.752	25.968
Not treated	70.521	36.886	36.489	27.439	18.999	6.593	70.521	58.954	44.179	38.005	31.789	6.593
Not UTI												
Treated with UTI antibiotics	0.000	4.686	1.592	3.490	7.919	88.132	0.000	1.181	2.680	4.653	9.405	88.132
Treated with non-UTI antibiotics	29.479	28.132	29.004	28.433	27.116	5.266	29.479	29.132	28.708	28.113	26.730	5.266
Not treated	70.521	67.182	69.404	68.077	64.965	6.601	70.521	69.686	68.613	67.234	63.865	6.601

CJ, clinical judgement; DUTY5%, highly specific DUTY algorithm; DUTY10%, moderately sensitive and specific DUTY algorithm; DUTY20%, highly sensitive DUTY algorithm.

Clinical judgement resulted in generally poorer targeting of treatment. For example, the clean-catch DUTY5% strategy collected fewer urine samples than clinical judgement, but a slightly higher proportion of children with UTI received non-resistant antibiotics (47.8% vs. 46.4%) and a lower proportion of children without UTI received UTI antibiotics (1.6% vs. 4.7%). Similarly in nappy pad samples, a much higher proportion of children received non-resistant antibiotic treatment in DUTY5% than when clinical judgement was used (38.3% vs. 27.5%), although this was partly offset by a higher proportion of children without UTI receiving UTI antibiotics (2.7% vs. 1.2%).

When comparing DUTY strategies, a lower percentage of children who did not have UTI were incorrectly prescribed antibiotics for UTI in the strategies with higher specificity. For example, in the clean-catch DUTY5% strategy, 1.6% of children who did not have UTI had a false-positive diagnosis and were treated for UTI with an antibiotic; the comparable figure in the clean-catch DUTY20% strategy was 7.9%.

### Short-term costs and benefits

Average sampling, testing and antibiotic costs per patient were highest in strategies that collected urine samples in more patients (e.g. clean-catch DUTY5% £1.10; clean-catch clinical judgement £1.84; clean-catch DUTY20% £3.49). In general, short-term costs of treating acute illness were also higher in strategies that sampled more patients (*Table 60*).

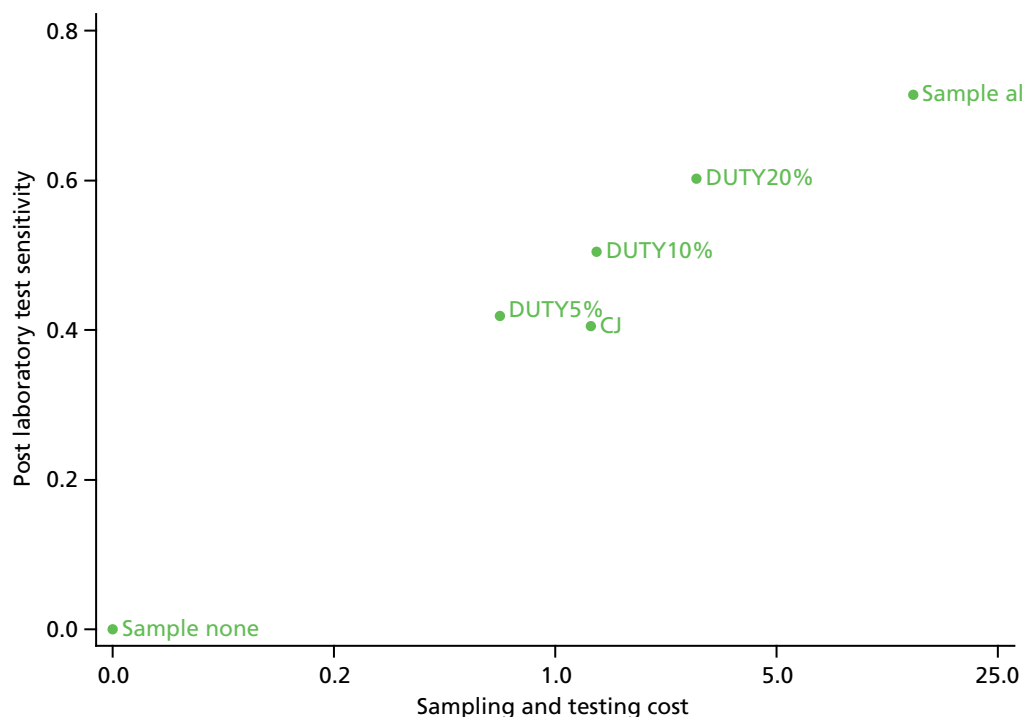
In clean-catch samples clinical judgement was both more costly and no more effective than the DUTY5% urine sampling strategy in terms of the proportion of UTIs correctly identified (*Figure 22*) and QALDs (*Figure 23*). In nappy pad samples clinical judgement was less costly than DUTY5% but had a substantially lower number of UTIs correctly identified (*Figure 24*) and QALDs (*Figure 25*). Across both collection methods, the DUTY20% urine sampling strategy achieved a very marginally higher number of QALDs than other DUTY strategies; however, the higher use of sampling and antibiotic prescription also resulted in higher costs.

There were only very small differences between strategies in short-term net benefit, which summarises the costs and outcomes of each strategy. The relatively low cost of urine sampling and an antibiotic prescription, the high rate of serendipitous antibiotic prescriptions, and the low prevalence of UTI within the DUTY population all contributed to the narrow range of estimated net benefits (ranging from £1075 for the 'sample all' strategy to £1091 for the 'sample none' strategy in the clean-catch samples, and from £1073 for the 'sample all' strategy to £1091 for the 'sample none' strategy in the nappy pad samples). The results of the model slightly favoured conservative (i.e. high specificity, low cost) urine sampling strategies such as DUTY5%. Furthermore, the PSA indicated that there was a very high probability that the 'test none' strategy was more cost-effective than even the DUTY5% strategy at a WTP of £20,000 per QALY (see *Table 60*).

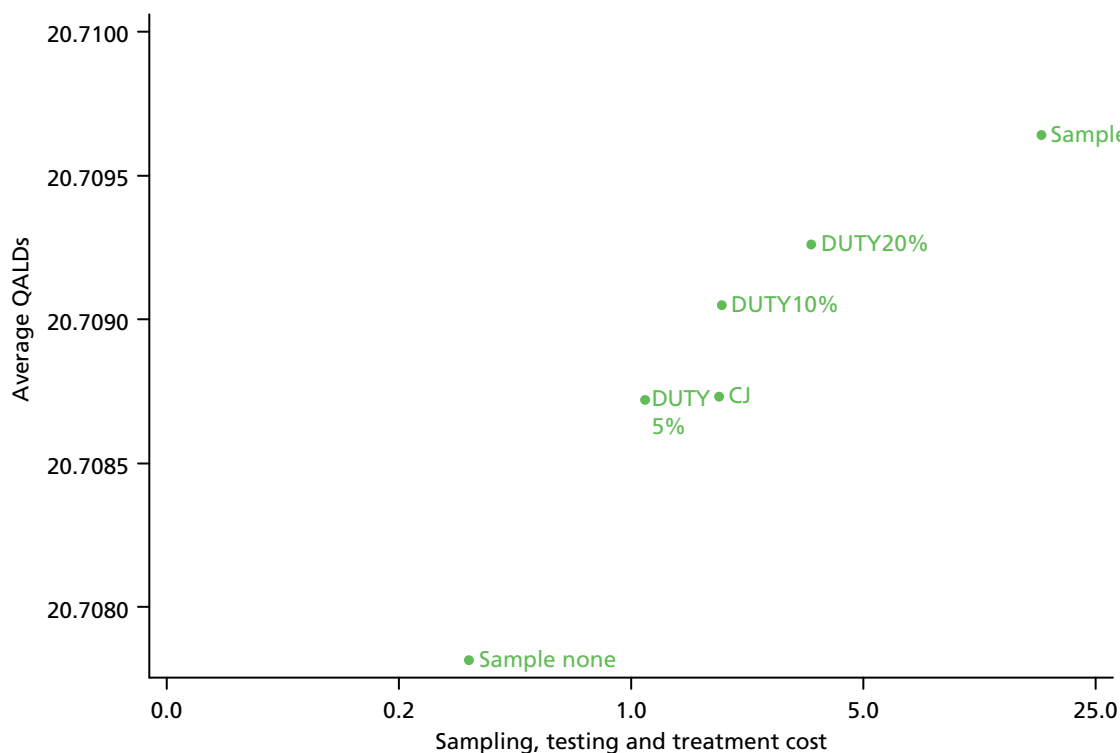
### Medium- and long-term outcomes

The magnitude of differences between strategies for all the medium- and long-term outcomes was also very small for both collection methods (*Table 61*). Urine sampling strategies with the highest sensitivity identified and treated VUR in a higher proportion of cases and hence had the lowest rates of UTI recurrence. However, as the treatment for VUR is only moderately effective and ESRD is rare, the difference between strategies in ESRD incidence, life expectancy and QALYs is negligible. Lifetime net benefits were generally marginally higher in urine sampling strategies where fewer children were sampled. Once again, when considering lifetime net benefits, the model suggested that conservative sampling strategies were more cost-effective and there was a high probability that the 'test none' strategy was the most cost-effective strategy at a WTP of £20,000 per QALY (see *Table 61*).

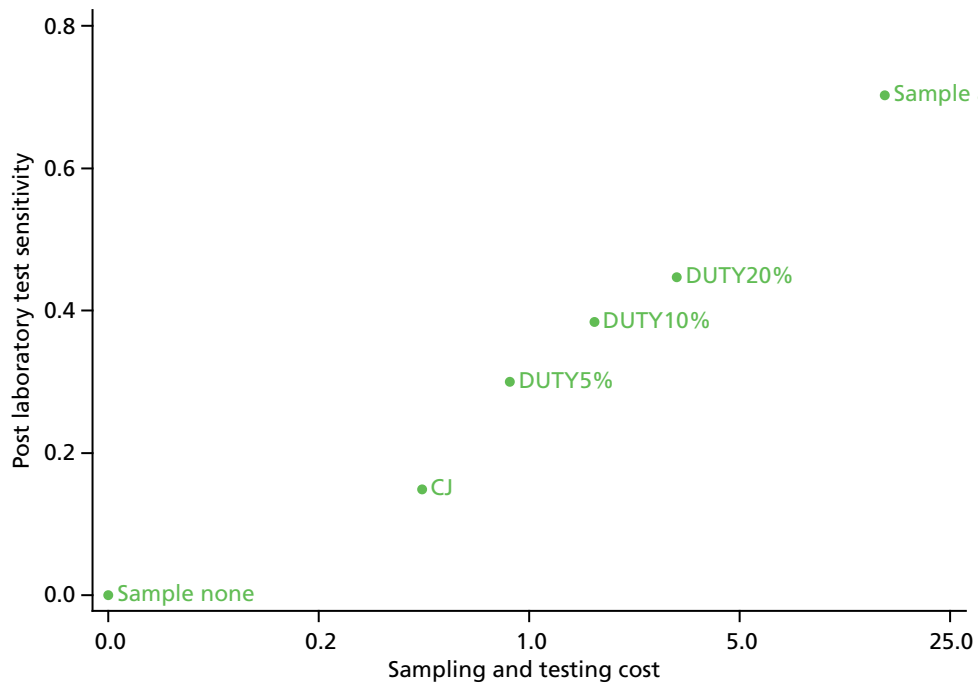




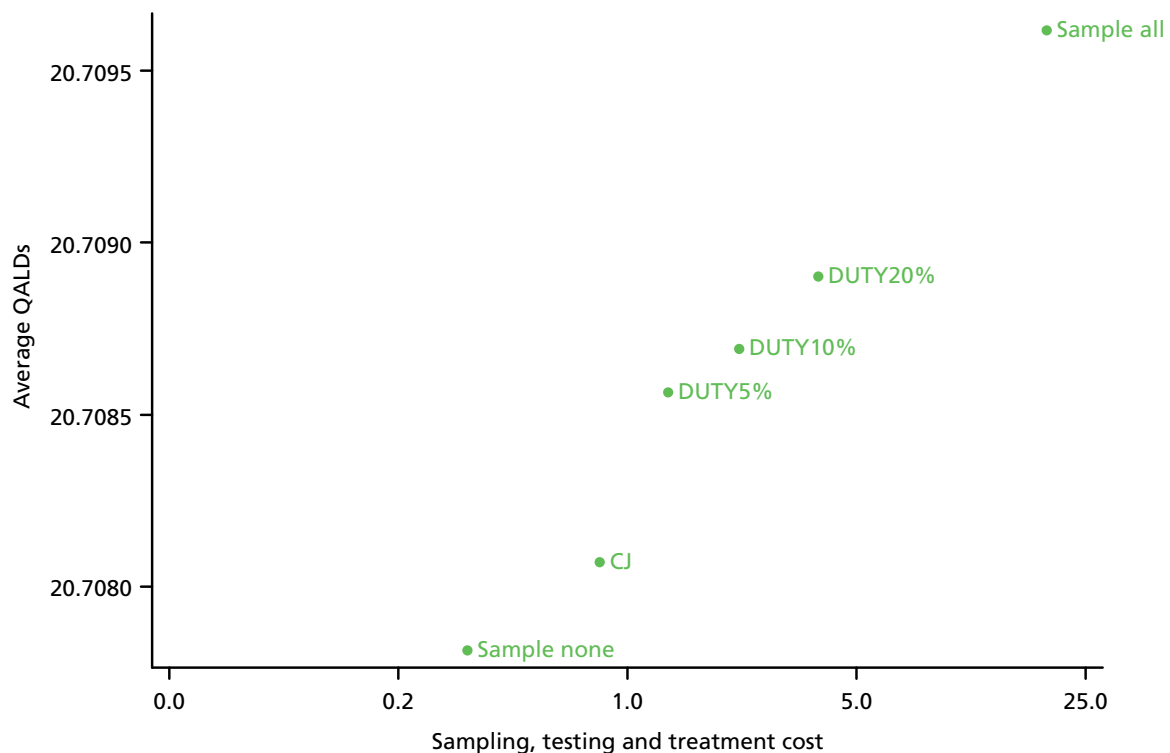
**FIGURE 22** Clean-catch sampling and testing costs vs. percentage of laboratory-confirmed UTI. CJ, clinical judgement; DUTY5%, highly specific DUTY algorithm; DUTY10%, moderately sensitive and specific DUTY algorithm; DUTY20%, highly sensitive DUTY algorithm.



**FIGURE 23** Clean-catch short-term costs vs. average QALDs, comparison of sampling strategies. CJ, clinical judgement; DUTY5%, highly specific DUTY algorithm; DUTY10%, moderately sensitive and specific DUTY algorithm; DUTY20%, highly sensitive DUTY algorithm.



**FIGURE 24** Nappy pad sampling and testing costs vs. percentage of laboratory-confirmed UTI. CJ, clinical judgement; DUTY5%, highly specific DUTY algorithm; DUTY10%, moderately sensitive and specific DUTY algorithm; DUTY20%, highly sensitive DUTY algorithm.



**FIGURE 25** Nappy pad short-term costs vs. average QALDs, comparison of sampling strategies. CJ, clinical judgement; DUTY5%, highly specific DUTY algorithm; DUTY10%, moderately sensitive and specific DUTY algorithm; DUTY20%, highly sensitive DUTY algorithm.

TABLE 61 Medium- and long-term outcomes

Outcome	Clean catch				Nappy pad						
	Sample none	CJ	DUTY5%	DUTY10%	DUTY20%	Sample all	CJ	DUTY5%	DUTY10%	DUTY20%	Sample all
Average number UTI at 3 years/10,000 patients	171.09	171.00	170.99	170.97	170.95	170.92	171.09	171.06	171.00	170.99	170.92
% ESRD	0.288	0.288	0.288	0.288	0.288	0.288	0.288	0.288	0.288	0.288	0.288
Average years lived	72.86	72.86	72.86	72.86	72.86	72.86	72.86	72.86	72.86	72.86	72.86
Average lifetime cost, £	196.13	200.16	197.92	200.10	204.85	245.99	196.13	197.47	201.31	206.23	252.44
Average lifetime QALY	25.722	25.722	25.722	25.722	25.722	25.722	25.722	25.722	25.722	25.722	25.722
Net benefit (£20,000)	514,240	514,236	514,239	514,237	514,232	514,191	514,240	514,239	514,235	514,230	514,184
INMB (95% CI) <sup>a</sup>	3.94 (3.90 to 3.96)	-	2.24 (2.22 to 2.26)	0.09 (0.08 to 0.11)	-4.63 (-4.67 to -4.59)	-45.73 (-45.99 to -45.41)	1.31 (1.29 to 1.32)	-	-3.78 (-3.81 to -3.74)	-8.68 (-8.74 to -8.62)	-54.81 (-55.17 to -54.44)
PCE (£20,000)	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000

CJ, clinical judgement; DUTY5%, highly specific DUTY algorithm; DUTY10%, moderately sensitive and specific DUTY algorithm; DUTY20%, highly sensitive DUTY algorithm; INMB, incremental net monetary benefit; PCE, probability cost-effective.  
 a Incremental net monetary benefit compared with clinical judgement (bootstrapped 95th percentile CIs).

### Comparison of testing and treatment strategies for the clean-catch DUTY5% strategy

In a secondary analysis we compared three alternative urine testing and treatment strategies (laboratory testing-based treatment; dipstick test-based treatment; and presumptive treatment) to explore the role of dipstick testing in guiding laboratory testing and antibiotic prescriptions assuming that urine sampling was based on the DUTY5% strategy. Samples are obtained in 4.35% of children in the DUTY5% strategy; of these, 2.14% are considered low risk and 2.22% are considered higher risk. In the laboratory testing strategy, the 2.14% of children considered low risk have their treatment delayed until the laboratory test is returned. In the dipstick testing strategy, a positive dipstick test result expedites treatment in approximately 26% of these low-risk children, increasing the overall proportion treated immediately with antibiotics from 2.14% to 2.72%. In the presumptive treatment strategy, all low-risk children are treated presumptively and hence 4.35% receive immediate antibiotics (*Table 62* and *Figure 26*).

A greater percentage of children were treated with immediate non-resistant antibiotics in the strategies with the highest presumptive treatment rates (laboratory testing 31.2%; dipstick testing 41.8%; presumptive treatment 45.2%); the laboratory testing strategy has the largest proportion of patients treated with delayed non-resistant antibiotics (16.6%) due to a larger proportion of children who had their treatment delayed until the laboratory result was known. More aggressive treatment strategies led to a higher proportion of children incorrectly treated for UTI (e.g. LT 1.6%; dipstick testing 1.9%; presumptive treatment 3.4%) (see *Table 62*).

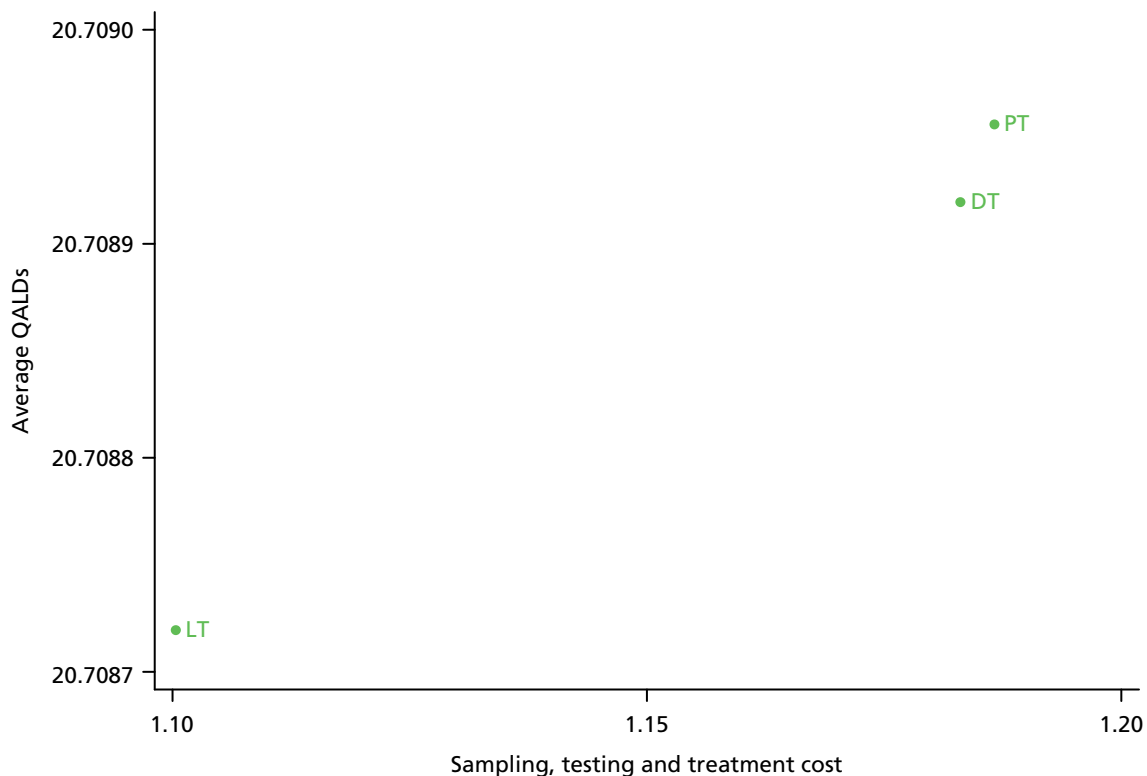
**TABLE 62** Comparison of testing and treatment strategies for clean-catch DUTY5% algorithm

Outcome	LT-DUTY5%	DT-DUTY5%	PT-DUTY5%
Diagnostic pathway			
% dipstick sample	0.000	2.217	0.000
% prescribe antibiotics immediately	2.135	2.720	4.352
% wait for laboratory test before antibiotic prescription	2.217	1.631	0.000
Treatment pathway			
% UTI, immediate (NR)	31.192	41.804	45.174
% UTI, delayed treatment (NR)	16.617	4.469	0.922
% UTI, other treatment	15.701	19.655	20.857
% UTI, not treated	36.489	34.072	33.047
% not UTI, treated with UTI antibiotics	1.592	1.884	3.416
Short-term costs and benefits			
Average cost of sampling, testing (£)	0.668	0.726	0.668
Average cost of sampling, testing, treatment (£)	1.100	1.183	1.187
Short-term average QALDs	20.709	20.709	20.709
Short-term net benefit (£20,000)	1090.44	1090.38	1090.40
INMB, (95% CI) <sup>a</sup>	–	–0.05 (–0.07 to –0.04)	–0.04 (–0.05 to –0.03)

DT, dipstick-based treatment; LT, laboratory-based treatment; PT, presumptive treatment.

<sup>a</sup> Incremental net monetary benefit compared with clinical judgement (bootstrapped 95th percentile CIs).





**FIGURE 26** Clean-catch short-term costs vs. average QALDs, comparison of treatment and testing strategies. LT, laboratory-based treatment; DT, dipstick-based treatment; PT, presumptive treatment.

Average sampling and testing costs are higher in the DT strategy (£0.73) than in the laboratory testing and presumptive treatment strategies (£0.67) due to the additional cost of the dipstick test. When including antibiotic treatment costs, the presumptive treatment strategy (£1.187) was more expensive than both the dipstick testing (£1.183) and laboratory testing (£1.100) strategies. Strategies with higher levels of immediate treatment had marginally larger QALDs, but net benefits were very similar across all three testing and treatment strategies (see *Table 62*).

### Comparison of the full DUTY risk score with points-based algorithm

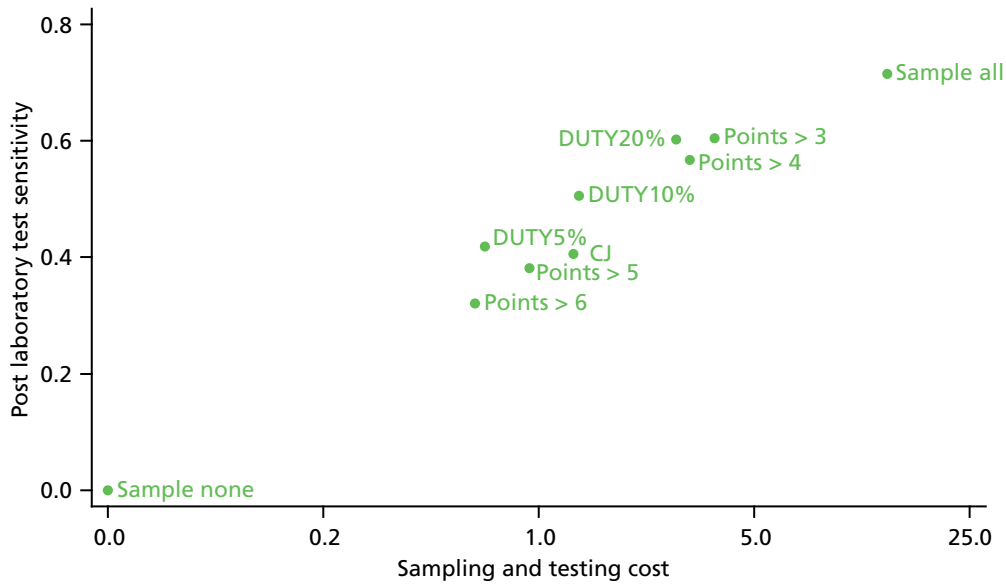
In a further analysis we compared the outcomes of the full DUTY risk score to the simpler points-based algorithm. We attempted to choose cut points at 3, 4, 5 and 6 on the DUTY points score. Owing to nature of the points score, with only a small number of possible values, a secondary cut-point could not always be chosen to allow close to half of the sample children to be treated presumptively and allow maximum comparability with the DUTY coefficient model. For example, the proportion of children who were sampled who were treated with immediate treatment varied from 16% to 65% in the points-based strategies, compared with between 48% and 49% in the full risk score strategies (*Table 63*).

The diagnostic accuracy of the points-based algorithms was lower than the full risk score. For example, in clean-catch samples the full model DUTY5% strategy had similar sampling and testing costs (£0.67 vs. £0.62) to the points-based DUTY  $\geq 6$  strategy, but much higher post-laboratory test sensitivity (41.9% vs. 32.1%) (*Figure 27*). Despite the loss of accuracy when compared with the full model, the points-based model still outperformed clinical judgement. For example, the points-based algorithm DUTY  $\geq 5$  strategy had a similar average number of short-term QALDs to the clinical judgement strategy despite substantially lower costs (£1.43 vs. £1.84) (*Figure 28*). Short-term net benefits were similar across all of the full model and points-based strategies, ranging from £1087 to £1090, but were generally lower in strategies which sampled more patients.

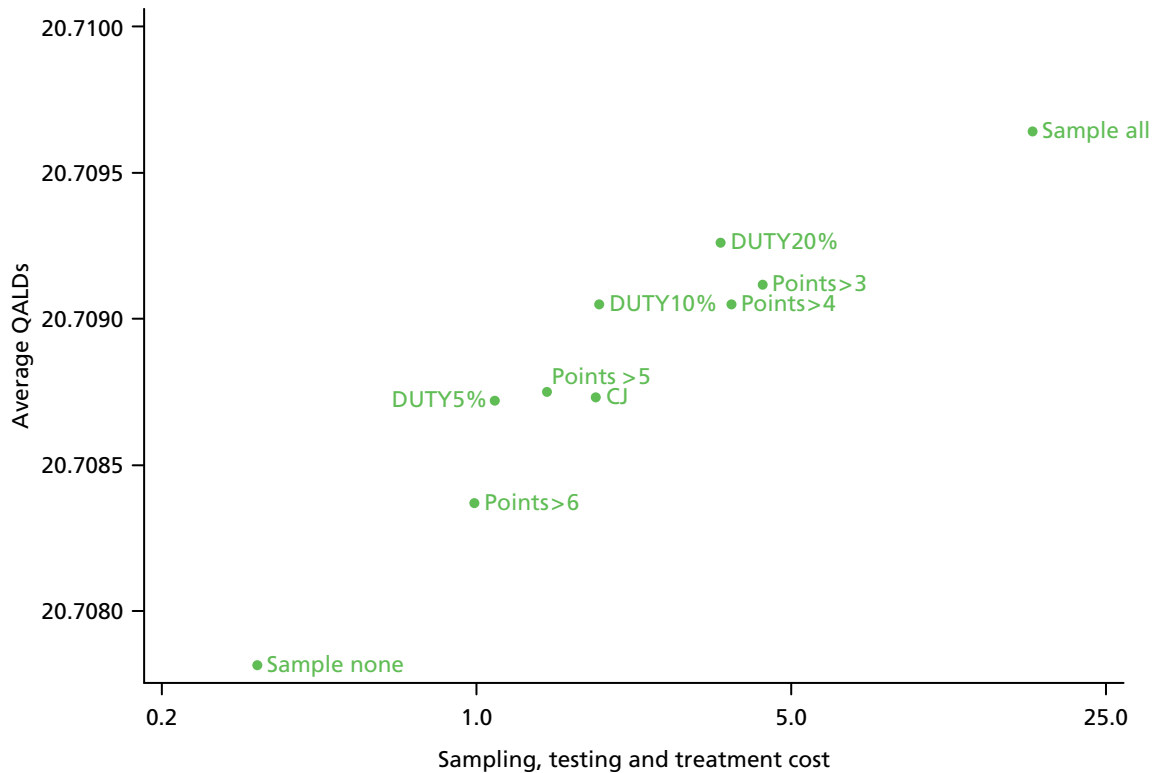
TABLE 63 Comparison of full and points-based models for clean-catch samples<sup>a</sup>

Outcome	CJ	DUTY5%	DUTY10%	DUTY20%	DUTY points ≥ 6	DUTY points ≥ 5	DUTY points ≥ 4	DUTY points ≥ 3
Diagnostic pathway								
% attempt sample	9.218	4.755	9.636	19.909	4.424	6.649	22.017	26.540
% obtain sample	8.437	4.352	8.819	18.221	4.049	6.085	20.150	24.290
% prescribe antibiotics immediately	5.202	2.135	4.294	8.718	0.655	3.973	6.024	6.016
% wait for laboratory test before antibiotic prescription	3.234	2.217	4.525	9.504	3.394	2.112	14.126	18.274
Diagnostic accuracy								
Sensitivity (laboratory)	0.405	0.419	0.505	0.603	0.321	0.382	0.567	0.604
Specificity (laboratory)	0.998	0.999	0.998	0.996	0.999	0.999	0.995	0.994
PPV (laboratory)	0.805	0.904	0.835	0.734	0.880	0.848	0.699	0.671
NPV (laboratory)	0.988	0.989	0.990	0.992	0.987	0.988	0.991	0.992
Short-term costs and benefits								
Average cost of sampling, testing (£)	1.295	0.668	1.354	2.798	0.622	0.934	3.094	3.729
Average cost of sampling, testing, treatment (£)	1.842	1.100	1.875	3.492	0.990	1.434	3.685	4.322
Short-term average QALDs	20.709	20.709	20.709	20.709	20.708	20.709	20.709	20.709
Short-term net benefit (£20,000)	1089.70	1090.44	1089.72	1088.16	1090.49	1090.12	1087.93	1087.30
INMB (95% CI) <sup>a</sup>	-	0.74 (0.72 to 0.76)	0.02 (0.01 to 0.04)	-1.54 (-1.56 to -1.51)	0.79 (0.77 to 0.81)	0.42 (0.40 to 0.44)	-1.76 (-1.79 to -1.74)	-2.40 (-2.42 to -2.37)

Outcome	CJ	DUTY5%	DUTY10%	DUTY20%	DUTY points $\geq 6$	DUTY points $\geq 5$	DUTY points $\geq 4$	DUTY points $\geq 3$
Medium-term costs and benefits								
Average number with UTI at 3 years/10,000 patients	171.00	170.99	170.97	170.95	171.02	171.00	170.96	170.95
% ESRD	0.288	0.288	0.288	0.288	0.288	0.288	0.288	0.288
Average years lived	72.86	72.86	72.86	72.86	72.86	72.86	72.86	72.86
Average lifetime cost (£)	200.16	197.92	200.10	204.85	197.77	198.94	205.54	207.43
Average lifetime QALY	25.722	25.722	25.722	25.722	25.722	25.722	25.722	25.722
INMB	-	2.24 (2.22 to 2.26)	0.09 (0.08 to 0.11)	-4.63 (-4.67 to -4.59)	2.35 (2.33 to 2.37)	1.22 (1.20 to 1.24)	-5.34 (-5.38 to -5.29)	-7.23 (-7.29 to -7.17)
CJ, clinical judgement; DUTY5%, highly specific DUTY algorithm; DUTY10%, moderately sensitive and specific DUTY algorithm; DUTY20%, highly sensitive DUTY algorithm; INMB, incremental net monetary benefit. a Incremental net monetary benefit compared with clinical judgement (bootstrapped 95th percentile CIs).								



**FIGURE 27** Clean-catch sampling and testing costs vs. post-laboratory test sensitivity, comparison of full and points-based models. CJ, clinical judgement; DUTY5%, highly specific DUTY algorithm; DUTY10%, moderately sensitive and specific DUTY algorithm; DUTY20%, highly sensitive DUTY algorithm.



**FIGURE 28** Clean-catch short-term costs vs. average QALDs, comparison of full and points-based models. CJ, clinical judgement; DUTY5%, highly specific DUTY algorithm; DUTY10%, moderately sensitive and specific DUTY algorithm; DUTY20%, highly sensitive DUTY algorithm; PB, points-based algorithm.

### Deterministic sensitivity analysis

We undertook one-way deterministic analysis to assess the generalisability of our results and to identify how robust our estimates were to large changes in the point estimate of key parameters (Tables 64 and 65). None of the scenarios tested changed our conclusions that 'sample none' or the conservative DUTY5% sampling strategy were likely to be most cost-effective. Of all short-term scenarios tested (see Table 64), an assumption that the prevalence of UTI could be as high as 10% had the biggest impact on the difference in short-term net benefits between the conservative (DUTY5%) urine sampling strategy and more liberal sampling strategies (e.g. DUTY20%). However, the conservative urine sampling strategies remained most cost-effective. Similar results were found in sensitivity analyses on nappy pad outcomes.

**TABLE 64** Sensitivity analysis, short-term parameters

Parameter	Base case	Sensitivity analysis	Short-term NMB, £			
			'Sample none'	DUTY5%	DUTY10%	DUTY20%
UTI prevalence	0.019	0.100	1088.96	1088.30	1087.67	1086.20
Pyelonephritis prevalence	0.157	0.570	1090.73	1090.22	1089.53	1087.97
Serendipitous treatment rates	0.267	0.000	1091.67	1091.13	1090.42	1088.83
Diagnostic accuracy of laboratory culture (sensitivity, specificity)	0.780, 0.978	1.000, 1.000	1091.45	1090.88	1090.16	1088.59
Antibiotic treatment effect	RR <sup>a</sup> = 0.550	RR = 0.250	1091.01	1090.61	1089.96	1088.44
Impact of antibiotic resistance on symptom resolution	RR <sup>b</sup> = 0.700	RR = 0.550	1091.42	1090.84	1090.13	1088.55
Utility of UTI (UTI, PA) <sup>c</sup>	0.943, 0.711	0.688, 0.500	1021.61	1021.14	1020.47	1018.93
Resistance rates, non-UTI antibiotics	0.531	0.750	1091.43	1090.86	1090.14	1088.56
Resistance rates, UTI antibiotics	0.281	0.000	1091.45	1090.90	1090.21	1088.65

PA, pyelonephritic attack.  
a Risk ratio comparing symptom resolution rates for children not treated with antibiotics with those in children treated with antibiotics.  
b Risk ratio comparing symptom resolution rates for children treated with antibiotics where the uropathogen was resistant compared with where the uropathogen was sensitive.  
c Utility for patients with UTI and pyelonephritis, respectively.

**TABLE 65** Sensitivity analysis, long-term parameters

Parameter	Base case	Sensitivity analysis	Short-term NMB, £			
			'Sample none'	DUTY5%	DUTY10%	DUTY20%
Probability of PRS (0 PA, 1 PA, 2 PA, 3 PA, 4 PA) <sup>a</sup>	0.05, 0.09, 0.16, 0.34, 0.58	0.10, 0.18, 0.32, 0.68, 1.00	512,740	512,738	512,736	512,731
Cost of ESRD, £ (dialysis, transplant) <sup>b</sup>	21,655, 19,456	43,310, 38,912	514,594	514,592	514,590	514,585
Discount rate (%)	3.5	1.5	889,820	889,818	889,816	889,811

PA, pyelonephritic attack.  
a Probability of PRS given 0–4 pyelonephritic attacks.  
b Cost of dialysis (annual) and transplant, respectively.

## Discussion

### Summary of findings

The economic analyses described in this chapter evaluated the potential to use the clinical symptoms and signs predictive of UTI to improve upon current practice in selecting children for urine sampling and targeting antibacterial therapy. We found that in older children, when a clean-catch urine sample is feasible, the full DUTY risk score outperformed GPs' clinical judgement, even taking account of the reduced reliability and accuracy of NHS laboratories compared with the research laboratory and the need to repeat contaminated samples. For example, by choosing a high specificity threshold (DUTY5%), GPs would request a urine sample from fewer children (4.8% vs. 9.2%) with equivalent or higher sensitivity and specificity, achieving marginally better costs and patient outcomes in the short and long term.

In younger children, if urine was collected using a nappy pad, the distinction in cost-effectiveness between the DUTY risk score and clinical judgement was not clear-cut. This is a result of the lower diagnostic value of the DUTY risk score in younger children, noted in *Chapter 5*, the higher contamination rates necessitating repeat urine sampling and the lower accuracy of NHS laboratory results in urine collected using nappy pads, raising the possibility that a correct clinical diagnosis is overturned by an incorrect laboratory test result.

The choice of a threshold on the DUTY risk score which will optimally identify patients whose risk of UTI is high enough to justify collecting a urine sample is not straightforward. We explored three possible thresholds which represented a range of trade-offs between high sensitivity and high specificity. The results of our decision analysis model slightly favoured conservative (i.e. high specificity) urine sampling strategies. However, the absolute difference in average short- and long-term net benefits between the three DUTY thresholds evaluated was very small.

Using a DUTY points-based algorithm in clean-catch samples, which can be calculated simply, achieves lower short-term costs of testing and treatment and equivalent patient outcomes to clinical judgement. For example, the incremental net benefit when comparing the DUTY points  $\geq 5$  and clinical judgement algorithms was £0.42 (95% CI £0.40 to £0.44). However, this simplification came at the expense of lower diagnostic accuracy than the comparable full DUTY risk score.

The additional diagnostic information provided by dipstick tests has the potential to play a role in the efficient diagnosis and treatment of children with suspected UTI. We found that, if used in children thought to be at 'lower' risk of UTI to determine whether or not immediate antibacterial therapy should be started, a dipstick test strategy could increase the proportion of children with UTI treated immediately (41.8% vs. 31.2%) compared with a strategy based on laboratory test results, and decrease the proportion of children without UTI treated with UTI antibacterials (1.9% vs. 3.4%) compared with a presumptive treatment strategy. These potential benefits are counterbalanced by the additional costs of the dipstick test.

### Strengths and limitations

Our model was based on individual patient data from a large diagnostic cohort study which avoided many of the potential biases of diagnostic accuracy studies. Our analysis is based on high-quality evidence about the potential for GPs to use clinical symptoms and signs predictive of UTI to improve upon current practice in selecting children for urine sampling. However, our estimate of current practice in the clinical judgement strategy is based on primary care clinicians' responses to questions about working diagnoses and hypothetical testing and treatment actions if the child had not been in DUTY. This could have introduced a type of Hawthorne effect,<sup>170</sup> whereby clinicians participating in DUTY had a higher index of suspicion of UTI or were more careful in reaching their working diagnosis because they were participating in a research study on UTI. If this were the case, clinical judgement in routine practice might have lower sensitivity than that which we observed in DUTY. We believe that it is unlikely that any such effect would alter our finding that the DUTY risk scores in clean-catch samples were a more cost-effective diagnostic strategy than clinical judgement.

We did not attempt to quantify the wider societal costs of antimicrobial resistance, beyond the treatment of childhood UTI. A recent review of the literature concluded that current estimates of the costs of resistance are likely to be underestimates and that an accurate forecast of costs may not be possible.<sup>104</sup> The economic evidence provided by our study should be viewed in the wider context of increasingly resistant organisms and fewer new antimicrobial drugs reaching the market. Furthermore, we did not model the impact of urine sampling strategies on children who do not have UTI. Strategies that promptly and accurately 'rule out' UTI may help children by allowing clinicians to quickly diagnose and treat other illnesses; however, given the myriad of other illnesses that children may have, this is extremely difficult to quantify.

Our model relies on incomplete and imperfect evidence. For example, the antibiotic treatment effect in the short-term model is based on a small RCT conducted in adults and the utility scores are based on children presenting to primary care with other acute conditions. The evidence underlying much of the medium- and long-term model is based on observational associations which may be subject to bias. In some areas the evidence is evolving. The Randomised Intervention for Children with Vesicoureteral Reflux (RIVER) study (ClinicalTrials.gov number NCT00405704), which at the time of writing is yet to publish results, will provide useful information on the benefits of long-term antibiotic prophylaxis and has the potential to change the results of our model. We have explored the uncertainty in our model introduced by limited and imperfect evidence by conducting a range of deterministic and probabilistic sensitivity analyses. These analyses suggest that our conclusions, favouring conservative (i.e. high specificity) urine sampling strategies, are not affected by relatively large but plausible changes in parameters such as UTI prevalence and the diagnostic accuracy of NHS laboratories. Because of the very small differences in predicted average patient outcomes between strategies, our findings support low-cost, high-specificity strategies across the range of scenarios tested in sensitivity analyses.

We modelled a limited number of urine sampling, testing and treatment strategies. The large number of risk score thresholds which could be used to decide which children to sample or treat presumptively and the multiple permutations of dipstick test results which might be used to expedite treatment or determine which samples to send to the laboratory produce an almost unlimited number of possible strategies. We evaluated those which most closely reflected the range of current practice among GPs on the project team, but these might not be representative of all strategies currently used by GPs, and other strategies might potentially prove to be more cost-effective.

### **Results in context with other studies**

As far as we are aware, this is the first study to evaluate the cost-effectiveness of prediction rules to optimise the selection of children with suspected UTI for urine sampling in any primary or secondary care setting. Previous work has assessed the most cost-effective test or series of tests for diagnosing UTI rather than evaluating which children should be selected for urine sampling. A NIHR HTA programme systematic review and economic model evaluating 79 testing and treatment strategies for children with suspected UTI concluded that the optimal strategy was dependent on the age and sex of the child.<sup>56</sup> Presumptive treatment had the highest probability of being cost-effective at lower willingness-to-pay thresholds. At higher thresholds, up to £20,000 per QALY, testing and treatment based on positive dipstick nitrites and LE results became the most probably cost-effective strategy.

### **Clinical and research implications**

For an individual child, who may or may not have UTI, attending primary care with non-specific symptoms, the expected costs and benefits of all six diagnostic strategies were very similar. This is because the cost of urine sampling is relatively low and for the vast majority of children, who do not have UTI, outcome is not dependent on the diagnostic strategy chosen. However, for the NHS's provision of care to the population of children visiting primary care throughout their early years, differences in the aggregate costs of the diagnostic strategies will be important. Equally, for the individual child who does have UTI, the choice of diagnostic strategy will have a significant impact on outcome. Our findings demonstrate the need for clinicians to base the decision to collect a urine sample on symptoms and signs known to be predictive of UTI in primary care rather than on personal judgements or mixed evidence derived from secondary care settings.

Our results tended to favour low-cost conservative (higher specificity) urine sampling strategies. Indeed, in both the clean-catch and nappy pad models, the 'sample none' strategy yielded the highest expected net benefit, although the difference between it and the conservative DUTY5% strategy was very small. The reasons for this are manifold and include the high prevalence of antibiotic prescribing for other illnesses leading to serendipitous treatment of undiagnosed UTI, the relatively short duration of antibiotic effects in children with UTI and the rare and uncertain nature of long-term sequelae of untreated UTI. Assuming that a 'sample none' strategy is untenable, then a reasonably conservative strategy for a GP concerned about the overprescription of antibiotics would be similar to the DUTY5% strategies. For example, a GP requesting urine samples in children scoring  $> 5$  on the DUTY points-based score or a probability of more than 0.085 on the DUTY full risk model would sample approximately 4.4% of all children with non-specific constitutional symptoms and request a sample in  $> 43\%$  of children who have UTI. A GP more concerned about missing a diagnosis of UTI could use the more liberal DUTY10% or DUTY20% thresholds.

Where possible, the DUTY risk score should be calculated based on the full model rather than the points-based model because of the increased diagnostic accuracy. For electronic records where the relevant symptoms and signs are routinely recorded, this process should be inexpensive and automated. However, in settings where resources do not permit this, a hand calculation based on the DUTY points-based model will allow clinicians to more efficiently identify children at highest risk of UTI than through using clinical judgement alone.

The poorer diagnostic and economic performance of the DUTY risk score in younger children with nappy pad samples underlines the importance of obtaining a clean-catch sample whenever practical. The initial inconvenience and time costs of collecting a clean-catch sample in a busy clinical setting is likely to be more than outweighed by the cost and health implications of contaminated samples, repeated samples and false-positive laboratory results associated with nappy pad samples.

Future research could further define the role of dipstick testing in diagnosis and treatment strategies for childhood UTI. We have demonstrated the potential for positive nitrite and leucocyte results to expedite treatment in children at intermediate risk of UTI. However, the many permutations of dipstick results and the multiple ways in which it might inform clinical decision-making should be explored further. In particular, dipstick tests may also be useful in children considered, based on symptoms and signs, to be at higher risk of UTI. In this group, a negative dipstick test result might be an efficient way to prevent hasty presumptive treatment in children who do not have UTI. Additional work is needed to also quantify the benefits of prompt diagnosis of UTI. Studies of parent-reported quality of life and disutility of UTI symptoms in young children would enable more precise estimations of the short-term benefits of antibiotics. Long-term epidemiological study designs are also needed to better quantify the strength of association between childhood UTI and eventual renal disease. However, because some aspects of the decision problem, such as the societal impact of antibacterial resistance, are essentially intangible, the choice of diagnostic strategy is likely to remain a balance between evidence and judgement.

### Conclusions

For older children, we found that the full DUTY coefficient and the simpler DUTY points-based model were both marginally more cost-effective than GPs' clinical judgement in selecting children in whom to collect a clean-catch sample and test for UTI. Small differences between strategies in cost-effectiveness are important given the large number of urine samples collected in children. The difference in cost-effectiveness between the DUTY risk score and clinical judgement for younger children with nappy pad samples was not clear-cut, underlining the importance of obtaining a clean-catch sample whenever practical. Our findings suggest that high specificity thresholds, such as DUTY5%, are likely to be more cost-effective, particularly for GPs concerned about the societal impact of antibacterial resistance.



# Chapter 7 Determinants of urinary contamination

## Introduction

Diagnosis of UTI is difficult to make and may be missed in up to 50% of children presenting to primary care.<sup>48,107</sup> Establishing a diagnosis in pre- or early school-aged children is challenging, because a high proportion are preverbal, there is a lack of specific symptoms, and collection of uncontaminated urine samples is difficult, particularly in primary care settings where the provision of time and private space for young children is challenging.<sup>64</sup> UK NICE guidelines suggest that samples should be collected using a method suitable for the age of the infant or child.<sup>2</sup> 'Clean-catch' samples are preferred, but urine collection pads (nappy pads) are suggested if this is not possible. In the USA, SPA is the gold standard and has been used along with urethral catheterisation, but both are invasive and can be painful for infants, with limited success rate.<sup>102,120,146,147</sup> US guidelines recognise ultrasonography increases success but at increased cost.

Laboratory diagnosis is based on colony counts after culture, using quantity and type of bacteria present in the urine sample as the main criteria. Culture methods differ between NHS laboratories but all are derivatives of the UK standards for microbiology investigations (B41),<sup>82</sup> which describe criteria for significant bacteruria for children as  $\geq 10^5$  CFU/ml of a single species from a clean-catch specimen,  $10^4$ – $10^5$  CFU/ml single species or  $\geq 10^5$  CFU/ml of any bacteria from a SPA. These guidelines also describe colony counts of  $\leq 10^4$  CFU/ml and  $\geq 10^5$  CFU/ml from bag urines as diagnostically useful for a negative result and as an indicator of a poor-quality sample, respectively.

It is recognised that difficulty in specimen collection and the interpretation of specimens potentially contaminated prior to culture, from skin, faeces and other sources, may contribute to the misdiagnosis of UTI.<sup>89</sup> Overdiagnosing UTI can lead to unnecessary investigations and treatment, which entail risks of complications and psychological stress to the child and his or her family. Discriminating between contamination with faecal organisms and potential UTI pathogens in the laboratory is difficult, as the most common faecal organism and the most common pathogen causing UTI is *E. coli*, and high colony counts may be found in children without UTI.

Urine samples in the DUTY study were collected by either clean catch or nappy pad, with samples split between a local NHS laboratory and, for a more detailed examination, a research laboratory. The aim of this chapter is to compare the prevalence of contamination by urine collection method, based on the research laboratory result, and to identify the independent factors associated with contamination using the local laboratory results.

## Methods

### *Participants, urine collection and microbiological methods*

These have all been described in detail in previous chapters and our protocol paper.<sup>105</sup> For a full description of the measures taken to minimise contamination of urine specimens, see *Chapter 2*.

## Definition of contamination

Interpretation of laboratory results from microscopy and culture is not straightforward, with definitions, collection methods and criteria for contamination and UTI inextricably linked. The HPA in its 2012 guideline<sup>82</sup> does not use differing cut levels for the diagnosis of UTI in children, but UTI criteria are dependent upon all patient information, with different criteria for various patient groups. Generally, UTI is suspected if culture counts of a pure or predominant growth of organism are  $\geq 10^5$  CFU/ml. Any counts below this with  $\geq 2$  organisms present may be indicative of contamination. In children, the guidelines suggest that cultures  $< 10^4$  CFU/ml from bag urines may be regarded as negative for UTI and that cultures  $\geq 10^5$  CFU/ml may be contaminated and should be repeated with a reliable specimen. There is no mention of any use of nappy pads even though this has been established as a recommended collection method by NICE.<sup>2</sup> Squamous epithelial cells present in the urine are a useful indicator of the degree of contamination, but no clear cut-off is recommended in the HPA guidelines.

The European Urinalysis Guidelines of 2000<sup>136</sup> consider 'most acute uncomplicated urinary tract infections result from one bacterial species' and 'the isolation of more than one organism from a single specimen of urine must be interpreted in the light of (potential) contamination'. The European Association of Urology Guidelines of 2009 regard colony counts of between  $10^3$  and  $10^5$  CFU/ml as UTI in symptomatic adults and  $\geq 10^5$  CFU/ml in asymptomatic adults.<sup>83</sup> In children, the UTI cut-off criteria are based on collection method: any growth from SPA,  $10^3$  to  $10^4$  CFU/ml from catheter urine,  $\geq 10^4$  CFU/ml in children with UTI symptoms and  $\geq 10^5$  CFU/ml without UTI symptoms. Neither mention of contamination nor any mention of nappy pad as a collection method is made. A variety of different definitions of contamination have been proposed, differences being in the number of different organisms present and the quantity, derived from colony counts (Table 66).

For the purposes of this chapter we defined contamination using the research laboratory results, with a criterion of count  $\geq 10^5$  and  $> 2$  organisms, equivalent to Jackson *et al.*'s<sup>94</sup> definition of frankly contaminated. We also excluded any urines collected via bags ( $n = 73$ ) and any urines with an unknown collection method ( $n = 17$ ), leaving 5017 urines to analyse.

We also considered a number of other definitions of contamination, firstly equivalent to Feasey<sup>90</sup> or Rao *et al.*<sup>91</sup> and secondly equivalent to Bekeris *et al.*<sup>96</sup> or to Jackson *et al.*<sup>93</sup> of probable contamination or higher. Although we did not carry out extensive analyses using these definitions, we considered their effect

**TABLE 66** Comparing other research definitions of contamination (in date order)

Paper authors	Contamination definition
Feasey <sup>90</sup>	$> 10^5$ growth and $\geq 2$ organisms
Rao <i>et al.</i> <sup>91</sup>	$> 10^5$ growth and $\geq 2$ organisms
Alam <i>et al.</i> <sup>92</sup>	Any growth including $< 10^5$ (that is not UTI)
Jackson <i>et al.</i> <sup>93</sup>	$10^4$ to $10^5$ 1 or 2 species possible contamination $10^4$ to $10^5$ $> 2$ species probable contamination $> 10^5$ $> 2$ species frank contamination
Blake <i>et al.</i> <sup>94</sup>	Growth of a non-uropathogen
Unlu <i>et al.</i> <sup>95</sup>	$10^3$ to $10^4$ growth and $\geq 2$ organisms
Vaillancourt <i>et al.</i> <sup>89</sup>	Single organism growth $< 10^5$ OR $\geq 2$ organisms
Bekeris <i>et al.</i> <sup>96</sup>	$> 10^4$ growth and $> 2$ organisms
Wingerter <i>et al.</i> <sup>97</sup>	Multiple pathogens OR non-uropathogens OR any growth $< 10^4$
Tosif <i>et al.</i> <sup>98</sup>	$\geq 2$ organisms (and a collection method specific growth)

on the numbers of contaminated urines by collection method. We selected these definitions from *Table 66* to most closely reflect what we thought a NHS laboratory would report in practice and also which did not include too many borderline growths (lower than  $10^4$  CFU/ml) that could be a UTI.

## Statistical analysis

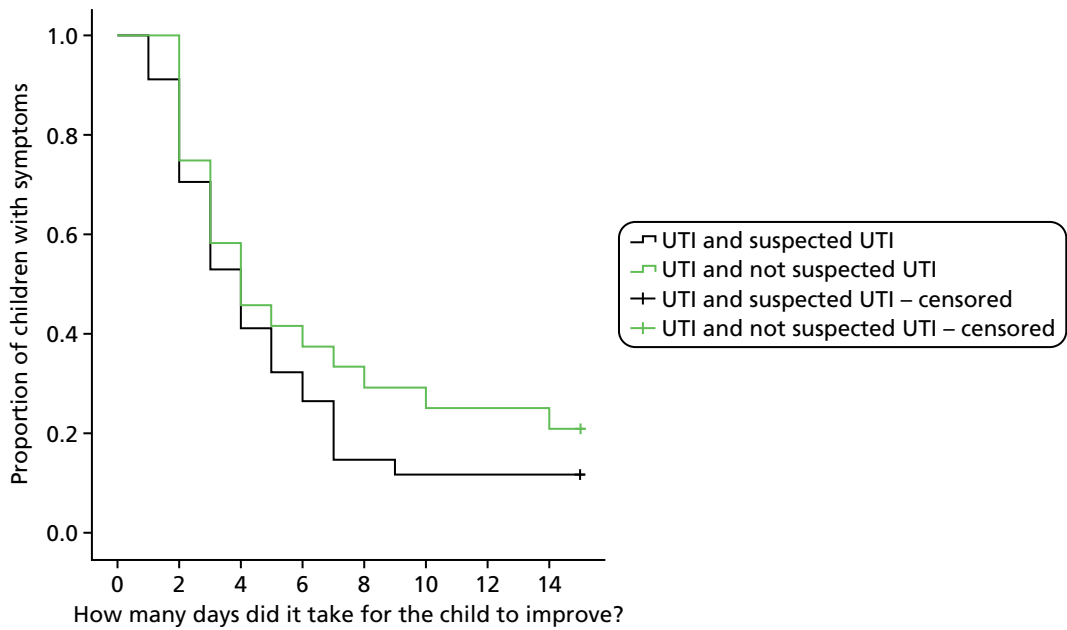
All analyses used results from the research laboratory to define contamination. We compared contamination rates by urine collection method. To create a model showing factors associated with contamination, we selected a priori a number of variables. The variables were grouped in order to reduce the potential for spurious associations into those associated with laboratory and clinical findings, with faeces, skin variables and variables from our prediction rule (see *Chapter 5*). The laboratory and clinical findings included were dipstick tests results, leucocytes, pH, specific gravity, proteins, ketones, blood, nitrites, urine collection method, and the urine being smelly, dark or cloudy. In addition, we examined where the sample was collected, whether the nurse was an in-house or a DUTY nurse, and the time taken for the sample to reach the laboratory. The faecal variables were diarrhoea in the past 24 hours, the child taking laxatives and constipation in the last week. The skin variables were sex, circumcision, wearing nappies, dehydration, number of nappies used in the previous 24 hours, number of showers/baths in a week and nappy rash. The variables from our prediction rule (see *Chapter 5*) included were cough, global impression of child, pain/crying when passing urine, abdominal tenderness on examination, any acute abnormality on ear examination and history of UTI. We then examined the frequency of variable categories, blind to their associations with contamination, and merged the least frequent categories prior to analyses. Univariable associations with contamination were estimated using logistic regression. *P*-values were derived from a LR test of adding each single variable. In the case where the variable was ordinal a single heterogeneity *p*-value was reported and another for the trend. These univariable associations were stratified by urine collection method. We selected any variable for our multivariable model if it had a *p*-value  $< 0.01$  considering either heterogeneity or trend *p*-values. These two multivariable models were then displayed using ORs and their 95% CIs. All analyses were completed using a complete case approach.

We carried out a number of comparisons: the number of contaminated urines in clean catch and nappy pads; and for the contaminated samples only, the numbers of urines with organisms (coagulase-negative staphylococci, *E. coli*, enterococci and staphs) present in clean catch and nappy pads. These comparisons included the risk difference and risk ratio along with their 95% CIs and a *p*-value derived from a chi-squared test.

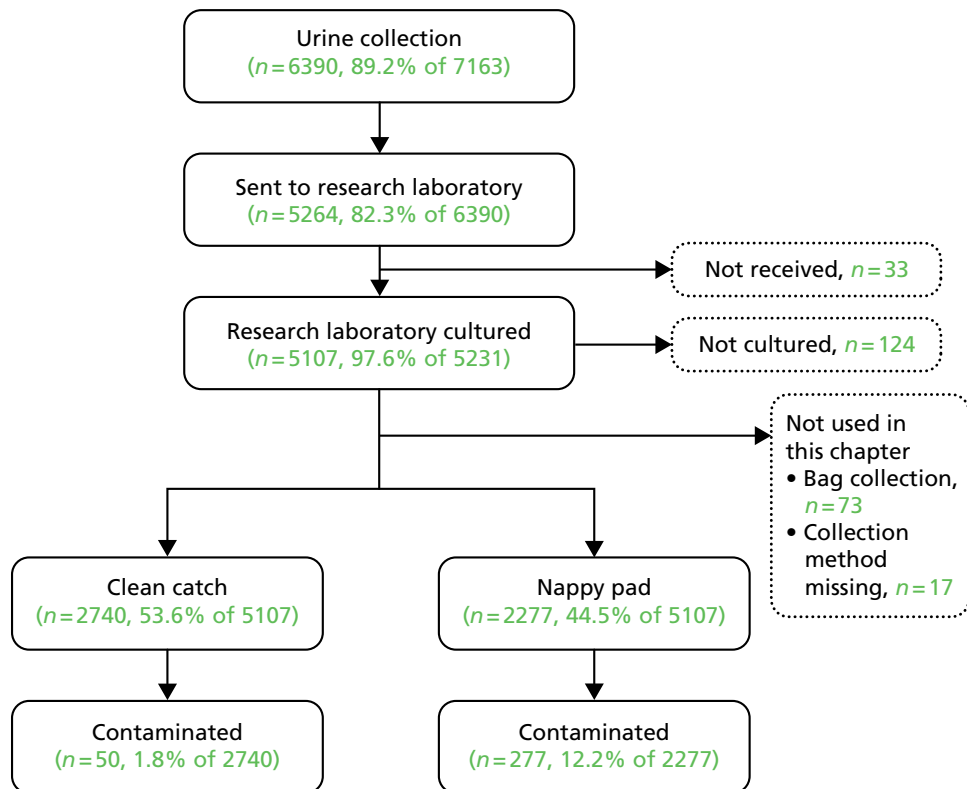
Finally, we compared the number of squamous epithelial cells (using 10 or more squamous epithelial cells as a cut-point) in contaminated and uncontaminated samples; and the number of urines with organisms [*E. coli*, enterococci and other Enterobacteriaceae (e.g. *Proteus*)] present in contaminated and UTI-positive urines. We looked at both the risk difference and risk ratio with their 95% CIs and *p*-values derived from a chi-squared test. These analyses were stratified by collection method.

## Results

Of the 7163 children in the study as a whole, we received urine from 6390. The local laboratories cultured 6079 samples and the research laboratory 5107 samples, with 4910 urines cultured at both laboratories (*Figure 29*). Thirty-three urines were not received at the research laboratory. Of the 5107 research laboratory results, 5017 were collected using either clean catch or nappy pads with 73 bag urines, and 17 with missing collection method were removed from the analysis due to the unsuitability of the collection method. This is why the numbers are not the same as in *Chapter 5*. Similar numbers of urines from each collection method were included: 55% clean catch and 45% nappy pad. *Figure 30* shows the flow of participants in this chapter which relates only to urines received and processed at the research laboratory.



**FIGURE 29** Effect of clinical suspicion of UTI on the number of days until symptoms improved among children with UTI.



**FIGURE 30** Participant flow in contamination chapter (research laboratory).

### Microbiological predictors of contamination

Table 67 shows that both clean-catch and nappy pad samples have increasing prevalence of contamination as the definition is relaxed according to the criteria outlined in Table 66 (although note that we restricted this to those definitions that could be reported by a NHS laboratory) and that, for all definitions used, the nappy pad contamination rate consistently exceeds that seen in the clean-catch samples.

Table 68 shows that there was a 1.8% chance that clean-catch urines were contaminated, compared with 12.2% in nappy pad urines. The risk was 10% (95% CI 8.9 to 11.8%) greater and the risk of contamination 6.7 times larger in nappy pad than in clean catch.

The most dominant organisms found in contaminated urines from 10 research laboratories are shown in Tables 69–71. To verify which organisms were associated with contamination, the prevalence of all organisms present in the contaminated urines were compared with those present in urines from patients considered to have a UTI from the same study (see Chapter 4). Tables are divided by collection method of urine to evaluate any difference between clean catch and nappy pad.

*E. coli* was present in 85% out of 86.7% of clean-catch/nappy pad urines from patients diagnosed with a UTI, compared with 76% out of 84.1% in contaminated urines (see Tables 69 and 70). The numbers of

**TABLE 67** Numbers of contaminated urines by different definition (research laboratory)

Contamination definition	Clean-catch contaminated, <i>n</i> (%)	Nappy pad contaminated, <i>n</i> (%)
≥ 10 <sup>5</sup> > 2 organisms <sup>a</sup>	50 (1.8)	277 (12.2)
≥ 10 <sup>5</sup> ≥ 2 organisms	78 (2.9)	426 (18.7)
≥ 10 <sup>4</sup> > 2 organisms	175 (6.4)	599 (26.3)
Total	2740	2277

a Definition used for subsequent analyses.

**TABLE 68** Number of contaminated urines by collection method (clean catch/nappy pad)

Clean catch, <i>n/N</i> (%)	Nappy pad, <i>n/N</i> (%)	Risk difference (95% CI)	Risk ratio (95% CI)	<i>p</i> -value
50/2740 (1.8)	277/2277 (12.2)	0.10 (0.08 to 0.11)	6.66 (4.95 to 8.96)	< 0.001

**TABLE 69** Organisms prevalent in clean-catch contaminated urines compared with UTI-positive urines

Organism identity	UTI positive, <i>n/N</i> (%)	Contaminated, <i>n/N</i> (%)	Risk difference (95% CI)	Risk ratio (95% CI)	<i>p</i> -value
<i>E. coli</i>	51/60 (85.0%)	38/50 (76.0%)	−0.090 (−0.239 to 0.059)	0.894 (0.740 to 1.080)	0.232
<i>Enterococcus</i> species	14/60 (23.3%)	41/50 (82.0%)	0.587 (0.436 to 0.738)	3.514 (2.182 to 5.661)	< 0.001
KES <sup>a</sup>	4/60 (6.7%)	26/50 (52.0%)	0.453 (0.301 to 0.606)	7.800 (2.917 to 20.855)	< 0.001
<i>Proteus</i> species	6/60 (10.0%)	7/50 (14.0%)	0.040 (−0.083 to 0.163)	1.400 (0.503 to 3.897)	0.518

a *Klebsiella* species, *Enterobacter* species and *Serratia* species.

**TABLE 70** Organisms prevalent in nappy pad contaminated urines compared with UTI-positive urines

Organism identity	UTI positive, n/N (%)	Contaminated, n/N (%)	Risk difference (95% CI)	Risk ratio (95% CI)	p-value
<i>E. coli</i>	26/30 (86.7%)	233/277 (84.1%)	-0.026 (-0.155 to 0.104)	0.971 (0.836 to 1.127)	0.715
<i>Enterococcus</i> species	16/30 (53.3%)	255/277 (92.1%)	0.387 (0.206 to 0.569)	1.726 (1.233 to 2.417)	<0.001
KES <sup>a</sup>	4/30 (13.3%)	114/277 (41.2%)	0.278 (0.143 to 0.413)	3.087 (1.226 to 7.769)	0.003

<sup>a</sup> *Klebsiella* species, *Enterobacter* species and *Serratia* species.

**TABLE 71** Organisms prevalent in contaminated urines

Organism identity	Clean catch, n/N (%)	Nappy pad, n/N (%)	Risk difference (95% CI)	Risk ratio (95% CI)	p-value
Coagulase-negative staphylococci	20/50 (40.0%)	196/277 (70.8%)	0.308 (0.12 to 0.45)	1.769 (1.24 to 2.50)	<0.001
<i>E. coli</i>	38/50 (76.0%)	233/277 (84.1%)	0.081 (-0.04 to 0.20)	1.107 (0.93 to 1.30)	0.161
<i>Enterococcus</i> species	41/50 (82.0%)	255/277 (92.1%)	0.101 (-0.01 to 0.22)	1.123 (0.98 to 1.28)	0.025

urines with *E. coli* present were similar for UTI-positive and contaminated samples in both clean-catch and nappy pad urines (with p-values of 0.232 and 0.715, respectively). *Proteus* species was also present at similar levels in clean catch (p-value of 0.518) from patients diagnosed with a UTI and contaminated urines. However, looking at other Enterobacteriaceae and *Enterococcus* species showed evidence of a difference between UTI-positive and contaminated samples in both clean-catch and nappy pad samples, that is 6.7% of clean-catch and 13.3% of nappy pad urines from patients diagnosed with a UTI contained one of three other Enterobacteriaceae (*Klebsiella*, *Enterobacter* and *Serratia* species), compared with 52% and 41.1% in contaminated urines, respectively. Enterococci were present in 23.3% of clean-catch and 53.3% of nappy pad urines from patients diagnosed with a UTI, compared with 82% and 92.1% in contaminated urines, respectively.

In order to determine if any particular organism was associated with contamination in nappy pad samples, the most dominant organisms present in the contaminated urines were compared according to collection method.

Nappy pad urines are 30.8% more likely to be contaminated with coagulase-negative staphylococci than clean catch (risk difference) or, equivalently, the risk was 1.8 times higher in nappy pad urine. In both nappy pad and clean-catch urines, contamination caused by *E. coli* was similar, with a p-value of 0.161. However, there is some evidence to suggest that enterococci are more likely to be present in contaminated nappy pad samples, 92.1%, than in contaminated clean-catch samples, 82% (p-value of 0.025) (see Table 71).

Squamous epithelial cells are routinely used in the laboratory as a guide to contamination. The numbers of contaminated and all other urines from each collection method to exhibit  $\geq 10$  squamous epithelial cells present are shown in Table 72.

In clean-catch urines, the number of urines with  $\geq 10$  squamous epithelial cells was significantly higher in contaminated (10%) than in non-contaminated urines (2.7%). However, in nappy pad urines the levels of squamous epithelial cells were similar for contaminated (10.5%) and non-contaminated (9%) urines.

**TABLE 72** Numbers of urines from nappy pad and clean-catch collection methods with  $\geq 10$  SECs

Urine collection method	Contaminated/ non-contaminated urine	Number of urines with $\geq 10$ SECs, n/N (%)	Risk difference (95% CI)	Risk ratio (95% CI)	p-value
Clean catch	Contaminated	5/50 (10)	0.073 (-0.010 to 0.156)	3.701 (1.563 to 8.766)	0.002
	Non-contaminated	72/2690 (2.7)			
Nappy pad	Contaminated	29/277 (10.5)	0.011 (-0.028 to 0.050)	1.112 (0.768 to 1.612)	0.575
	Non-contaminated	180/2000 (9)			

SEC, squamous epithelial cell.

### Clinical predictors of contamination

Univariable associations with contamination were analysed in groups, laboratory and clinical (faecal and skin) variables, to link potential contamination sources and variables associated with the DUTY algorithm identified in *Chapter 5*.

*Table 73* shows the variables associated with contamination in clean-catch samples. Other variables analysed, but not found to be associated, were all other dipstick results, urine colour, odour and opacity, presence of nurse during collection, urine transport time, laxative use, constipation, presence of diarrhoea, circumcision, dehydration, number of baths/showers per week, nappy rash (important as this was shown to be inversely associated with UTI in the nappy pad samples), presence of cough, global impression of child by clinician, pain/crying on passing urine, abdominal tenderness, ear abnormality and history of UTI.

In clean-catch urines, a positive dipstick test for nitrites, collection of urine at home, sex of child (female) and a higher number of nappies used in the previous 24 hours were predictors of contamination, the last exhibiting a consistent gradient.

In nappy pad urines, a positive dipstick test for leucocytes, higher pH, protein and blood, fewer hours taken in transport and sex (male) were predictors of contamination (*Table 74*).

**TABLE 73** Predictors of contamination in clean-catch urines

Predictive criterion	Category	Number in contaminated/ all urines, n/N (%)	Crude OR (95% CI) <sup>a</sup>	Adjusted OR (95% CI) <sup>a</sup>
Dipstick: nitrites	Negative	44/2658 (1.7)	1 (ref.)	1 (ref.)
	Positive	6/74 (8.1)	5.24 (2.16 to 12.72)	3.73 (1.48 to 9.40)
Location of sample collection?	Surgery	25/2103 (1.2)	1 (ref.)	1 (ref.)
	Home	25/637 (3.9)	3.40 (1.94 to 5.95)	2.88 (1.62 to 5.11)
Gender	Male	10/1267 (0.8)	1 (ref.)	1 (ref.)
	Female	40/1473 (2.7)	3.51 (1.75 to 7.05)	3.72 (1.83 to 7.56)
How many nappies has the child used in the past 24 hours?	Missing	10/1552 (0.6)	1 (ref.)	1 (ref.)
	1–4	28/963 (2.9)	4.62 (2.23 to 9.55)	4.55 (2.19 to 9.48)
	5–9	8/181 (4.4)	7.13 (2.78 to 18.31)	7.06 (2.70 to 18.47)
	$\geq 10$	4/40 (10.0)	17.13 (5.13 to 57.21)	18.11 (5.22 to 62.82)

Ref., reference.

<sup>a</sup> Crude OR and adjusted OR are from univariable models and a single multivariable model, respectively.

TABLE 74 Predictors of contamination in nappy pad urines

Variable	Category	Number in contaminated/ all urines, n/N (%)	Crude OR (95% CI)	Adjusted OR (95% CI)
Dipstick: leucocytes	Negative	193/1759 (11.0)	1 (ref.)	1 (ref.)
	Trace	17/125 (13.6)	1.28 (0.75 to 2.18)	1.13 (0.65 to 1.97)
	+	15/119 (12.6)	1.17 (0.67 to 2.05)	1.16 (0.64 to 2.10)
	++	36/177 (20.3)	2.07 (1.40 to 3.08)	2.16 (1.38 to 3.38)
	+++	16/91 (17.6)	1.73 (0.99 to 3.03)	1.53 (0.82 to 2.84)
Dipstick: pH	5	60/580 (10.3)	1 (ref.)	0.99 (0.70 to 1.42)
	6	87/824 (10.6)	1.02 (0.72,1.45)	1 (ref.)
	6.5	53/364 (14.6)	1.48 (0.99 to 2.19)	1.48 (1.02 to 2.16)
	7	27/203 (13.3)	1.33 (0.82 to 2.16)	1.20 (0.75 to 1.93)
	7.5	21/149 (14.1)	1.42 (0.83 to 2.42)	1.26 (0.75 to 2.13)
	8	16/95 (16.8)	1.76 (0.96 to 3.20)	1.59 (0.87 to 2.88)
	8.5	13/56 (23.2)	2.62 (1.33 to 5.15)	2.18 (1.10 to 4.31)
Dipstick: protein	Negative	180/1649 (10.9)	1 (ref.)	1 (ref.)
	Positive	97/622 (15.6)	1.51 (1.16 to 1.97)	1.24 (0.93 to 1.66)
Dipstick: blood	Negative	236/2060 (11.5)	1 (ref.)	1 (ref.)
	Positive	41/211 (19.4)	1.86 (1.29 to 2.69)	1.44 (0.95 to 2.17)
Dipstick: nitrites	Negative	214/1916 (11.2)	1 (ref.)	1 (ref.)
	Positive	63/355 (17.7)	1.72 (1.26 to 2.33)	1.55 (1.12 to 2.15)
Hours from urine sample taken to arriving at the research laboratory	< 24 hours	112/652 (17.2)	1 (ref.)	1 (ref.)
	24 to < 48 hours	107/816 (13.1)	0.73 (0.55 to 0.97)	0.69 (0.52 to 0.93)
	48 to < 72 hours	30/347 (8.6)	0.46 (0.30 to 0.70)	0.43 (0.28 to 0.66)
	72 to < 96 hours	18/252 (7.1)	0.37 (0.22 to 0.62)	0.36 (0.21 to 0.61)
	≥ 96 hours	10/210 (4.8)	0.24 (0.12 to 0.47)	0.25 (0.13 to 0.48)
Gender	Male	166/1183 (14.0)	1 (ref.)	1 (ref.)
	Female	111/1094 (10.1)	0.69 (0.54 to 0.89)	0.58 (0.43 to 0.76)

Ref., reference.



## Discussion

### Summary of main results

Using a definition selected to reflect 'frank contamination', this study showed a 10% greater risk of contamination when urine was collected using nappy pads than with a clean-catch sample. Depending on the definition selected, contamination was present in as many as 26% of nappy pad samples, compared with 6% for clean catch. This corroborates the findings in *Chapters 4 and 5* in addition to other reports, stating that a clean-catch sample is preferred for diagnosing UTI in young children. The two most common sources of contamination in urine from young children are faeces and skin. *E. coli* and enterococci are common faecal organisms and as such would be represented highly in contaminated urines; however, *E. coli* is also the most common cause of UTI. In this study, as expected, *E. coli* was not exclusively associated with contamination of urine, but *E. coli* was dominant in urines from patients diagnosed with a UTI. Enterococci were more prevalent in contaminated urines than UTI positive urines. Contamination of the urine with both *E. coli* and enterococci is almost certainly via faeces. In contrast, coagulase-negative staphylococci, part of normal skin flora, were found to be a dominant contaminant. The data also show no association between presence of these faecal organisms and nappy pad use, which is corroborated in the lack of association by univariate analysis of presence of diarrhoea.

Contamination of urine can happen with all collection methods in the young, but nappy pads have been implicated with increased contamination rates.<sup>171</sup> Only coagulase-negative staphylococci showed a significant association with the nappy pad collection method. Coagulase-negative staphylococci are considered skin contaminants and the close association with nappy pad reflects the close contact between nappy pads and skin. The clean-catch samples showed an increasing risk of contamination with use of an increasing number of nappies per day, but there was no relationship with nappy rash. Relaxing the criteria for contamination would increase the number of potentially contaminated samples, but we chose to base our analysis on a conservative definition.

Urinalysis by dipstick test is commonly used in primary care to inform diagnosis of UTI. This study found that in both clean-catch and nappy pad urines the variables showing association with contaminated urines are NOT exclusive to contaminated urines. Thus, they are not reliable indicators of when urines are contaminated. A positive leucocyte test is generally considered to be an indicator of pyuria. Reports have shown that pyuria can be present in feverish children without UTI<sup>172</sup> and so this is probably not a good predictor of contamination. The pH of urine is dependent upon the patient's acid-base status, but bacteria present in the urine can raise pH, as can a recent meal.<sup>173</sup> A positive blood dipstick result was associated with contaminated urines, a finding reported previously.<sup>174</sup> Similarly, positive protein dipstick results have been shown to be caused by bacteriuria and fever in children.<sup>173</sup> A positive nitrite dipstick test is used to detect bacteriuria by common pathogens known to cause UTI. In this study, high numbers of squamous epithelial cells were an indicator of contamination only in clean-catch samples. Squamous epithelial cells should be ignored if samples are derived from nappy pads.

The study considered if the time taken for the urine sample to reach the laboratory was a factor in contamination. Probability of contamination was not increased by increased time taken to arrive at the laboratory, but home sampling was a risk for contamination with clean-catch samples, perhaps reflecting less controlled collection technique.

### Comparison with existing literature

Contamination rates in collection methods vary, as indeed do the definitions of contamination. Contamination rates have been shown to vary from 0% in clean-catch urine of 23 samples to 48% of bag urines.<sup>98</sup> In a retrospective observational cohort study, contamination in clean catch, catheter specimen of urine and bag was 1%, 12% and 26%.<sup>98</sup> Definitions of contamination vary from single organism growth  $< 10^5$  OR  $\geq 2$  organisms to  $\geq 2$  organisms present at  $> 10^5$  CFU/ml of urine.<sup>89,91</sup>

In a US study, no institutional factors, such as access to refrigeration, were found to associate with either low or high contamination rates.<sup>99</sup> Similarly, sex and diarrhoeal symptoms have been shown to have no association with higher contamination rates.<sup>100</sup> Perineal cleansing in female adults had no association with contamination rates, while urine contamination rates were higher in midstream urine collected from toilet-trained children when obtained without perineal/genital cleaning.<sup>89,94</sup> One of the only factors shown to reduce contamination rates published was changing nappy pads every 30 minutes.<sup>91</sup>

### *Clinical implications*

To reduce the risk of contamination, urine samples should ideally be clean catch, taken in the GP surgery and not taken from children where there has been more than usual nappy use within 24 hours of presenting at surgery. Further, our data suggest that contamination is more likely in clean-catch urines if the child is female, the sample was taken at home as opposed to the GP surgery and if the child had increased nappy use 24 hours prior to presenting at the surgery. These factors may be useful in addition to microbiological criteria in clinical decisions.

## Chapter 8 Other results

### Introduction

In this chapter, we describe a number of other results from the DUTY study which have not been covered in the preceding chapters, including the revisions to objective 5 as detailed in *Chapter 1* (see *Planned change to research objective 5*). This introduction describes which results we will present and briefly summarises the methods used to generate these results.

Detailed information (rates and 95% CIs) is presented on the organisms cultured and antibiotic sensitivities for those individuals with a UTI in both the local and the research laboratories.

As in *Chapter 5*, we investigated if there are different illness trajectories according to clinical suspicion of UTI. Children with UTIs and contaminated samples who were or were not suspected of having a UTI by the clinician were compared for symptom improvement, symptom resolution (at 2 weeks) and numbers of NHS consultations (at 3 months) using chi-squared tests for categorical data (and, where there were insufficient data, Fisher's exact tests), Mann–Whitney *U*-tests for non-parametric comparisons and log-rank (Mantel–Cox) tests for survival data.

The main development of the clinical prediction rule was focused on children for whom a result was available at the research laboratory. In order to consider the generalisability of these results, we compare those with and without a research laboratory result in more detail than that presented in *Chapter 5* (see *Results, Participants*). Those without a research laboratory result may have not produced a urine sample, may have only provided enough urine for a sample to have been sent to the NHS laboratory or may have had a urine sample sent to the research laboratory which could not be cultured. Those with and without a result are tested for association with a chi-squared test, and a Mantel–Haenszel test for trend where the chi-squared test is significant.

We also considered if any practice level or demographic factors were associated with UTI prevalence. To do this, a two-level logistic regression model was fitted to model if patient demographics such as deprivation (from postcode) were associated with UTI and also to assess the degree to which UTI status was clustered by practice. This was done separately for samples collected by clean catch and nappy pad and is presented in addition to the prevalence and 95% CI.<sup>175</sup> Finally, this chapter reports the potential for incorporation bias.

### Microbiological results (sensitivities and susceptibilities)

*Table 75* presents the sensitivities of the research laboratory-cultured Enterobacteriaceae that were considered to be causing a UTI. Of these isolates, 79 (84.0%; 95% CI 75.3% to 90.1%) were *E. coli*, of which 50.6% (95% CI 39.8% to 61.4%) were sensitive to amoxicillin; 89.9% (95% CI 81.3% to 94.8%) were sensitive to co-amoxiclav [at a 2 : 1 ratio breakpoint (BP) with BP8; 93.6% (95% CI 86.0% to 97.3%) at BP32], 83.5% (95% CI 75.9% to 90.1%) were sensitive to co-amoxiclav [at a fixed concentration of 2 mg/l with BP8; 94.9% (95% CI 87.7% to 98.0%) at BP32], 98.7% (95% CI 93.2% to 99.8%) were sensitive to cephalexin, 70.9% (95% CI 60.1% to 79.8%) were sensitive to trimethoprim, 100.0% (95% CI 95.4% to 100.0%) were sensitive to nitrofurantoin and 96.2% (95% CI 89.4% to 98.7%) were sensitive to ciprofloxacin. The range is that of minimum inhibitory concentration (MIC) values within the test population, so, for example, for the 79 *E. coli* tested against amoxicillin, there were one or more isolates that had a MIC of 0.5 mg/l and one or more that had a MIC of > 128 mg/l, and all of the other isolates had MICs in between. It is a standard measure that we report to give an indication of the population susceptibility that is not determined by the BP or expert rules.

TABLE 75 Antimicrobial sensitivities for the 94 organisms considered positive for UTI at the research laboratory

Species	Amoxicillin		Co-amoxiclav 2:1 ratio		Co-amoxiclav fixed 2 mg/l		Cephalexin		Trimethoprim		Nitrofurantoin		Ciprofloxacin			
	Range <sup>a</sup>	Sensitive (%)	Range <sup>a</sup>	Sensitive (%) (BP8) <sup>b</sup>	Range <sup>a</sup>	Sensitive (%) (BP8) <sup>b</sup>	Range <sup>a</sup>	Sensitive (%) (BP32) <sup>c</sup>	Range <sup>a</sup>	Sensitive (%)	Range <sup>a</sup>	Sensitive (%)	Range <sup>a</sup>	Sensitive (%)		
<i>Citrobacter</i> spp. (n=1)	> 128	0.0	4	0.0 <sup>d</sup>	4	0.0 <sup>a</sup>	4	0.0 <sup>d</sup>	4	0.0 <sup>d</sup>	0.5	100.0	16	100.0	0.008	100.0
<i>E. coli</i> (n=79)	0.5 to >128	50.6	0.5 to 16	89.9	100.0	83.5	94.9	2 to 32	98.7	0.125 to >128	70.9	1 to 16	100.0	<0.008 to 16	96.2	
<i>Enterobacter</i> spp. (n=4)	16 to >128	0.0	2 to 64	0.0 <sup>d</sup>	1 to >128	0.0 <sup>d</sup>	0.0 <sup>d</sup>	4 to 64	0.0 <sup>a</sup>	0.25 to 1	100.0	4 to 16	100.0	<0.008 to 0.03	100.0	
<i>Klebsiella</i> spp. (n=4)	32 to >128	0.0	1 to 4	100.0	100.0	100.0	100.0	4	100.0	0.5 to -1	100.0	4 to 32	100.0	0.015 to 0.03	100.0	
<i>Morganella</i> spp. (n=1)	2	0.0 <sup>d</sup>	2	0.0 <sup>d</sup>	2	0.0 <sup>d</sup>	0.0 <sup>d</sup>	16	0.0 <sup>d</sup>	4	0.0	32	0.0 <sup>d</sup>	<0.008	100.0	
<i>Proteus</i> spp. (n=5)	0.5 to >128	80.0	0.5 to 4	100.0	100.0	100.0	100.0	8 to 16	100.0	2 to >128	60.0	32 to 64	0.0 <sup>d</sup>	0.03 to 0.06	100.0	
All	0.5 to >128	46.8 <sup>d</sup>	0.5 to 64	85.1 <sup>d</sup>	93.6 <sup>d</sup>	79.8 <sup>d</sup>	89.4 <sup>d</sup>	2 to 64	92.6 <sup>a</sup>	0.125 to >128	72.3	1 to 64	93.6 <sup>d</sup>	<0.008 to 16	96.8	

a Ranges expressed as MIC range (mg/l).

b Per cent sensitivity using BSAC/EUCAST systemic BP of S ≤ 8 mg/l.

c Per cent sensitivity using BSAC urinary BP of S ≤ 32 mg/l.

d Per cent sensitivity edited according to EUCAST expert rules.

Table 76 presents the NHS laboratory sensitivities of those same 94 that were considered to be causing a UTI in the research laboratory. In the NHS laboratory, 55 of these isolates were *E. coli*, of which 51.5% (95% CI 35.2% to 67.5%) were sensitive to amoxicillin, 91.1% (95% CI 79.3% to 96.5%) were sensitive to co-amoxiclav, 100.0% (95% CI 88.3% to 100.0) were sensitive to cephalexin, 88.9% (95% CI 76.5% to 95.2%) were sensitive to trimethoprim, 96.3% (95% CI 87.5% to 99.0%) were sensitive to nitrofurantoin and 93.3% (95% CI 82.1% to 97.7%) were sensitive to ciprofloxacin. Local laboratories only assessed sensitivities to antibiotics within their usual practice and, therefore, this is not a comprehensive set of results.

## Follow-up outcomes

Of the 94 children with a UTI, 57 (60.6%) were suspected as having a UTI by the GP. GPs also suspected that 38 (10.9%) of the children whose sample was contaminated had a UTI. The definition we used for contamination here is as described in Chapter 7, that is using the research laboratory results, with a criterion of count  $\geq 10^5$  and  $> 2$  organisms, equivalent to Jackson *et al.*'s<sup>93</sup> definition of frankly contaminated. In this chapter, all of the different urine collection methods are included, and so the number differs from that in Chapter 7. A similar proportion of children in all groups, regardless of GP suspicion of UTI or actual urine culture result (positive or contaminated), had improved by the time of the 2-week interview (92.9% vs. 85.7% and 88.2% vs. 79.2%; Table 77). The median number of days to improvement was similar in all groups (3 to 4 days).

The group with the highest proportion of children who had fully recovered by the time of the 2-week interview was children with a contaminated culture result who had been suspected of having a UTI by clinicians (85.7% vs. 72.3% and 79.4% vs. 75.0%; see Table 77), but there were no statistically significant differences between groups. Median recovery time was longest (9.5 days) for children who had a UTI that was not suspected by clinicians. However, there was no statistically significant difference between this group and those suspected of UTI ( $p = 0.464$ ).

Of those with UTI, antibiotics were prescribed at the index visit in 73.7% of children in whom clinicians suspected UTI, compared with 56.8% of those in whom UTI was not suspected ( $p = 0.09$ ). Trimethoprim or nitrofurantoin were prescribed at the index visit in 45.6% of children with UTI in whom UTI was suspected and also in 14.0% during the subsequent 2 weeks. Trimethoprim or nitrofurantoin were significantly more likely to be prescribed at the index visit if clinicians suspected UTI in both UTI and contaminated urine culture groups ( $p < 0.001$  and  $p < 0.001$ ). Children with a contaminated culture result, in whom clinicians suspected UTI, were also more likely to be prescribed trimethoprim or nitrofurantoin during the 2 weeks following the index visit (19.2% vs. 1.3%;  $p < 0.001$ ).

Survival curves are presented (see Figure 29 and Figures 31–33) for the comparisons made in Table 77 regarding symptom improvement and patient recovery. We found no evidence of differences but the  $p$ -values show that these are not significant, although the power for these comparisons is not high.

Children who had a UTI were more likely to have further urine samples in the subsequent 3 months (19.1%) than children who had contaminated (3.9%) or negative or indeterminate (8.0%) urine cultures (Table 78).

There was no difference in the proportion of children prescribed antibiotics in the subsequent 3 months between the three groups (45.7% of those who had UTI, 47.8% of those who had contaminated urine cultures and 46.7% of those who had negative or indeterminate urine cultures). However, children who had a UTI were more likely to be prescribed trimethoprim or nitrofurantoin in the subsequent 3 months (10.6%) than children who had a contaminated urine culture (2.7%) or those who had a negative or indeterminate urine culture (3.6%).

TABLE 76 NHS local laboratory antimicrobial sensitivities for the 94 organisms considered positive for UTI at the research laboratory

Species	Amoxicillin		Ampicillin		Cephalexin		Cephpodoxime		Cephhradine		Cephuroxime		Ciprofloxacin		Co-amoxiclav		Gentamicin		Nitrofurantoin		Trimethoprim		
	Sensitive Tests (%)	Tests	Sensitive Tests (%)	Tests	Sensitive Tests (%)	Tests	Sensitive Tests (%)	Tests	Sensitive Tests (%)	Tests	Sensitive Tests (%)	Tests	Sensitive Tests (%)	Tests	Sensitive Tests (%)	Tests	Sensitive Tests (%)	Tests	Sensitive Tests (%)	Tests	Sensitive Tests (%)	Tests	
Unknown coliform (n=14)	14.3	7	25.0	4	100.0	5	100.0	7	100.0	3	100.0	1	100.0	10	90.0	12	83.3	7	100.0	13	92.3	14	85.7
<i>E. coli</i> (n=55)	33	51.5	8	75.0	29	100.0	28	100.0	12	100.0	12	100.0	45	93.3	45	91.1	24	100.0	54	96.3	55	88.9	
<i>Proteus</i> spp (n=4)	3	66.7	0	N/A	3	100.0	1	100.0	1	100.0	0	N/A	2	100.0	2	100.0	2	100.0	4	0.0	4	50.0	

N/A, not applicable.



TABLE 77 Early outcomes index visit and first 2 weeks (continued)

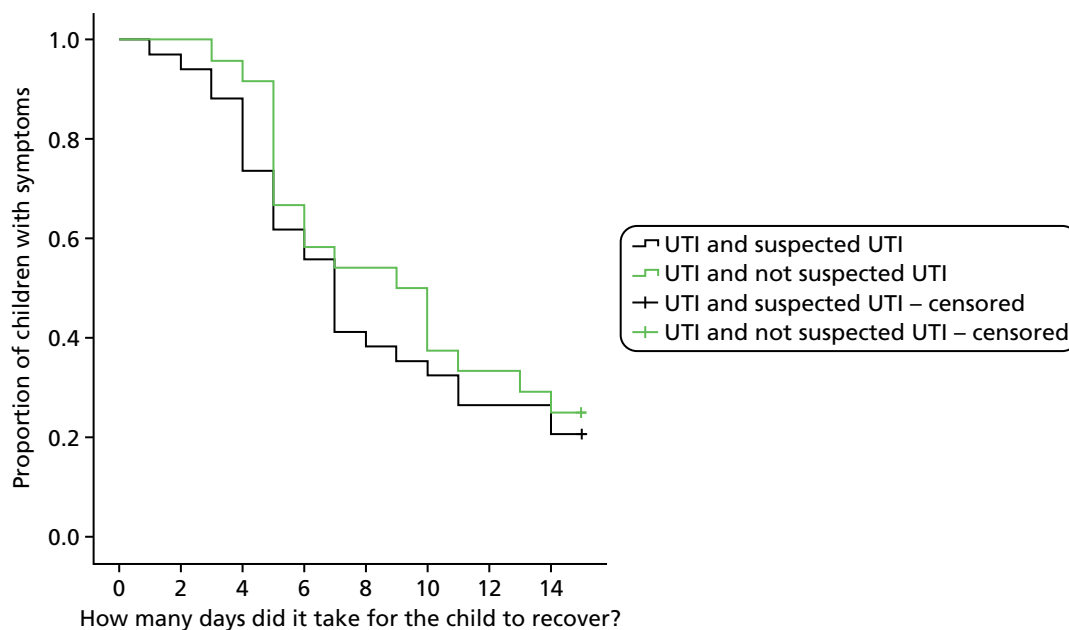
Variable	Group		p-value	Contaminated		p-value	Indeterminate and negative	
	UTI							
n in group with notes review at 3 months	56	37		26	229		979	
	n (%)	n (%)		n (%)	n (%)		n (%)	
Any antibiotics at the index visit? (Yes)	42 (73.7)	21 (56.8)	0.088	13 (50.0)	101 (44.1)	0.567	476 (48.6)	
Any trimethoprim or nitrofurantoin at the index visit? (Yes)	26 (45.6)	2 (5.4)	<0.001 <sup>a</sup>	7 (26.9)	3 (1.3)	<0.001 <sup>a</sup>	56 (5.7)	
Any urine tests in the first 2 weeks after the index visit? (Yes)	8 (14.0)	7 (18.9)	0.528	1 (3.8)	6 (2.6)	0.533 <sup>a</sup>	50 (5.1)	
Any antibiotics in the first 2 weeks after the index visit? (Yes)	17 (29.8)	15 (40.5)	0.284	7 (26.9)	48 (21.0)	0.484	229 (23.4)	
Any trimethoprim or nitrofurantoin in the first 2 weeks after the index visit? (Yes)	8 (14.0)	5 (13.5)	0.943	5 (19.2)	3 (1.3)	<0.001 <sup>a</sup>	47 (4.8)	
Any secondary care consultations in the first 2 weeks? (Yes)	5 (8.8)	3 (8.1)	1.000 <sup>a</sup>	1 (3.8)	10 (4.4)	1.000 <sup>a</sup>	70 (7.2)	

<sup>a</sup> Fisher's exact tests.

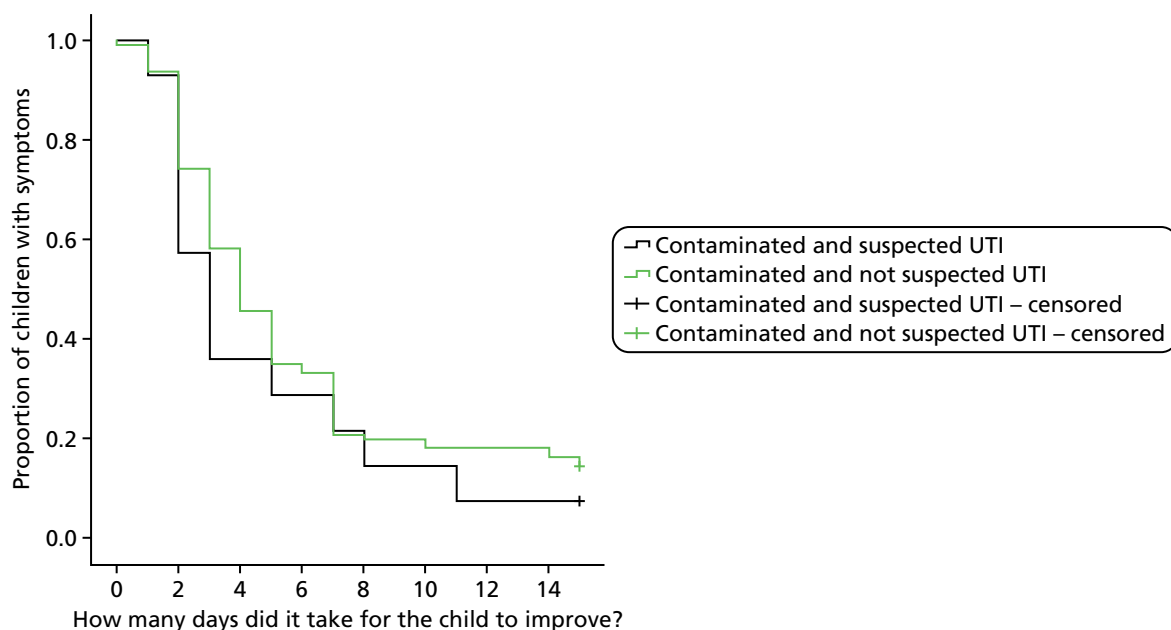
<sup>b</sup> Mann-Whitney U-tests.

<sup>c</sup> Log-rank (Mantel-Cox) tests.

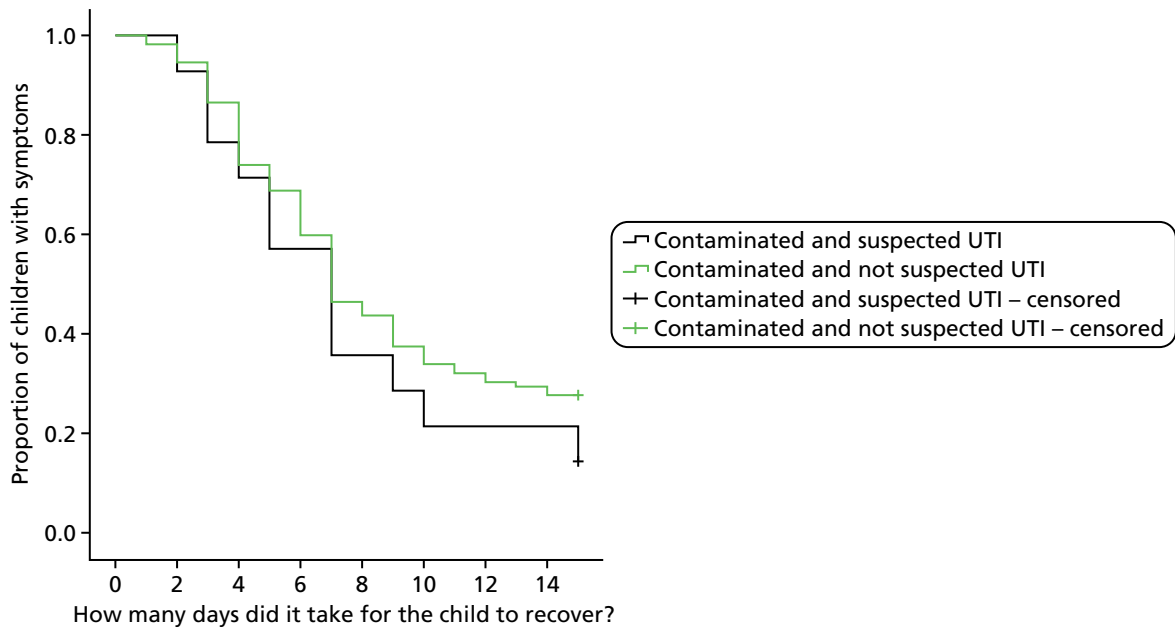




**FIGURE 31** Effect of clinical suspicion of UTI on the number of days until symptoms resolved among children with UTI.



**FIGURE 32** Effect of clinical suspicion of UTI on the number of days until symptoms improved among children with contaminated urine.



**FIGURE 33** Effect of clinical suspicion of UTI on the number of days until symptoms resolved among children with contaminated urine.

Trimethoprim or nitrofurantoin were significantly more likely to be prescribed in the 3 months following the index consultation if clinician had suspected UTI in both UTI and contaminated urine culture groups ( $p < 0.001$  and  $p < 0.001$ , respectively).

There was no difference in the median number of reconsultations in the subsequent 3 months between the groups. Slightly more children (25.5%) who had UTI were seen in secondary care than children with contaminated (19.2%) and negative or indeterminate (20.8%) urine cultures.

Seven (7.4%) children with UTI and four (1.6%) children with a contaminated urine culture had urine-related hospital tests (ultrasound scan, DMSA scan or MCUG) in the 3 months following the index consultation. Children in both groups were more likely to have received these tests if the clinician had suspected UTI, although the differences were not statistically significant ( $p = 0.040$  and  $p = 0.053$ , respectively).

## Generalisability and potential bias

### Verification bias

This section supplements data already presented in *Chapter 5* (see *Results, Participants*) and compares the characteristics of children with and without a culture result from the research laboratory (*Tables 79 and 80*). The statistical significance level for the nine tests performed here is set at a Bonferroni corrected  $\alpha = 0.05/9 = 0.005$ .

*Table 79* shows that as age increases, there is a higher likelihood of having a result ( $p < 0.001$ ). In addition, the test for trend gives a  $p < 0.001$  for trend. For clinician impression of illness severity, children with a result were more likely to have a mild illness as defined by the clinician than those without (82.5% vs. 79.9% clinician global impression  $\leq 3$ ;  $p = 0.001$ ), although this is a relatively small actual difference. Clinicians were more likely to suspect UTI in children in whom UTI status was known (6.8% vs. 4.1%;  $p < 0.001$ ). Children with results were similar in terms of sex (50.9% vs. 50.4% female), ethnicity (83.7 vs. 82.3% white), parental highest qualification (50.4% vs. 49.2% diploma, degree or equivalent), cost of living (11.8% vs. 10.4% quite comfortably off), deprivation (21.1% vs. 20.6% most deprived), parental

TABLE 78 Outcomes up to 3 months

Variable	Group		Contaminated		Indeterminate and negative	
	UTI		Suspected UTI		Not suspected UTI	
	Suspected UTI	Not suspected UTI	Suspected UTI	Not suspected UTI	Suspected UTI	Not suspected UTI
	n (%)	n (%)	n (%)	n (%)	p-value	p-value
n in group	57	37	38	310		4665
n in group with notes review at 3 months	56	37	26	229		979
Variable	n (%)	n (%)	n (%)	n (%)	p-value	n (%)
Any urine tests? (Yes)	18 (32.1)	13 (35.1)	2 (7.7)	14 (6.1)	0.764	122 (12.5)
Any urine tests after the first 2 weeks? (Yes)	11 (19.3)	7 (18.9)	2 (7.7)	8 (3.5)	0.964	78 (8.0)
Any secondary care consultations? (Yes)	14 (24.6)	13 (35.1)	3 (11.5)	53 (23.1)	0.268	243 (24.8)
Any secondary care consultations after the first 2 weeks? (Yes)	14 (24.6)	10 (27.0)	2 (7.7)	47 (20.5)	0.789	204 (20.8)
Any hospital tests? (Yes)	8 (14.0)	3 (8.1)	2 (7.7)	12 (5.2)	0.518 <sup>a</sup>	63 (6.4)
Any urine-related hospital tests? (Yes)	7 (12.3)	0 (0.0)	2 (7.7)	2 (0.9)	0.040 <sup>a</sup>	23 (2.3)
Any antibiotics? (Yes)	52 (91.2)	31 (83.8)	19 (73.1)	167 (72.9)	0.332 <sup>a</sup>	755 (77.1)
Any antibiotics after the first 2 weeks? (Yes)	25 (43.9)	18 (48.6)	8 (30.8)	114 (49.8)	0.649	457 (46.7)
Any trimethoprim or nitrofurantoin? (Yes)	37 (64.9)	8 (21.6)	13 (50.0)	11 (4.8)	<0.001	125 (12.8)
Any trimethoprim or nitrofurantoin after the first 2 weeks? (Yes)	8 (14.0)	2 (5.4)	2 (7.7)	5 (2.2)	0.185	35 (3.6)
Variable	Median	IQR	Median	IQR	p-value	Median
Number of reconsultations	1.0	2.0	1.0	2.0	0.515 <sup>b</sup>	1.0
IQR, interquartile range.						
<sup>a</sup> Fisher's exact tests.						
<sup>b</sup> Mann-Whitney U-tests.						
<sup>c</sup> Log-rank (Mantel-Cox) tests.						
Variable	Median	IQR	Median	IQR	p-value	Median
Number of reconsultations	1.0	2.0	1.0	2.0	0.101 <sup>b</sup>	1.0
IQR, interquartile range.						
<sup>a</sup> Fisher's exact tests.						
<sup>b</sup> Mann-Whitney U-tests.						
<sup>c</sup> Log-rank (Mantel-Cox) tests.						

**TABLE 79** Comparison of those with a UTI status with those without a UTI status

Criterion	Does the patient have a sample in the research laboratory from which a UTI result can be ascertained?		<i>p</i> -value (chi-squared test)
	No ( <i>N</i> = 2056), <i>n</i> (%)	Yes ( <i>N</i> = 5107), <i>n</i> (%)	
Age			
< 6 months	238 (11.6)	412 (8.1)	< 0.001
6 months to < 12 months	465 (22.6)	675 (13.2)	
1 year to < 2 years	676 (32.9)	1006 (19.7)	
2 years to < 3 years	359 (17.5)	989 (19.4)	
3 years to < 4 years	195 (9.5)	1138 (22.3)	
≥ 4 years	123 (6.0)	887 (17.4)	
Gender			
Male	1019 (49.6)	2507 (49.1)	0.717
Female	1037 (50.4)	2600 (50.9)	
Ethnicity groupings			
White	1660 (82.3)	4235 (83.7)	0.020
Mixed	110 (5.5)	261 (5.2)	
Asian	98 (4.9)	203 (4.0)	
Black	146 (7.2)	325 (6.4)	
Other	3 (0.1)	33 (0.7)	
Highest level of qualification			
Degree (or equivalent)	511 (33.3)	1318 (32.0)	0.838
Diploma (or equivalent)	262 (17.1)	709 (17.2)	
'A' level	202 (13.2)	540 (13.1)	
GCSE/'O' level	458 (29.9)	1252 (30.4)	
None	101 (6.6)	299 (7.3)	
Cost of living			
Find it a strain to get by week to week	89 (5.7)	292 (7.1)	0.141
Have to be careful about money	793 (50.6)	2101 (50.9)	
Able to manage without much difficulty	501 (32.0)	1305 (31.6)	
Quite comfortably off	185 (11.8)	429 (10.4)	
Deprivation quintile <sup>a</sup>			
1	359 (17.6)	908 (18.0)	0.869
2	377 (18.5)	962 (19.1)	
3	411 (20.2)	981 (19.4)	
4	471 (23.1)	1132 (22.4)	
5	419 (20.6)	1065 (21.1)	

TABLE 79 Comparison of those with a UTI status with those without a UTI status (continued)

Criterion	Does the patient have a sample in the research laboratory from which a UTI result can be ascertained?		p-value (chi-squared test)
	No (N = 2056), n (%)	Yes (N = 5107), n (%)	
Please rate your overall impression of your child's current illness when it is at its worst (parents' score)			
0	7 (0.3)	40 (0.8)	0.069
1	39 (1.9)	148 (2.9)	
2	83 (4.1)	256 (5.0)	
3	184 (9.0)	500 (9.8)	
4	282 (13.8)	659 (12.9)	
5	395 (19.3)	931 (18.3)	
6	371 (18.1)	879 (17.2)	
7	371 (18.1)	906 (17.8)	
8	223 (10.9)	537 (10.5)	
9	62 (3.0)	154 (3.0)	
10	30 (1.5)	91 (1.8)	
Please give your global impression of the child on a scale of 0–10 (clinicians' score)			
0	128 (6.4)	476 (9.3)	0.001
1	492 (24.4)	1333 (26.2)	
2	553 (27.5)	1413 (27.7)	
3	435 (21.6)	982 (19.3)	
4	203 (10.1)	472 (9.3)	
5	88 (4.4)	196 (3.8)	
6	67 (3.3)	125 (2.5)	
7	33 (1.6)	68 (1.3)	
8	10 (0.5)	25 (0.5)	
9	4 (0.2)	4 (0.1)	
10	0 (0.0)	2 (0.0)	
Clinician's working diagnosis?			
No UTI	1925 (95.8)	4748 (93.2)	<0.001
UTI	63 (3.1)	298 (5.9)	
UTI + another diagnosis	21 (1.0)	46 (0.9)	

a For deprivation, quintile 1 is least deprived and quintile 5 is most deprived.

**TABLE 80** Comparison of days unwell prior to consulting for those with a UTI status with those without a UTI status

	Does the patient have a sample in the research laboratory from which a UTI result can be ascertained?						<i>p</i> -value (Mann–Whitney <i>U</i> -test)
	No ( <i>N</i> = 2056)			Yes ( <i>N</i> = 5107)			
	<i>n</i>	Median	IQR	<i>n</i>	Median	IQR	
Days unwell	2047	4.0	5.0	5101	4.0	4.0	0.039

IQR, interquartile range.

impression of overall illness severity (18.5% vs. 15.3% score of  $\leq 3$ ) and the number of days unwell prior to recruitment (both have a median of 4.0 days; see *Table 80*). In summary, for age, clinician impression and working diagnosis, there were significant associations with whether or not UTI status was known. This suggests that the chance of gaining a sample which was large enough for splitting to go to the research lab and be processed was higher for older children, who were not as generally unwell and where UTI was suspected.

### **Incorporation bias**

Although not important to the development of the algorithm using the research laboratory, in order to assess how often the sending of clinical information to the local laboratory might be occurring in DUTY, a sample of 34 clinicians, a mix of GPs and nurses, using a mix of option 1 and 2 recruitment methods, were asked what, if any, clinical information they provided on the request form to the local NHS laboratory. The results of this showed that 10 (29.4%) did not provide any information and eight (23.5%) did mention the possibility of UTI. The extent to which laboratories use such clinical information is also likely to vary, with the large number of urine samples processed by many NHS laboratories prohibiting use of any clinical information in many.

### **Prevalence in urinary tract infection and its variation**

We had 5107 children with a known UTI status from the research laboratory, of whom 94 were UTI positive, giving an overall UTI prevalence of 1.8% (95% CI 1.5% to 2.3%). Of the 2740 samples collected via clean catch, 60 were positive, giving a prevalence of 2.2% (95% CI 1.7% to 2.8%) and of the 2277 collected via nappy pads, 30 were positive, giving a prevalence of 1.3% (95% CI 0.9% to 1.9%). We also had 73 samples collected via bags, of which three were positive, and 17 children (one positive) in whom the collection method was not recorded.

The children were recruited from 229 different sites. This gives an average of 22.3 children per site, varying between 1 and 337 children per site.

When considering only clean-catch collection method, the children were recruited from 224 different sites; on average there were 12.2 children per site, with the number varying between 1 and 169. We considered the degree of variation in UTI prevalence between practices using a two-level random-effect logistic regression model with practice/site as a random effect and no fixed effects. A LR test between this multilevel model and logistic regression showed no evidence of clustering ( $p = 1.000$ ) and hence we used logistic regression instead of a multilevel model in the analysis, which included recruitment centre, recruitment site, age, sex and deprivation. *Table 81* shows that there was no evidence of variation in UTI prevalence by recruitment centre or site. Prevalence varied by age (highest prevalence in 6- to 12-month-old children), though numbers were too small to be significant.

**TABLE 81** Variation in UTI prevalence (clean-catch samples)

Variable	Category	UTI-ve/UTI+ve, N/n	Prevalence	Crude OR (95% CI)	Adjusted OR (95% CI) <sup>a</sup>
Centre	Bristol	1157/25	2.1%	1 (ref.)	1 (ref.)
	Cardiff	593/12	2.0%	0.94 (0.47 to 1.88)	0.90 (0.44 to 1.82)
	London	499/13	2.5%	1.21 (0.61 to 2.38)	1.12 (0.53 to 2.37)
	Southampton	431/10	2.3%	1.07 (0.51 to 2.25)	1.00 (0.46 to 2.14)
Recruitment site	GP surgery	2452/55	2.2%	1 (ref.)	1 (ref.)
	ED	190/4	2.1%	0.94 (0.34 to 2.62)	0.60 (0.16 to 2.16)
	Walk-in centre	38/1	2.6%	1.17 (0.16 to 8.70)	1.18 (0.15 to 9.43)
Age (months)	0 to < 6	33/1	2.9%	1.51 (0.20 to 11.60)	1.86 (0.21 to 16.9)
	≥ 6 to < 12	51/3	5.6%	2.94 (0.85 to 10.17)	3.66 (0.90 to 14.8)
	≥ 12 to < 24	89/2	2.2%	1.12 (0.26 to 4.86)	1.41 (0.31 to 6.36)
	≥ 24 to < 36	599/16	2.6%	1.33 (0.69 to 2.58)	1.30 (0.67 to 2.52)
	≥ 36 to < 48	1049/21	2.0%	1 (ref.)	1 (ref.)
	≥ 48	859/17	1.9%	0.99 (0.52 to 1.89)	0.99 (0.52 to 1.89)
Sex	Male	1254/13	1.0%	0.31 (0.17 to 0.58)	0.32 (0.17 to 0.60)
	Female	1426/47	3.2%	1 (ref.)	1 (ref.)
Deprivation quintile <sup>b</sup>	1	504/6	1.2%	1 (ref.)	1 (ref.)
	2	514/12	2.3%	1.96 (0.73 to 5.26)	2.04 (0.76 to 5.52)
	3	517/15	2.8%	2.44 (0.94 to 6.33)	2.37 (0.91 to 6.25)
	4	604/18	2.9%	2.50 (0.99 to 6.35)	2.45 (0.93 to 6.46)
	5	506/9	1.8%	1.49 (0.53 to 4.23)	1.56 (0.53 to 4.59)

ED, emergency department; ref., reference.

<sup>a</sup> Ratio adjusted for factors in table only and hence  $n = 2705$  owing to missing deprivation quintiles.

<sup>b</sup> For deprivation, quintile 1 is least deprived and quintile 5 is most deprived.

As seen in *Chapter 5* (see *Results, Clean-catch models*), girls were more likely than boys to have a UTI, even after adjustment for the other model factors such as deprivation and recruitment centre. On the whole, children from more deprived areas had higher prevalence of UTI, though this association could have occurred by chance as there were wide CIs which included no association.

While considering the nappy pad collection method, the children were recruited from 188 different sites; on average there were 12.1 children per site, with the number varying between 1 and 165 children. We considered the degree of variation in prevalence between practices using a two-level random-effect logistic regression model with practice/site as a random effect and no fixed effects. A LR test between this multilevel model and logistic regression showed little evidence of clustering ( $p = 0.315$ ), and so we used logistic regression instead of a multilevel model for the following analyses, which included recruitment centre, recruitment site, age, sex and deprivation.

*Table 82* shows that there was no evidence of variation in UTI prevalence by recruitment centre (though the Bristol centre rate appears higher than the other centres) and that the number of UTI events was too small to compare sites. There was less variation by age, reflecting the more narrow age range of children providing nappy pad samples. Prevalence varied by age (most prevalent in 6- to 12-month-old children), though numbers were too small to be significant. As seen in *Chapter 5* (see *Results, Nappy pad models*), girls were more likely than boys to have a UTI, even after adjustment for the other model factors such as deprivation and recruitment centre. There did not appear to be as much UTI prevalence variation by deprivation as seen in the clean-catch samples.

**TABLE 82** Variation in UTI prevalence (nappy pad samples)

Variable	Category	UTI-ve/UTI+ve, N/n	Prevalence, %	Crude OR (95% CI)	Adjusted OR (95% CI) <sup>a</sup>
Centre	Bristol	1001/20	2.0	1 (ref.)	1 (ref.)
	Cardiff	615/5	0.8	0.41 (0.15 to 1.09)	0.44 (0.16 to 1.11)
	London	335/3	0.9	0.45 (0.13 to 1.52)	0.44 (0.10 to 1.35)
	Southampton	296/2	0.7	0.34 (0.08 to 1.45)	0.32 (0.07 to 1.35)
Recruitment site	GP surgery	2206/30	1.3	N/A <sup>b</sup>	N/A <sup>b</sup>
	ED	8/0	0	N/A	N/A
	Walk-in centre	33/0	0	N/A	N/A
Age (months)	0 to < 6	364/5	1.4	1.72 (0.54 to 5.46)	2.11 (0.66 to 6.79)
	≥ 6 to < 12	592/11	1.8	2.33 (0.90 to 6.04)	2.40 (0.92 to 6.28)
	≥ 12 to < 24	877/7	0.8	1 (ref.)	1 (ref.)
	≥ 24 <sup>c</sup>	414/7	1.6	2.11 (0.74 to 6.08)	2.03 (0.70 to 5.85)
Gender	Male	1174/9	0.8	0.39 (0.18 to 0.86)	0.39 (0.18 to 0.86)
	Female	1073/21	1.9	1 (ref.)	1 (ref.)
Deprivation quintile <sup>d</sup>	1	382/6	1.6	1 (ref.)	1 (ref.)
	2	423/4	0.9	0.60 (0.17 to 2.14)	0.65 (0.18 to 2.35)
	3	419/11	2.6	1.67 (0.61 to 4.56)	1.92 (0.69 to 5.33)
	4	481/5	1.0	0.66 (0.20 to 2.19)	0.81 (0.24 to 2.76)
	5	519/4	0.8	0.49 (0.14 to 1.75)	0.62 (0.17 to 2.30)
	Missing	23/0	0		

ED, emergency department; N/A, not applicable; ref., reference.

a Ratio adjusted for factors in table only and hence  $n = 2254$  owing to missing deprivation quintiles.

b Crude and adjusted ORs could not be calculated for this variable as all the UTI positives are within one category.

c The older age categories were merged (compared with earlier table) owing to there being no UTI positives over the age of 36 months in nappy pads.

d For deprivation, quintile 1 is least deprived and quintile 5 is most deprived.

## Discussion

The local and research laboratories report marked differences in the proportions of organisms grown at the various concentration categories. These are likely to represent differences in laboratory processes and reporting procedures. A range of antimicrobial susceptibilities are observed between the local and research laboratories, with nitrofurantoin sensitivity appearing to be higher than other commonly used antibiotics in primary care, such as trimethoprim and amoxicillin. No differences were observed in terms of symptom improvement or resolution at 2 weeks, or number of NHS consultations at 3 months in children with or without clinically suspected UTI at the index consultation. Unsurprisingly, clinicians were more likely to use a UTI-specific antibiotic (trimethoprim or nitrofurantoin) if they suspected that the child had a UTI. UTI prevalence did not vary by recruitment centre or with deprivation. There were variations by sex (female more common than male), which are in keeping with the literature.<sup>9</sup>

The main clinical implication of the results presented in this chapter is that, based on bacterial sensitivities, clinicians may wish to consider nitrofurantoin as the first-line antibiotic for suspected UTI in young children. There does seem to be little difference in recovery from this illness episode regardless of accurate diagnosis of UTI; however, without understanding the potential long-term sequelae of this self-limiting illness in young children, it is hard to draw firm clinical practice conclusions about this.



## Chapter 9 Discussion

### Summary of objectives and main results

The primary aim of the DUTY study was to derive and validate a clinical algorithm for the selection of children aged < 5 years presenting to primary care with an acute illness of up to 28 days' duration who warrant urine sampling. Additional aims were to identify the additional value of a point-of-care dipstick urine test; model cost-effectiveness of one or more diagnostic algorithm guided strategies; compare contamination rates for two urine sampling methods; compare the results obtained from local NHS laboratories with a single research laboratory; and explore the clinical significance of clinician-suspected UTI.

In total, 7163 children were recruited. Urine samples were obtained from 6390 (89%). Where there was sufficient urine, samples were split, with 6079 (95%) analysed in a NHS microbiology laboratory and 5107 (80%) analysed in a research laboratory. More than half of the samples analysed by the research laboratory were obtained by clean catch [ $n = 2740$  (55%)], with 94% of these samples from children aged  $\geq 2$  years. Samples were obtained by the Newcastle nappy pad method in 2277 (45%), and 82% of these samples were from children < 2 years.

We showed that the agreement of microbiological diagnosis of UTI in routine NHS laboratories and a research laboratory was lower than expected and worse for urine samples collected using nappy pads than for clean-catch samples. Associations of microbiological positivity with prespecified symptoms, signs and urine dipstick test results were lower for NHS laboratories than the research laboratory and for nappy pad samples than clean-catch samples. Urines giving a 'positive' result in a NHS laboratory but not in the research laboratory had only modest associations with the preselected symptoms, signs and dipstick test results. These findings did not appear to be attributable to the younger age of the children providing nappy pad samples.

We used urine culture results obtained from the research laboratory for our reference standard because samples were analysed in a single laboratory by a limited number of well-trained individuals using more detailed urine processing and reporting techniques than routinely available in NHS laboratories, and because we found that the research laboratory was both more reliable and more accurate than the NHS laboratories. We defined our reference standard as  $\geq 10^5$  CFU/ml of a single uropathogen ('pure growth') or  $\geq 10^5$  CFU/ml of a uropathogen with  $\geq 3 \log_{10}$  (1000-fold) difference between the growth of this and the next species ('predominant growth'). We defined uropathogens as members of the Enterobacteriaceae group for the purpose of our analyses.

Although our results suggest the research laboratory was providing more accurate and reliable results than the NHS laboratories, the reasons for this and the differences in UTI prevalence between research and NHS laboratories are not clear and could include differences in transport, laboratory methods and laboratory interpretation. For example, the more rigorous research laboratory methods (including sample processing by only two laboratory technicians) and UTI definition could have reduced UTI and increased contamination prevalence. Equally, the NHS laboratories could be classifying contaminated samples as UTI – the clinically safer policy where doubt exists.

The prevalence of ' $\geq 10^5$  CFU/ml pure/predominant', ' $\geq 10^5$  CFU/ml > 2 species' and 'no growth' was lower in the research laboratory than the local NHS laboratories, possibly reflecting the variable use of boric acid in samples transported to the NHS laboratories. Higher prevalence was seen in all categories with  $10^3$  to  $10^5$  CFU/ml colony counts, probably due to more accurate bacterial quantification in the research laboratory.

Of the samples analysed in the research laboratory, 2.2% and 1.3% of urine samples obtained by clean-catch and nappy pad methods, respectively, met criteria for a microbiological diagnosis of UTI. There was no evidence of variation in UTI prevalence by recruitment centre or site. Prevalence did vary by age (most prevalent in 6- to 12-month-old children), but this variation was not statistically significant. Girls were more likely than boys to have a UTI, even after adjustment for a range of variables. Regarding the primary objective of developing a diagnostic algorithm to identify the children warranting urine collection, we used bootstrapping to validate the algorithms, which were developed separately for the clean-catch and nappy pad samples. Two forms of algorithm were developed for each urine collection method: 'full' models, using the adjusted coefficients from index tests with all gradations of severity; and points-based models, using integer points based on the coefficients from models where index tests were dichotomised into present or absent. We found that history of UTI, pain/crying on passing urine, smelly urine, absence of cough, more severe clinician global impression of illness, abdominal tenderness on examination and normal ear examination were independently associated with UTI in the clean-catch full model. This gave a validated AUROC of 0.876. UTI in nappy pad samples was positively associated with female sex, smelly urine, darker urine and absence of nappy rash, giving a full model validated AUROC of 0.778.

Both the clean-catch and nappy pad algorithms performed better than clinicians' judgement in predicting the likelihood of a UTI on urine culture. The diagnostic utility increased with the addition of urine dipstick testing in the clean-catch and, to a lesser degree, in the nappy pad full models, giving validated AUROCs ( $p$ -value for change from symptoms and signs full models) of 0.903 ( $p = 0.009$ ) and 0.821 ( $p = 0.036$ ), respectively. In the clean-catch samples, we were also able to demonstrate that the added value of dipstick testing increased with increasing UTI probability post symptoms and signs. We therefore found that parent-reported symptoms and clinical signs could be useful in identifying the preschool children presenting to primary care who should have their urine sampled and that dipstick tests provide additional diagnostic information to improve the targeting of empirical antibiotics. Diagnostic utility was better for clean-catch samples, but nappy pad samples did provide better diagnostic accuracy than clinical assessment alone for children < 2 years.

With and without the dipstick tests, the diagnostic utility of the points-based models were only marginally lower than the full models, although the clean-catch points-based models continued to outperform the nappy pad points-based models.

Our health economic analyses showed that for older children, the symptoms and signs clean-catch (full) model would result in fewer urine samples, with equivalent or higher sensitivity and specificity and marginally better costs and patient outcomes than clinical judgement in the short term and long term.

The distinction in cost-effectiveness between clinical judgement and the symptom and sign nappy pad full model for younger children was not clear-cut, underlining the importance of obtaining a clean-catch sample whenever practical. Although the points-based symptoms and signs algorithms were found to be less accurate than the full model algorithm, they had a higher probability of being cost-effective than clinical judgement alone.

The choice of a threshold on the DUTY risk score which will optimally identify children whose risk of UTI is high enough to justify collecting a urine sample is not straightforward. We used decision-analytic models to explore three possible thresholds which represented a range of trade-offs between high sensitivity and high specificity. The results slightly favoured low-cost, high-specificity urine sampling strategies, particularly for GPs concerned about the societal impact of antibacterial resistance. The absolute difference in average short- and long-term net benefits between the three DUTY thresholds evaluated was very small, but likely to be important for the NHS in aggregate owing to the implications for the volume of urine testing.

The additional diagnostic information provided by dipstick tests has the potential to play a role in the efficient diagnosis and treatment of children with suspected UTI, which should be explored in further research. In children with intermediate risk of UTI dipstick results may help differentiate those where immediate antibacterial therapy is indicated from those at lower risk where laboratory confirmation of UTI is needed.

Contamination was seven times more common in the nappy pad than the clean-catch samples, and was identified more often by the NHS than research laboratories. Probability of contamination was not increased by increased time taken for samples to arrive at laboratories, but being female, home sampling, and increased frequency of nappy use did increase risk of contamination. Of note, contamination was no more likely among children with a nappy rash.

A similar proportion of children who had culture results meeting criteria for UTI, contamination and no growth/indeterminate results had improved symptomatically by 2 weeks, and the median number of days to improvement was similar in all of these groups. Children who had a UTI were more likely to have further urine samples in the subsequent 3 months. There was no difference in the proportion of children prescribed antibiotics or in the median number of reconsultations in the subsequent 3 months between the groups. Slightly more children who had UTI were seen in secondary care.

## Strengths

This is the largest prospective study of UTI in individually recruited, acutely unwell young children presenting to primary care. It was conducted in a large number of primary care sites supported locally by four university-supported centres. We recruited ahead of schedule, and included additional patients to help to account for fewer than expected numbers of children with UTI diagnosed in the research laboratory.

At some sites, we were able to record details of eligible children invited but declining to participate, and although these data were incomplete, they suggest that, compared with the recruited sample, there were no important differences in age or sex. We therefore consider that participating children are likely to be representative of children presenting to primary care in the UK with an acute illness.

All index tests were measured according to routine clinical practice using standardised reporting forms and equipment, and nearly all were completed within 24 hours of urine sample retrieval, minimising the impact of disease progression bias. We obtained a urine sample from a high percentage of recruited children and were able to describe certain demographic differences between those children for whom we were able to obtain a urine sample and those for whom a urine sample was not obtained.

We asked clinicians to provide a working diagnosis and estimate the likelihood of a UTI based on their clinical opinion alone, information which was crucial in demonstrating the added diagnostic value of the diagnostic algorithm and technologies such as dipstick testing over and above 'clinician diagnosis'.

Two members of staff carried out all urine analyses in the research laboratory, blind to all of the index tests and clinical features apart from the children's age. They performed the microbiological cultures and interpreted results using standardised processes, which included spiral plating, a more accurate method of quantifying bacterial counts and differentiating mixed cultures.<sup>126</sup> We prospectively collected cost data, allowing us to model cost-effectiveness. Our statistical techniques are well described, and the clinical rules we produced have clinical face validity, and should be easy use in everyday general practice.<sup>114</sup>

## Weaknesses

### Design weaknesses

We were not able to obtain a sufficient volume of urine to send a large enough fraction to the research laboratory for all children who submitted a urine sample, as we prioritised the NHS laboratory fraction in order to ensure that clinicians were sent laboratory results for clinical purposes. There was no mechanism to provide a result from a research laboratory to all the general practices that participated in DUTY.

Although we knew which NHS laboratories routinely used boric acid containing specimen pots, we did not record which individual NHS samples contained boric acid and so were unable to perform exploratory analyses of how boric acid in urine containers may have influenced NHS laboratory culture results. All samples were sent via routine transport (NHS laboratories) and post (research laboratory). This introduced delays before research laboratory processing, but means that our results reflect those likely to be obtained if such a system was introduced into routine clinical practice.

Although we piloted our CRF with, and invited comments from practising clinicians and parents, our CRF did not ask about infectious contacts, and so we were unable to determine if infectious contacts were protective or otherwise for UTI.

While participants were asked to provide clean-catch samples whenever possible, we received a large number of nappy pad samples (our second-choice sampling method). Nappy pad and bag samples are often the only feasible methods (outside hospital settings) for obtaining urine samples from young children in the UK; there is usually insufficient space or time for parents and children to wait to provide clean-catch specimens and most primary care clinicians (other than those with specialist paediatric or emergency department training) are not trained in SPA or catheterisation techniques. SPA and catheterisation sampling methods are invasive and are unlikely to be acceptable to parents and children in the context of low risk for UTI. Nappy pads have been shown to be acceptable to parents<sup>63</sup> and endorsed by NICE.<sup>2</sup>

Our reference standard defined uropathogens as members of the Enterobacteriaceae group at the UK guidelines<sup>2,82</sup> threshold of a pure/predominant growth of  $\geq 10^5$  CFU/ml. We used a rigorous criterion (minimum 3-log difference between the predominant and next most concentrated organism) for defining predominance. This definition could have reduced estimated prevalence if some UTIs were incorrectly classified as contamination; the only other UK primary care-based study of which we are aware estimated prevalence to be 6% (based on NHS laboratory culture results).<sup>9</sup> Collecting an uncontaminated urine specimen is most difficult in the youngest children, and no study has yet reliably distinguished pathogen from contaminant, especially when they coexist. Our definition of UTI excluded atypical bacteria which are also thought to be more common in younger children.<sup>137,138</sup>

There is some evidence that false-positive samples taken from children can be identified by taking sequential urine samples.<sup>81</sup> We considered, however, that the increased burden on patients and the health services associated with taking two separate samples from each child would lead to fewer children being sampled. Thus, the advantages of sequential sampling in terms of reduced false positives would probably be outweighed by more UTIs being identified at an early stage of the child's illness through a single-sample approach. Single samples are the norm in routine UK primary care.

Clinicians' assessment of the likelihood of UTI may have been influenced by a Hawthorne effect from participating in a study that they knew was about diagnosing UTI in young children.

### **Analytic weaknesses**

A major limitation of our study is the relatively small number of UTI microbiological diagnoses, particularly for nappy pad samples. It is plausible that contamination of urine specimens in these samples led to underdiagnosis, limiting our ability to identify the symptoms, signs and dipstick results associated with UTI in children aged < 2 years. Because of lower than expected urine samples positive for UTI on culture, and the unanticipated differences between clean-catch and nappy pad samples in the reliability and accuracy of the laboratory results that meant developing separate algorithms by urine collection method, we had insufficient numbers of children with UTI to both derive and externally validate the algorithms. Although bootstrap validation is an accepted technique to use in this situation, leading to reduced AUROC for our models, even these estimates may be optimistic in relation to values that would have been achieved had we externally validated.

Our economic analyses did not attempt to capture possible impact on antibiotic resistance rates in the population. Taking these into account is likely to increase the cost-effectiveness of more conservative antibiotic treatment. Similarly, our analysis does not include benefits for false-negative UTI diagnosis where an antibiotic serendipitously prescribed for a presumed infection might be effective in treating a true underlying UTI. We focused on NHS rather than societal cost. Our Bayesian framework required the specification of a prior distribution for each parameter. We used diffuse priors to ensure that the posterior distribution was dominated by the likelihood; however, for parameters where the number of events was small the prior might have had an effect.

## Results in the context of research literature

Our clean-catch models include clinically intuitive items. Previous investigation of malodorous urine has shown conflicting results,<sup>67,139</sup> but our study strongly supports its diagnostic value. We investigated, but did not find evidence for, a number of non-specific symptoms (including fever, vomiting, lethargy, irritability and poor feeding) previously found to be diagnostic of UTI<sup>116</sup> and recommended for clinical use by NICE.<sup>2</sup> It remains possible that such symptoms are of use in the secondary care settings in which studies reporting their utility were conducted, or in children with a different illness spectrum. We found that presence of symptoms and signs suggestive of alternative diagnoses to UTI were associated with a reduced risk of UTI. Such inverse associations are unlikely to reflect biological mechanisms, but to arise because both they and UTI are causes of children attending primary health care. Such 'conditioning on a common effect' induces inverse associations between factors that are independent in the source population of well and unwell children.<sup>140</sup> These associations are, nonetheless, of diagnostic utility.

Our nappy pad model, and to some extent the clean-catch model for children aged between 2 and 3 years, is the first primary care study to identify parent-reported symptoms that can be used to select preverbal children warranting urine sampling and presumptive antibiotic treatment. Female sex and parent-reported smelly or darker urine all appear biologically plausible as contributing to the diagnosis of UTI. However, we believe that the apparently substantial reduction in the risk of UTI associated with presence of a nappy rash should be interpreted with caution. The inverse association may arise through conditioning on the common effect of primary care attendance,<sup>140</sup> but this is unlikely to produce such a substantial association. Alternative candidate explanations are that rash may be a risk factor for contamination of urine and hence mask the presence of a UTI, or that skin products used to treat nappy rash could render the urine sterile. However, we did not find nappy rash to be associated with contamination in either the clean-catch or nappy pad samples, and antimicrobial substances were not present for the vast majority of samples. An increased likelihood of contamination of nappy pad samples might also explain the more modest associations of symptoms and dipstick test results with UTI than were found in clean-catch samples. These differences could also arise from differences in illness profiles between older and younger children.

As far as we are aware, this is the first study to evaluate the cost-effectiveness of prediction rules to optimise the selection of children with suspected UTI for urine sampling. Previous work has assessed the most cost-effective test or series of tests for diagnosing UTI, rather than evaluating which children should be selected for urine sampling.

## Clinical and research implications

Parent-reported symptoms and clinical signs can be used to identify preschool children presenting to primary care in whom urine should be collected. The diagnostic utility is better for urine collected using clean catch than nappy pads, but despite higher contamination rates in nappy pad samples they still provide better diagnostic accuracy than current clinical practice for children < 2 years in whom the diagnosis of UTI is most challenging. That said, clinicians should be cautious about using the nappy pad collection method in children with a nappy rash, in whom they should try to collect urine via clean catch.

The addition of dipstick testing can help to decide whether or not to prescribe an immediate antibiotic. Dipsticks have been considered unhelpful in young children<sup>2</sup> until now. We found that dipstick testing compared with a strategy based on laboratory test results could increase the proportion of children with UTI treated immediately and decrease the proportion of children without UTI treated with antibiotics.

By choosing a high specificity threshold for older children in whom clean catch is possible, primary care clinicians would request a urine sample from fewer children with equivalent or higher sensitivity and specificity, achieving marginally better costs and patient outcomes in the short and long term. However, in younger children whose urine was collected using a nappy pad, the distinction in cost-effectiveness between the DUTY risk score and clinical judgement was not clear-cut.

We believe that our results can be applied to other resource rich nations with similar 'first point of contact' health service provision, but may not be applicable to the spectra of illness in preschool children presenting to primary care in resource-poor settings or those referred to secondary care following an initial primary care assessment.

In adult medicine, results from urine microbiology can be interpreted in the clinical context of the patient's presentation. However, in young children the significant difficulties in obtaining uncontaminated samples, together with the non-specific nature of the presenting symptoms, mean there is greater reliance on the laboratory result. More detailed routine microbiological examination of paediatric urine samples would have resource implications that could be better justified if urines were selected for testing through an algorithm that increased the prior probability of positivity. Our results suggest that NHS laboratories should distinguish primary care paediatric (aged < 5 years) samples from adult samples and consider reporting these in more detail, and that national procedures should, correspondingly, be updated.

Even for samples processed in the research laboratory, the diagnostic utility of microbiology based on nappy pad samples was less than for clean-catch samples. Therefore, primary care clinicians should try to obtain clean-catch samples in even very young children in whom they suspect a UTI,<sup>127</sup> for example by providing the time and space to support urine collection. If an algorithm based on parent-reported symptoms can provide earlier ID of the children at greatest risk of UTI, parents could be advised to obtain a urine sample prior to attending primary care.

To reduce the risk of contamination, urine samples should ideally be clean catch, taken in the GP surgery and not taken from children where there has been extensive nappy use within 24 hours of presenting at surgery. These factors may be useful in addition to microbiological criteria in clinical decisions.

Further research is needed to (1) distinguish pathogens from contaminants when bacteria are found in significant concentrations in urine; (2) improve our understanding of the reasons for the discrepancies between research and NHS laboratories; (3) establish the cost-effectiveness of different sensitivity/specificity cut-points for the clean-catch and nappy pad models for use in routine clinical practice, using routine health service laboratories; and (4) to assess the impact on clinical behaviour and patient outcomes, the third of the three steps in the development of a clinical algorithm,<sup>128</sup> using a RCT.

The first of these may be achievable through the use of new diagnostic methods involving sequencing of genetic material in urine samples and investigating the contribution of proteomics and the value of immunological markers in the urine.

# Acknowledgements

## Contributions of authors

**Alastair D Hay** (GP, NIHR Research Professor and Professor of Primary Care) was the lead applicant and co-chief investigator for the DUTY study. Substantial contributions were made to the overall design of the study, the statistical analysis plan, and the writing of the background, methods, results and discussion sections of the report.

**Kate Birnie** (Research Associate in Medical Statistics) contributed to the data analysis in the derivation of the definition of positive UTI, and contributed to the writing of the results section of the report.

**John Busby** (Research Assistant in Health Economics) contributed to the health economic data analysis, and in the writing of the methods and results sections of the report.

**Brendan Delaney** (Guy's and St Thomas' Charity Chair in Primary Care Research) was a co-applicant and contributed to the overall study design, data analysis and writing of the methods and results sections of the report.

**Harriet Downing** (research project manager) was the lead for study management and co-ordination of study implementation and regional recruitment, and commented on the final draft report.

**Jan Dudley** (consultant paediatric nephrologist) was a co-applicant and provided specialist paediatric nephrology advice and commented on final draft of the report.

**Stevo Durbaba** (e-resources developer) was responsible for the development of the DUTY electronic database and contributed to the writing of the methods section of the report.

**Margaret Fletcher** (Professor of Clinical Nursing) was a co-applicant and contributed to the overall study design, contributed to and commented on the final draft of the report.

**Kim Harman** (Research Project Manager) contributed to study management and to the writing of the methods and results sections of the report. She also contributed to the compilation of the final report.

**William Hollingworth** (Professor in Health Economics) was a co-applicant, lead health economist and contributed to the health economic analysis plan and to the writing of the background, methods, results and discussion sections of the report.

**Kerenza Hood** (Professor in Statistics, director of SEWTU) was a co-applicant and contributed to the overall study design, data analysis and writing of the background, methods, results and discussion sections of the report.

**Robin Howe** (consultant microbiologist and head of the Wales Public Health Laboratory Reference Laboratory) was a co-applicant, lead microbiologist and contributed to the overall study design, data analysis and writing of the methods and results sections of the report.

**Michael Lawton** (statistician) contributed to the statistical data analysis, development of the algorithm, and to the writing of the methods and results sections of the report.

**Catherine Lises** (research assistant) contributed to the data management of the study and to the writing of the methods section, and proofreading of the report.

**Paul Little** (Professor of Primary Care Research) was a co-applicant and contributed to the overall study design and reviewing the final draft of the report.

**Alasdair MacGowan** (Professor of Clinical Microbiology and Antimicrobial Therapeutics) was a co-applicant, lead microbiologist and contributed to the overall study design, data analysis and reviewing the final draft of the report.

**Kathryn O'Brien** (GP and clinical lecturer) was a co-applicant and contributed to the overall study design, data analysis and writing of the background and result sections of the report.

**Timothy Pickles** (statistician) contributed to the data management, statistical analysis and to the writing of the results section of the report. He also contributed to the proofreading of the final report.

**Kate Rumsby** (project co-ordinator) contributed to the study design and implementation, regional recruitment and to the writing of the methods and results sections of the report.

**Jonathan AC Sterne** (Professor of Medical Statistics and Epidemiology) was a co-applicant and lead study statistician. Substantial contributions were made to the overall design of the study, the statistical analysis plan, and the writing of the methods, results and discussion sections of the report.

**Emma Thomas-Jones** (research fellow) was the lead for study data management and regional study management in Wales, contributed to the study design and implementation, and to the writing of the methods and results sections of the report. She also contributed to the compilation, formatting and proofreading of the final report.

**Judith van der Voort** (consultant paediatrician and paediatric nephrologist) was a co-applicant and provided specialist paediatric nephrology advice and commented on final draft of the report.

**Cherry-Ann Waldron** (research associate) contributed to the study management in Wales, and in the writing of the methods and results section of the report. She also contributed to the compilation, formatting and proofreading of the final report.

**Penny Whiting** (research fellow) was a co-applicant and contributed to the writing of the background section of the report.

**Mandy Wootton** (lead scientist) was responsible for overseeing the quality of the microbiology data and contributed to the overall study design, data analysis and writing of the methods and results sections of the report.

**Christopher C Butler** (GP, professor of primary care medicine at Cardiff University and professor of primary care at the Nuffield Department of Primary Care Health Sciences at the University of Oxford) was a joint lead applicant and co-chief investigator for the study. Substantial contributions were made to the overall design of the study, the statistical analysis plan, and the writing of the background, methods, results and discussion sections of the report.

### *Other members of the study team*

We would like to thank Jonathan Benger (Professor of Emergency Care, University of the West of England, Bristol) who was a co-applicant and was responsible for recruitment at the Children's Emergency Department, Bristol. We would also like to thank the following for their contribution to the conduct of the study: Theresa Bowes (Southampton University), Lisa Calver (University of Bristol), Lewis Darmanin (NISCHR-CRC), Miceala Gal (research fellow, Cardiff University), Susan George (University of Bristol), Andrea Jarman (University of Bristol), Lyn Liddiard (University of Bristol), Ruth Munn (University of Bristol), Carolyn Powell (University of Bristol), Jennifer Richards (SACU Lab, Cardiff), Elizabeth Thomas (University of Bristol), Tessa Wade (Southampton University) and Stana Williams (King's College London).



### Other contributors

Peter Brindle (co-applicant, GP, Director of the Primary Care Research Network-South West and Research and Development Lead, Bristol Clinical Commissioning Group).

Alan Montgomery (co-applicant, reader and director of the Bristol Randomised Trials Collaboration, University of Bristol).

We would also like to thank Information and Communications Technology at SLaM for hosting and supporting the online database during the project.

### Independent members of the Study Steering Committee

We would like to thank our Study Steering Committee for their support: Professor Frank Sullivan (Trial Steering Committee Chairperson, Dundee University), Dr Rafael Perera (University of Oxford), Dr Cliodna McNulty (HPA) and Dr Matthew Thompson (University of Oxford).

### Administrative team

We would like to thank all the administrative team for all four centres, namely Steven Beech (Assistant Study Administrator, University of Bristol), Christina Curry (Study Administrator, University of Bristol), Catherine Derrick (Study Administrator, University of Bristol), Veronica Dunning (Study Administrator, Cardiff University), Margaret Hague (Study Administrator, University of Southampton), Marilyn Peters (Study Administrator, King's College London), Victoria Roberts (Study Administrator, Cardiff University), Annie Sadoo (Study Administrator, University of Bristol) and Jane Woodhead (Study Administrator, Cardiff University).

### Research networks

We also wish to thank the providers of nursing/CSO support in all four centres. For London, these include staff at the Primary Care Research Network in Greater London, for Southampton in Kent and Medway, Sussex, Surrey, Thames Valley, in Bristol the South-west, Cumbria and Lancashire, Northumberland and Tyne and Wear, and in Cardiff the NISCHR-CRC in Wales (all three regions (North Wales, South West Wales and South East Wales). Additionally, thanks to the CLRNs of Central and East London, Western, Hampshire and Isle of Wight, Peninsula and the NIHR Biomedical Research and Development Department, Guy's and St Thomas' NHS Foundation Trust.

### Recruitment sites

Accident & Emergency Department, Southampton General Hospital.

Afan Valley Group Practice, Neath.

Albion Street Group Practice, London.

Argyle Medical Group Practice, Pembroke Dock.

Argyll House Surgery.

Ash Vale Health Centre, Aldershot.

Ashtrees Surgery, Carnforth.

Avicenna Medical Centre.

Ball Tree Surgery, Lancing.

Bampton Medical Centre, Bampton.

Barnfield Hill Surgery, Exeter.

Batheaston Medical Centre, Bath.

Belle Vue Surgery, Fleetwood.

Bellevue Surgery, Newport, Gwent.

Bere Regis Surgery, Wareham.

Bethesda Medical Centre, Margate.

Binfield Road, London.

Bitterne WIC, Southampton.

Blackfriars Medical Practice, London.

Blaen Y Cwm Surgery, Newport, Gwent.

Blithehale Health Centre, London.

Bosmere Medical Practice, Havant.

Bradford on Avon Health Centre, Bradford on Avon.

Bradford Road Medical Centre, Trowbridge.

Bradgate Surgery, Bristol.

Branch End, Stocksfield.

Brannam Medical Centre, Barnstaple.

Brigstock and South Norwood Medical Partnership, Croydon.

Brixton Hill Group Practice.

Broadway Medical Centre, Lancashire.

Broadway Surgery, Brighton.

Brockworth Surgery, Gloucester.

Brook Lane Surgery, Southampton.

Brookside Health Centre, Freshwater, Isle of Wight.

Brunswick Surgery, Swansea.

Bunny Hill MIU, Sunderland.

Burn Brae Medical Group, Hexham.

Burnham Health Centre, Slough.

Cape Cornwall Surgery, Penzance.

Cardiff Road Surgery, Mountain Ash.

Carshalton Fields Surgery, Carshalton Beeches.

Cathedral Medical Group, Chichester.

Chandlers Ford Family Practice, Chandlers Ford.

Chatfield Medical Centre, London.

Chessel Practice.

Chew Medical Practice, Bristol.

Children's Emergency Department BRI, Bristol.

Chippenham Surgery, Monmouth.

Church Street Practice, Wantage.

Church View Surgery, Broadway, Somerset.

Claremont Medical Practice, Exmouth.

Clarence House, Rhyl.

Clevedon Riverside Group, Clevedon.

Cleveleys Group Practice, Thornton-Cleveleys.

Clifton Surgery, Cardiff.

Coastal Medical Group, Morecombe.

Cobtree Medical Practice, Maidstone.

Coleford Health Centre, Coleford.

Coleridge Medical Centre, Ottery St Mary.

Combe Down Surgery, Bath.

Connor Downs Surgery, Hayle.

Cossington House Surgery, Canterbury.

Cotswold Medical Practice, Bourton on the Water.

Court House Medical Centre, Caerphilly.

Cranleigh Medical Centre, Cranleigh.

Crown Dale Medical Centre, London.

Crown Medical Centre, Taunton.

Deerbrook Surgery, London.

Didcot Health Centre, Didcot.

DMC Silverlock, London.

Dr Ali's Surgery, Fleetwood.

Dr Gunasuntharam's Surgery, London.

St Luke's Surgery, Cwmcam, Newbridge.

St Mark's Dee View Surgery, Connah's Quay.

St Melor Surgery, Amesbury.

St Richard's Road Surgery, Deal.

St Thomas' Hospital, London.

Staines & Thameside Medical Group, Staines.

Station Road Surgery, Penarth.

Stockwell Group Practice, London.

Stonedean Practice, Milton Keynes.

Streatham Common Group Practice, London.

Summertown Health Centre, Oxford.

Summervale Medical Centre, Ilminster.

Sunbury Medical Centre, Sunbury on Thames.

Tawstock Medical Centre, Tawstock.

The Apples Medical Centre, Sherborne.

The Avenue Surgery, Warmisnter.

The Bellingham Practice, Hexham.

The Butchery Practice, Sandwich.

The Cranborne Practice, Cranborne.

The Jenner Practice, London.

The Laurels Surgery, Flint.

The Malthouse Surgery, Abingdon.

The Merrywood Practice.

The New East Quay Medical Centre, Bridgewater.

The Norwood Surgery, London.

The Practice of Health, Barry.

The Redhouse Surgery, Milton Keynes.

The Sele Medical Practice, Hexham.

The Station Practice, Hastings.

The Surgery (Drs Robb & Robb), Preston.

The Surgery, Gwalchmai, Ynys Mon.

The Surgery, Lyminge, Folkestone.

The Thornton Practice, Thornton-Cleveleys.

The Three Spires Medical Practice, Truro.

The Vale of Neath Practice, Neath.

The Village Surgery, Caerphilly.

Thornton Health Centre, London.

Three Swans Surgery, Salisbury.

Tonyfelin Medical Centre, Caerphilly.

Tudor Lodge Surgery, Bristol.

Tulse Hill Practice, London.

Ty Bryn Surgery, Caerphilly.

University of Canterbury Medical Centre, Canterbury.

Upper Norwood Surgery, London.

Vassall Road, Brixton.

Vauxhall Surgery, London.

Village Practice Thornton, Thornton-Cleveleys.

Vine Surgery, Street.

Walnut Tree Health Centre, Milton Keynes.

Wandle Valley Health Centre, London.

Wandsworth Medical Centre, London.

Waterfront Surgery, Barry.

Wateringbury Surgery, Maidstone.

Waterloo Medical Centre, Blackpool.

Wells City Practice, Wells.

Wellspring Surgery, Bristol.

West Farm Surgery, Newcastle upon Tyne.

West Quay Medical Centre.

Westlake Surgery, West Coker.

Westongrove Research Centre, Aston Clinton.

Wilton Health Centre, Wilton.

Wincanton Health Centre, Wincanton.

Winchcombe Medical Centre, Cheltenham.

Windermere and Bowness Medical Practice.

Wish Park Surgery, Hove.

Woodlands Surgery, Bridgend.

Woolston Lodge Surgery, Southampton.

Woolwell & Lisson Grove Medical Centre, Plymouth.

Woosehill Surgery, Wokingham.

Worden Medical Centre.

Wrythe Green Surgery, London.

Wye Valley Practice, St Briavels.

Y Feddygfa Wen, Porthmadog.

## **NHS laboratories**

Ashford & St Peter's Hospitals NHS Foundation Trust, Surrey.

Barts and The London NHS Trust, London.

Blackpool Teaching Hospitals NHS Foundation Trust, Blackpool.

Conquest Hospital, East Sussex NHS Trust.

Crawley Hospital, NHS West Sussex.

Darent Valley Hospital, Dartford & Gravesham NHS Trust.

Derriford Hospital, Plymouth Hospital NHS Trust.

Dorset County Hospital NHS Foundation Trust, Dorchester.

Eastbourne District General Hospital, East Sussex Hospitals NHS Trust.

Epsom & St Helier University Hospitals NHS Trust.

Freeman Hospital, Newcastle upon Tyne Hospitals NHS Foundation Trust.

Frimley Park Hospital NHS Trust, Frimley.

Gateshead NHS Foundation Trust, Gateshead.

Gloucester Royal Hospital Gloucestershire NHS Foundation Trust, Gloucester.

Great Western Hospital NHS Foundation Trust, Swindon.

Guy's & St Thomas' NHS Foundation Trust, London.

Hampshire Hospital NHS Foundation Trust, Basingstoke.

John Radcliffe Hospital Oxford Radcliffe Hospitals NHS Foundation Trust, Oxford.

KingsPath, King's College Hospital NHS Foundation Trust, London.

Kingston Hospital NHS Trust, London.

Lewisham and Greenwich NHS Trust, London.

Maelor Microbiology Lab, Wrexham, Betsi Cadwaladr University Health Board.

Maidstone Hospital, Maidstone & Tunbridge Wells NHS Trust.

Mayday Hospital, Croydon Health Services NHS Trust.

Medway Maritime Hospital, Medway NHS Foundation Trust, Gillingham.

Milton Keynes Hospital, Milton Keynes Hospital NHS Foundation Trust.

Musgrove Park Hospital, Taunton & Somerset NHS Foundation Trust.

North Bristol NHS Trust, Bristol.

North Devon District Hospital, Northern Devon Healthcare NHS Foundation Trust, Barnstaple.

North Tyneside General Hospital, Northumbria Healthcare NHS Foundation Trust, North Shields.

Public Health Wales Microbiology Aberystwyth, Hywel Dda Health Board.

Public Health Wales Microbiology Bangor, Betsi Cadwaladr University Health Board.

Public Health Wales Microbiology Cardiff, Cardiff and Vale University Health Board.

Public Health Wales Microbiology Carmarthen, Hywel Dda University Health Board.

Public Health Wales Microbiology Princess of Wales, Bridgend, Abertawe Bro Morgannwg University Health Board.

Public Health Wales Microbiology Rhyl, Betsi Cadwaladr University Health Board, Rhyl.

Public Health Wales Microbiology Swansea, Abertawe Bro Morgannwg University Health Board.

Poole Hospital, Poole Hospital NHS Foundation Trust.

Prince Charles Hospital, Merthyr Tydfil, Cwm Taf University Health Board.

Queen Alexandra Hospital, Portsmouth Hospitals NHS Trust.

Royal Berkshire NHS Foundation Trust, Reading.

Royal Cornwall Hospitals NHS Foundation Trust, Truro.

Royal Devon & Exeter NHS Foundation Trust, Exeter.

Royal Glamorgan Hospital, Llantrisant, Cwm Taf University Health Board.

Royal Gwent Microbiology Lab, Newport, Aneurin Bevan Health Board.

Royal Hampshire County Hospital, Hampshire Hospitals NHS Foundation Trust, Winchester.



Royal Lancaster Infirmary, University Hospitals of Morecambe Bay NHS Foundation Trust, Lancaster.

Royal Preston Hospital, Lancashire Teaching Hospitals NHS Trust, Preston.

Royal Sussex County Hospital, Brighton & Sussex University Hospitals NHS Trust.

Royal United Hospitals Bath NHS Foundation Trust, Bath.

Salisbury NHS Foundation Trust, Salisbury.

South Tyneside District Hospital, South Tyneside NHS Foundation Trust, South Shields.

Southampton General Hospital, University Hospital Southampton NHS Foundation Trust.

St George's Healthcare NHS Trust, London.

St Helier Hospital, Epsom & St Helier University Hospitals NHS Trust.

St Mary's Hospital, Isle of Wight PCT.

St Richard's Hospital, Western Sussex Hospitals NHS Foundation Trust, Chichester.

Stoke Mandeville Hospital, Buckinghamshire Healthcare NHS Trust.

Sunderland Royal Hospital, City Hospitals Sunderland NHS Foundation Trust.

Torbay Hospital, South Devon Healthcare NHS Foundation Trust, Torbay.

University Hospitals Bristol NHS Foundation Trust, Bristol.

Weston General Hospital, Weston Area Health NHS Trust.

Withybush Microbiology Lab, Haverfordwest, Hywel Dda Health Board.

Worthing Hospital, Western Sussex Hospitals NHS Foundation Trust, Worthing.

Wexham Park Hospital, Heatherwood & Wexham Park Hospitals NHS Foundation Trust, Slough.

Finally, we would like to thank the children and families who participated in the study, without whom this study would not have been possible.

## Publications

Downing H, Thomas-Jones E, Gal M, Waldron CA, Sterne J, Hollingworth W, *et al.* 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* 2012;**12**:158.

Butler CC, O'Brien K, Pickles T, Hood K, Wootton M, Howe R, *et al.*, on behalf of the DUTY study team. Childhood urinary tract infection in primary care: a prospective observational study of prevalence, diagnosis, treatment, and recovery. *Br J Gen Pract* 2015;**65**(633):7. doi: 10.3399/bjgp15X684361.

Bulter CC, Sterne JAC, Lawton M, O'Brien K, Wootton M, Hood K, *et al.* 'Nappy pad' urine samples for investigation and treatment of urinary tract infection in young children: the 'DUTY' prospective diagnostic cohort study. *Br J Gen Pract* 2016; in press.

Hay AD, Sterne JAC, Hood K, Little P, Delaney B, Hollingworth W, *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; in press.

## References

1. Craig JC, Williams GJ, Jones M, Codarini M, Macaskill P, Hayen A, *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.
2. National Institute for Health and Care Excellence. *Urinary Tract Infection in Children: Diagnosis, Treatment and Long Term Management*. London: NICE; 2007.
3. Shaikh N, Morone NE, Lopez J, Chianese J, Sangvai S, D'Amico F, *et al.* Does this child have a urinary tract infection? *JAMA* 2007;**298**:2895–904. <http://dx.doi.org/10.1001/jama.298.24.2895>
4. Shaikh N, Morone NE, Bost JE, Farrell MH. Prevalence of urinary tract infection in childhood: a meta-analysis. *Pediatr Infect Dis J* 2008;**27**:302–8. <http://dx.doi.org/10.1097/INF.0b013e31815e4122>
5. Bachur R, Harper MB. Reliability of the urinalysis for predicting urinary tract infections in young febrile children. *Arch Pediatr Adolesc Med* 2001;**155**:60–5. <http://dx.doi.org/10.1001/archpedi.155.1.60>
6. Newman TB, Bernzweig JA, Takayama JI, Finch SA, Wasserman RC, Pantell RH. Urine testing and urinary tract infections in febrile infants seen in office settings: the Pediatric Research in Office Settings' Febrile Infant Study. *Arch Pediatr Adolesc Med* 2002;**156**:44–54. <http://dx.doi.org/10.1001/archpedi.156.1.44>
7. O'Brien K. The prevalence of urinary tract infection (UTI) in children under five years old presenting with an acute illness in UK general practice. Ph.D. thesis. Cardiff: Cardiff University; 2013. URL: <http://orca.cf.ac.uk/47309/>
8. Singh-Grewal D, Macdessi J, Craig J. Circumcision for the prevention of urinary tract infection in boys: a systematic review of randomised trials and observational studies. *Arch Dis Child* 2005;**90**:853–8. <http://dx.doi.org/10.1136/adc.2004.049353>
9. 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**:91–2. <http://dx.doi.org/10.3399/bjgp13X663127>
10. Vernon SJ, Coulthard MG, Lambert HJ, Keir MJ, Matthews JN. New renal scarring in children who at age 3 and 4 years had had normal scans with dimercaptosuccinic acid: follow up study. *BMJ* 1997;**315**:905–8. <http://dx.doi.org/10.1136/bmj.315.7113.905>
11. Sacks SH, Verrier-Jones K, Roberts R, Asscher AW, Ledingham JG. Effect of symptomless bacteriuria in childhood on subsequent pregnancy. *Lancet* 1987;**2**:991–4. [http://dx.doi.org/10.1016/S0140-6736\(87\)92558-X](http://dx.doi.org/10.1016/S0140-6736(87)92558-X)
12. Ferry SA, Holm SE, Stenlund H, Lundholm R, Monsen TJ. Clinical and bacteriological outcome of different doses and duration of pivmecillinam compared with placebo therapy of uncomplicated lower urinary tract infection in women: the LUTIW project. *Scand J Prim Health Care* 2007;**25**:49–57. <http://dx.doi.org/10.1080/02813430601183074>
13. Christiaens TC, De Meyere M, Verschraegen G, Peersman W, Heytens S, De Maeseneer JM. Randomised controlled trial of nitrofurantoin versus placebo in the treatment of uncomplicated urinary tract infection in adult women. *Br J Gen Pract* 2002;**52**:729–34.

14. Hannan TJ, Totsika M, Mansfield KJ, Moore KH, Schembri MA, Hultgren SJ. Host-pathogen checkpoints and population bottlenecks in persistent and intracellular uropathogenic *Escherichia coli* bladder infection. *Microbiol Rev* 2012;**36**:616–48. <http://dx.doi.org/10.1111/j.1574-6976.2012.00339.x>
15. Coulthard M, Flecknell P, Orr H, Manas D, O'Donnell H. Renal scarring caused by vesicoureteric reflux and urinary infection: a study in pigs. *Pediatr Nephrol* 2002;**17**:481–4. <http://dx.doi.org/10.1007/s00467-002-0878-2>
16. Cardiff, Oxford Bacteriuria Study Group. Sequelae of covert bacteriuria in schoolgirls. A four-year follow-up. *Lancet* 1973;**1**:889–93.
17. Lindberg U, Claesson I, Hanson LA, Jodal U. Asymptomatic bacteriuria in schoolgirls:VIII. Clinical course during a 3-year follow-up. *J Pediatr* 1978;**92**:194–9. [http://dx.doi.org/10.1016/S0022-3476\(78\)80003-1](http://dx.doi.org/10.1016/S0022-3476(78)80003-1)
18. Newcastle Covert Bacteriuria Research Group. Covert bacteriuria in schoolgirls in Newcastle upon Tyne. *Arch Dis Child* 1981;**56**:585–92. <http://dx.doi.org/10.1136/adc.56.8.585>
19. Savage DCL, Howie G, Adler K, Wilson MI. Controlled trial of therapy in covert bacteriuria in childhood. *Lancet* 1975;**1**:358–61. [http://dx.doi.org/10.1016/S0140-6736\(75\)91277-5](http://dx.doi.org/10.1016/S0140-6736(75)91277-5)
20. Fitzgerald A, Mori R, Lakhanpaul M. Interventions for covert bacteriuria in children. *Cochrane Database Syst Rev* 2012;**2**:CD006943. <http://dx.doi.org/10.1002/14651858.CD006943.pub2>
21. Zhanel GG, Harding GK, Guay DR. Asymptomatic bacteriuria: which patients should be treated? *Arch Intern Med* 1990;**150**:1389–96. <http://dx.doi.org/10.1001/archinte.1990.00390190055007>
22. Wettergren B, Hellstrom M, Stokland E, Jodal U. Six year follow up of infants with bacteriuria on screening. *BMJ* 1990;**301**:845–8. <http://dx.doi.org/10.1136/bmj.301.6756.845>
23. Savage DCL, Wilson MI, Mc Hardy M, Dewar DAE, Fee WM. Covert bacteriuria of childhood: a clinical and epidemiological study. *Arch Dis Child* 1973;**48**:8–20. <http://dx.doi.org/10.1136/adc.48.1.8>
24. Davies JM, Littlewood JM, Gibson GL, Meadow SR. Prevalence of bacteriuria in infants and preschool children. *Lancet* 1974;**304**:7–10. [http://dx.doi.org/10.1016/S0140-6736\(74\)91345-2](http://dx.doi.org/10.1016/S0140-6736(74)91345-2)
25. McLachlan MS, Meller ST, Jones ER, Asscher AW, Fletcher EW, Mayon-White RT, et al. Urinary tract infection in schoolgirls with covert bacteriuria. *Arch Dis Child* 1975;**50**:253–8. <http://dx.doi.org/10.1136/adc.50.4.253>
26. Newcastle Covert Bacteriuria Research Group. Asymptomatic bacteriuria in children in Newcastle upon Tyne. *Arch Dis Child* 1975;**50**:90–102. <http://dx.doi.org/10.1136/adc.50.2.90>
27. Saxena SR, Collis A, Laurance BM. The prevalence of asymptomatic urinary-tract infection in pre-school children. *Practitioner* 1975;**214**:257–80.
28. Silverberg DS, Jackson FL, Bryan LE. Antibody-coated bacteria in the urine of preschool and school-aged girls with asymptomatic bacteriuria. *CMAJ* 1975;**115**:1091–3.
29. Siegel S, Siegel B, Sokoloff B. Urinary infection in infants and preschool children: five year follow-up. *Am J Dis Child* 1980;**134**:369–72. <http://dx.doi.org/10.1001/archpedi.1980.04490010027010>
30. Goossens H, Mol P, Hall M, Butzler J. Prevalence of asymptomatic bacteriuria and comparison between different screening methods for its detection in infants. *Eur J Epidemiol* 1985;**1**:301–4. <http://dx.doi.org/10.1007/BF00237106>
31. Wettergren B, Jodal U, Jonasson G. Epidemiology of bacteriuria during the first year of life. *Acta Paediatr Scand* 1985;**74**:925–33. <http://dx.doi.org/10.1111/j.1651-2227.1985.tb10059.x>
32. Coulthard MG, Lambert HJ, Keir MJ. Occurrence of renal scars in children after their first referral for urinary tract infection. *BMJ* 1997;**315**:918–19. <http://dx.doi.org/10.1136/bmj.315.7113.918>

33. Shaikh N, Ewing AL, Bhatnagar S, Hoberman A. Risk of renal scarring in children with a first urinary tract infection: a systematic review. *Pediatrics* 2010;**126**:1084–91. <http://dx.doi.org/10.1542/peds.2010-0685>
34. Berg UB, Johansson SB. Age as a main determinant of renal functional damage in urinary tract infection. *Arch Dis Child* 1983;**58**:963–9. <http://dx.doi.org/10.1136/adc.58.12.963>
35. Sinha MD, Gibson P, Kane T, Lewis MA. Accuracy of ultrasonic detection of renal scarring in different centres using DMSA as the gold standard. *Nephrol Dial Transplant* 2007;**22**:2213–16. <http://dx.doi.org/10.1093/ndt/gfm155>
36. Coulthard MG. Is reflux nephropathy preventable, and will the NICE childhood UTI guidelines help? *Arch Dis Child* 2008;**93**:196–9. <http://dx.doi.org/10.1136/adc.2006.100529>
37. Venhola M, Hannula A, Huttunen N-P, Renko M, Pokka T, Uhari M. Occurrence of vesicourethral reflux in children. *Acta Paediatr* 2010;**99**:1875–8. <http://dx.doi.org/10.1111/j.1651-2227.2010.01909.x>
38. Dick P, Feldman W. Routine diagnostic imaging for childhood urinary tract infections: a systematic overview. *J Pediatr* 1996;**128**:15–22. [http://dx.doi.org/10.1016/S0022-3476\(96\)70422-5](http://dx.doi.org/10.1016/S0022-3476(96)70422-5)
39. Doganis D, Siiafas K, Mavrikou M, Issaris G, Martirosova A, Perperidis G, *et al.* Does early treatment of urinary tract infection prevent renal damage? *Pediatrics* 2007;**120**:e922–8. <http://dx.doi.org/10.1542/peds.2006-2417>
40. Jodal U. The natural history of bacteriuria in childhood. *Infect Dis Clin North Am* 1987;**1**:713–29.
41. Hiraoka M, Hashimoto G, Tsuchida S, Tsukahara H, Ohshima Y, Mayumi M. Early treatment of urinary infection prevents renal damage on cortical scintigraphy. *Pediatr Nephrol* 2003;**18**:115–18.
42. Smellie JM, Poulton A, Prescod NP. Retrospective study of children with renal scarring associated with reflux and urinary infection [see comment]. *BMJ* 1994;**308**:1193–6. <http://dx.doi.org/10.1136/bmj.308.6938.1193>
43. Winberg J, Andersen HJ, Bergstrom T, Jacobsson B, Larson H, Lincoln K. Epidemiology of symptomatic urinary tract infection in childhood. *Acta Paediatr Scand Suppl* 1974;1–20. <http://dx.doi.org/10.1111/j.1651-2227.1974.tb05718.x>
44. Oh MM, Kim JW, Park MG, Kim JJ, Yoo KH, Moon DG. The impact of therapeutic delay time on acute scintigraphic lesion and ultimate scar formation in children after their first febrile UTI. *Eur J Pediatr* 2012;**171**:565–70. <http://dx.doi.org/10.1007/s00431-011-1614-3>
45. Jacobson SH, Eklof 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**:703–6. <http://dx.doi.org/10.1136/bmj.299.6701.703>
46. Round J, Fitzgerald AC, Hulme C, Lakhanpaul M, Tullus K. Urinary tract infections in children and the risk of ESRF. *Acta Paediatr* 2012;**101**:278–82. <http://dx.doi.org/10.1111/j.1651-2227.2011.02542.x>
47. Smellie JM, Prescod NP, Shaw PJ, Risdon RA, Bryant TN. Childhood reflux and urinary infection: a follow-up of 10–41 years in 226 adults. *Pediatr Nephrol* 1998;**12**:727–36. <http://dx.doi.org/10.1007/s004670050535>
48. Coulthard MG, Vernon SJ, Lambert HJ, Matthews JNS. A nurse led education and direct access service for the management of urinary tract infections in children: prospective controlled trial. *BMJ* 2003;**327**:656. <http://dx.doi.org/10.1136/bmj.327.7416.656>
49. van der Voort J, Edwards A, Roberts R, Verrier JK. The struggle to diagnose UTI in children under two in primary care. *Fam Pract* 1997;**14**:44–8. <http://dx.doi.org/10.1093/fampra/14.1.44>

50. Jadresic L, Cartwright K, Cowie N, Witcombe B, Stevens D. Investigation of urinary tract infection in childhood. *BMJ* 1993;**307**:761–4. <http://dx.doi.org/10.1136/bmj.307.6907.761>
51. Rowlands S, Moser K. Consultation rates from the general practice research database. *Br J Gen Pract* 2002;**52**:658–60.
52. Fallon UB, Murphy AW, Majawit E, Oriordan C, Bury G, O'Mahony D. Primary care utilisation rates in pre-school children. *Ir Med J* 2007;**100**(Suppl.):23–7.
53. Saxena S, Majeed A, Jones M. Socioeconomic differences in childhood consultation rates in general practice in England and Wales: prospective cohort study. *BMJ* 1999;**318**:642–6. <http://dx.doi.org/10.1136/bmj.318.7184.642>
54. Cunningham AM, Edwards A, Jones KV, Bourdeaux K, Willock J, Barnes R. Evaluation of a service development to increase detection of urinary tract infections in children. *J Eval Clin Pract* 2005;**11**:73–6. <http://dx.doi.org/10.1111/j.1365-2753.2004.00507.x>
55. Office for National Statistics. *Census Statistical Bulletin: Population and Household Estimates for Wales, March 2011*. Office for National Statistics; 2013. URL: [www.ons.gov.uk/ons/rel/census/2011-census/population-and-household-estimates-for-wales/stb-2011-census-wales.html](http://www.ons.gov.uk/ons/rel/census/2011-census/population-and-household-estimates-for-wales/stb-2011-census-wales.html) (accessed 6 October 2013).
56. Whiting P, Westwood M, Bojke L, Palmer S, Richardson G, Cooper J, et al. Clinical effectiveness and cost-effectiveness of tests for the diagnosis and investigation of urinary tract infection in children: a systematic review and economic model. *Health Technol Assess* 2006;**10**(36). <http://dx.doi.org/10.3310/hta10360>
57. Hardy JD, Furnell PM, Brumfitt W. Comparison of sterile bag, clean catch and suprapubic aspiration in the diagnosis of urinary tract infection in early childhood. *Br J Urol* 1976;**48**:279–83. <http://dx.doi.org/10.1111/j.1464-410X.1976.tb10222.x>
58. Braude H, Fofar JO, Gould JC, McLeod JW. Diagnosis of urinary tract infection in childhood based on examination of paired non-catheter and catheter specimens of urine. *BMJ* 1976;**4**:702–5. <http://dx.doi.org/10.1136/bmj.4.5581.702>
59. Benito Fernandez J, Sanchez Echaniz J, Mintegui Raso S, Montejo F. Urinary tract infection in infants: use of urine specimens obtained by suprapubic bladder aspiration in order to determine the reliability of culture specimen of urine collected in perineal bag. *An Esp Pediatr* 1996;**45**:149–52.
60. Cohen HA, Woloch B, Linder N, Vardi A, Barzilai A. Urine samples from disposable diapers: an accurate method for urine cultures. *J Fam Pract* 1997;**44**:290–2.
61. Etoubleau C, Reveret M, Brouet D, Badier I, Brosset P, Fourcade L, et al. Moving from bag to catheter for urine collection in non-toilet-trained children suspected of having urinary tract infection: a paired comparison of urine cultures. *J Pediatr* 2009;**154**:803–6. <http://dx.doi.org/10.1016/j.jpeds.2009.01.008>
62. Krahenbuhl JD, Beaulieu C, Gehri M. Evaluation of a novel in-vitro diagnostic device for the detection of urinary tract infections in diaper wearing children. *Swiss Med Wkly* 2012;**142**:w13560. <http://dx.doi.org/10.4414/smw.2012.13560>
63. Liaw LCT, Nayar DM, Pedler SJ, Coulthard MG. Home collection of urine for culture from infants by three methods: survey of parents' preferences and bacterial contamination rates. *BMJ* 2000;**320**:1312–13. <http://dx.doi.org/10.1136/bmj.320.7245.1312>
64. Mori R, Lakhanpaul M, Verrier-Jones K. Diagnosis and management of urinary tract infection in children: summary of NICE guidance. *BMJ* 2007;**335**:395–7. <http://dx.doi.org/10.1136/bmj.39286.700891.AD>

65. Gorelick MH, Shaw KN. Clinical decision rule to identify febrile young girls at risk for urinary tract infection. *Arch Pediatr Adolesc Med* 2000;**154**:386–90. <http://dx.doi.org/10.1001/archpedi.154.4.386>
66. 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**:162–4. <http://dx.doi.org/10.1097/01.pec.0000081238.98249.40>
67. 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**:6. <http://dx.doi.org/10.1542/peds.2011-2856>
68. Craig JC, Irwig LM, Knight JF, Sureshkumar P, Roy LP. Symptomatic urinary tract infection in preschool Australian children. *J Paediatr Child Health* 1998;**34**:154–9. <http://dx.doi.org/10.1046/j.1440-1754.1998.00190.x>
69. Hoberman A, Chao HP, Keller DM, Hickey R, Davis HW, Ellis D. Prevalence of urinary tract infection in febrile infants. *J Pediatr* 1993;**123**:17–23. [http://dx.doi.org/10.1016/S0022-3476\(05\)81531-8](http://dx.doi.org/10.1016/S0022-3476(05)81531-8)
70. 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**:e16. <http://dx.doi.org/10.1542/peds.102.2.e16>
71. Schwartz S, Raveh D, Toker O, Segal G, Godovitch N, Schlesinger Y. A week-by-week analysis of the low-risk criteria for serious bacterial infection in febrile neonates. *Arch Dis Child* 2009;**94**:287–92. <http://dx.doi.org/10.1136/adc.2008.138768>
72. Cathcart P, Nuttall M, van der Meulen J, Emberton M, Kenny SE. Trends in paediatric circumcision and its complications in England between 1997 and 2003. *Br J Surg* 2006;**93**:885–90. <http://dx.doi.org/10.1002/bjs.5369>
73. Sukanya D, Williams G, Hayen A, Macaskill P, McCaskill M, Isaacs D, *et al.* Accuracy of the “traffic light” clinical decision rule for serious bacterial infections in young children with fever: a retrospective cohort study. *BMJ* 2013;**346**:f866. <http://dx.doi.org/10.1136/bmj.f866>
74. Roberts KB, Charney E, Sweren RJ, Ahonkhai VI, Bergman DA, Coulter MP, *et al.* Urinary tract infection in infants with unexplained fever: a collaborative study. *J Pediatr* 1983;**103**:864–7. [http://dx.doi.org/10.1016/S0022-3476\(83\)80702-1](http://dx.doi.org/10.1016/S0022-3476(83)80702-1)
75. Kass E. Asymptomatic infections of the urinary tract. *Trans Assoc Am Physicians* 1956;**69**:56–64.
76. Stamm W. Quantitative urine cultures revisited. *Eur J Clin Microbiol Infect Dis* 1984;**3**:279–81.
77. Pryles CV. The diagnosis of urinary tract infection. *Pediatrics* 1960;**26**:441–51.
78. Hoberman A, Wald ER, Reynolds EA, Penchansky L, Charron M. Pyuria and bacteriuria in urine specimens obtained by catheter from young children. *J Pediatr* 1994;**124**:513–19. [http://dx.doi.org/10.1016/S0022-3476\(05\)83127-0](http://dx.doi.org/10.1016/S0022-3476(05)83127-0)
79. Kanellopoulos T, Vassiliakos P, Kantis M, Ellina A, Kolonitsiou F, Papanastasiou D. Low bacterial count urinary tract infections in infants and young children. *Eur J Pediatr* 2005;**164**:355–61. <http://dx.doi.org/10.1007/s00431-005-1632-0>
80. Coulthard MG, Kalra M, Lambert HJ, Nelson A, Smith T, Perry JD. Redefining urinary tract infections by bacterial colony counts. *Pediatrics* 2010;**125**:335–41. <http://dx.doi.org/10.1542/peds.2008-1455>
81. Kass E. Pyelonephritis and bacteriuria. A major problem in preventive medicine. *Ann Intern Med* 1962;**56**:46–53. <http://dx.doi.org/10.7326/0003-4819-56-1-46>

82. Public Health England. *Investigation of Urine. UK Standards for Microbiology Investigations. B 41 Issue 7.2*. 2014. URL: <http://www.hpa.org.uk/SMI/pdf>
83. Grabe M, Bishop M, Bjerklund-Johansen T, Botto H, Çek M, Lobel B, et al. *Guidelines on Urological Infections*. European Association of Urology; 2010.
84. Messi G, Peratoner L, Paduano L, Marchi AG. Epidemiology of urinary tract infections and vesico-ureteral reflux in children. *Helv Paediatr Acta* 1989;**43**:389–96.
85. Nayir A. Circumcision for the prevention of significant bacteriuria in boys. *Pediatr Nephrol* 2001;**16**:1129–34. <http://dx.doi.org/10.1007/s004670100044>
86. Franz M, Horl WH. Common errors in diagnosis and management of urinary tract infection. I. Pathophysiology and diagnostic techniques. *Nephrol Dial Transplant* 1999;**14**:2746–53. <http://dx.doi.org/10.1093/ndt/14.11.2746>
87. Kahlmeter G. An international survey of antimicrobial susceptibility of pathogens from uncomplicated urinary tract infections: ECO SENS project. *J Antimicrob Chemother* 2003;**51**:69–76. <http://dx.doi.org/10.1093/jac/dkg028>
88. Varga J. *Investigation of Urine BSOP 178*. National Health Service for Wales; 2008.
89. Vaillancourt S, McGillivray D, Zhang X, Kramer MS. To clean or not to clean: effect on contamination rates in midstream urine collections in toilet-trained children. *Pediatr* 2007;**119**:e1288–93. <http://dx.doi.org/10.1542/peds.2006-2392>
90. Feasey S. Are Newcastle urine collection pads suitable as a means of collecting specimens from infants? *Paediatr Nurs* 1999;**11**:17–21.
91. Rao S, Bhjatt J, Houghton C, Macfarlane P. An improved urine collection pad method: a randomised clinical trial. *Arch Dis Child* 2004;**89**:773–5. <http://dx.doi.org/10.1136/adc.2003.037770>
92. Alam MT, Coulter JBS, Pacheo J, Correla JB, Ribeiro MG, Coelho MF, et al. Comparison of urine contamination rates using three different methods of collection: clean-catch, cotton wool pad and urine bag. *Ann Trop Paediatr* 2005;**25**:29–34. <http://dx.doi.org/10.1179/146532805X23326>
93. Jackson SR, Dryden M, Gillett P, Kearney P, Weatherall R. A novel midstream urine-collection device reduces contamination rates in urine cultures amongst women. *BJU Int* 2005;**96**:360–4. <http://dx.doi.org/10.1111/j.1464-410X.2005.05631.x>
94. Blake DR, Doherty LF. Effect of perineal cleansing on contamination rate of mid-stream urine culture. *J Pediatr Adolesc Gynecol* 2006;**19**:31–4. <http://dx.doi.org/10.1016/j.jpag.2005.11.003>
95. Unlu H, Sardan YC, Ulker S. Comparison of sampling methods for urine cultures. *J Nurs Scholarsh* 2007;**39**:325–9. <http://dx.doi.org/10.1111/j.1547-5069.2007.00188.x>
96. Bekker LG, Jones BA, Walsh MK, Wagar EA. Urine culture contamination: a College of American Pathologists Q-Probes study of 127 laboratories. *Arch Pathol Lab Med* 2008;**132**:913–17. [http://dx.doi.org/10.1043/1543-2165\(2008\)132\[913:UCCACO\]2.0.CO;2](http://dx.doi.org/10.1043/1543-2165(2008)132[913:UCCACO]2.0.CO;2)
97. Wingerter SM, Bachur RM. Risk factors for contamination of catheterized urine specimens in febrile children. *Pediatr Emerg Care* 2011;**27**:1–4. <http://dx.doi.org/10.1097/PEC.0b013e3182037c20>
98. Tosif S, Baker A, Oakley E, Donath S, Babl FE. Contamination rates of different urine collection methods for the diagnosis of urinary tract infections in young children: an observational cohort study. *J Paediatr Child Health* 2012;**48**:659–64. <http://dx.doi.org/10.1111/j.1440-1754.2012.02449.x>
99. Valenstein P, Meier F. Urine culture contamination: a College of American Pathologists Q-Probes study of contaminated urine cultures in 906 institutions. *Arch Pathol Lab Med* 1998;**122**:123–9.



100. Choi WH, Lim IS. Factors affecting the contamination of bag urine culture in febrile children under two years. *Korean J Pediatr* 2009;**52**:346–50. <http://dx.doi.org/10.3345/kjp.2009.52.3.346>
101. Stamm WE, Counts GW, Fihn S, Turck M, Holmes KK, Running KR. Diagnosis of coliform infection in acutely dysuric women. *N Engl J Med* 1982;**307**:463–8. <http://dx.doi.org/10.1056/NEJM198208193070802>
102. Lau AY, Wong SN, Yip KT, Fong KW, Li SP, Que TL. A comparative study on bacterial cultures of urine samples obtained by clean-void technique versus urethral catheterization. *Acta Paediatr* 2007;**96**:432–6. <http://dx.doi.org/10.1111/j.1651-2227.2006.00146.x>
103. Petersen I, Hayward AC, Subgroup SS. Antibacterial prescribing in primary care. *J Antimicrob Chemother* 2007;**60**(Suppl. 1):i43–7. <http://dx.doi.org/10.1093/jac/dkm156>
104. Smith R, Coast J. The true cost of antimicrobial resistance. *BMJ* 2013;**346**:f1493. <http://dx.doi.org/10.1136/bmj.f1493>
105. Downing H, Thomas-Jones E, Gal M, Waldron CA, Sterne J, Hollingworth W, *et al.* 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* 2012;**12**:158. <http://dx.doi.org/10.1186/1471-2334-12-158>
106. Armon K, Stephenson T, Gabriel V, MacFaul R, Eccleston P, Werneke U, *et al.* Audit: determining the common medical presenting problems to an accident and emergency department. *Arch Dis Child* 2001;**84**:390–2. <http://dx.doi.org/10.1136/adc.84.5.390>
107. Zorc JJ, Levine DA, Platt SL, Dayan PS, Macias PG, Krief W, *et al.* Clinical and demographic factors associated with urinary tract infection in young febrile infants. *Pediatrics* 2005;**116**:644–8. <http://dx.doi.org/10.1542/peds.2004-1825>
108. Watson AR. Management of urinary tract infection in children. *BMJ* 2007;**335**:356–7. <http://dx.doi.org/10.1136/bmj.39309.423542.80>
109. 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**:19–22. <http://dx.doi.org/10.3109/02813432.2011.554268>
110. Fekkes M, Theunissen NCM, Brugman E, Veen S, Verrips EG, Koopman HM, *et al.* Development and psychometric evaluation of the TAPQOL: a health-related quality of life instrument for 1–5-year-old children. *Qual Life Res* 2000;**9**:961–72. <http://dx.doi.org/10.1023/A:1008981603178>
111. Bunge EM, Essink-Bot ML, Kobussen MPH, van Suijlekom-Smit LWA, Moll HA, Raat H. Reliability and validity of health status measurement by the TAPQOL. *Arch Dis Child* 2005;**90**:351–8. <http://dx.doi.org/10.1136/adc.2003.048645>
112. Howe RA, Andrews JM. BSAC standardized disc susceptibility testing method (version 11). *J Antimicrob Chemother* 2012;**67**:2783–4. <http://dx.doi.org/10.1093/jac/dks391>
113. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med* 1998;**17**:857–72. [http://dx.doi.org/10.1002/\(SICI\)1097-0258\(19980430\)17:8<857::AID-SIM777>3.0.CO;2-E](http://dx.doi.org/10.1002/(SICI)1097-0258(19980430)17:8<857::AID-SIM777>3.0.CO;2-E)
114. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. *JAMA* 1997;**277**:488–94. <http://dx.doi.org/10.1001/jama.1997.03540300056034>
115. Elm E, Altman D, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, *et al.* Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007;**335**:806–8. <http://dx.doi.org/10.1136/bmj.39335.541782.AD>

116. Hay AD, Whiting P, Butler CC. How best to diagnose urinary tract infection in preschool children in primary care? *BMJ* 2011;**343**:d6316. <http://dx.doi.org/10.1136/bmj.d6316>
117. Marple CD. The frequency and character of urinary tract infections in an unselected group of women. *Ann Intern Med* 1941;**14**:2220–39. <http://dx.doi.org/10.7326/0003-4819-14-12-2220>
118. Sanford JP, Favour CB, Mao FH. Evaluation of the 'positive' urine culture – an approach to the differentiation of significant bacteria from contaminants. *Am J Med* 1956;**20**:88–93. [http://dx.doi.org/10.1016/0002-9343\(56\)90175-9](http://dx.doi.org/10.1016/0002-9343(56)90175-9)
119. Gillespie WA, Linton KB, Miller A, Slade N. The diagnosis, epidemiology and control of urinary infection in urology and gynaecology. *J Clin Pathol* 1960;**13**:187–94. <http://dx.doi.org/10.1136/jcp.13.3.187>
120. Kass EH. Bacteriuria and the diagnosis of infections of the urinary tract. *Arch Intern Med* 1957;**100**:709–14. <http://dx.doi.org/10.1001/archinte.1957.00260110025004>
121. 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**:595–610. <http://dx.doi.org/10.1542/peds.2011-1330>
122. Cattell WR, Lefford MJ. Bacteriological examination of urine. *Br Med J* 1963;**1**:97–100. <http://dx.doi.org/10.1136/bmj.1.5323.97>
123. Cheetham P, Brown SE. Technique for the culture and direct sensitivity testing of large numbers of urine specimens. *J Clin Pathol* 1986;**39**:335–7. <http://dx.doi.org/10.1136/jcp.39.3.335>
124. Leigh DA, Williams JD. Method for the detection of significant bacteriuria in large groups of patients. *J Clin Pathol* 1964;**17**:498–503. <http://dx.doi.org/10.1136/jcp.17.5.498>
125. Miles AA, Misra SS, Irwin JO. The estimation of the bactericidal power of the blood. *J Hyg* 1938;**38**:732–49. <http://dx.doi.org/10.1017/S002217240001158X>
126. Hedges AJ, Shannon R, Hobbs RP. Comparison of the precision obtained in counting viable bacteria and the spiral plate maker, the droplette and the Miles and Misra methods. *J Appl Bacteriol* 1978;**45**:57–65. <http://dx.doi.org/10.1111/j.1365-2672.1978.tb04198.x>
127. Herreros Fernández ML, González Merino N, Tagarro Garcia A, Perez Seoane B, de la Serna Martínez M, Contreras Abad MT, et al. A new technique for fast and safe collection of urine in newborns. *Arch Dis Child* 2013;**98**:27–9. <http://dx.doi.org/10.1136/archdischild-2012-301872>
128. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. *JAMA* 2000;**284**:79–84. <http://dx.doi.org/10.1001/jama.284.1.79>
129. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ* 2003;**326**:41–4. <http://dx.doi.org/10.1136/bmj.326.7379.41>
130. Wood AM, White IR, Royston P. How should variable selection be performed with multiply imputed data? *Stat Med* 2008;**27**:3227–46. <http://dx.doi.org/10.1002/sim.3177>
131. Rubin D. Inference and missing data. *Biometrika* 1976;**63**:581–92. <http://dx.doi.org/10.1093/biomet/63.3.581>
132. 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**:774–81. [http://dx.doi.org/10.1016/S0895-4356\(01\)00341-9](http://dx.doi.org/10.1016/S0895-4356(01)00341-9)
133. Sullivan LM, Massaro JM, D'Agostino RB. Presentation of multivariate data for clinical use: the Framingham Study risk score functions. *Stat Med* 2004;**23**:1631–60. <http://dx.doi.org/10.1002/sim.1742>

134. Kinlin LM, Kirchner C, Zhang H, Fisman DN. Derivation and validation of a clinical prediction rule for nosocomial pneumonia after coronary artery bypass graft surgery. *Clin Infect Dis* 2010;**50**:493–501. <http://dx.doi.org/10.1086/649925>
135. de Salis I, Whiting P, Sterne JAC, Hay AD. Using qualitative research to inform development of a diagnostic algorithm for UTI in children. *Fam Pract* 2012;**30**:7. <http://dx.doi.org/10.1093/fampra/cms076>
136. European Confederation of Laboratory Medicine. European Urinalysis Guidelines: summary. *Scand J Clin Lab Invest* 2000;**60**:1–96.
137. Friedman S, Reif S, Assia A, Levy I. Clinical and laboratory characteristics of non-*E. coli* urinary tract infections. *Arch Dis Child* 2006;**91**:845–6. <http://dx.doi.org/10.1136/adc.2005.080721>
138. Paschke AA, Zaoutis T, Conway PH, Xie D, Keren R. Previous antimicrobial exposure is associated with drug-resistant urinary tract infections in children. *Pediatrics* 2010;**125**:664–72. <http://dx.doi.org/10.1542/peds.2009-1527>
139. Struthers S, Scanlon J, Parker K, Goddard J, Hallett R. Parental reporting of smelly urine and urinary tract infection. *Arch Dis Child* 2003;**88**:250–2. <http://dx.doi.org/10.1136/adc.88.3.250>
140. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999;**10**:37–48. <http://dx.doi.org/10.1097/00001648-199901000-00008>
141. Coulthard MG, Lambert HJ, Keir MJ. Do systemic symptoms predict the risk of kidney scarring after urinary tract infection? *Arch Dis Child* 2009;**94**:278–81. <http://dx.doi.org/10.1136/adc.2007.132290>
142. Downs SM. Technical report: urinary tract infections in febrile infants and young children. The Urinary Tract Subcommittee of the American Academy of Pediatrics Committee on Quality Improvement. *Pediatrics* 1999;**103**:e54. <http://dx.doi.org/10.1542/peds.103.4.e54>
143. Harmsen M, Adang EM, Wolters RJ, van der Wouden JC, Grol RP, Wensing M. Management of childhood urinary tract infections: an economic modeling study. *Value Health* 2009;**12**:466–72. <http://dx.doi.org/10.1111/j.1524-4733.2008.00477.x>
144. Hay AD, Fahey T, Peters TJ, Wilson A. Predicting complications from acute cough in pre-school children in primary care: a prospective cohort study. *Br J Gen Pract* 2004;**54**:9–14.
145. Falagas ME, Kotsantis IK, Vouloumanou EK, Rafailidis PI. Antibiotics versus placebo in the treatment of women with uncomplicated cystitis: a meta-analysis of randomized controlled trials. *J Infect* 2009;**58**:91–102. <http://dx.doi.org/10.1016/j.jinf.2008.12.009>
146. Christiaens TC, De Meyere M, Verschraegen G, Peersman W, Heytens S, De Maeseneer JM. Randomised controlled trial of nitrofurantoin versus placebo in the treatment of uncomplicated urinary tract infection in adult women. *Br J Gen Pract* 2002;**52**:729–34.
147. Hay AD, Heron J, Ness A, and the ALSPAC team. The prevalence of symptoms and consultations in pre-school children in the Avon Longitudinal Study of Parents and Children (ALSPAC): a prospective cohort study. *Fam Pract* 2005;**22**:367–74. <http://dx.doi.org/10.1093/fampra/cmi035>
148. Marild S, Jodal U. Incidence rate of first-time symptomatic urinary tract infection in children under 6 years of age. *Acta Paediatr* 1998;**87**:549–52. <http://dx.doi.org/10.1111/j.1651-2227.1998.tb01502.x>
149. Dickinson JA. Incidence and outcome of symptomatic urinary tract infection in children. *Br Med J* 1979;**1**:1330–2. <http://dx.doi.org/10.1136/bmj.1.6174.1330>
150. Dighe AM, Grace JF. General practice management of childhood urinary tract infection. *Br J Gen Pract* 1984;**34**:324–7.

151. Nagler EV, Williams G, Hodson EM, Craig JC. Interventions for primary vesicoureteric reflux. *Cochrane Database Syst Rev* 2011;**6**:CD001532. <http://dx.doi.org/10.1002/14651858.CD001532.pub4>
152. Arant BS Jr. Vesicoureteric reflux and renal injury. *Am J Kidney Dis* 1991;**17**:491–511. [http://dx.doi.org/10.1016/S0272-6386\(12\)80490-2](http://dx.doi.org/10.1016/S0272-6386(12)80490-2)
153. Jacobson SH, Hansson S, Jakobsson B. Vesico-ureteric reflux: occurrence and long-term risks. *Acta Paediatr Suppl* 1999;**88**:22–30. <http://dx.doi.org/10.1111/j.1651-2227.1999.tb01315.x>
154. Office for National Statistics. *Mortality in England and Wales: Average Life Span*. London: Office for National Statistics; 2012.
155. Mowatt G, Vale L, Perez J, Wyness L, Fraser C, MacLeod A, et al. Systematic review of the effectiveness and cost-effectiveness, and economic evaluation, of home versus hospital or satellite unit haemodialysis for people with end-stage renal failure. *Health Technol Assess* 2003;**7**(2). <http://dx.doi.org/10.3310/hta7020>
156. Kaufman DB. *Assessment and Management of the Renal Transplant Patient*. 2013. URL: <http://emedicine.medscape.com/article/429314-overview> (accessed 9 August 2013).
157. Alexander SR, Arbus GS, Butt KM, Conley S, Fine RN, Greifer I, et al. The 1989 report of the North American Pediatric Renal Transplant Cooperative Study. *Pediatr Nephrol* 1990;**4**:542–53. <http://dx.doi.org/10.1007/BF00869842>
158. Curtis L. *Unit Costs of Health and Social Care*. York: Personal Social Services Research Unit, University of York; 2011.
159. Department of Health. *Report of the Second Phase of the Review of NHS Pathology Services in England*. London: Department of Health; 2008.
160. Health and Social Care Information Centre. *Prescription Cost Analysis, England –2011*. Health and Social Care Information Centre; 2012.
161. Department of Health. *2010–11 Reference Costs*. London: Department of Health; 2011.
162. Baboolal K, McEwan P, Sondhi S, Spiewanowski P, Wechowski J, Wilson K. The cost of renal dialysis in a UK setting—a multicentre study. *Nephrol Dial Transplant* 2008;**23**:1982–9. <http://dx.doi.org/10.1093/ndt/gfm870>
163. Churchill DN, Torrance GW, Taylor DW, Barnes CC, Ludwin D, Shimizu A, et al. Measurement of quality of life in end-stage renal disease: the time trade-off approach. *Clin Invest Med* 1987;**10**:14–20.
164. Brisson M, Senecal M, Drolet M, Mansi JA. Health-related quality of life lost to rotavirus-associated gastroenteritis in children and their parents: a Canadian prospective study. *Pediatr Infect Dis J* 2010;**29**:73–5. <http://dx.doi.org/10.1097/INF.0b013e3181b41506>
165. Center for the Evaluation of Value and Risk in Health. *The Cost-Effectiveness Analysis Registry*. URL: [www.cearegistry.org](http://www.cearegistry.org) (accessed 29 November 2012).
166. Barry HC, Ebell MH, Hickner J. Evaluation of suspected urinary tract infection in ambulatory women: a cost-utility analysis of office-based strategies. *J Fam Pract* 1997;**44**:49–60.
167. Higgins JPT, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.0.2, updated September 2009. The Cochrane Collaboration; 2009. URL: [www.cochrane-handbook.org](http://www.cochrane-handbook.org) (accessed 9 October 2013).
168. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput* 2000;**10**:325–37. <http://dx.doi.org/10.1023/A:1008929526011>

169. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*. London: NICE; 2008.
170. Mangione-Smith R, Elliott MN, McDonald L, McGlynn EA. An observational study of antibiotic prescribing behavior and the Hawthorne effect. *Health Serv Res* 2002;**37**:1603–23. <http://dx.doi.org/10.1111/1475-6773.10482>
171. Macfarlane PI, Houghton C, Hughes C. Pad urine collection for early childhood urinary-tract infections. *Lancet* 1999;**35**:571. [http://dx.doi.org/10.1016/S0140-6736\(99\)02270-9](http://dx.doi.org/10.1016/S0140-6736(99)02270-9)
172. Turner GM, Coulthard MG. Fever can cause pyuria in children. *BMJ* 1995;**311**:924. <http://dx.doi.org/10.1136/bmj.311.7010.924>
173. Patel HP. The abnormal urinalysis. *Pediatr Clin North Am* 2006;**53**:325–37. <http://dx.doi.org/10.1016/j.pcl.2006.02.004>
174. Lam MH. False 'hematuria' due to bacteriuria. *Arch Pathol Lab Med* 1995;**119**:717–21.
175. Wilson EB. Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc* 1927;**22**:209–12. <http://dx.doi.org/10.1080/01621459.1927.10502953>
176. Rodriguez MJ, Rodriguez A, Maranon R. Gram stain as a predictor of urinary infections in children under 2 years. *Indian Pediatr* 2011;**48**:816–17.
177. Shaw KN, McGowan KL, Gorelick MH, Schwartz JS. Screening for urinary tract infection in infants in the emergency department: which test is best? *Pediatrics* 1998;**101**:e1. <http://dx.doi.org/10.1542/peds.101.6.e1>
178. Dayan PS, Bennett J, Best R, Bregstein JS, Levine D, Novick MK, et al. Test characteristics of the urine Gram stain in infants up to 60 days of age with fever. *Pediatr Emerg Care* 2002;**18**:12–14. <http://dx.doi.org/10.1097/00006565-200202000-00004>
179. Sharief N, Hameed M, Petts D. Use of rapid dipstick tests to exclude urinary tract infection in children. *Br J Biomed Sci* 1998;**55**:242–6.
180. Armengol CE, Hendley JO, Schlager TA. Should we abandon standard microscopy when screening for urinary tract infections in young children? *Pediatr Infect Dis J* 2001;**20**:1176–7. <http://dx.doi.org/10.1097/00006454-200112000-00018>
181. Lejeune B, Baron R, Guillois B, Mayeux D. Evaluation of a screening test for detecting urinary tract infection in newborns and infants. *J Clin Pathol* 1991;**44**:1029–30. <http://dx.doi.org/10.1136/jcp.44.12.1029>
182. Shaw KN, Hexter D, McGowan KL, Schwartz JS. Clinical evaluation of a rapid screening test for urinary tract infections in children. *J Pediatr* 1991;**118**:733–6. [http://dx.doi.org/10.1016/S0022-3476\(05\)80035-6](http://dx.doi.org/10.1016/S0022-3476(05)80035-6)
183. Ramlakhan SL, Burke DP, Goldman RS. Dipstick urinalysis for the emergency department evaluation of urinary tract infections in infants ages less than 2 years. *Eur J Emerg Med* 2011;**18**:221–4. <http://dx.doi.org/10.1097/MEJ.0b013e3283440e88>
184. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ* 2004;**329**:168–9. <http://dx.doi.org/10.1136/bmj.329.7458.168>
185. Joint Formulary Committee. *British National Formulary* (online). London: BMJ Group and Pharmaceutical Pres. URL: [www.medicinescomplete.com](http://www.medicinescomplete.com) (accessed 23 February 2015).
186. Automobile Association. *Motoring Costs 2011*. 2011. URL: [www.theaa.com/motoring\\_advice/running\\_costs/petrol2011.pdf](http://www.theaa.com/motoring_advice/running_costs/petrol2011.pdf) (accessed 23 February 2015).
187. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;**21**:271–92. [http://dx.doi.org/10.1016/S0167-6296\(01\)00130-8](http://dx.doi.org/10.1016/S0167-6296(01)00130-8)



# Appendix 1 National Institute for Health Research Health Technology Assessment brief



NIHR Health Technology Assessment programme

**NHS**  
National Institute for  
Health Research

## *The diagnosis of UTI in children under 2 years of age in primary care*

*HTA no 08/66*

### Introduction

The aim of the HTA programme is to ensure that high quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage, provide care in or develop policy for the NHS. Topics for research are identified and prioritised to meet the needs of the NHS. Health technology assessment forms a substantial portfolio of work within the National Institute for Health Research and each year about fifty new studies are commissioned to help answer questions of direct importance to the NHS. The studies include both primary research and evidence synthesis.

### Question

*Which clinical features of potential infection are useful in making a preliminary diagnosis of UTI in children less than 2 years of age and indicate the need for a urine specimen to be taken?*

- 1 **Technology:** Urine sampling methods and tests.
- 2 **Patient group:** Children under about 2 years of age.
- 3 **Setting:** Primary care and other settings in the community.
- 4 **Control or comparator:** Reference standard is microbiology and culture (or to be specified by researcher).
- 5 **Design:** Primary research to inform the development of a diagnostic algorithm and to determine which clinical features of infection (febrile, poor feeding, vomiting, irritability, etc.) indicate the need to undertake urine sampling for MC & S and whether urine dipstick testing contributes usefully to the decision making process.
- 6 **Primary outcomes/quality of life:** Develop and validate a decision aid to help in the identification of children at risk of UTI in primary care settings. Secondary outcomes: diagnostic accuracy, changes in patient management, timeliness of testing, cost-effectiveness.
- 7 **Minimum duration of follow-up:** To be specified and justified in the proposal.

### Background to commissioning brief:

*The question of when to test for UTI has been evaluated in a number of studies. The presence of specific symptoms of UTI, including dysuria, frequency, suprapubic discomfort, and flank pain, should lead to testing and investigation. However, young children with UTI may present with non-specific symptoms such as poor feeding, vomiting, irritability, jaundice (in newborns) or fever alone, and a broader approach to testing may be appropriate.*

*Studies in emergency departments suggest that up to 5% of children under the age of 2 presenting with fever have UTI, and over half of these would have been given alternative diagnoses such as otitis media had the urine not been tested as part of the study. Acute complications in young infants include systemic infection.*

*Urine collection methods include: clean catch, suprapubic aspiration, catheterization, pads and sterile urine bag. Suprapubic aspiration and catheterization are considered better methods to minimize contamination, but are not done in primary care, are painful and invasive. Urinalysis is done with a dipstick to determine levels of different constituent in the urine, e.g. blood, protein, white cells, leukocyte esterase and nitrites.*





## Appendix 2 Systematic review (update) for the DUTY study: accuracy of symptoms and signs and dipstick tests for diagnosing UTI in children < 5 years old in primary care and choice of urine sampling method

### Methods

A HTA-funded review that evaluated all tests for the diagnosis of UTI in children was published in 2006; we updated the searches conducted for this review to inform a publication on diagnosing UTI in children published in the *British Medical Journal (BMJ)* in 2011.<sup>56,116</sup> In 2007, NICE published clinical guidelines on the diagnosis, treatment and long-term management of UTI in children.<sup>2</sup> We conducted a systematic literature search to identify studies published since these reports.<sup>2,56,116</sup> We searched MEDLINE from 2010 (the end date of the *BMJ* paper searches) to January 2013 using a sensitive search strategy combining terms related to UTI with terms relating to 'clinical signs and symptoms', 'dipstick testing' or 'urine sampling'. We aimed to find evidence to address three questions on the diagnosis of UTI in children aged less than 5 years in primary care:

1. What is the accuracy of clinical signs and symptoms?
2. What is the accuracy of the combination of nitrite and LE dipstick tests?
3. What is the evidence on nappy pad or nappy bag urine samples compared with SPA or catheter samples?

We included primary studies or systematic reviews that addressed one of the above questions, used urine culture as the reference standard, enrolled children aged 5 years or less, and were conducted in a primary care or emergency department setting in the developed world. Full details of the inclusion criteria are summarised in *Table 83*. We also included relevant studies from our previous reviews,<sup>56,116</sup> and recommendations from the NICE guidelines.<sup>2</sup>

**TABLE 83** Inclusion criteria

Question	Clinical signs and symptoms	Dipstick tests	Urine sampling
Population	Children aged < 5 years		
Index test	Any clinical sign, symptom, or combination of signs/symptoms/tests	Nitrite or LE dipstick test	Nappy or bag sample
Reference standard	Culture	Culture	CVU, catheter or SPA sample
Target condition	UTI	UTI	Contaminated sample
Study design	Diagnostic cohort		
Setting	Primary care or ED		

ED, emergency department.

## Results

The update literature searches identified 368 hits; of these 11 were identified as being potentially relevant and full texts were obtained. Five studies and one systematic review fulfilled the inclusion criteria. The systematic review evaluated dipstick testing in children with suspected UTI and examined the relationship with age.<sup>64</sup> Two studies were retrospective cohort studies that assessed dipstick testing in children aged less than 2 years presenting to paediatric emergency departments,<sup>176</sup> two were prospective diagnostic cohort studies and assessed clinical features for the prediction of UTI.<sup>67</sup> The final prospective cohort study assessed a novel device which combined a dipstick test for nitrite and LE with a urine collection pad and was designed to be inserted into the nappy.<sup>62</sup> Combined with the studies identified by our previous reviews<sup>56,116</sup> this gives a total of six included studies for urine sampling, five primary studies and one systematic review for clinical signs and symptoms, and eight primary studies and one systematic review for dipstick testing (Figure 34).

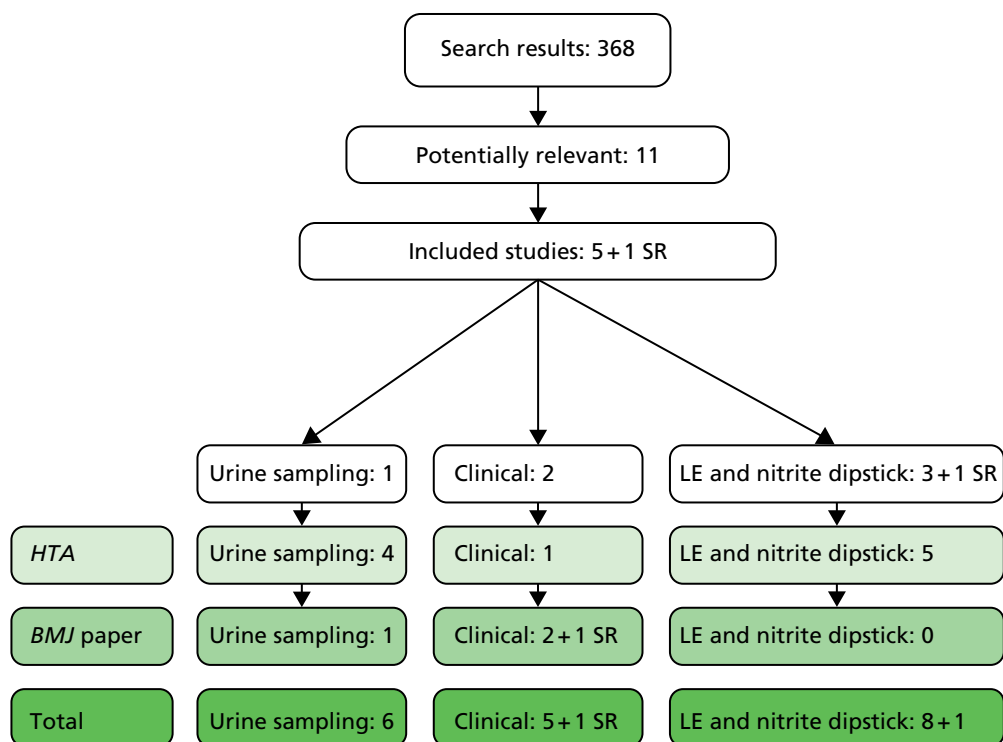


FIGURE 34 Flow of studies through the review. SR, systematic review.

## Urine sampling (six primary studies)

We identified a total of six studies that assessed urine sampling methods: four from the HTA review,<sup>57-60</sup> one from the BMJ paper<sup>61</sup> and one from our update searches.<sup>62</sup>

### Bag specimens

Two studies compared culture of urine bag specimens with culture of SPA samples.<sup>57,59</sup> There were considerable differences in the results from these studies with one reporting a sensitivity of 100% and the other of 50%. Both studies found specificity to be around 90%. Two studies compared culture results from urine samples obtained by bag specimens with those obtained by catheter.<sup>58,61</sup> One reported sensitivity of 81% and specificity of 87%;<sup>58</sup> the other reported sensitivity of 88% and specificity of 80%.<sup>61</sup> The appropriateness of a catheter specimen as the reference standard is questionable, and this, along with the small number of data available, means that these results are of limited value.

### Nappy pads

One study compared culture of a pad/nappy specimen with culture of SPA samples. This study reported a sensitivity of 100% and specificity of 94% suggesting excellent agreement between the two sampling methods.<sup>60</sup> A further recently published study identified by our update searches assessed a device known as the 'U-test' which is a nappy pad incorporating a urine dipstick.<sup>62</sup> The accuracy results are therefore a combination of the nappy pad and the dipstick but show good accuracy with a sensitivity of 100% and specificity of 79%. However, these were compared with a reference standard consisting of a variety of urine collection methods (clean catch, bag, catheter or SPA) and the study only had results available for 25 participants.

The NICE guidelines found 'insufficient data to draw conclusions about urine collection bags and urine collection pads'. There have only been two studies published since these guidelines and these do not provide sufficiently strong data to change these conclusions although the very limited data suggest that pad specimens may be a more accurate method of urine collection than bag specimens.

## Clinical signs and symptoms (five primary studies and one systematic review)

The update searches identified two new primary studies.<sup>9,67</sup> This gave a total of five primary studies<sup>1,65-67</sup> ( $n = 18\,390$ ) and one systematic review<sup>3</sup> that included eight primary studies in children aged < 5 years ( $n = 7892$ ). Of these 13 studies, 10 were conducted in hospital emergency departments, and two in paediatricians' offices; only one was conducted in GP practices. *Table 83* shows positive and negative LRs for each clinical sign and symptom together with estimates of the post-test probability of disease for the presence and absence of each symptom based on the prevalences (pre-test probabilities) of UTI of 2% and 5.9% seen in the DUTY research laboratory and O'Brien<sup>9</sup> study, respectively.

These data show that no individual, or any combination of, symptom/s or sign/s were sufficient to rule in a diagnosis of UTI, though some post-test probabilities (e.g. 25% for increased capillary refill time, no fluid intake and supra-pubic tenderness) appear high enough to mandate urine testing and empirical treatment while awaiting culture confirmation. A number of symptoms and signs did not appear to have diagnostic value, including some recommended for the diagnosis of UTI by NICE<sup>2</sup> (e.g. poor feeding and vomiting). Some symptoms, signs (for example respiratory) and proposed clinical prediction rules did reduce the probability of UTI to below 2% (given a pre-test probability of 6%) and these may be considered low enough to rule out UTI and avoid the need to obtain urine. These are summarised in *Table 84*.

TABLE 84 Diagnostic utility of symptoms, signs and urine dipstick testing for UTI in children &lt; 5 years

Clinical signs/symptoms	Age (months)	Number of studies	Number of children	Feature present		Feature absent			
				LR (95% CI)	Post-test probability based on 2% <sup>a</sup>	LR (95% CI)	Post-test probability based on 6% <sup>a</sup>		
<b>Individual signs and symptoms</b>									
Capillary refill time > 3 seconds	0-60	1	15,781	4.8 (2.2 to 10.6)	8.9%	23.5%	0.99 (0.98 to 1.00)	2.0%	5.9%
No fluid intake	0-60	1	15,781	4.4 (1.7 to 11.2)	8.2%	21.9%	0.99 (0.98 to 1.00)	2.0%	5.9%
Urinary symptoms/ increased frequency	0-60	1	15,781	4.4 (3.2 to 6.1)	8.2%	21.9%	0.94 (0.91 to 0.96)	1.9%	5.7%
	0-60	1	597	2.4 (1.4 to 4.0)	4.6%	13.2%	0.79 (0.63 to 0.99)	1.6%	4.8%
Supra-pubic tenderness	0-24	1 (SR)	2411	4.4 (1.6 to 12.4)	8.2%	21.9%	0.96 (0.90 to 1.01)	1.9%	5.8%
Age < 3 months	0-60	1	15,781	3.9 (3.2 to 4.8)	7.4%	19.9%	0.87 (0.83 to 0.90)	1.7%	5.3%
Temperature > 40°C	0-12	1 (SR)	945	3.3 (1.3 to 8.3)	6.3%	17.4%	0.66 (0.35 to 1.25)	1.3%	4.0%
	0-3	1 (SR)	182	3.2 (0.7 to 15.6)	6.1%	17.0%	0.93 (0.80 to 1.08)	1.9%	5.6%
	0-60	1	15,781	1.1 (0.9 to 1.4)	2.2%	6.6%	0.99 (0.95 to 1.02)	2.0%	5.9%
Circumcised males	0-24	6 (SR)	6835	0.33 (0.18 to 0.63)	0.7%	2.1%	2.8 (1.9 to 4.3)	5.4%	15.2%
	0-36	1	331	0.07 (0.00 to 1.16)	0.1%	0.4%	1.1 (1.1 to 1.2)	2.2%	6.6%
History of prior UTI	0-24	1 (SR)	2411	2.9 (1.2 to 7.1)	5.6%	15.6%	0.95 (0.89 to 1.02)	1.9%	5.7%
	0-12	1 (SR)	945	2.3 (0.3 to 17.4)	4.5%	12.8%	0.97 (0.89 to 1.07)	1.9%	5.8%
Wetting when previously dry	0-60	1	597	2.5 (1.0 to 6.0)	4.5%	13.0%	0.91 (0.79 to 1.04)	1.8%	5.5%
Bulging fontanelle	0-60	1	15,781	2.3 (0.8 to 6.3)	4.5%	12.8%	0.99 (0.99 to 1.00)	2.0%	5.9%
Prolonged fever (> 24 hours)	0-3	1 (SR)	1608	2.0 (1.4 to 2.9)	3.9%	11.3%	0.90 (0.83 to 0.97)	1.8%	5.4%

Clinical signs/symptoms	Age (months)	Number of studies	Number of children	Feature present		Feature absent			
				Post-test probability based on 2% <sup>a</sup>	Post-test probability based on 6% <sup>a</sup>	LR (95% CI)	LR (95% CI)		
Parent report of smelly urine	0-36	1	331	3.5%	10.3%	1.8 (1.3 to 2.4)	0.63 (0.46 to 0.88)	1.3%	3.9%
Painful urination	0-36	1	331	3.5%	10.3%	1.8 (1.0 to 3.1)	0.87 (0.74 to 1.02)	1.7%	5.3%
Respiratory symptom	0-60	1	597	5.4%	15.2%	3.1 (1.3 to 7.6)	0.90 (0.78 to 1.03)	1.8%	5.5%
Ill appearance	0-60	1	15,781	1.4%	4.2%	0.68 (0.63 to 0.74)	1.9 (1.7-2.0)	3.7%	10.8%
	0-2	1(SR)	1005	1.2%	3.6%	0.59 (0.22 to 1.6)	1.03 (0.99 to 1.08)	2.1%	6.2%
	0-3	1(SR)	1608	2.2%	6.6%	1.1 (0.9 to 1.3)	0.95 (0.84 to 1.08)	1.9%	5.7%
	3-24	1(SR)	2411	3.7%	10.8%	1.9 (1.5 to 2.4)	0.68 (0.53 to 0.88)	1.4%	4.2%
Chronic disease	0-60	1	15,781	3.4%	9.8%	1.7 (1.4 to 2.1)	0.93 (0.90 to 0.97)	1.9%	5.6%
Cough	0-60	1	15,781	3.2%	9.3%	1.6 (1.3 to 1.9)	0.92 (0.88 to 0.96)	1.8%	5.5%
Non-black race	0-60	1	15,781	1.1%	3.5%	0.57 (0.50 to 0.65)	1.5 (1.4 to 1.6)	3.0%	8.7%
No other source of fever	0-24	6 (SR)	6858	2.8%	8.2%	1.4 (1.1 to 1.8)	0.52 (0.29 to 0.73)	1.1%	3.2%
Temperature > 39 °C	0-24	3 (SR)	4964	2.8%	8.2%	1.4 (1.1 to 1.8)	0.69 (0.55 to 0.80)	1.4%	4.2%
	0-24	4 (SR)	5969	2.8%	8.2%	1.4 (1.2 to 1.7)	0.78 (0.65 to 0.81)	1.6%	4.7%
Meningococcal unvaccinated	0-60	1	15,781	2.6%	7.7%	1.3 (1.2 to 1.4)	0.80 (0.73 to 0.88)	1.6%	4.9%
Crying	0-60	1	15,781	2.8%	8.2%	1.4 (1.3 to 1.5)	0.60 (0.52 to 0.69)	1.2%	3.7%
Normal ENT examination	0-60	1	15,781	2.8%	8.2%	1.4 (1.3 to 1.5)	0.81 (0.75 to 0.87)	1.6%	4.9%
Pneumococcal unvaccinated	0-60	1	15,781	2.6%	7.7%	1.3 (1.1 to 1.3)	0.67 (0.60 to 0.75)	1.3%	4.1%
Female	0-60	1	15,781	2.6%	7.7%	1.3 (1.1 to 1.3)	0.83 (0.74 to 0.93)	1.7%	5.0%

continued

TABLE 84 Diagnostic utility of symptoms, signs and urine dipstick testing for UTI in children &lt; 5 years (continued)

Clinical signs/symptoms	Age (months)	Number of studies	Number of children	Feature present		Feature absent			
				LR (95% CI)	Post-test probability based on 2% <sup>a</sup>	Post-test probability based on 6% <sup>a</sup>	LR (95% CI)	Post-test probability based on 2%	Post-test probability based on 6%
Prolonged fever > 48 hours	0–24	1 (SR)	2411	1.3 (0.8 to 1.9)	2.6%	7.7%	0.95 (0.85 to 1.06)	1.9%	5.7%
Prolonged fever > 72 hours	0–36	1	331	1.4 (1.0 to 1.9)	2.8%	8.2%	0.79 (0.60 to 1.05)	1.6%	4.8%
Elevated heart rate	0–60	1	15,781	1.2 (1.1 to 1.3)	2.6%	7.7%	0.84 (0.77 to 0.92)	1.7%	5.1%
Age > 3 years	0–60	1	15,781	1.2 (1.1 to 1.2)	2.4%	7.1%	0.47 (0.37 to 0.61)	1.0%	2.9%
Temperature > 38°C	0–60	1	15,781	1.1 (1.1 to 1.2)	2.2%	6.6%	0.56 (0.44 to 0.70)	1.1%	3.5%
Irritable/grouchy	0–60	1	597	1.5 (0.9 to 2.2)	2.9%	8.6%	0.81 (0.60 to 1.08)	1.6%	4.9%
Muscle aches or pains	0–36	1	331	1.1 (1.0 to 1.3)	2.2%	6.6%	0.12 (0.00 to 1.95)	0.2%	0.8%
Abdominal pain	0–36	1	331	1.3 (1.1 to 1.5)	2.5%	7.5%	0.54 (0.28 to 1.06)	1.1%	3.3%
Dysuria	0–60	1	597	0.14 (0.0 to 2.2)	0.3%	0.9%	1.1 (1.0 to 1.1)	2.2%	6.6%
Infectious contacts	0–36	1	331	0.91 (0.64 to 1.3)	1.8%	5.5%	1.1 (0.84 to 1.14)	2.2%	6.6%
Diarrhoea	0–36	1	331	0.90 (0.50 to 1.6)	1.8%	5.4%	1.0 (0.89 to 1.29)	2.0%	6.0%
	0–60	1	15,781	0.72 (0.61 to 0.86)	1.4%	4.4%	1.1 (1.1 to 1.2)	2.2%	6.6%
	0–12	1 (SR)	945	0.64 (0.32 to 1.26)	1.3%	3.9%	1.3 (1.0 to 1.7)	2.6%	7.7%
	0–60	1	15,781	0.83 (0.71 to 0.98)	1.7%	5.0%	1.1 (1.0 to 1.1)	2.2%	6.6%
Rash	0–36	1	331	1.1 (0.69 to 1.85)	2.2%	6.6%	0.96 (0.80 to 1.15)	1.9%	5.8%
Audible wheeze	0–60	1	15,781	0.71 (0.56 to 0.88)	1.4%	4.3%	1.1 (1.0 to 1.1)	2.2%	6.6%
Abnormal chest sounds	0–60	1	15,781	0.25 (0.13 to 0.48)	0.5%	1.6%	1.1 (1.0 to 1.1)	2.2%	6.6%
Chest crackles	0–60	1	15,781	0.43 (0.31 to 0.58)	0.9%	2.7%	1.1 (1.1 to 1.1)	2.2%	6.6%
Breathing difficulty	0–60	1	15,781	0.46 (0.30 to 0.70)	0.9%	2.9%	1.1 (1.0 to 1.1)	2.2%	6.6%
	0–60	1	15,781	0.48 (0.35 to 0.66)	1.0%	3.0%	1.1 (1.1 to 1.1)	2.2%	6.6%

Clinical signs/symptoms	Age (months)	Number of studies	Number of children	Feature present		Feature absent			
				LR (95% CI)	Post-test probability based on 2% <sup>a</sup>	Post-test probability based on 6% <sup>a</sup>	LR (95% CI)	Post-test probability based on 2%	Post-test probability based on 6%
Vomiting	0-3	1(SR)	1608	0.85 (0.58 to 1.23)	1.7%	5.1%	1.0 (1.0 to 1.1)	2.0%	6.0%
	0-12	1(SR)	945	1.1 (0.6 to 1.9)	2.2%	6.6%	1.0 (0.7 to 1.4)	2.0%	6.0%
	0-36	1	331	0.82 (0.53 to 1.27)	1.6%	5.0%	1.1 (0.9 to 1.4)	2.2%	6.6%
Felt hot	0-60	1	15,781	1.0 (1.0 to 1.1)	2.0%	6.0%	0.47 (0.31 to 0.73)	1.0%	2.9%
Focal bacterial infection	0-60	1	15,781	0.67 (0.48 to 0.93)	1.3%	4.1%	1.0 (1.0 to 1.1)	2.0%	6.0%
Stridor	0-60	1	15,781	0.20 (0.05 to 0.81)	0.4%	1.3%	1.0 (1.0 to 1.0)	2.0%	6.0%
Poor feeding	0-3	1 (SR)	1608	0.98 (0.8 to 1.2)	2.0%	5.9%	1.0 (0.9 to 1.1)	2.0%	6.0%
	0-12	1 (SR)	945	1.00 (0.72 to 1.39)	2.0%	6.0%	1.0 (0.5 to 1.8)	2.0%	6.0%
Irritability	0-60	1	597	1.26 (1.00 to 1.60)	2.5%	7.4%	0.69 (0.42 to 1.13)	1.4%	4.2%
	0-12	1 (SR)	945	0.94 (0.7 to 1.2)	1.9%	5.7%	1.3 (0.52 to 3.0)	2.5%	7.4%
<b>Clinical prediction rules</b>									
1. Temperature > 39 °C, > 48 hours, no apparent sources of fever	0-24	1 (SR)	2411	4.0 (1.2 to 13.0)	7.5%	20.3%	Not reported	Not reported	Not reported
2. Temperature < 39 °C with a potential source of fever	0-12	1 (SR)	945	Not reported	Not reported	Not reported	0.37 (0.16 to 0.85)	0.7%	2.3%
3. < 12 months, white race, temperature ≥ 39 °C, no apparent source of fever, fever > 48 hours	0-24	2	1681	1.4 (1.3 to 1.5)	2.8%	8.2%	0.15 (0.05 to 0.46)	0.3%	0.9%
				2.9 (2.2 to 3.9)	5.6%	15.6%	0.17 (0.10 to 0.46)	0.3%	1.1%

continued

**TABLE 84** Diagnostic utility of symptoms, signs and urine dipstick testing for UTI in children < 5 years (continued)

Clinical signs/symptoms	Age (months)	Number of studies	Number of children	Feature present		Feature absent			
				LR (95% CI)	Post-test probability based on 2% <sup>a</sup>	Post-test probability based on 6% <sup>a</sup>	LR (95% CI)	Post-test probability based on 2%	Post-test probability based on 6%
<b>Dipstick testing</b>									
Leucocyte esterase and nitrite positive	0-2	1	145	6.2 (1.1 to 34.2)	11.3%	28.5%	0.88 (0.71 to 1.1)	1.8%	5.3%
	0-18	1	243	36.7 (15.3 to 88.0)	42.9%	70.1%	0.11 (0.04 to 0.28)	0.2%	0.7%
	0-12	1	195	107.7 (6.3 to 1844)	68.7%	87.3%	0.69 (0.52 to 0.92)	1.4%	4.2%
	0-24	1	321	7.5 (4.0 to 14.0)	13.3%	32.4%	0.66 (0.56 to 0.79)	1.3%	4.1%
	0-24	1	980	27.4 (10.2 to 73.4)	35.9%	63.6%	0.75 (0.71 to 0.79)	1.5%	4.6%
	0-60	1	25	5.8 (1.8 to 17.9)	10.5%	26.8%	0.05 (0.00 to 0.76)	0.1%	0.3%
	POOLED	6		22.8 (11.1 to 46.5)	31.8%	59.3%	0.46 (0.19 to 1.13)	0.9%	2.9%
	0-12	1	193	10.5 (6.2 to 17.9)	17.6%	40.1%	0.16 (0.06 to 0.46)	0.3%	1.0%
	0-24	1	321	1.8 (1.6 to 2.1)	3.5%	10.3%	0.18 (0.09 to 0.37)	0.4%	1.1%
	0-24	1	980	4.7 (3.8 to 5.8)	8.8%	23.1%	0.24 (0.20 to 0.29)	0.5%	1.5%
LE or nitrite positive	0-24	1	145	9.4 (4.7 to 18.5)	16.1%	37.5%	0.31 (0.14 to 0.71)	0.6%	1.9%
	0-24	1	3394	72.6 (50.6 to 104.2)	59.7%	82.3%	0.28 (0.20 to 0.38)	0.6%	1.8%
	0-60	1	260	32.2 (13.1 to 79.0)	39.7%	67.3%	0.31 (0.18 to 0.53)	0.6%	1.9%
	POOLED	6	5293	10.5 (3.4 to 32.2)	17.6%	40.1%	0.22 (0.16 to 0.30)	0.4%	1.4%

SR, study was included in the systematic review by Shaikh *et al.*<sup>3</sup>  
 a Pre-test probabilities selected based on O'Brien *et al.* (6%)<sup>9</sup> and the DUTY research laboratory (2%).  
 Data from the table come from a systematic review by Shaikh *et al.* (2007)<sup>3</sup> and primary studies by Craig *et al.* (2010),<sup>1</sup> Gorelick and Shaw (2000)<sup>65</sup> and Gorelick *et al.* (2003),<sup>66</sup> Gauthier *et al.* (2012),<sup>67</sup> (clinical signs and symptoms), Armengol *et al.* (2001),<sup>180</sup> Dayan *et al.* (2002),<sup>178</sup> Lejeune *et al.* (1991),<sup>181</sup> Shaw *et al.* (1998)<sup>177</sup> and Shaw *et al.* (1991),<sup>182</sup> Krahenbuehel *et al.* (2012),<sup>62</sup> Ramlakhan *et al.* (2011)<sup>183</sup> and Rodriguez *et al.* (2011)<sup>176</sup> (dipstick testing).



TABLE 85 Symptoms and signs and dipstick test results helpful for ruling out UTI in children &lt; 5 years

Clinical signs/symptoms	Age (years)	Number of studies	Number of children	Prevalence of sign/symptom <sup>a</sup>	LR (95% CI) <sup>b</sup>	Post-test probability based on 2% prevalence <sup>c</sup>	Post-test probability based on 6% prevalence <sup>c</sup>
Two or less of the following: < 12 months, white race, temperature $\geq 39^{\circ}\text{C}$ , no apparent source of fever, fever > 48 hours	0–2	2	1681	NA	0.15 (0.05 to 0.46) <sup>65</sup> 0.17 (0.10 to 0.46) <sup>66</sup>	0.3%	0.9%
Stridor	0–5	1 <sup>1</sup>	15,781	1.8%	0.20 (0.05 to 0.81)	0.4%	1.3%
Audible wheeze	0–5	1 <sup>1</sup>	15,781	6.4%	0.25 (0.13 to 0.48)	0.5%	1.6%
Circumcised males	0–2	6 (SR) <sup>3</sup>	6835	NA	0.33 (0.18 to 0.63)	0.7%	2.1%
Temperature < 39 °C with a potential source of fever	0–3	1	331	NA	0.07 (0.00 to 1.16)	0.1%	0.4%
	0–1	1 (SR) <sup>3</sup>	945	NA	0.37 (0.16 to 0.85)	0.7%	2.3%
Abnormal chest sounds	0–5	1 <sup>1</sup>	15,781	15.7%	0.43 (0.31 to 0.58)	0.9%	2.7%
Chest crackles	0–5	1 <sup>1</sup>	15,781	8.3%	0.46 (0.30 to 0.70)	0.9%	2.9%
Age $\leq 3$ years	0–5	1 <sup>1</sup>	15,781	78.5%	0.47 (0.37 to 0.61)	1.0%	2.9%
Did not feel hot	0–5	1 <sup>1</sup>	15,781	7.7%	0.47 (0.31 to 0.73)	1.0%	2.9%
Breathing difficulty	0–5	1 <sup>1</sup>	15,781	13.7%	0.48 (0.35 to 0.66)	1.0%	3.0%
Dipstick: LE and nitrite negative	0–5	6	5293	NA	0.22 (0.16 to 0.30)	0.4%	1.4%

NA, not available; SR, study was included in the systematic review by Shaikh *et al.*<sup>3</sup>

<sup>a</sup> Data on the prevalence of the individual features is taken from the study by Craig *et al.*<sup>1</sup> which enrolled children presenting to an emergency department with a fever. Data on the prevalence of features not assessed in this study were not available. It should be noted that the prevalence of these features is likely to vary according to setting.

<sup>b</sup> The further the LR is away from 1, the stronger the evidence for the absence of disease. A LR of < 0.1 is considered to provide strong evidence to rule out a UTI.<sup>184</sup>

<sup>c</sup> Pretest probabilities selected based on O'Brien *et al.* (6%)<sup>9</sup> and the DUTY research laboratory (2%).

The largest study, which included almost 16,000 children,<sup>1</sup> derived a clinical prediction rule based on a combination of 27 signs and symptoms. Results were not reported for specific thresholds, but the model was found to have an AUROC of 0.80 (95% CI 0.78 to 0.82) leading the authors to conclude that a computer-assisted diagnostic decision tool based on this model could improve decision-making in the emergency department.

### Leucocyte esterase and nitrite dipstick testing (eight primary studies and one systematic review)

The systematic review included six primary studies; only three of these reported data separately for children aged < 5 years. Two of these were also included in our review,<sup>177,178</sup> and one was excluded<sup>179</sup> as it was based on secondary care. The review is therefore not considered further and this synthesis focuses instead on the eight primary studies.

Six studies evaluated the combination of both LE and nitrite positive (*Figure 35*). These data showed substantial heterogeneity in estimates of sensitivity which varied from 14% to 100%; specificity was more homogeneous varying from 86% to 100%. Negative LRs were too heterogeneous to permit conclusions regarding the utility of this combination for ruling out a diagnosis of UTI ranging from < 0.01 to 0.88 with a pooled estimate of 0.46 (95% CI 0.18 to 1.13). Positive LRs ranged from 6 to 108 with a pooled estimate of 22.8 (95% CI 11.1 to 46.5) suggesting that a dipstick positive for both LE and nitrite may be useful for ruling in a diagnosis of UTI.

Data were also heterogeneous for the combination of either nitrite or LE positive (*Figure 36*). Specificity ranged from 50% to 99% and sensitivity ranged from 70% to 91%. Positive LRs ranged from 1.8 to 73 with a pooled estimate of 10.5 (95% CI 3.4 to 32.2) making it difficult to draw conclusions regarding the utility of this combination for ruling in a diagnosis of UTI. Negative LRs ranged from 0.16 to 0.32 with a pooled estimate of 0.22 (95% CI 0.16 to 0.30) suggesting that a dipstick negative for both nitrite and LE may be useful in ruling out a diagnosis of UTI.

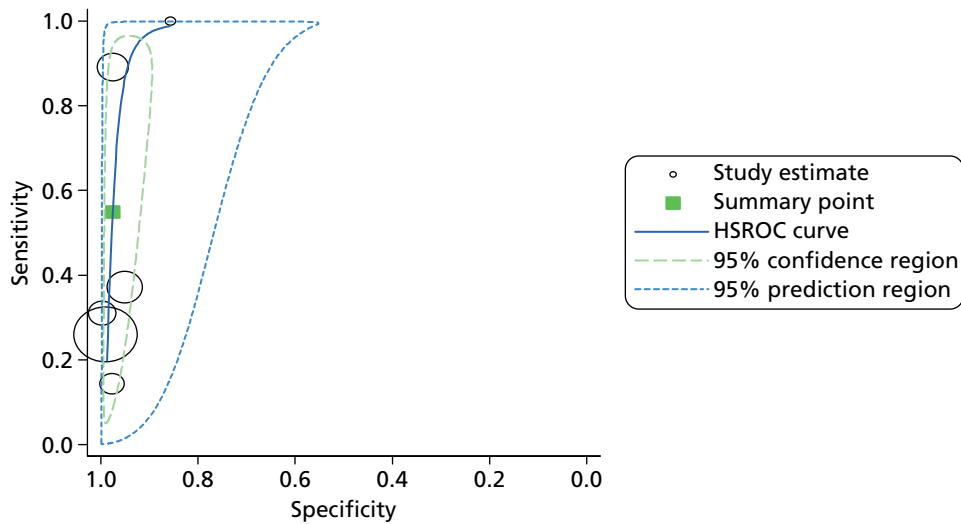
Overall the data were too heterogeneous to draw firm conclusions regarding the accuracy of dipstick testing, however, the data suggest that a dipstick positive for both nitrite and LE may be useful for ruling in a diagnosis of UTI, while dipstick negative for both nitrite and LE may be useful for ruling out a UTI. The NICE guidelines stated that 'further investigation of leucocyte esterase and nitrite dipstick tests alone and in combination, stratified by age and method of urine collection, is required to determine their accuracy in diagnosing UTI.'<sup>2</sup>

## Conclusions

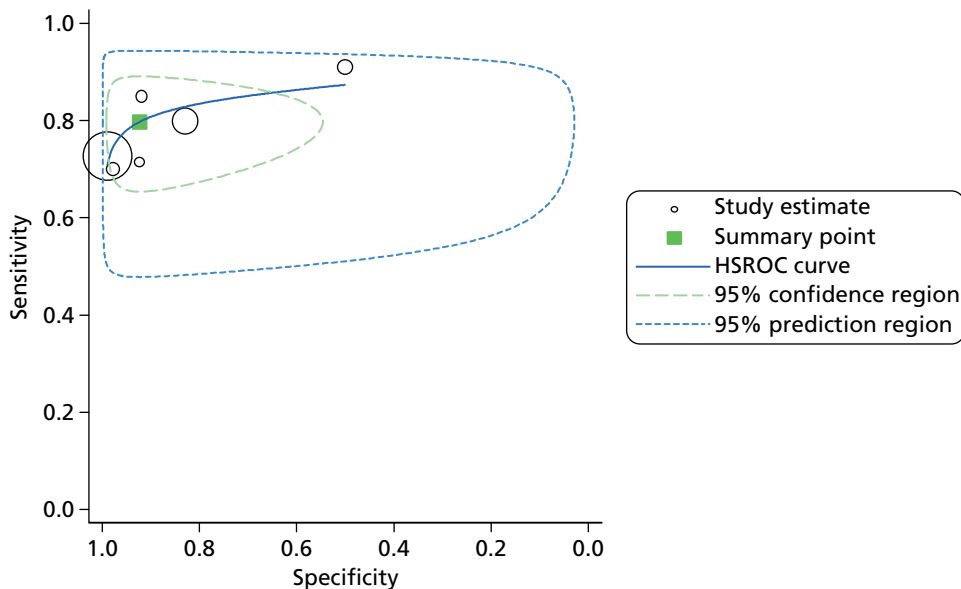
There were insufficient data to draw firm conclusions about urine collection bags and urine collection pads although limited data suggest that pad specimens may be a more accurate method of urine collection than bag specimens.

Most previous studies on clinical signs and symptoms for diagnosing UTIs in children were conducted in hospital EDs; none were conducted in primary care. No individual, or any combination of, symptoms or signs were sufficient to rule in a diagnosis of UTI, though some post-test probabilities appear high enough to mandate urine testing and empirical treatment while awaiting culture confirmation. A number of symptoms and signs did not appear to have diagnostic value. Some symptoms, signs (for example respiratory) and proposed clinical prediction rules reduced the probability of UTI to a threshold that may be considered low enough to rule out UTI and avoid the need to obtain urine.

Heterogeneous data on dipstick testing suggest that a dipstick positive for both nitrite and LE may be useful for ruling in a diagnosis of UTI, while a dipstick negative for both nitrite and LE may be useful for ruling out a UTI.



**FIGURE 35** Summary ROC plot showing estimates of sensitivity and specificity from studies assessing dipstick positive for both nitrite and LE together with summary ROC curve, summary estimate and 95% confidence and prediction regions. HSROC, hierarchical summary receiver operating characteristic.



**FIGURE 36** Summary ROC plot showing estimates of sensitivity and specificity from studies assessing dipstick positive for either nitrite and LE together with summary ROC curve, summary estimate and 95% confidence and prediction regions. HSROC, hierarchical summary receiver operating characteristic.

## Primary study data extraction tables

### Urine sampling

Study details	Study design and setting	Population	Index test	Reference standard	Sensitivity	Specificity
Cohen <i>et al.</i> (1997) <sup>60</sup>	<b>Study design:</b> prospective <b>Country:</b> Israel <b>Setting:</b> primary care	<b>Number (number girls):</b> 38 (24) <b>Age:</b> < 2 years <b>Urine sampling:</b> catheter; SPA; nappy <b>Patient spectrum:</b> query UTI	Pad vs.	SPA	100	94
Hardy <i>et al.</i> (1976) <sup>57</sup>	<b>Study design:</b> prospective <b>Country:</b> UK <b>Setting:</b> secondary care	<b>Number (number girls):</b> 30 (10) <b>Age:</b> < 5 years <b>Urine sampling:</b> clean catch; bag; SPA <b>Patient spectrum:</b> query UTI	Bag vs.	SPA	50	92
Benito Fernandez <i>et al.</i> (1996) <sup>59</sup>	<b>Study design:</b> prospective <b>Country:</b> Spain <b>Setting:</b> secondary care	<b>Number (number girls):</b> 61 [48 included in analysis (20)] <b>Age:</b> < 2 years <b>Urine sampling:</b> bag; SPA <b>Patient spectrum:</b> query UTI	Bag vs.	SPA	100	89
Braude <i>et al.</i> (1967) <sup>58</sup>	<b>Study design:</b> prospective <b>Country:</b> UK <b>Setting:</b> secondary care	<b>Number (number girls):</b> 68 (49) <b>Age:</b> 3.2 years (9 days to 11 years), results for subgroup < 5 years <b>Urine sampling:</b> bag; SPA <b>Patient spectrum:</b> symptomatic or query UTI	Bag	Catheter	81	87
Etoubleau <i>et al.</i> (2009) <sup>61</sup>	<b>Study design:</b> prospective <b>Country:</b> France <b>Setting:</b> emergency department	<b>Number (number girls):</b> 192 (138) <b>Age:</b> < 3 years (non-toilet trained) <b>Urine sampling:</b> bag and catheter <b>Patient spectrum:</b> positive bag sample <b>Results also reported for polybacterial samples:</b> there were 16 using catheter and 84 using bag; 14 of these were polybacterial using both samples	Bag	catheter	88	80

Study details	Study design and setting	Population	Index test	Reference standard	Sensitivity	Specificity
Krahenbuhl <i>et al.</i> (2012) <sup>62</sup>	<p><b>Study design:</b> Prospective</p> <p><b>Country:</b> Switzerland</p> <p><b>Setting:</b> children's hospital and 15 paediatricians' offices</p> <p><b>Reference standard:</b> culture</p>	<p><b>Number (number girls):</b> 75 (45); only 25 had culture and U-test results available</p> <p><b>Age:</b> 0.9 years (9 days to 3.2 years)</p> <p><b>Urine sampling:</b> mid-stream clean catch, urine bag, catheter or SPA for urine collection for reference standard</p> <p><b>Patient spectrum:</b> suspected UTI</p>	Pad combined with dipstick	Mid-stream clean catch, urine bag, catheter or SPA	100%	79%

### Clinical signs and symptoms

Study details	Study design and setting	Population	Index test	Sensitivity	Specificity
Craig <i>et al.</i> (2010) <sup>1</sup>	<p><b>Study design:</b> prospective cohort</p> <p><b>Country:</b> Australia</p> <p><b>Setting:</b> emergency department</p>	<p><b>Number of children (number of girls):</b> 15,781 (8814)</p> <p><b>Age:</b> &lt; 5 years</p> <p><b>Urine sampling:</b> mixed</p> <p><b>Patient spectrum:</b> query infection</p>	<p>The following were associated with a significant increase in the odds of UTI: urinary symptoms, general appearance (mild, moderately or very unwell), fluid intake (none but not small or moderate decrease), highest temperature &gt; 38 °C, chronic disease, felt hot, unvaccinated by meningococcal vaccine, crying, elevated heart rate</p> <p>The following showed no significant association: capillary refill time, chest crackles, pneumococcal vaccine status, difficult breathing, elevated respiratory rate, abnormal chest sounds, bulging fontanelle, wheezing and stridor</p> <p>The following were associated with a significant decrease in the odds of UTI: infectious contacts, male, respiratory symptoms, diarrhoea, abnormal ear, nose and throat signs, cough, focal bacterial infection, rash, age &gt; 3 months, duration of illness</p> <p>A multivariable model was constructed based on the above signs and symptoms (final model not shown). This had diagnostic discrimination with an AUROC of 0.80 (95% CI 0.78 to 0.82). Validation of the model in 5584 illnesses also showed good performance (AUROC 0.78, 95% CI 0.74 to 0.81). The model performed significantly better than clinician estimation</p>		
Gauthier <i>et al.</i> (2012) <sup>67</sup>	<p><b>Study design:</b> prospective diagnostic cohort</p> <p><b>Country:</b> Canada</p> <p><b>Setting:</b> paediatric emergency department</p> <p><b>Reference standard:</b> culture</p>	<p><b>Number (number girls):</b> 331 (189)</p> <p><b>Age:</b> median 12 (range 1–36) months</p> <p><b>Urine sampling:</b> 'usual method'; bad specimens excluded</p> <p><b>Patient spectrum:</b> suspected UTI</p>	<p>Circumcision</p> <p>Presence of fever</p> <p>Duration of fever (≥ 72 hours)</p> <p>Parent report of:</p> <p>Smelly urine</p> <p>Vomiting</p> <p>Diarrhoea</p> <p>Abdominal pain</p> <p>Dysuria</p> <p>Painful urination</p>	<p>0</p> <p>100%</p> <p>49%</p> <p>57%</p> <p>31%</p> <p>27%</p> <p>41%</p> <p>19%</p> <p>26%</p>	<p>71%</p> <p>8%</p> <p>44%</p> <p>68%</p> <p>62%</p> <p>76%</p> <p>55%</p> <p>78%</p> <p>86%</p>

Study details	Study design and setting	Population	Index test	Sensitivity	Specificity
Gorelick and Shaw (2000) <sup>65</sup>	<b>Study design:</b> prospective cohort <b>Country:</b> USA <b>Setting:</b> emergency department	<b>Number of children (number of girls):</b> 1469 (1469) <b>Age:</b> < 2 years <b>Urine sampling:</b> catheter specimen <b>Patient spectrum:</b> query UTI	CPR: < 12 months, white race, temperature $\geq 39^{\circ}\text{C}$ , absence of source of fever, fever for $\geq 2$ days  Presence of $\geq 2/5$ variables	95	31
<sup>a</sup> Gorelick <i>et al.</i> (2003) <sup>66</sup>  Validation of above	<b>Study design:</b> nested case-control <b>Country:</b> unclear <b>Setting:</b> emergency department	<b>Number of children (number of girls):</b> 212 (212) <b>Age:</b> < 2 years <b>Urine sampling:</b> not stated <b>Patient spectrum:</b> query UTI	CPR  $\geq 3$ risk factors	88	70
O'Brien <i>et al.</i> (2013) <sup>9</sup>	<b>Study design:</b> prospective cohort <b>Country:</b> Wales <b>Setting:</b> general practice	<b>Number of children (number of girls):</b> 597 (284) <b>Age:</b> < 5 years <b>Urine sampling:</b> clean catch or nappy pad <b>Patient spectrum:</b> acute illness < 28 days' duration	Increased urinary frequency Wetting when previously dry Pain/crying when passing urine Irritable/grouchy Temperature $\geq 38^{\circ}\text{C}$ Muscle aches or pains Poor feeding/off food	31 14 14 80 43 0 69	87 94 95 37 71 90 45

a It is unclear what was meant by dipstick normal.

**Dipstick testing**

Study details	Study design and setting	Population	Threshold	Sensitivity	Specificity
Armengol <i>et al.</i> (2001) <sup>180</sup>	<b>Study design:</b> retrospective  <b>Country:</b> USA  <b>Setting:</b> outpatient paediatric clinic or emergency department	<b>Number (number girls):</b> 260 (not reported)  <b>Age:</b> < 4 years  <b>Urine sampling:</b> catheter specimen  <b>Patient spectrum:</b> query UTI	Nitrite or LE	70	98
Dayan <i>et al.</i> (2002) <sup>178</sup>	<b>Study design:</b> prospective  <b>Country:</b> USA  <b>Setting:</b> emergency department	<b>Number (number girls):</b> 246 [232 (118)]  <b>Age:</b> ≤ 60 days  <b>Urine sampling:</b> catheter or SPA  <b>Patient spectrum:</b> query UTI	Nitrite or LE  Nitrite and LE	85  30	92  100
Krahenbuhl <i>et al.</i> (2012) <sup>62</sup>	<b>Study design:</b> prospective diagnostic cohort  <b>Country:</b> Switzerland  <b>Setting:</b> children's hospital and 15 paediatricians' offices  <b>Reference standard:</b> culture	<b>Number (number girls):</b> 75 (45); only 25 had culture and U-test results available  <b>Age:</b> 0.9 years (9 days to 3.2 years)  <b>Urine sampling:</b> mid-stream clean catch, urine bag, catheter or SPA for urine collection for reference standard  <b>Patient spectrum:</b> suspected UTI	U-Test® device: nappy pad incorporating dipstick tests for nitrite and LE  Standard dipstick: nitrite and LE	100%  100%	79%  86%
Lejeune <i>et al.</i> (1991) <sup>181</sup>	<b>Study design:</b> prospective  <b>Country:</b> France  <b>Setting:</b> unclear	<b>Number (number girls):</b> 243 (not reported)  <b>Age:</b> < 18 months  <b>Urine sampling:</b> not clear  <b>Patient spectrum:</b> query UTI	Nitrite and LE	89	98

Study details	Study design and setting	Population	Threshold	Sensitivity	Specificity
Ramlakhan <i>et al.</i> (2011) <sup>183</sup>	<b>Study design:</b> retrospective diagnostic cohort  <b>Country:</b> UK  <b>Setting:</b> paediatric emergency department  <b>Reference standard:</b> culture (> 10 <sup>5</sup> CFU/ml of a single pathogen for clean void urines or 10 <sup>4</sup> CFU/ml for other specimens)	<b>Number (number girls):</b> 321 (202)	Nitrite and LE	37	95
		<b>Age:</b> mean 9.3 months (range 2 days to 2 years)	Nitrite or LE	91	50
Rodriguez <i>et al.</i> (2011) <sup>176</sup>	<b>Study design:</b> retrospective diagnostic cohort  <b>Country:</b> Spain  <b>Setting:</b> paediatric emergency department  <b>Reference standard:</b> culture (> 10,000 CFU/ml of a single pathogen)	<b>Number (number girls):</b> 980 (430)	LE and nitrite	26	99
		<b>Age:</b> mean 6 months (< 2 years)	LE or nitrite	80	83
Shaw <i>et al.</i> (1998) <sup>71</sup>	<b>Study design:</b> prospective  <b>Country:</b> USA  <b>Setting:</b> emergency department	<b>Number (number girls):</b> 3873 (2363)	Nitrite or LE: trace LE	73	99
		<b>Age:</b> < 2 years			
Shaw <i>et al.</i> (1991) <sup>182</sup>	<b>Study design:</b> prospective  <b>Country:</b> USA  <b>Setting:</b> emergency department	<b>Number (number girls):</b> 491 (309)	< 2 years: nitrite or LE	71	92
		<b>Age:</b> < 18 years	< 2 years: nitrite and LE	14	98
		<b>Urine sampling:</b> most clean void urine; otherwise, catheter or SPA			
		<b>Urine sampling:</b> bladder catheterisation			
		<b>Patient spectrum:</b> febrile children who had urine dipstick and quantitative culture			
		<b>Patient spectrum:</b> possible UTI			
		<b>Urine sampling:</b> clean catch; catheter			
		<b>Patient spectrum:</b> query UTI			
		<b>Urine sampling:</b> clean catch; bag; catheter			
		<b>Patient spectrum:</b> query UTI			



## Appendix 3 DUTY study protocol

Downing et al. *BMC Infectious Diseases* 2012, **12**:158  
<http://www.biomedcentral.com/1471-2334/12/158>



### STUDY PROTOCOL

### Open Access

# 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

Harriet Downing<sup>1†</sup>, Emma Thomas-Jones<sup>2††</sup>, Micaela Gal<sup>3</sup>, Cherry-Ann Waldron<sup>2</sup>, Jonathan Sterne<sup>4</sup>, William Hollingworth<sup>4</sup>, KerENZA Hood<sup>2</sup>, Brendan Delaney<sup>5</sup>, Paul Little<sup>6</sup>, Robin Howe<sup>7</sup>, Mandy Wootton<sup>7</sup>, Alastair Macgowan<sup>8</sup>, Christopher C Butler<sup>3</sup> and Alastair D Hay<sup>1</sup> DUTY study team

#### Abstract

**Background:** Urinary tract infection (UTI) is common in children, and may cause serious illness and recurrent symptoms. However, obtaining a urine sample from young children in primary care is challenging and not feasible for large numbers. Evidence regarding the predictive value of symptoms, signs and urinalysis for UTI in young children is urgently needed to help primary care clinicians better identify children who should be investigated for UTI. This paper describes the protocol for the Diagnosis of Urinary Tract infection in Young children (DUTY) study. The overall study aim is to derive and validate a cost-effective clinical algorithm for the diagnosis of UTI in children presenting to primary care acutely unwell.

**Methods/design:** DUTY is a multicentre, diagnostic and prospective observational study aiming to recruit at least 7,000 children aged before their fifth birthday, being assessed in primary care for any acute, non-traumatic, illness of  $\leq 28$  days duration. Urine samples will be obtained from eligible consented children, and data collected on medical history and presenting symptoms and signs. Urine samples will be dipstick tested in general practice and sent for microbiological analysis. All children with culture positive urines and a random sample of children with urine culture results in other, non-positive categories will be followed up to record symptom duration and healthcare resource use. A diagnostic algorithm will be constructed and validated and an economic evaluation conducted.

The primary outcome will be a validated diagnostic algorithm using a reference standard of a pure/predominant growth of at least  $>10^3$ , but usually  $>10^5$  CFU/mL of one, but no more than two uropathogens.

We will use logistic regression to identify the clinical predictors (i.e. demographic, medical history, presenting signs and symptoms and urine dipstick analysis results) most strongly associated with a positive urine culture result. We will then use economic evaluation to compare the cost effectiveness of the candidate prediction rules.

(Continued on next page)

\* Correspondence: [Thomas-jones@cf.ac.uk](mailto:Thomas-jones@cf.ac.uk)

<sup>†</sup>Equal contributors

<sup>2</sup>South East Wales Trials Unit (SEWTU), Institute for Translation, Innovation, Methodologies and Engagement, School of Medicine, Cardiff University, 7th Floor Neuadd Meirionnydd, Heath Park, Cardiff CF14 4XN, UK

Full list of author information is available at the end of the article

(Continued from previous page)

**Discussion:** This study will provide novel, clinically important information on the diagnostic features of childhood UTI and the cost effectiveness of a validated prediction rule, to help primary care clinicians improve the efficiency of their diagnostic strategy for UTI in young children.

**Keywords:** Urinary Tract Infection, Children, Primary care, Point-of-care-test, Dipstick test, Near-patient testing, Diagnosis, Economic models

## Background

Acute illness in young children is one of the commonest reasons for seeking health care worldwide. Reported rates of urinary tract infection (UTI) in children consulting for any acute condition vary widely (from 2% to 20% depending on setting and inclusion criteria) and most of this research has been hospital based [1,2]. Only one study has systematically sampled urine from sequentially presenting acutely unwell children in primary care, and found UTI prevalence of 4% [3]. However, that study was not adequately powered to estimate the predictive value of symptoms and signs with adequate precision.

UTI may be missed in as many as 50% of young children presenting to primary care [4,5]. The clinical diagnosis of UTI in young children is difficult because: (1) pre-verbal (predominantly under 3 years) children cannot articulate symptoms and present with the same non-specific symptoms (e.g. fever, irritability, vomiting and poor feeding) when suffering from a wide range of illnesses [6]; (2) identifying dysuria and changes in urination frequency in children wearing nappies (diapers) is difficult; (3) obtaining urine samples is often challenging and time consuming for parents [1] and costly to the health service [7,8]; and (4) the National Institute for Health and Clinical Excellence (NICE) does not recommend routine urine dipstick testing in children under 3 years because of a lack of research evidence as to its diagnostic value [9]. UTI diagnosis is therefore often delayed [6], missed [4] or symptoms attributed to other causes (such as otitis media) [10].

UTIs in young children may cause acute morbidity and recurrent symptoms that may indicate functional and anatomical abnormalities. In some young children, UTI may lead to renal scarring [11], leading to poor renal growth, recurrent pyelonephritis, impaired glomerular function, early hypertension, end stage renal disease [12] and pre-eclampsia [13-15]. Some experts, therefore, recommend aggressive, early antibiotic treatment for symptoms suggestive of UTI in young children to prevent renal scarring [16].

Dramatic reductions during the second half of the 1990s and early 2000s in antibiotic prescribing for children with upper respiratory infections [17] may have reduced serendipitous treatment of undiagnosed UTI and the consequent prevention of renal scarring.

However, antibiotic prescribing in the UK may now be on the increase [18]. NICE promotes early recognition and treatment to prevent short-term suffering and possibly serious long-term complications [9]. However, increased urine sampling will increase costs, consultation length and frequency of consultations in primary care. Clinicians will therefore only increase their sampling rates if evidence shows this really does improve the identification of UTI among the many acutely unwell children consulting primary care.

To date there is a lack of evidence as to the clinical predictors of UTI in young children. A meta-analysis [1] including 8,837 mostly pre-verbal children from 12 studies, showed that fever, non black race, a history of a previous UTI, temperature higher than 40°C, and suprapubic tenderness were the findings most useful for identifying those with a UTI. Uncircumcised boys were also more likely to have a UTI. While individual symptoms and signs were helpful in the diagnosis of a UTI, they were not sufficiently accurate to definitely rule it in, however a combination of findings could identify infants with a low probability of UTI [1]. The applicability of this review to UK general practice is limited because: (1) included studies were set in the US private and emergency care system where consultation and investigation threshold differs from UK primary care and other health care systems that are free at the point of delivery; (2) children had to either already have symptoms of UTI or fever  $\geq 38^{\circ}\text{C}$ , so many subtle symptoms and signs may not have been considered; (3) urine sampling was by catheter or suprapubic aspiration (which is not conducted in most primary care settings world-wide and from which any bacterial growth is regarded as significant); (4) diagnostic criteria used were different ( $\geq 10^4$  colony forming units per millilitre (CFU/mL)) to UK practice ( $\geq 10^5$  CFU/mL); (5) the relationship between ethnicity and UTI could be confounded; and (6) none of the studies included in the meta-analysis checked the external validity [19] of the findings, meaning that estimates of association could be inflated [20].

A more recent review of primary care based paediatric studies using urine culture as the reference standard found that no individual symptom or sign, or any combination of symptoms or signs, was sufficient to rule in a diagnosis of UTI, although some (e.g. increased capillary

refill time, no fluid intake, and supra-pubic tenderness) appear to warrant urine testing and empirical treatment while awaiting culture confirmation [2]. Furthermore, a number of symptoms and signs did not appear to have diagnostic value, including some of those included in the NICE guidelines (e.g. poor feeding and vomiting). Some symptoms, signs and proposed clinical prediction rules were associated with a sufficiently low UTI probability to rule out UTI, thereby removing the need to obtain a urine sample [1-3].

We found only one clinical algorithm derived from primary research for the diagnosis of UTI in young children. The research included febrile girls aged under two years in one US Emergency Department [21] and was validated in a case-control study in a different Emergency Department [22]. They found that more than three findings of: aged less than 12 months; white race; temperature of  $\geq 39^{\circ}\text{C}$ ; absence of any other likely source of fever; or fever for 2 or more days gave an area under the curve of 0.72, a sensitivity of 88% (95% CI: 79% to 94%) with a false-positive rate of 70% (95% CI: 61% to 79%).

An additional issue for the diagnosis of UTI in children is uncertainty as to the best criteria for microbiological diagnosis of UTI in this age group. Historically this has been based on a colony count of bacteria in the urine with a cut-off of  $\geq 10^5$  colony forming units (cfu)/mL of a uropathogen indicating infection. However this was derived from studies in adult women [23] and its applicability to children has been questioned by NICE [9] and others. Current guidance for microbiological diagnosis of UTI in children is at variance. The UK National Standard Method suggests that colony counts of  $\geq 10^3$  cfu/mL of a single species may be diagnostic of UTI in voided urine and that a pure growth of between 104–105 cfu/mL is indicative of UTI in a carefully taken specimen [24]. For midstream specimens, the European Association of Urology, suggests a cut-off of  $\geq 10^4$  cfu/mL if associated with symptoms, but  $\geq 10^5$  cfu/mL if symptoms are absent; lower cut-offs are suggested for PSA and bladder catheterisation samples [25]. However guidance from the American Academy of Pediatrics suggests that both urinalysis suggesting infection (pyuria and/or bacteriuria) plus the presence of  $\geq 5 \times 10^4$  cfu/mL of a uropathogen are required for a diagnosis of UTI, although these guidelines are for urine specimens obtained through catheterization or an SPA, which would not be routine in the UK [26].

#### Additional value of dipstick testing in young children

A 2006 Health Technology Assessment (HTA) review found there was inadequate evidence on the diagnostic performance of dipstick tests for protein or blood for children aged under 5 years old. The combination of a

positive test for both nitrite and leucocyte esterase (LE) was most accurate for ruling in UTI (pooled LR +28.2 (95% CI: 17.3 to 46.0)), and a negative test for both nitrite and LE was most accurate for ruling out UTI (pooled LR- 0.20 (95% CI: 0.16 to 0.26)) [27]. The NICE UTI guideline development group concluded that there was insufficient evidence to recommend the use of dipstick urine tests for children under 3 years [9].

#### Economic impact of UTI

UTI is the fourth most common reason overall for prescribing antibiotics in UK general practice, accounting for approximately 8% of all antibacterial prescriptions. However, UTI is currently infrequently diagnosed in children [28]. Whilst the unit costs of laboratory testing and antibiotic prescribing are relatively low [27], the *economic implications* of new clinical algorithms for urine sampling and testing may be substantial in young children because of: (1) the large numbers of children who present with non-specific symptoms who might be candidates for urine sampling and testing; (2) the cost, to healthcare services and to patients, of subsequent diagnostic tests (e.g. ultrasound, Micturating Cystourethrogram (MCUG) and Dimercaptosuccinic Acid (DMSA) scans) used to further evaluate children with recurrent/atypical UTI [9]; (3) the substantial societal costs and utility detriments of a missed diagnosis that leads to rare but serious complications of UTI; and (4) the wider, long-term population impact of diagnostic algorithms on antibiotic prescribing and resistance [29].

The few economic evaluations of methods for diagnosing UTI in young children [27,30] have primarily evaluated 'which tests to use?' rather than 'who to test?' The 2006 HTA review [27] evaluated 79 permutations of dipstick, cultures, ultrasound, and MCUG, and identified four testing strategies most likely to be cost-effective, although the optimal strategy differed by gender and age group. Current NICE guidance on testing strategies for UTI in children under 3 years is not based on evidence of cost-effectiveness [9].

#### Research objectives

In summary, rigorous evidence regarding the predictive value of symptoms, signs and urinalysis for UTI in young children is urgently needed to help primary care clinicians better identify UTI. Furthermore, since obtaining urine samples is especially challenging in children aged before their fifth birthday, the resulting algorithm will be constructed to answer two separate questions: first, which children warrant urine sampling? And second, can point of care dipstick urinalysis help clinicians determine which samples should be sent for laboratory culture? The algorithm will then be the subject of a validation study. Furthermore, since changes in the

frequency with which urine samples are requested has implications for parents and for healthcare services, analyses will model the economic impact (from the NHS and societal perspectives) of GP judgement versus diagnostic algorithm guided diagnosis and management with respect to the cost per correctly identified UTI, cost per symptomatic day avoided and the cost per quality adjusted life year.

The Diagnosis of Urinary Tract Infection in Young Children (DUTY) study protocol has the following research objectives:

1. To develop candidate clinical prediction rules that accurately identify children presenting in primary care with an acute illness in whom a urine sample should be obtained, based on socio-demographic factors, medical history, symptoms and signs.
2. To assess whether dipstick urinalysis for nitrite, leukocyte esterase, protein, blood and glucose gives additional diagnostic information to objective (1) in the identification of urine samples that should be sent for laboratory analysis.
3. To model cost-effectiveness from NHS and societal perspectives of the candidate clinical prediction rules.
4. To compare contamination rates for different urine sampling methods.

## Methods/Design

### Ethical and governance approval

Multi-centre approval was granted by the South West Southmead Research Ethics Committee (previously Southmead Research Ethics Committee, then South West 4 REC), Ref #09/H0102/64. Research and Development (R&D) approval has been granted for all sites taking part in the study.

### Design

DUTY is a 3-year, multicentre, diagnostic accuracy study to derive and validate a cost effective algorithm for the diagnosis of UTI in children under 5 presenting to primary care with an acute illness.

Children are eligible if they are aged before their fifth birthday and present to primary care with a new acute illness episode of less than or equal to 28 days duration. A Case Report Form (CRF) will be completed for all eligible, consented children and a urine sample obtained. The prevalence of UTI will be determined on laboratory culture. An algorithm will be derived and validated in separate samples of children.

### Setting

This UK study will be implemented from four research centres at the Universities of Bristol, Cardiff, Southampton and King's College London. Each centre will recruit

children from primary care, defined as any NHS facility providing first-point-of-contact face-to-face advice for parents of unwell children (GP practices, Walk-in-Centres (WIC), and Children's Emergency Departments (CED)).

## Study procedures

### Primary care sites/GP practice recruitment

Primary care sites will be recruited by each study centre covering both urban and rural areas across England and Wales. Two models of recruitment will be offered: (1) Option 1, in which the majority of the recruitment procedures will be undertaken by a dedicated Research Nurse (RN) or Clinical Studies Officer (CSO) external to the site; and (2) Option 2, in which recruitment will be undertaken entirely 'in house' by the primary care site's clinical team. From now on, members of staff taking informed consent for the DUTY study will be referred to as "recruiting clinicians". Dedicated RNs/CSOs providing external support for sites will be referred to as "DUTY recruiters", while members of practice staff involved in option 2 recruitment will be referred to as "site-based recruiters".

### Recruiting staff

The study grant will provide full-time equivalent DUTY recruiter posts across all four study centres, which will be supplemented by additional DUTY recruiter posts provided by local Primary Care Research Networks (PCRN) and Comprehensive Local Research Networks (CLRNs) (in England) and by the National Institute for Social Care and Health Research – Coordinating Research Centre (NISCHR-CRC) (in Wales). These DUTY recruiters will be available to provide Option 1 support to primary care sites, and to support autonomously recruiting Option 2 sites through the provision of expert training, mentoring and problem-solving.

### NHS microbiology laboratory recruitment

The participation of any primary care site in recruitment to the study will depend on the support and participation of the local NHS microbiology laboratory to which the site routinely sends urine samples. In each area of recruitment, the local NHS laboratory will be approached and service level agreements put in place prior to involvement in the study.

### Participant recruitment

The recruitment process is summarised in Figure 1.

### Registration and consent

Parents and children may be invited to take part in the study in a number of ways:

1. Where possible, primary care sites will mention the study to parents of children under 5 when they

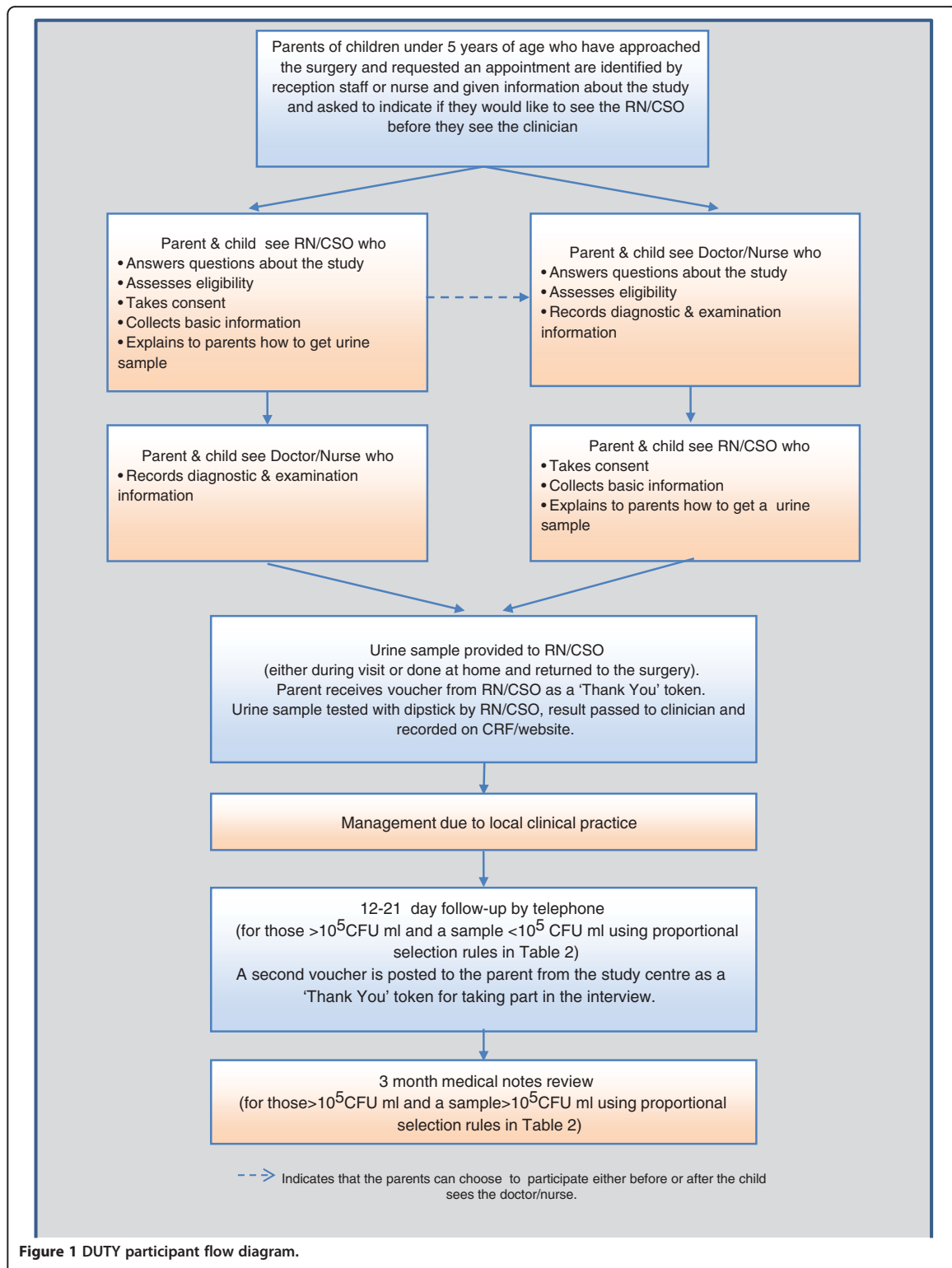


Figure 1 DUTY participant flow diagram.

phone for an appointment, and ask them to come to the surgery 15 minutes early to receive further information.

2. Where the study was not raised at the time of making the appointment, parents of children already booked in may be phoned and told about the study and invited to attend a little earlier.
3. If they cannot be contacted by telephone, they will be approached on arrival at the site, given information sheets and asked if they would be happy to see a recruiting clinician to discuss participation.

Once the parent has indicated they are happy to discuss the study, the recruiting clinician will explain study participation answer any questions that the parent may have, ensure that they fully understand the implications of participation, and check the child's eligibility. If the child is eligible and the parent agrees to participate, written informed consent will be obtained from the parent. If the parent wishes to see the GP before consenting, the recruiting clinician will arrange for this to happen, with the parent and child returning to the recruiting clinician afterwards to complete recruitment. Where possible the recruiting clinician will recruit the participant while they are waiting to see the GP, in order that the parent and child are not delayed. However, if more practical or convenient, the recruiting clinician may offer, with the parent's permission, to visit the family later the same day at their home to complete recruitment. If the parent is not interested in hearing more about the study, no further approach will be made.

#### Non-registration

A screening log of all children aged before their fifth birthday, who are attending for care and whose parents are approached by the recruiting clinician to invite participation in the study, will be compiled. Details will be recorded as to their eligibility, whether consent was given or declined, and reason for declining to participate.

#### Participant eligibility

Table 1 details the inclusion and exclusion criteria for the study. Since ruling out UTI may be as important as ruling it in, the study inclusion criteria were designed to be as broad as possible. Therefore, children consulting with other 'obvious' causes for their symptoms such as otitis media or bronchiolitis, as well as those with a history of previous UTI and known abnormalities of the urinary tract, learning difficulties, or re-consulting for an existing illness are all included as long as none of the exclusion criteria apply.

The study will include parents who speak other (non-English) languages. Parent information sheets and consent forms will be translated into other languages as required by participating GP practices (e.g. Welsh, Polish and Brazilian Portuguese). For languages less commonly spoken in the UK, particularly for those in which oral translation is more useful than written translation (e.g. Somali), translational services will be accessed, where possible, via interpreters employed by recruiting primary care sites to support patient-clinician communications. Where these services are not available, translational services will be provided via Language Line.

#### Collecting urine samples and dipstick testing

The recruiting clinician will attempt to obtain a urine sample from the children of consenting parents during the recruitment visit. If this is not possible, the recruiting clinician will ask the parent to collect the sample at home, give them the appropriate equipment and explain how to collect it. To minimise contamination, urine samples will be obtained using the 'clean catch' where this is possible. Where this is not possible, the 'nappy pad' method (which involves cleaning the child's perineum and inserting a sterile pad into the nappy to soak up urine, for a maximum period of 1 hour) will be used, as described by Liaw *et al.* [31] and as recommended by the recent NICE guidelines [9]. Urine sampling can be underway whilst the recruiting clinician completes the study CRF.

The recruiting clinician will retrieve the urine sample, test it with a urine dipstick (Siemens/Bayer multistix 8SG) provided by the study and record the urine sampling method and dipstick result on the CRF. The urine sample will then be split, if sufficient quantity is available, with the priority fraction being sent to the local NHS laboratory for routine diagnostic processing, and the second 'research' fraction being sent to the SACU (Specialist Antimicrobial Chemotherapy Unit, Public Health Wales Microbiology Cardiff, University Hospital of Wales) reference laboratory: for more in-depth analysis. As only small volumes of urine (minimum 1 mL) are required for each laboratory, it is expected that for most urine samples, it will be possible to split the urine into the two fractions.

If a sample is not obtained during this visit the parent will be asked if they could obtain one at home, refrigerate it and return it to the primary care site within 24 hours. Where possible, the DUTY recruiter will offer to collect the urine sample from the patient's home.

In a sub-sample of children recruited from a handful of sites linked to the Bristol study centre, we will use time-motion techniques to measure the additional time (parent and healthcare professional) taken to collect the urine sample and to perform dipstick testing during the

**Table 1 DUTY eligibility criteria**

Children will be included if:	Children will be excluded if:
Aged before their fifth birthday.	Aged 5 years and above.
Presenting at a participating NHS primary care site.	Parents are unable or unwilling to assist with study.
Presenting with an acute ( $\leq 28$ days) illness as the main reason for the parent to have requested an appointment.	Illness longer than 28 days duration. Presenting with trauma as a predominant concern.
Presenting with at least one 'constitutional' symptom or sign identified by NICE [9] as a potential marker for UTI – that is, fever, vomiting, lethargy/malaise, irritability, poor feeding and failure to thrive <i>and/or</i> at least one urinary symptom identified by NICE [9] as a potential marker of UTI – that is, abdominal pain, jaundice (children $< 3$ months only), haematuria, offensive urine, cloudy urine, loin tenderness, frequency, apparent pain on passing urine and changes to continence.	No urinary or constitutional symptoms as defined by NICE [9] and listed in the left hand column. Known neurogenic (e.g. spina bifida) or surgically reconstructed bladder or urinary permanent or intermittent catheterisation (for whom different bacterial concentration cut points are used). Taking any antibiotics in the last 7 days. Taking immunosuppressant medication (e.g. anti-rejection drugs, oral or intramuscular steroids or chemotherapy). Already recruited into the DUTY study.
	Involved in current research or have recently (within 28 days) been involved in any research prior to recruitment.
	There will be no recruitment to the study after the last NHS laboratory transport of the day has departed from that primary care site on Fridays. For recruitment at A&E settings only: children will not be eligible if their presentation at A&E is a direct result of GP referral.

primary care appointment. These data will be used to inform the economic analysis.

#### Maximising urine samples

Obtaining the urine samples will be challenging, and a suboptimal return rate will diminish power and increase risk of bias. Therefore: (1) we will monitor the location of urine specimens using a web-based database. Clinical data for recruited children will be logged onto the secure study website. Dipstick urinalysis data may be added after the clinical data and will provide a record of the urine having been obtained. This will allow the research team to identify children for whom urine samples have not been provided, and to check with the relevant recruiter as to whether this should be followed up. Both the Research and NHS laboratories will also record the arrival of, and results from, the specimens on the website; and (2) each centre will provide dedicated DUTY recruiter resource to assist Option 1 practices with obtaining urines.

#### Laboratory processing of urine samples

The NHS 'clinical' fraction will be labelled with the child's unique DUTY study identification (ID) number on DUTY specific labels as provided in patient packs. Similar DUTY labels will be adhered to the DUTY study specific microbiology form and the sample sent to the local laboratories using the site's normal method of transport. Any samples not collected within 4 hours will be refrigerated at the site and processed within 36 hours. Clinicians will receive and act on reports from their local laboratory as in the course of usual clinical care.

The remaining portion of urine will be decanted into a sterile monovette container containing boric acid. This will be labelled with the child's study ID number and sent by 1<sup>st</sup> Class Royal Mail using Post Office approved Safeboxes<sup>TM</sup> to the central research laboratory.

#### Minimising effects of sample contamination and assessment of asymptomatic bacteriuria

Contamination of urine (a cultured organism from a source other than the urinary tract) can lead to false positives: a potential false positive rate of 7.2% has been identified in one study by comparing pairs of urines from 203 children [32]. All nine (5.4%) children in this study with a mixed culture  $\geq 10^5$  CFU/mL of uropathogens (a heavy mixed growth) in their first sample had a UTI excluded in the second [32]. In addition, bacteria at  $\geq 10^6$  CFU/mL (have also been found in the urine of approximately 1.5% of young, asymptomatic, children when screened using the 'gold standard method', supra-pubic aspiration [33], and most did not experience long term sequelae [34]. Therefore, distinguishing UTI from asymptomatic bacteriuria and bacterial contamination is difficult, and could lead to spurious associations between symptoms (e.g. diarrhoea) and apparent 'UTI' that is in reality contamination or potential harmLess asymptomatic carriage.

Clinicians use the presence of UTI symptoms to help interpret culture positive results but this leads to incorporation bias. In DUTY, we could restrict recruitment to those children with currently recognised symptoms of UTI, but since the purpose of DUTY is to determine the strength of association between currently recognised as

well as *currently unrecognised* symptoms/signs and UTI, it is important that eligibility criteria are as 'open' as possible (and that a prospective cohort, as opposed to retrospective case-control, design is used), but without including children in whom a positive culture is unlikely to be clinically relevant (e.g. a well child with conjunctivitis). Therefore, DUTY will recruit children with constitutional and/or urinary symptoms and make the assumption that the presence of no more than two pathological bacteria of at least  $>10^3$ , but usually  $>10^5$  CFU/mL on culture of their urine is clinically significant. This could result in more urine samples being tested and more children receiving antibiotics than is strictly necessary, but carries the benefit that more UTIs would be identified and treated promptly.

#### Data collection

Unique study identification numbers will be sequentially generated and used on pre-printed consent forms, paper CRFs, urine sample labels and test request forms (for local NHS and central research laboratories).

A CRF will be completed for all consented patients. This will include a short medical history including recent antibiotic use and other potential risk factors for UTIs and resistance, and clinical examination findings. An outline of the domains covered in the CRF can be found below in the next section.

#### Case report form

The CRF will contain as many of the known and potential features associated with UTI as are feasible without overly compromising the speed and simplicity of completion.

Five sections will facilitate data entry by different personnel (recruiter taking consent or responsible clinician) so as to minimise the burden to healthcare professionals undertaking same day primary care:

1. Eligibility screening and consent (to be completed by recruiting clinician within the recruitment interview with the parent).
2. Registration (to be completed by recruiting clinician as above): Socio-demographic data (to include: date of consultation, name, address, contact telephone number/s, ethnicity [21], date of birth and gender [9]). We will also ask about parent's highest educational attainment level and their financial well-being in order to assign a measure of socio-economic deprivation.
3. Presenting Symptoms and Medical History (to be completed by recruiting clinician as above): child's presenting symptoms will be recorded, along with known previous medical history (e.g. previous UTI,

circumcision [5,35], child or family history of vesico-ureteric reflux [36], other abnormalities of the urinary tract, learning difficulties, details of prior surgery, other co-morbidities, recent and previous long-term use of medicines, including antibiotics).

4. Clinical Examination and Management (to be completed by child's responsible clinician within a standard consultation): in addition to the 'constitutional' and 'urinary' study eligibility symptoms defined by NICE [9], we will collect information regarding the clinician's global assessment of illness severity [37], respiratory and gastro-intestinal symptoms and signs, and the symptoms and signs proposed by NICE to distinguish 'typical' from 'atypical' UTI, such as poor urinary flow and abdominal mass. Clinicians will be asked to record the child's management, including antibiotic use and immediate referral to secondary care. To assess the diagnostic value of the urine dipstick test, we will ask clinicians to record their working diagnosis before having seen the dipstick results, and to record whether their working diagnosis has changed after they have seen them. Finally, for the economic analysis, we will ask clinicians to state what their management would be if the patient were not enrolled in the DUTY study (e.g. no urine test/not treated for UTI or urine test/treat for UTI). This will provide information on the 'clinician judgement' diagnostic strategy that will be a comparator in the economic evaluation.
5. Urine collection and processing (to be completed by recruiting clinician): Urine sampling method (clean catch or nappy pad) and urinalysis results with date, time of testing, with a prompt to inform the responsible clinician of dipstick result and confirmation that the sample has been sent to the local NHS and central research laboratories.

In addition to the CRF, the study web-based data collection platform will include additional sections to cover data entry for the following: (1) NHS microbiology laboratory microscopy and culture; (2) SACU research laboratory microscopy and culture; (3) patient follow-up at Day 14 following recruitment, and (4) patient notes review at 3 months from recruitment. Further detail of the content of these domains is provided below.

#### Processing of urine samples by NHS laboratories

NHS laboratories will be informed of the study and the agreement of the lead consultant microbiologist, laboratory manager and the NHS hospital Research and Development approval obtained before patient recruitment begins. The laboratories will be asked to complete the following tasks for DUTY study samples:



1. Log the date and time of specimen arrival on the secure DUTY web-based database.
2. Process the urine and report the result back to the requestor using their own Standard Operating Procedures (SOPs) and Laboratory Information Management Systems (LIMS).
3. Enter the results of urine culture onto the DUTY study web-based database. Since laboratories vary in their SOPs, not all of the following will be available however, microscopy for white and red cells; quantification and purity of bacterial growth; and speciation will be requested. Laboratories will be asked to transcribe this information onto the DUTY web-based database in order to activate laboratory payment.
4. Store any isolates from urines with  $>10^5$  CFU/mL in pure/predominant ureopathogen growth for referral onto the central research laboratory at the end of the study. These should be stored, ideally on cryogenic beads, at temperatures of  $-70$  C, or on slopes at  $-20$  C if  $-70$  C storage facilities are not available.

#### Processing of urine samples by the research laboratory

The central research laboratory has experience in supporting other primary care UTI studies and performed a similar role to that described below in the previous EUR-ICA study (Epidemiology of urinary tract infection in children with acute illness in primary care) [3].

1. Urines will be sent overnight by Royal Mail SafeBoxes™ by the participating sites. Boric acid will be used to stabilise bacterial counts.
2. On receipt at the central research laboratory, the urine sample will be spiral-plated on blood agar and UTI Chromogenic agar will be used to quantify bacteria  $>2 \times 10^1$  CFU/mL and  $<10^{10}$  CFU/mL.
3. The bacteria will be identified to species level and stored on cryogenic beads at  $-80^\circ\text{C}$ . The urine will be stored frozen.
4. Results will be recorded on a designated laboratory worksheet and entered into the DUTY web-based database.

Where the urine culture result produced by the central research laboratory is positive and the local NHS

laboratory result is negative or not processed, and if this discrepancy is considered by the lead SACU microbiologist to be clinically significant, we will inform the child's responsible clinician. This will enable the child's responsible clinician to consider the future management of the child in the light of the additional information arising from the study.

#### Patient follow-Up

##### Telephone follow-up at Day 14

Each centre will telephone parents of all children selected for follow-up according to the proportional selection rules in Table 2, to record symptom duration and healthcare resource use (e.g. repeat primary care contacts other community care, secondary care contacts, prescribed and over-the-counter medications) during the 14 day period after recruitment. Parents will also be asked to detail expenses and time off work due to their child's illness and rate the child's quality of life (including symptoms, sleeping, feeding, behaviour and wellbeing) using a previously validated measure (TAPQoL) [38].

##### 3 Months note review

Each research centre will conduct a primary care notes review for all children who were selected for follow-up. Primary care contacts, medications and secondary care utilisation during the 3 months after study recruitment will be recorded during the review.

##### Withdrawal & loss to follow-Up

In the majority of cases the only active participation of participants is at the initial consultation, and withdrawal from the study in most cases is unlikely. Attrition in those selected for 14 day follow-up due to the challenges of making contact with busy parents will be minimised by making several attempts to contact parents/guardians by telephone and, if this is unsuccessful, a postal version of the resource use questionnaire will be posted to participants with a stamped addressed envelope for return. Parents will be offered a £5 voucher, by post, on completion of follow-up either by telephone or post.

**Table 2 Proportional selection rules for DUTY follow-up**

Category	Definition	Location	Proportion to be sampled at Day 14
$> 10^5$ CFU/ml	Pure or 1 predominant species	BOTH NHS lab and Central research lab	100% (All)
$>10^3$ and $< 10^5$ CFU/ml	Pure or 1 predominant species	Central research lab	20% in total (combination of both categories)
$>10^5$ CFU/ml	2 or more species	BOTH NHS lab and Central research lab	
$< 10^3$ CFU/ml and 'No Growth'		BOTH NHS lab and Central research lab	10%

**Electronic data entry**

The DUTY data collection process is complex and involves input from a number of different personnel at different sites: (1) CRF data by a combination of study DUTY recruiters, practice nurses, GPs and dedicated staff at recruiting primary care sites; (2) clinical and research urine culture results by laboratory technicians and managers; and (3) follow-up data at Day 14 and 3 months by study centre administrators and research nurses. To optimise the quality of the data entry and to enable effective data collection from multiple sites across England and Wales, we decided to use a secure, web-based electronic data collection platform.

We will use a secure web 2.0 clinical study management system (The electronic Primary Care Research Network (ePCRN)). Hosted by South London and Maudsley (SLAM) NHS Foundation Trust, the ePCRN implementation is of a separate domain and a Citrix farm serving published applications, with a Structured Query Language (SQL) server providing clinical based study application databases. The system avoids potential data loss, duplication and security issues with laptops and portable media and has been approved by ethics and by the SLAM Caldicott and Executive Committees.

Web-forms for data collection will be created in ASP.net (a dynamic web application framework) on top of a dedicated SQL data management server, with data variables forced to comply with entry and validation rules defined in the data element definitions. The SQL data management server will incorporate auditing, backup and recovery facilities. The study workflow and algorithms will be enforced using the same methods, and a visual algorithm on the web pages will guide users. The web-based system will be piloted for ease of use prior to data entry go-live.

**Electronic data protection and confidentiality**

The ePCRN safeguards the legal and ethic rights of service users through a fully integrated research security management system consisting of two component parts: (1) technical specifications built into the DUTY study database during the development phase, and (2) procedural standards governing the launch and day-to-day use of the application by DUTY study researchers.

Access to users will be provided through study-specific logon points in Citrix Access Gateway Advanced Access Control. Citrix software establishes a secure, encrypted, connection with the user's PC, allowing access from the Internet uniquely to the Citrix Access Gateway and enabling access to identifiable study data for authorized users.

**Data entry in primary care sites**

In order to maximise the acceptability and ease of use of DUTY data collection tools, clinicians working in settings without web access or whose working practice made web data entry an unwelcome burden, will be able to opt for paper-based data collection with the support of the local study centre in entering data, or making alternative arrangements for data entry, on their behalf.

The web-based data collection system will be presented as the preferred method of data collection, and practice-based recruiting staff will be strongly encouraged and supported to enter CRF data onto the database directly or, if using paper-based CRFs in the recruitment interviews, to retrospectively enter the data in a timely way (consent and registration within 24 hours, and full eCRF data within 5 working days).

**Data entry in the local NHS laboratory and central research laboratory**

Once in the local NHS and SACU research laboratories, staff will be able to access an anonymised data collection page, where only study numbers and the data collection forms for the urine samples can be seen. Laboratory staff will be asked to log the samples on receipt and enter the results when available.

**Follow-up data entry in research centres**

At day 14 from recruitment, and at 3 months, research staff will enter symptom duration, healthcare resource utilisation and expenses data from telephone interviews and practice records respectively onto web-based data collection forms.

**Analysis**

The overall aim of the analyses is to derive a validated clinical prediction rule for UTI among acutely ill children presenting to primary care.

**Sample size calculation**

To estimate the required sample size we drew on our experience with the EURICA study, which found a UTI prevalence rate in children aged before their fifth birthday of 4% [3]. We considered first the strength of association between candidate predictors (symptoms, signs or dipstick results) and UTI as well as the precision of the final algorithm's sensitivity for the detection of UTI. Taking the most conservative assumptions, i.e. candidate predictors present in 10% of children and an overall UTI prevalence of 2%, 3,000 urine sample results are required to detect an odds ratio of 2.4 with 80% power and a two-sided alpha of 5%. With an overall prevalence of UTI of 2%, an algorithm sensitivity of 80% and 3,100 urines, the 95% confidence interval (CI) will be no more than +/-10%. We propose to recruit 4,000 children with

a target of recovering urines from at least 77.5% for algorithm derivation and a further 2,000 children for validation.

Given the complexity of the statistical analyses, large number of possible predictors and the need to account for some missing data in predictor variables, we propose to recruit at least 7000 children in total (two-thirds for algorithm derivation; one-third for validation), in order to maximise the statistical power of the sample).

### Statistical analysis

#### Defining the primary outcome

The first stage of the analysis will investigate the best combination of microbiology data from the local (NHS) laboratories and the central research laboratory that can be used to define urine samples as positive for UTI. We will agree definitions of UTI positivity for data from each laboratory (using culture results and white cell counts) and will cross-tabulate these. Possible disagreements will be examined, and samples classified after discussion as: (1) "Agree UTI negative"; (2) "Agree UTI positive"; (3) "Disagree (CL positive, LL negative)"; (4) "Disagree (CL negative, LL positive)". If overall between-laboratory reliability (classified on kappa statistics) is good, the primary outcome will be defined as positive for samples classified as (2) "Agree UTI positive". If overall agreement is moderate or poor, we will explore reasons for this. We will stratify according to age of the child and method of collection (clean catch or nappy pad), and investigate whether reliability varies between strata. We will also examine the inference of laboratory methods (e.g. whether samples are processed in boric acid) and time from sample collection to laboratory processing.

If the best definition of UTI positivity remains unclear, then we will select a small number of signs and symptoms that the literature suggests are clearly associated with presence of a UTI. We will examine associations of UTI positivity with these symptoms, using different definitions of positivity (e.g. based on central or local laboratory, threshold for amount of growth, evidence of growth of other species, method of sample collection). We will select the best definition of microbiological positivity based on the magnitude of associations with the selected signs and symptoms.

#### Descriptive analyses

We will use methods appropriate for small proportions [39] to estimate the prevalence (with 95% confidence interval) of culture positive urines in acutely unwell children aged before their fifth birthday presenting in primary care. This will be undertaken on the whole dataset. The degree of variation in prevalence between practices and geographical areas will be explored using two level

random-effects logistic regression models (with practice/site as a random effect and area as a fixed effect). This analysis will also explore difference by recruitment site type (general practice, WICs and CEDs). Children in whom urine samples are obtained will be compared to those who are recruited, but no urine sample is obtained in terms of clinical presentation and demographics.

We will compare the probability of contamination in samples that are retrieved via a 'clean catch' method with those using nappy pads, controlling for patient and practice factors in a two level random-effects logistic regression model (objective 4). We will examine the impact of timing of sample in relation to the time between obtaining the urine transportation (including day of the week) and laboratory analysis on the rates of positive and contaminated urine samples (e.g. exploring if delayed samples such as those taken after daily laboratory collection have an impact on contamination rates).

The sample will then be sub-divided into algorithm derivation and validation datasets, comprising two-thirds to one-third of the data set, respectively. This will be done by randomly selecting practices: all of their patients will then contribute to one of the two datasets.

#### Development of clinical prediction rules

We will develop a clinical prediction rule based on the linear predictor in a logistic regression model in which the outcome variable is a culture-positive urine result. Candidate diagnostic variables will be categorised into demographic background and medical history (for example, gender, previous UTI); both specific and general systemic presenting symptoms and signs (for example overall illness severity, fever, vomiting); and results from urine dipstick analysis (nitrite, leukocyte esterase, protein, blood and glucose). Because of concerns that some aspects of medical history or demographics may be associated with asymptomatic carriage rather than active infection, we will also develop a 'signs and symptoms' only prediction rule. Variables will be included in logistic regression models based on an "inclusive" p value threshold of 0.1. We will check for nonlinear effects of continuous variables, and will examine candidate interactions specified *a priori*. Any further candidate interactions will be agreed before analyses commence. Such effects will be included in the final models as necessary.

We will begin by examining the predictive value (based on diagnostic odds ratios and C statistics) of the best predictors from each of the three categories (socio-demographic and previous medical history, clinical assessment, and dipstick urinalysis) of variables. We will then examine the additional diagnostic value of presenting signs and symptoms (compared with socio-demographic and medical history alone) and of dipstick results (compared with the other two categories). We

will examine whether it is possible to identify subgroups of children in whom dipstick testing is and is not justified based on their signs and symptoms. The final algorithm will be characterised based on its sensitivity and specificity, and positive and negative likelihood ratios.

#### Validation of clinical prediction rule

Diagnostic models that are developed using *p*-value-based variable selection will inevitably suffer from statistical over-optimism. Therefore, the final models will be validated using the second dataset, and the published rule will be based on the linear predictors from the model re-estimated in this validation dataset. A comparison will be made between the results obtained from the validation and the use of shrinkage based approaches applied to the original development dataset [40]. A comparison will be made between the results obtained from the validation dataset and the use of shrinkage based approaches applied to the original development dataset. The magnitudes of regression coefficients, and overall diagnostic value of the linear predictor, will be compared between the primary outcome and other definitions of culture positivity.

#### Analyses of follow up data

Children with positive urine cultures ('contaminants' and 'UTIs') who the clinician felt at recruitment had a suspected UTI will be compared to those who the clinician felt there was little probability of a UTI in terms of their subsequent illness course and resource usage over the next three months.

#### Risk stratification for clinical practice

In the final stages of analysis, we will examine the sensitivity and specificity of the linear predictor, based on a set of chosen thresholds for positivity. This will be used to identify several candidate clinical prediction rules for comparison in the economic evaluation. We will select thresholds that provide a range of clinical prediction rules from high sensitivity/low specificity to low sensitivity/high specificity.

#### Minimising Bias

The following design and analytic strategies will be employed to minimise bias:

- (1) Selection bias: where possible we will recruit consecutive children. We will ask sites to keep a screening log of patients approached but who did not take part in the study and reasons for this;
- (2) Index test technology: all tests (symptoms, signs, nappy pads, dipstick tests) will be carried out using standardised equipment and protocols;
- (3) Incorporation bias: the reference standard will consist of culture alone and will not incorporate any of the index tests;
- (4) Review bias: observers assessing the index tests will differ from and be blind to those assessing the reference standard (and vice versa);
- (5) Verification bias: all children who contribute to the study will have a urine sample sent to assess the reference standard. Children in whom it is not possible to obtain a sample will be excluded from the analysis. It is unlikely that reasons for failure to obtain urine samples will be related to the index tests but we will compare children with and without urine cultures;
- (6) Disease progression bias: we expect the time between clinical assessment and obtaining the urine samples to be minimal (no more than 24 hours);
- (7) Treatment paradox: for most children, antibiotic treatment will be started after the urine sample has been obtained, but we will record where this has not been possible;
- (8) Handling of indeterminate or uninterpretable results or withdrawals: these parameters will be measured and considered in the analysis, and;
- (9) Appropriateness of the reference standard: use of  $>10^3$ , but usually  $>10^5$  CFU/mL of one, but no more than two uropathogens is likely to detect the majority of children with UTI, but the second 'research' urine result from the SACU laboratory will allow for sensitivity analyses around different bacterial concentrations. Where possible, we will measure all threats to validity (e.g. time between clinical assessment and obtaining and culturing the urine sample) that could influence results.

#### Economic analysis

The aim of the economic analysis is to compare candidate clinical prediction rules (CPRs) on: (a) incremental cost per correct diagnosis of bacteruria, b) incremental cost per symptomatic day avoided, and c) incremental lifetime cost per quality adjusted life year, including potential long term complications of UTI from NHS and societal perspectives.

The cost-effectiveness of each candidate prediction rule will be compared to a 'clinical judgement' testing strategy and two 'boundary strategies': 1) Performing a urine test in every child meeting the DUTY eligibility criteria, and 2) No testing, the diagnosis of UTI is made clinically. The two boundary strategies are not intended to reflect clinical reality, but provide a reference point against which other diagnostic strategies can be compared. The cost-effectiveness of each candidate CPR will be assessed against these three strategies using a decision analytical model. The face validity of the model

structure will be reviewed by clinicians on the DUTY study team before being finalised.

#### **Cost per correct diagnosis of UTI (diagnostic model)**

The diagnostic cost-effectiveness model requires information on cost (cost of sample collection, cost of NHS laboratory testing), probabilities (probability of sample being obtained and probability of sample being contaminated) and diagnostic accuracy parameters (sensitivity and specificity of the various diagnostic strategies). The diagnostic accuracy parameters will be derived directly from the DUTY CRF and the statistical analysis described above. The probability of sample collection and contamination will be observed in the DUTY study. The cost of sample collection and laboratory testing will be derived from a combination of surveys of NHS laboratories and GPs and a time motion study observing primary care clinical staff as they collect and process urine samples. The model will estimate the cost per patient correctly diagnosed. The model will also include diagnostic strategies that incorporate dipstick testing (e.g. in those children thought to be a moderate risk of bacteriuria) with information on a small number of additional parameters (i.e. cost, sensitivity and specificity of dipstick testing).

#### **Cost per symptomatic day avoided (short term model)**

The model will then be extended to estimate cost per symptomatic day avoided at day 14. A Markov model will estimate the short-run cost-effectiveness of each diagnostic strategy. This extension will require information on additional parameters including the probabilities of receiving antibiotics and hospital admission in children diagnosed with and without UTI, the cost of antibiotics and hospital admission/testing, and the daily recovery probability for children with bacteriuria treated with antibiotics, children with bacteriuria not treated with antibiotics and children without bacteriuria. The probabilities of antibiotic treatment and hospital admission will be based on 14 day interviews and review of patients' primary care medical notes at 3 months. The daily recovery probabilities will be based on 14 day interviews with parents of children selected for follow-up. As all children in DUTY will receive a urine test, we will not be able to observe the recovery of children with bacteriuria not treated with antibiotics (i.e. false negatives). For this transition probability we will use literature estimates of the effect of antibiotic therapy on symptom duration.

#### **Cost per quality adjusted life year (QALY -long term model)**

The structure of the long-run Markov model will be based on the a previously developed model [27]. The

model will provide a link between the number of UTI attacks that a child will experience, the proportion that are pyelonephritic, the prevalence of vesicoureteral reflux, the probability of progressive renal scarring, the risk of end stage renal disease (ESRD) and disease management. Outcome parameters such as probability of UTI recurrence, renal scarring, ESRD, survival and utility values of renal disease will be based estimates for the sensitivity and specificity of imaging tests for VUR (e.g. ultrasound, MCUG), costs of treatment (e.g. antibiotic prophylaxis, cost of pyelonephritic treatment, dialysis, transplant), recurrence of UTI, disutilities (pyelonephritic, dialysis, transplant) and survival [27].

All analyses will be probabilistic, as all parameters will be entered into the model as distributions. Therefore the results will be presented as cost-effectiveness acceptability curves (CEACs). The CEAC demonstrates which of several testing strategies is most likely to be cost-effective at any fixed willingness to pay for a correct diagnosis. Costs and outcomes occurring after the first year will be discounted at 3.5%. We will use net monetary benefits and cost-effectiveness acceptability curves, at plausible willingness to pay thresholds (e.g. £0 to £50,000 per QALY) to identify the most cost-effective diagnostic strategies. Deterministic sensitivity analyses will also be used to evaluate the impact of key parameters on results and the influence of various CFU/mL thresholds on the choice of clinical prediction rule.

#### **Discussion**

This paper describes a diagnostic and prospective observational study in primary care, that aims to recruit at least 7,000 children aged before their fifth birthday, who are assessed for any acute, non traumatic, illness of  $\leq 28$  days duration. Urine samples will be obtained from eligible consented children and tested with a dipstick before being sent to a local NHS laboratory and a central research laboratory for microbiological analysis.

This study will provide novel, clinically important information on the diagnostic features of childhood UTI to help primary care clinicians improve their diagnostic efficiency. New insight into the diagnostic value of dipstick urinalysis and urine sampling methods will also be provided.

The observational design of the study will minimise disruption to normal practice and reduce the research burden on healthcare professionals, thereby maximising the applicability and generalisability of findings. In addition, this study will increase awareness of UTI as a possible diagnosis in appropriate children under 5 years old. We will collect full clinical information, including presenting symptoms and signs, medical history and clinical diagnosis; urine dipstick data and culture data from both usual routine local NHS laboratories and

central research laboratory. All the above will be utilised for derivation and validation of the diagnostic algorithm. In addition the follow-up data at 14 days and 3 months will allow for full health economic analysis, providing cost effectiveness outcomes. Additional useful information such as the best way of sampling urine from young children in primary care, the species and sensitivities of the infecting organisms, and contamination rates will also inform care in the future.

The main challenge to the DUTY study is the large number of participants and urine samples needed. A substantial number of practices across each of the four centres will need to be recruited, to provide a large enough population of potential participants. In addition to this, parental consent will be required at the same general practice visit when they are invited to participate, meaning there is a reliance on parents being willing and having the time to participate immediately. Despite the dedicated DUTY recruiting staff on the ground, who will provide support to practices, study buy-in is needed from all primary care staff to ensure adequate resources and infrastructure are in place to facilitate the conduct of the study. This will require training of all staff involved at each participating primary care site. A major contribution from NHS microbiology laboratories will also be necessary, for processing the increased number of urine samples received from primary care sites.

This study will rely on sampling methods most commonly used currently in UK primary care, namely nappy pad and clean catch. Suprapubic aspiration and catheterisation methods are neither feasible nor appropriate for primary care. Nappy pad sampling may carry an increased risk of contamination or un-interpretable results. This may lead to the exclusion of some results from the main analysis. However, this will be minimised by obtaining clean catch samples where possible. Lastly, the electronic data capture on the database has been designed to allow for all parties to enter data separately. A challenge will be to ensure all data is entered into the database in a timely way, to allow for both the real time monitoring of recruitment and urine sample location, as well as the conduction of the follow-up interviews within the tight timeframe.

In summary, this will be one of the largest studies of its kind undertaken in primary care, involving obtaining clinical samples from children, and will help guide management of the acutely unwell child, which is a common and important aspect of primary health care delivery. Improved assessment and diagnosis may lead to more appropriate microbiological and point of care testing and more timely treatment and investigation of those children who are most likely to benefit, while reducing unnecessary treatment and investigation among those that are most unlikely to benefit. This is likely to

improve outcomes for individual patients and may help prevent long-term sequelae. The overall outcome of the study will be to achieve a more consistent approach in the clinical care of a common condition, based on accurate diagnosis and effective clinical management.

## Endnotes

<sup>a</sup> This paper will use the term 'parent' to refer to the person with legal responsibility for the child, therefore as applied in this paper the term also encompasses carers (foster parents, legal guardians etc.).

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

CCB and ADH are co-chief investigators and act as guarantors of the study in its entirety. CCB and ADH jointly led the development of the research question, study design and implementation of the study protocol, along with KH, JS, WH, PL, BD, RH, MW and AM. HD is Study Manager and coordinated the operational delivery of the study protocol across the UK. ETJ and CAW are study managers at Cardiff, and coordinated the delivery for the study protocol in Wales. PL and BD are principal investigators, responsible for study oversight at Southampton and London, respectively. RH, MW and AM have provided expert microbiology input. HD, ETJ, MG and CAW jointly drafted the manuscript. All authors listed provided critical review and final approval of the manuscript.

## Acknowledgements

In addition to the authors, the DUTY study team comprises: Kate Birnie, Kim Harman, Marilyn Peters, Kate Martinson, Catherine Lises, Timothy Pickles, Michael Lawton, Jonathan Bengler, John Busby, Margaret Fletcher, Jan Dudley, Judith van der Voort, Kathryn O'Brien, Penny Whiting, Catherine Derrick, Steven Beech, Elizabeth Thomas, Victoria Roberts, Lewis Darmanin, Stevo Durbaba, Margaret Hague, Susan George, Teresa Wade, Annie Sadoo, Peter Brindle, Lisa Carver.

The authors would like to acknowledge the funder of this study: National Institute for Health Research Health Technology Assessment (HTA) programme.

The authors would also like to acknowledge the contribution of the Study Steering Committee members, namely Frank Sullivan, Matthew Thompson, Rafael Perera, Clodna McNulty and Malcolm Coulthard, and the support from the local Primary Care Research Networks (PCRN), Comprehensive Local Research Networks (CLRNs) and National Institute for Social Care and Health Research – Coordinating Research Centre (NISCHR-CRC).

## Author details

<sup>1</sup>Academic Unit of Primary Health Care, School of Social and Community Medicine, University of Bristol, Canynge Hall, 39 Whatley Road, Clifton, Bristol BS8 2PS, UK. <sup>2</sup>South East Wales Trials Unit (SEWTU), Institute for Translation, Innovation, Methodologies and Engagement, School of Medicine, Cardiff University, 7th Floor Neuadd Meirionnydd, Heath Park, Cardiff CF14 4XN, UK. <sup>3</sup>Wales School of Primary Care Research (WSPCR), Institute of Primary Care & Public Health, School of Medicine, Cardiff University, 5th Floor Neuadd Meirionnydd, Heath Park, Cardiff CF14 4XN, UK. <sup>4</sup>School of Social and Community Medicine, University of Bristol, Canynge Hall, 39 Whatley Road, Clifton, Bristol BS8 2PS, UK. <sup>5</sup>Department of Primary Care and Public Health Sciences, King's College London, School of Medicine, 5th Floor Capital House, 42 Weston Street, London SE1 3QD, UK. <sup>6</sup>Department of Primary Medical Care, University of Southampton, Aldermeer Close, Southampton SO16 5ST, UK. <sup>7</sup>Specialist Antimicrobial Chemotherapy Unit, Public Health Wales Microbiology Cardiff, University Hospital Wales, Heath Park, Cardiff CF14 4XW, UK. <sup>8</sup>North Bristol NHS Trust, Southmead Hospital, Westbury-on-Trym, Bristol BS10 5NB, UK.

Received: 24 May 2012 Accepted: 4 July 2012

Published: 19 July 2012

## References

1. Shaikh N, Morone NE, Bost JE, Farrell MH: **Prevalence of urinary tract infection in childhood: A meta-analysis.** *Paediatr Infect Dis J* 2008, **27**:302–308.
2. Hay A, Whiting P, Butler C: **How Best to diagnose urinary tract infection in preschool children in primary care?** *Br Med J* 2011, **343**:d6316.
3. 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**:19–22.
4. Coulthard MG, Vernon SJ, Lambert HJ, Matthews JNS: **A nurse led education and direct access service for the management of urinary tract infections in children: prospective controlled trial.** *Br Med J* 2003, **327**:656.
5. Zorc JJ, Levine DA, Platt SL, Dayan PS, Macias CG, Krief W, *et al*: **Clinical and demographic factors associated with urinary tract infection in young febrile infants.** *Pediatrics* 2005, **116**:644–648.
6. van der Voort J, Edwards AG, Roberts R, Newcombe RG, Jones KV: **Unexplained extra visits to general practitioners before the diagnosis of first urinary tract infection: a case-control study.** *Arch Dis Child* 2002, **87**:530–532.
7. van der Voort J, Edwards A, Roberts R, Verrier JK: **The struggle to diagnose UTI in children under two in primary care.** *Fam Pract* 1997, **14**:44–48.
8. Loane V: **Obtaining urine for culture from non-potty-trained children.** *Paediatr Nurs* 2005, **17**:39–42.
9. NICE: *Urinary tract infection in children: diagnosis, treatment and long term management.* London; 2007.
10. Hoberman A, Chao HP, Keller DM, Hickey R, Davis HW, Ellis D: **Prevalence of urinary tract infection in febrile infants.** *J Pediatr* 1993, **123**:17–23.
11. Vernon SJ, Coulthard MG, Lambert HJ, Keir MJ, Matthews JN: **New renal scarring in children who at age 3 and 4 years had had normal scans with dimercaptosuccinic acid: follow up study.** *Br Med J* 1997, **315**:905–908.
12. Fötter R: **Pediatric Uroradiology.** In *Medical Radiology: Diagnostic Imaging.* Edited by Baert AL, Sartor K. Berlin: Springer; 2001.
13. Sacks SH, Verrier JK, Roberts R, Asscher AW, Ledingham JG: **Effect of symptomless bacteriuria in childhood on subsequent pregnancy.** *Lancet* 1987, **2**:991–994.
14. Jacobson SH, Eklof O, Eriksson CG, Lins LE, Tidgren B, Winberg J: **Development of hypertension and uraemia after pyelonephritis in childhood: 27 year follow up.** *Br Med J* 1989, **299**:703–706.
15. Smellie JM, Prescod NP, Shaw PJ, Risdon RA, Bryant TN: **Childhood reflux and urinary infection: a follow-up of 10–41 years in 226 adults.** *Pediatr Nephrol* 1998, **12**:727–736.
16. Coulthard MG: **Is reflux nephropathy preventable, and will the NICE childhood UTI guidelines help?** *Arch Dis Child* 2008, **93**:196–199.
17. Sharland M, Kendall H, Yeates D, Randall A, Hughes G, Glasziou P, *et al*: **Antibiotic prescribing in general practice and hospital admissions for peritonsillar abscess, mastoiditis, and rheumatic fever in children: time trend analysis.** *Br Med J* 2005, **331**:328–329.
18. Meropol SB, Chen Z, Metlay JP: **Reduced antibiotic prescribing for acute respiratory infections in adults and children.** *Br J Gen Pract* 2009, **59**:e321–e328.
19. Royston P, Moons KG, Altman DG, Vergouwe Y: **Prognosis and prognostic research: Developing a prognostic model.** *Br Med J* 2009, **338**:b604.
20. Altman DG, Royston P: **What do we mean by validating a prognostic model?** *Stat Med* 2000, **19**:453–473.
21. Gorelick MH, Shaw KN: **Clinical decision rule to identify febrile young girls at risk for urinary tract infection.** *Arch Pediatr Adolesc Med* 2000, **154**:386–390.
22. 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**:162–164.
23. Kass EH: **Pyelonephritis and bacteriuria. A major problem in preventive medicine.** *Ann Intern Med* 1962, **56**:46–53.
24. Health Protection Agency: *Investigation of urine. National Standard Method BSOP 41 Issue 7.* 2009. [http://www.hpa.org.uk/srmd/div\\_esl\\_su/pdf\\_bacteriology.htm](http://www.hpa.org.uk/srmd/div_esl_su/pdf_bacteriology.htm).
25. European Association of Urology: *Guidelines on Urological Infections.* Arnhem, The Netherlands; 2009.
26. Subcommittee on Urinary Tract Infection and 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**:595–610.
27. Whiting P, Westwood M, Bojke L, Palmer S, Richardson G, Cooper J, *et al*: **Clinical effectiveness and cost-effectiveness of tests for the diagnosis and investigation of urinary tract infection in children: a systematic review and economic model.** *Health Technol Assess* 2006, **10**:No.36.
28. Petersen I, Hayward AC: **Antibacterial prescribing in primary care.** *J Antimicrob Chemother* 2007, **60**:i43–i47.
29. Costelloe C, Metcalfe C, Lovering A, Mant D, Hay AD: **Effect of antibiotic prescribing in primary care on antimicrobial resistance in individual patients: systematic review and meta-analysis.** *Br Med J* 2010, **340**:c2096.
30. Downs SM: **Technical report: urinary tract infections in febrile infants and young children.** Urinary Tract Subcommittee of the American Academy of Pediatrics Committee on Quality Improvement. *Pediatrics* 1999, **103**:e54.
31. Liaw LCT, Nayar DM, Pedler SJ, Coulthard MG: **Home collection of urine for culture from infants by three methods: survey of parents' preferences and bacterial contamination rates.** *Br Med J* 2000, **320**:1312–1313.
32. Coulthard MG, Kalra M, Lambert HJ, Nelson A, Smith T, Perry JD: **Redefining urinary tract infections by bacterial colony counts.** *Pediatrics* 2010, **125**:335–341.
33. Wettergren B, Jodal U, Jonasson G: **Epidemiology of bacteriuria during the first year of life.** *Acta Paediatr Scand* 1985, **74**:925–933.
34. Wettergren B, Hellstrom M, Stokland E, Jodal U: **Six year follow up of infants with bacteriuria on screening.** *Br Med J* 1990, **301**:845–848.
35. Prais D, Shoov-Furman R, Amir J: **Is ritual circumcision a risk factor for neonatal urinary tract infections?** *Arch Dis Child* 2009, **94**:191–194.
36. Watson AR: **Management of urinary tract infection in children.** *Br Med J* 2007, **335**:356–357.
37. Van den Bruel A, Aertgeerts B, Bruyninckx R, Aerts M, Buntinx F: **Signs and symptoms for diagnosis of serious infections in children: a prospective study in primary care.** *Br J Gen Pract* 2007, **57**:538–546.
38. Fekkes M, Theunissen NCM, Brugman E, Veen S, Verrips EGH, Koopman HM, Vogels T, Wit JM, Verloove-Vanhorick SP: **Development and psychometric evaluation of the TAPQOL: A health-related quality of life instrument for 1–5 year-old children.** *Qual Life Res* 2000, **9**:961–972.
39. Newcombe RG: **Two-sided confidence intervals for the single proportion: comparison of seven methods.** *Stat Med* 1998, **17**:857–872.
40. Copas JB: **Using regression models for prediction: shrinkage and regression to the mean.** *Stat Meth Med Res* 1997, **6**:167–183.

doi:10.1186/1471-2334-12-158

**Cite this article as:** Downing *et al.*: 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 Infectious Diseases* 2012 **12**:158.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)








# Appendix 4 Case report forms

DUTY: **D**iagnosis of **U**rinary **T**ract Infections in **Y**oung Children Study

DATE  /  /  Site ID  Patient ID  -



## SECTION 1: SCREENING FORM

(Research Nurse/CSO to complete)

### Inclusion Criteria

Please exclude if the answer is 'NO' to any of the following:

Aged before their fifth birthday?	<input type="checkbox"/> Yes <input type="checkbox"/> No
Acute illness $\leq$ 28 days?	<input type="checkbox"/> Yes <input type="checkbox"/> No
First time in DUTY study?	<input type="checkbox"/> Yes <input type="checkbox"/> No
Ability of parent/carer to understand & give informed consent?	<input type="checkbox"/> Yes <input type="checkbox"/> No
Does the child have any urinary tract or constitutional symptoms (generally unwell)? See questions below (* see also examples):	
<b>Answer Yes to at least one of the screening questions to be eligible:</b>	<input type="checkbox"/> Yes <input type="checkbox"/> No
<input type="checkbox"/> Has the child been generally unwell (e.g. fever/feeling hot, generally not right, tired, irritable, crying more than usual, recently 'failing to thrive')?	<input type="checkbox"/> Yes <input type="checkbox"/> No
<input type="checkbox"/> Has the child had abdominal/loin pain or colic or vomiting or not eating/feeding as normal?	<input type="checkbox"/> Yes <input type="checkbox"/> No
<input type="checkbox"/> Has the child 'not been themselves' (e.g. clingy; not interested in what's going on; not playing well; low energy/tired; sleeping more or less than usual; irritable or not settling; crying more than usual)?	<input type="checkbox"/> Yes <input type="checkbox"/> No
<input type="checkbox"/> Has the child had any changes relating to passing urine (e.g. pain, changes in frequency, changes in continence, changes in urine smell or appearance or blood noticed)	<input type="checkbox"/> Yes <input type="checkbox"/> No

### Exclusion Criteria

Please exclude if the answer is 'YES' to any of the following:

Recently ( $\leq$ 28 days) or currently in any research study	<input type="checkbox"/> Yes <input type="checkbox"/> No
Presenting with trauma/injury as the predominant concern	<input type="checkbox"/> Yes <input type="checkbox"/> No
Has taken oral or IV/IM antibiotics in the past 7 days*	<input type="checkbox"/> Yes <input type="checkbox"/> No
Currently taking systemic (oral/IV or IM) steroids for $\geq$ 2 weeks (e.g. prednisolone or dexamethasone inhaled steroids are acceptable)	<input type="checkbox"/> Yes <input type="checkbox"/> No
Currently or recently ( $\leq$ 28 days) on chemotherapy or other immunosuppression (e.g. anti-rejection medications following renal transplant)	<input type="checkbox"/> Yes <input type="checkbox"/> No
Current use of urinary catheter (including intermittent use within past month)	<input type="checkbox"/> Yes <input type="checkbox"/> No
Previous bladder surgery	<input type="checkbox"/> Yes <input type="checkbox"/> No
Spina bifida or neurogenic bladder	<input type="checkbox"/> Yes <input type="checkbox"/> No
Due to laboratory transport pick up times, urine sample would not have reached lab in the next 24 hours (e.g. Friday afternoon in GP surgeries).	<input type="checkbox"/> Yes <input type="checkbox"/> No

\* Use of topical antibiotics does not exclude participation


### Inclusion Criteria Examples for RNs/CSOs/clinicians\*

- e.g. 1) well child with a sticky eye/ear but no other symptoms – exclude
- e.g. 2) child with sticky eye/ear who is crying more than usual – include
- e.g. 3) child with cough but no other symptoms – exclude
- e.g. 4) child with cough and fever, or cough and difficulty breathing – include
- e.g. 5) child with cough, no fever, but not feeding well – include
- e.g. 6) child with diarrhoea, no fever, no abdominal pain, eating normally – exclude
- e.g. 7) child with diarrhoea and fever/abdominal pain/reduced appetite – include

If the child satisfies all inclusion and exclusion criteria please continue to the next section

DUTY CRF SECTION 1-Screening Form v1.3 (17 October 2011)

Page 1 of 1

DUTY: **D**iagnosis of **U**rinary **T**ract Infections in **Y**oung Children Study 

Site ID:   Patient ID:  -

**SECTION 2: PARTICIPANT REGISTRATION FORM**  
(Research Nurse/CSO to ask parent/carer)

**CRF Details**

Date:   /   /     Time (24 hr clock):  :  RN/CSO Study ID:

Consent obtained  Yes  No  
 Recruited at own GP surgery?  Yes  No

*If No, please enter name of GP/surgery & address*

**Child's Details**

First Name:             Surname:

Date of Birth:   /   /     Gender:  Female  Male

NHS Number:

**Carer's Details**

First Name:             Surname:

Relationship to child:  Mother  Father  Other please specify:

Address:

Postcode:

Contact Telephone Number:                 Mobile Number:

Contact Notes:

DUTY: **D**iagnosis of **U**rinary **T**ract Infections in **Y**oung Children Study

Site ID



Patient ID

 - 




### Ethnicity

Please describe your child's ethnic group?

**White**  British  Irish  Any other White background

**Mixed**  White & Black Caribbean  White & Asian  Any other Mixed background  
 White & Black African

**Asian or Asian British**  Indian  Bangladeshi  Any other Asian background  
 Pakistani

**Black or Black British**  African  Caribbean  Any other Black background

**Chinese or Other Ethnic Group**  Chinese  Any other

**Group**

**Prefer not to answer**  *please tick here*

### Additional Social Economic Questions:

What is your highest level of qualification? (Please cross one box only)

Degree (or equivalent)  Diploma (or equivalent)

'A' level  GCSE / 'O' level

None  Not given

Other, please specify:

Thinking about the cost of living and how it affects your household which of the following would best describe your situation?


find it a strain to get by from week to week  able to manage without much difficulty

have to be careful about money  quite comfortably off

Not given

DUTY: **D**iagnosis of **U**rinary **T**ract Infections in **Y**oung Children Study

DATE  /  /  Site ID  Patient ID  -



**SECTION 3: CASE REPORT FORM**  
(Research Nurse/CSO to ask parent/carer)

**Current Symptoms**

How many days (including today) has your child been unwell? Enter number (1- 28)

Compared to yesterday is your child same, better or worse?  Same  Better  Worse


Please rate your overall impression of your child's current illness when at its worst from 0-10

0  1  2  3  4  5  6  7  8  9  10  
 Completely Well Extremely Unwell

For each symptom, score symptom when it was at its **worst during this illness**.


Symptom	Severity				
	No Problem	Slight Problem	Moderate problem	Severe problem	Don't Know/ NA
Child 'not themselves' (e.g.: clingy; not interested in what's going on; not playing well; low energy/tired; irritable or not settling; crying more than usual).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Confused or Disorientated	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Disturbed sleep	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Fever at any time during this illness (fever is feeling hot or cold )	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Fever now or in the past 24 hours	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Chills or Shivering	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
New generalised rash with this illness (not worsening of existing skin conditions)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Nappy rash or similar	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Muscle aches or pains all over	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Headaches	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Refused feeds/eating less than normal	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Poor weight gain or weight loss (in the last month)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vomiting	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Diarrhoea (at any time)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Diarrhoea (in the past 24hrs)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Constipation in the last week	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Abdominal pain/tummy ache/pulling up legs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

DUTY: **D**iagnosis of **U**rinary **T**ract Infections in **Y**oung Children Study

DATE  /  /     Site ID   Patient ID  -     

Symptom	Severity				
	No Problem	Slight Problem	Moderate problem	Severe problem	Don't Know/NA
Passing urine more often	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Any changes in urine appearance (darker, cloudy, smelly or blood)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
If YES, please state:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Darker	Cloudy	Smelly	Bloody	Other
If Other, please describe:	<input type="text"/>				
Pain/crying when passing urine	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Day or bed wetting when previously dry *	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Blocked or runny nose	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Cough	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Wheeze	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Short of breath, difficulty breathing or grunting	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Chest pains	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Earache or holding ear/s	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sore throat	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
More unwell compared to similar previous illnesses	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Any other symptoms	<input type="checkbox"/>	Yes	<input type="checkbox"/>	No	
If YES, please state:	<input type="text"/>				

\* Not applicable for children wearing nappies day & night

DUTY: **D**iagnosis of **U**rinary **T**ract Infections in **Y**oung Children Study 

DATE   /   /    Site ID   Patient ID  -

### Past Medical & Family History

These questions relate to other medical problems your child has or has had in the past. There are also some questions concerning the pregnancy, birth and early life of the child.

Not counting today, approximately how many times has your child previously consulted a doctor or nurse for this episode of illness?

0  1  2  3  4  5  6  7  8  9  10  10+

Does your child have any on-going health problems?  Yes  No

If YES, please indicate:

asthma  diabetes  heart disease  high blood pressure  learning disabilities

other please specify:

Was the pregnancy full term for your child?  Yes  No

If NO, please indicate:  Born Late  Born Early If Early, estimate weeks:

Was your child breastfed?  Yes  No

If YES, for how long exclusively?  < 3months  ≥ 3months

Were you ever told that your child's kidney, bladder or urinary system was abnormal in any way after a pregnancy ultrasound scan?  Yes  No

If YES, please give details:

In boys ONLY – has your child been circumcised?  Yes  No

The next set of questions relate to your child and to close family members of your child (this includes mother, father and any brothers or sisters who are blood relatives of the child). We are asking these questions to see whether urine infections run in families.

Has your child or member of your family ever been diagnosed with vesico-ureteric or 'kidney' reflux?  Yes  No  Don't Know

If YES, please indicate:  Child  Mother  Father  Sibling


Has your child or any member of your family ever been diagnosed or treated for urine infections?  Yes  No  Don't Know

If YES, please indicate:  Child  Mother  Father  Sibling

Does your child or any member of your family have any other renal/urinary problem?  Yes  No  Don't Know

If YES, please indicate:  Child  Mother  Father  Sibling

If YES, please give details:

DUTY: **D**iagnosis of **U**rinary **T**ract Infections in **Y**oung Children Study 

DATE  /  /  Site ID  Patient ID  -

### Medications

Is your child currently taking any medication?  Yes  No

If **YES**, please indicate:

laxatives (e.g. lactulose, senna, modecol)  steroid inhaler (e.g. clenil)  beta<sub>2</sub> agonist inhaler (e.g. salbutamol)  antihistamine (e.g. piriton)

paracetamol  ibuprofen

other medication (please provide details):

### Toileting behaviour

Does your child use nappies/pull-ups?  Yes  No

If **YES**, please indicate:  Day  Night  Both

Approximately how many nappies/pull-ups has your child used in the last 24 hours? Estimate number

How many times do you usually bath or shower your child in a normal week? Estimate number

### Examination

To be completed by Research Nurse/CSO or responsible clinician

Please tick here if child refuses to be examined

Temperature  .  °C (range 35.0 – 42.0)

DUTY standard 'Thermoscan' thermometer used to check temp?  Yes  No

If **NO**, please indicate:  infrared ear  digital auxiliary  other:

O<sub>2</sub> saturation  % (range 80 – 100%)

Pulse rate  (range 80 – 250 bpm)

Respiratory rate  (range 20 – 80 rpm)


Capillary refill time\*  < 2 sec  2-5 sec  > 5 sec

\* The Advanced Paediatric Life Support manual<sup>1</sup> recommends that this should be done by pressing on the sternum for 5 seconds and then recording the time it takes for the skin colour to change from white back to pink.

Has the child's responsible clinician been informed of these results?  Yes **NB: this must be done for all children**

DUTY: **D**iagnosis of **U**rinary **T**ract Infections in **Y**oung Children Study

DATE  /  /  Site ID  Patient ID  -


**SECTION 4: CASE REPORT FORM**

(GP/responsible clinician to complete)

**Clinician's Details**

Clinician Name

Clinician Profession

 Doctor Nurse**Clinical Examination**

Please give your global impression of the child in a scale of 0-10

0	1	2	3	4	5	6	7	8	9	10	
Completely Well											Extremely Unwell


In order to for all clinicians, irrespective of profession (nurse vs. doctor) or their location (GP practice vs. Walk in Centre vs. Emergency Department) to complete the global impression scale in a similar fashion, the following guidance has been derived:

A zero score is appropriate for a child who is displaying no 'constitutional' upset. Child is fully conscious and alert, aware/interacting with their surroundings (e.g. smiling at parent/guardian), feeding/eating normally with normal tone, normal peripheral circulation. No fever, pain/discomfort, tachypnoea or respiratory distress at the time of assessment.

A score of 5 is appropriate for a child who is displaying 'constitutional' symptoms / signs warranting secondary care assessment but may not require admission. The primary care clinician would feel a secondary care assessment / period of observation was necessary. Emergency Department clinician would not feel happy to discharge without laboratory / radiographic tests or a period of observation.

A score of 10 is appropriate for a child who is displaying life-threatening signs requiring immediate resuscitation such as unconsciousness and / or circulatory collapse.




DUTY: **D**iagnosis of **U**rinary **T**ract Infections in **Y**oung Children Study 

DATE  /  /  Site ID  Patient ID  -

Please examine all of the following in all children.  
For each question, please indicate: 'normal' or 'abnormal'. If abnormal, tick all that apply.

Examination	Observations
<b>Hydration</b>	<input type="checkbox"/> Normal <input type="checkbox"/> Some dehydration <input type="checkbox"/> Severe dehydration <input type="checkbox"/> Not examined
<b>Conscious Level</b>	<input type="checkbox"/> Normal <input type="checkbox"/> Drowsy <input type="checkbox"/> Irritable <input type="checkbox"/> Not examined
<b>General</b>	<input type="checkbox"/> Normal <input type="checkbox"/> Abnormal <input type="checkbox"/> Not examined
<i>If abnormal please tick all that apply:</i>	
<input type="checkbox"/> Pallor <input type="checkbox"/> Flushed <input type="checkbox"/> Jaundice <input type="checkbox"/> Distressed <input type="checkbox"/> Lymphadenopathy	
<input type="checkbox"/> Other (please specify): <input type="text"/>	
<b>Throat</b>	<input type="checkbox"/> Normal <input type="checkbox"/> Abnormal <input type="checkbox"/> Not examined
<i>If abnormal please tick all that apply:</i>	
<input type="checkbox"/> Red or Inflamed <input type="checkbox"/> Swollen <input type="checkbox"/> Quinsy <input type="checkbox"/> Discharge or Pus	
<input type="checkbox"/> Other (please specify): <input type="text"/>	
<b>Ears</b>	<input type="checkbox"/> Normal <input type="checkbox"/> Abnormal <input type="checkbox"/> Not examined
<i>If abnormal please tick all that apply:</i>	
<input type="checkbox"/> Pink <input type="checkbox"/> Red or Bulging <input type="checkbox"/> Fluid Level	
<input type="checkbox"/> Acute Perforation <input type="checkbox"/> Chronic Perforation	
<input type="checkbox"/> Other (please specify): <input type="text"/>	
<b>Chest</b>	<input type="checkbox"/> Normal <input type="checkbox"/> Abnormal <input type="checkbox"/> Not examined
<i>If abnormal please tick all that apply:</i>	
<input type="checkbox"/> Bronchial Breathing <input type="checkbox"/> Unilateral <input type="checkbox"/> Bilateral	
<input type="checkbox"/> Wheeze <input type="checkbox"/> Unilateral <input type="checkbox"/> Bilateral	
<input type="checkbox"/> Crackles <input type="checkbox"/> Unilateral <input type="checkbox"/> Bilateral	
<input type="checkbox"/> Recession <input type="checkbox"/> Grunting <input type="checkbox"/> Nasal Flaring	
<small>(any intercostal/subcostal/supra-clavicular)</small>	
<input type="checkbox"/> Other - please specify: <input type="text"/>	
<b>Abdomen</b>	<input type="checkbox"/> Normal <input type="checkbox"/> Abnormal <input type="checkbox"/> Not examined
<i>If abnormal please tick all that apply:</i>	
<input type="checkbox"/> Mass or Organomegaly present <i>If Yes, please state:</i>	
<input type="checkbox"/> Loin Tenderness <input type="checkbox"/> Suprapubic Tenderness	
<input type="checkbox"/> Other - please specify: <input type="text"/>	

DUTY: <b>D</b> iagnosis of <b>U</b> rinary <b>T</b> ract Infections in <b>Y</b> oung Children Study					
<b>DATE</b>		<b>Site ID</b>	<b>Patient ID</b>		
<input type="text"/> / <input type="text"/> / <input type="text"/>		<input type="text"/>	<input type="text"/> - <input type="text"/>		

---

**Diagnosis and Planned Management Prior to Urine Dipstick Results**

---

We would prefer if you answer these questions prior to seeing the urine dipstick results:

---

Are you currently aware of the urine dipstick results?  Yes  No

---

What is your working diagnosis?

URTI  chest infection  bronchitis  bronchiolitis  pneumonia

exacerbation of asthma  
(infective or non-infective)  tonsillitis  otitis media  pharyngitis

UTI  gastroenteritis  viral illness

other *please state*:

---

How certain are you of this diagnosis at this point?

uncertain  fairly certain  certain  very certain

---

Before seeing dipstick results, are you planning on treating this child with antibiotics?

No  Yes [for suspected UTI]  Yes [for other reason]

If Yes:  immediate script  delayed script

---

Before seeing dipstick results, would you have referred this child to a paediatrician or admitted this child to hospital?

No  Yes [for suspected UTI]  Yes [for other reason]

N/A [recruited in ED]

---

If this child was **NOT** in the DUTY study would you have requested a urine sample?

Yes  No


---





DUTY: **D**iagnosis of **U**rinary **T**ract Infections in **Y**oung Children Study

DATE  /  /  Site ID  Patient ID  -



### Sample Processing

Has the clinician taking responsibility for this child's care been informed of the dipstick results?

Yes

*NB this must be done for all children*

Has the urine sample been divided, labelled with DUTY stickers and sent to the usual local NHS laboratory AND the central DUTY laboratory?

Yes: sent to NHS lab only.

Yes: sent to both labs.

No: dipstick test only.

\*NHS sample takes priority over the research lab sample if there is less than 2ml of urine.

Has the participant been given a £5 voucher?

Yes  No

Please record Voucher serial number here:



# Appendix 5 Day-14 data collection forms



University of  
**BRISTOL**

ACADEMIC UNIT OF PRIMARY HEALTH CARE  
25 Belgrave Road, Clifton, Bristol.

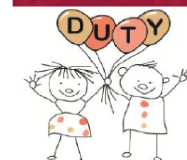


NIHR Health Technology Assessment programme

NIHR  
National Institute for  
Health Research



South East Wales  
Trials Unit  
Uned Ymchwil  
De-ddwyrain Cymru



DUTY: Diagnosis of Urinary Tract Infections in Young Children Study

Date: 

d	d	m	m	y	y
---	---	---	---	---	---

PID: 

--	--	--	--

\*\* once PID is confirmed, the following should appear on database

Child's Name 


Centre ID 

--

Parent/Carer's  
Details: 


Recruitment  
Location 

GP surgery, WIC, OOH, A&E
------------------------------

## Telephone administered Symptoms and Resource use data questionnaire (up to day 14)

*Instructions to project administrator: for this section, please ask the parent/guardian the following questions and complete on behalf of the parent/guardian.*

### Symptoms

I am going to ask you about how [name of child] has been since you went to the doctor and agreed that [he/she] would take part in the DUTY study.

1. How many days since [name of child] joined the DUTY study (day 0) was it until [his/her] symptoms improved?

<i>Fill in one box below</i>	
Number of days (enter number)	
OR symptoms not yet improved (tick box)	

2. How many days since [name of child] joined the DUTY study (day 0) was it until [he/she] was entirely well AND had returned to their normal activities of daily living (e.g. feeding, sleeping, playing, going to play group) for two consecutive days?

<i>Fill in one box below</i>	
Number of days (enter number)	
OR not yet returned to normal activities (tick box)	

**Resource Use – Primary Care Visits**

The next few questions are about health care use for [name of child] in the two weeks since you went to the doctor and agreed that [name of child] would take part in the DUTY study. Throughout this questionnaire, we use the term '14 days' to refer to this time period.

We are interested in your use of NHS services, and anything you have spent caring for your child, and any other costs you have incurred because of your child's illness.

Can you tell me if you've needed to use any of the following for [name of child]? If your child joined the DUTY study at the GP surgery, please **DO NOT** include the visit when you agreed to take part.

We would like to know **how many** consultations, you have had, **where** they took place, and **who** the consultation was with.

	Number of contacts (if none, write 'zero')
<b>3. PRACTICE- BASED CONTACTS IN- HOURS</b>	
General Practitioner at the surgery	
Practice Nurse/Nurse Practitioner at the surgery	
Telephone consultation with doctor at the surgery	
Telephone consultation with nurse at the surgery	
Home visit by the doctor	
<b>4. OUT- OF- HOURS CONTACTS FROM OOH CENTRE</b>	
Out-of Hours telephone contact - nurse	
Out-of Hours telephone contact - doctor	
Out-of Hours face-to-face contact - nurse	
Out-of Hours face-to-face contact - doctor	
Out-of-Hours home visit by doctor from OOH centre	
<b>5. COMMUNITY BASED AND OTHER CONTACTS</b>	
Walk-in Centre – nurse	
NHS Direct	
Health Visitor	
Baby Clinic	
Other (e.g. community nurse, midwife, please specify)	



**Resource Use – Hospital Visits**

Please can you now think about any contacts [name of child] has had at **any hospital** during the last **14 days**. If your child joined the DUTY study at the Accident and Emergency Department, please **DO NOT** include that visit when answering.

7. Has [name of child] visited an **Accident and Emergency** department during the last **14 days**?

No   
Yes  → Number of visits .....

Reason for visit 1:  
For same illness: No  → Please state: \_\_\_\_\_  
Yes

Reason for visit 2:  
For same illness: No  → Please state: \_\_\_\_\_  
Yes

8. Has [name of child] visited a **hospital clinic** during the last **14 days**?

No   
Yes  → Number of visits .....

Please describe the reason for visit/s and the type of clinic (e.g. paediatrics):

Reason for visit 1:  
For same illness: No  → Please state: \_\_\_\_\_/Clinic: \_\_\_\_\_  
Yes

Reason for visit 2:  
For same illness: No  → Please state: \_\_\_\_\_/Clinic: \_\_\_\_\_  
Yes

9. Has [name of child] stayed **overnight in hospital (admitted to hospital)** during the last **14 days**?

No   
Yes  → Please give the number of nights and reason for stay:

Admission 1: Number of nights: \_\_\_\_\_  
Reasons for stay: For same illness: No  → Please state: \_\_\_\_\_  
Yes

Admission 2: Number of nights: \_\_\_\_\_  
Reasons for stay: For same illness: No  → Please state: \_\_\_\_\_  
Yes

10. Has [name of child] needed an **ambulance** during the last **14 days**?

No   
Yes  → Please give details of journey and any treatment received:

---

Could you tell me about any expenses you have had because of these and how much these have cost you? For example, if you travelled by public transport or taxi, you will have paid a fare. If you travelled by private car you may have paid for car parking. If you travelled by private car, we would also like to know the approximately how far you travelled.

*[Note to research nurse: Please complete all that apply: each visit may involve more than one means of transport and may include any number of people. We are interested in the total cost of each visit, record the return journey. DO NOT include the recruitment visit]*

*[Please continue on a separate sheet if more than 4 visits]*

6		Primary Care Visit 1	Primary Care Visit 2	Primary Care Visit 3	Primary Care Visit 4
No Primary Care visits: tick here <input type="checkbox"/>					
Destination e.g. GP surgery, Walk-in Centre, Other (specify)					
Total cost of return fares paid for all people travelling together	Bus	£..... _ .....p	£..... _ .....p	£..... _ .....p	£..... _ .....p
	Taxi	£..... _ .....p	£..... _ .....p	£..... _ .....p	£..... _ .....p
	Train/Tube	£..... _ .....p	£..... _ .....p	£..... _ .....p	£..... _ .....p
Total distance of return journey	Car or other motor vehicle	.....miles	.....miles	.....miles	.....miles
Parking charge		£..... _ .....p	£..... _ .....p	£..... _ .....p	£..... _ .....p
Other cost (please specify)		£..... _ .....p	£..... _ .....p	£..... _ .....p	£..... _ .....p
If no cost incurred e.g. walk, cycle, please tick box		No Cost	No Cost	No Cost	No Cost

**Hospital Tests**

We would now like to know about any **special tests** [name of child] has received during the last **14 days**.

11. Has [he/she] had any x-rays, ultrasound or MRI scans during the last **14 days**?

No

Yes

→ Type of tests, [number] & location

..... [ ] .....

..... [ ] .....

..... [ ] .....

**Hospital Travel Costs\***

Please could you think about all the **journeys to hospital** that you or members of your immediate family have made during the **14 day** time period because of [name of child's] health. This includes taking a child to hospital for treatment plus any journeys made to visit a child who is in hospital. If your child joined the DUTY study at the Accident and Emergency Department, please **DO NOT** include that visit when answering

We would like to know about the cost of travel: If you travelled by public transport or taxi, you will have paid a fare. If you travelled by private car you may have paid for car parking. If you travelled by private car, we would also like to know the approximately how far you travelled.

*[Note to research nurse: Please complete all that apply: each visit may involve more than one means of transport and may include any number of people. We are interested in the total cost of each visit. Record the return journey. DO NOT include the recruitment visit if recruited in A&E]*

*[Please continue on a separate sheet if more than 4 visits]*

12		Visit 1	Visit 2	Visit 3	Visit 4
No Hospital visits: tick here <input type="checkbox"/>					
Destination e.g. Hospital, A&E Dept, Out-Patients Dept etc					
Total cost of return fares paid for all people travelling together	Bus	£..... _ .....p	£..... _ .....p	£..... _ .....p	£..... _ .....p
	Taxi	£..... _ .....p	£..... _ .....p	£..... _ .....p	£..... _ .....p
	Train/Tube	£..... _ .....p	£..... _ .....p	£..... _ .....p	£..... _ .....p
Total distance of return journey	Car or other motor vehicle	.....miles	.....miles	.....miles	.....miles
Parking charge		£..... _ .....p	£..... _ .....p	£..... _ .....p	£..... _ .....p
Other cost (please specify)		£..... _ .....p	£..... _ .....p	£..... _ .....p	£..... _ .....p
If no cost incurred e.g. walk, cycle, please tick box					

**Medication**

Please could you tell us about any **medication** [name of child] has received during the last **14 days**.

Firstly, has [name of child] taken any medications prescribed by a Doctor since taking part in the DUTY study? This may have been prescribed by your GP (family doctor) or by a hospital doctor. If possible, please copy the information from the label on the bottle or packet.

13	Name	Date started	Strength / Dose	Formulation e.g. suspension/ suppository	Quantity and frequency of dose	Total number of doses given in last 14 days
Example	Trimethoprim	01/01/10	50mg/5ml	Oral suspension	5ml twice daily	10
Medicine 1						
Medicine 2						
Medicine 3						
Medicine 4						
Medicine 5						
Medicine 6						

Secondly, has [name of child] taken any other **medication in the last 14 days** without a prescription e.g. bought over-the-counter at a chemist or supermarket?

14	Name	Size of packet / bottle	Proportion used	Cost of packet / bottle
Example	Calpol infant suspension	100 ml	One third	£2-78
Item 1				£..... _ .....p
Item 2				£..... _ .....p
Item 3				£..... _ .....p
Item 4				£..... _ .....p
Item 5				
Item 6				

**Expenses**

Finally, we'd like to ask you about any **other out-of-pocket expenses** you or your immediate family members have had to pay as a result of [name of child] illness during the past **14 days**.

15. Have you or any members of your immediate family taken **time off work** during the past **14 days** because of [name of child] illness?

No

Yes



Please give the number of days in total (to nearest half)

.....

16. Have you or any members of your immediate family suffered any **loss of earnings** as a result of taking time off work to care for [name of child] in the past **14 days**?

No

Yes



Please give the approximate amount of gross income that has been lost in total: £..... ..p

17. Have you incurred any extra costs of **childcare** or care of other dependents, during the past **14 days**, because of [name of child] illness?

No

Yes



Please give the approximate total amount: £..... ..p

**Feedback from Parent**

18. Please tell us about your overall experience of the study:

# TAPQOL

# Questionnaire

for parents of children aged 9 months to 6 yrs

Would you please answer the following questions first?

Is the child in question a boy or a girl?

boy

girl

What is the child's date of birth?

(day)

(month)

(year)

On what date was this questionnaire completed?

(day)

(month)

(year)

© TNO - Prevention and Health /LUMC, 2001

PID:

# INSTRUCTIONS

**Dear Sir / Madam,**

The questions in this questionnaire relate to all kinds of different aspects of your child's health. You can answer the questions by ticking the answer which best describes your child.

For example:

In the last 14 days, has your child had ..

**Ear-ache**  never  occasionally  often

**1**

At that time, my child felt:

fine  not so good  quite bad  bad

---

If things were not entirely satisfactory, you are also asked how your child felt when there was a problem. So, if you say that your child had ear-ache 'occasionally' or 'often', you can state, in the second part of the question, how your child felt at that time.

For example:

In the last 14 days, has your child had ..

**Ear-ache**  never  occasionally  often

**1**

At that time, my child felt:

fine  not so good  quite bad  bad

---

In the last 14 days, has your child had ..

**Stomach-ache or abdominal pain**       never       occasionally       often  
**1**      **At that time, my child felt:**  
 fine       not so good       quite bad       bad

---

**Colic (uncontrollable crying)**       never       occasionally       often  
**2**      **At that time, my child felt:**  
 fine       not so good       quite bad       bad

---

**Eczema**       never       occasionally       often  
**3**      **At that time, my child felt:**  
 fine       not so good       quite bad       bad

---

**Itchiness**       never       occasionally       often  
**4**      **At that time, my child felt:**  
 fine       not so good       quite bad       bad

---

**Dry skin**       never       occasionally       often  
**5**      **At that time, my child felt:**  
 fine       not so good       quite bad       bad

---

**Bronchitis**       never       occasionally       often  
**6**      **At that time, my child felt:**  
 fine       not so good       quite bad       bad

---

**Difficulty with breathing or lung problems**       never       occasionally       often  
**7**      **At that time, my child felt:**  
 fine       not so good       quite bad       bad

---



**In the last 14 days, has your child been ...**

Short of breath

 never  occasionally  often

8

At that time, my child felt:

 fine  not so good  quite bad  bad

Nauseous (feeling sick)

 never  occasionally  often

9

At that time, my child felt:

 fine  not so good  quite bad  bad**How did your child sleep in the last 14 days?**

Did your child sleep restlessly?

 never  occasionally  often

10

At that time, my child felt:

 fine  not so good  quite bad  bad

Was your child awake at night?

 never  occasionally  often

11

At that time, my child felt:

 fine  not so good  quite bad  bad

Did your child cry at night?

 never  occasionally  often

12

At that time, my child felt:

 fine  not so good  quite bad  bad

Did your child have difficulty sleeping through the night?

 never  occasionally  often

13

At that time, my child felt:

 fine  not so good  quite bad  bad

### How did your child eat and drink in the last 14 days?

Was your child's appetite poor?

never  occasionally  often

14

At that time, my child felt:

fine  not so good  quite bad  bad

Did your child have difficulty eating enough?

never  occasionally  often

15

At that time, my child felt:

fine  not so good  quite bad  bad

Did your child refuse to eat?

never  occasionally  often

16

At that time, my child felt:

fine  not so good  quite bad  bad

**Your child's behaviour in the last 14 days?****My child was short-tempered** never  occasionally  often**17**

---

**My child was aggressive** never  occasionally  often**18**

---

**My child was irritable** never  occasionally  often**19**

---

**My child was angry** never  occasionally  often**20**

---

**My child was restless or impatient with me** never  occasionally  often**21**

---

**My child was defiant / awkward with me** never  occasionally  often**22**

---

**I could not manage my child** never  occasionally  often**23**

---

**How was your child in the last 14 days?**

In good spirits

 never  occasionally  often**24**

---

Cheerful

 never  occasionally  often**25**

---

Happy

 never  occasionally  often**26**

---

Frightened

 never  occasionally  often**27**

---

Tense

 never  occasionally  often**28**

---

Anxious

 never  occasionally  often**29**

---

Energetic

 never  occasionally  often**30**

---

Active

 never  occasionally  often**31**

---

Lively

 never  occasionally  often**32**

---

If your child is aged below eighteen months, you do not have to complete the rest of this questionnaire.

If your child is older than eighteen months, you should continue with the questions on the following pages.

### How was your child's behaviour with other children in the last 14 days?

My child was able to play happily with other children

never  occasionally  often

33

---

My child was at ease with other children

never  occasionally  often

34

---

My child was confident with other children

never  occasionally  often

35

---

### In the last 14 days, did your child have, compared to other children of the same age ...

Difficulty with walking

no  yes, a little  yes, a lot  cannot walk

36

At that time, my child felt:

fine  not so good  quite bad  bad

---

Difficulty with running

no  yes, a little  yes, a lot  cannot walk

37

At that time, my child felt:

fine  not so good  quite bad  bad

---

Difficulty with walking up stairs without help?

no  yes, a little  yes, a lot  cannot walk

38

At that time, my child felt:

fine  not so good  quite bad  bad

---

Difficulty with balance

no  yes, a little  yes, a lot  cannot walk

39

At that time, my child felt:

fine  not so good  quite bad  bad

---

In the last 14 days, did your child have, compared to other children of the same age ...

Difficulty in understanding what others said?  never  occasionally  often  
**40**  
 At that time, my child felt:  
 fine  not so good  quite bad  bad

Difficulty in talking clearly?  never  occasionally  often  
**41**  
 At that time, my child felt:  
 fine  not so good  quite bad  bad

Difficulty in saying what he / she meant?  never  occasionally  often  
**42**  
 At that time, my child felt:  
 fine  not so good  quite bad  bad

Difficulty in making it clear what he / she wanted?  never  occasionally  often  
**43**  
 At that time, my child felt:  
 fine  not so good  quite bad  bad

This is the end of the questionnaire.

Thank you for completing it!






# Appendix 6 Three-month follow-up data collection form

DUTY: **D**iagnosis of **U**rinary **T**ract Infections in **Y**oung Children Study

DATE:   /   /

Site ID:

Patient ID:  -



## 3 MONTHS PRIMARY CARE NOTES REVIEW

(DUTY Research Nurse/Clinical Studies Officer (RN/CSO) to complete)

### Practice Details

Practice Name:

Practice Site ID:

Reviewers Name:

### Child's Details

Patient ID:  -

Date of Birth:

/   /

D D M M Y Y Y Y

Gender:

Female  Male

NHS Number:

Study entry date: (i.e. date of consultation):

/   /

D D M M Y Y Y Y

Note review end date: (i.e. 3 month from recruitment):

/   /

D D M M Y Y Y Y

Patient has left this surgery (tick if yes):

Date (if known):

/   /

D D M M Y Y Y Y

Patient has died (tick if yes):

Date (if known):

/   /

D D M M Y Y Y Y

### Please Note:

Data collected here refers to primary care contacts in the **3 MONTHS AFTER** the initial recruitment visit (for example, if recruited on 10/Sept/2010, 3month review end date is 10/Dec/2010).

Please do not include data recorded at the initial recruitment visit; any subsequent contacts on the day of recruitment can be included.


3 Months Case Note Review Form (v1.7) FINAL(03-05-2012) (ETJ).docx

DUTY: **D**iagnosis of **U**rinary **T**ract Infections in **Y**oung Children Study

DATE   /   /

Site ID

Patient ID  -



**Primary Care Visits**

Please record how many consultations (not including routine immunisations and screening checks) have taken place.

Please also include DNAs (e.g. appointments that were missed) in the total number of contacts.

1. PRACTICE- BASED CONTACTS IN- HOURS	Number of contacts
General Practitioner at the surgery	<input type="text"/> <input type="text"/>
Practice Nurse/Nurse Practitioner at the surgery	<input type="text"/> <input type="text"/>
Telephone consultation with doctor at the surgery	<input type="text"/> <input type="text"/>
Telephone consultation with nurse at the surgery	<input type="text"/> <input type="text"/>
Home visit by the doctor	<input type="text"/> <input type="text"/>
Walk in centres visits	<input type="text"/> <input type="text"/>


**2. OUT- OF- HOURS CONTACTS FROM OOH CENTRE**

Out-of Hours telephone contact – nurse	<input type="text"/> <input type="text"/>
Out-of Hours telephone contact – doctor	<input type="text"/> <input type="text"/>
Out-of Hours face-to-face contact – nurse	<input type="text"/> <input type="text"/>
Out-of Hours face-to-face contact – doctor	<input type="text"/> <input type="text"/>
Out-of-Hours home visit by doctor from OOH centre	<input type="text"/> <input type="text"/>

Please record the number of urine samples sent to the laboratory:

Please provide full results below, and attach a copy)

	Date of test	Laboratory Result (provide microscopy and culture information) (please tick if copy of results attached)
Urine test 1		<input type="checkbox"/>
Urine test 2		<input type="checkbox"/>
Urine test 3		<input type="checkbox"/>

DUTY: **D**iagnosis of **U**rinary **T**ract Infections in **Y**oung Children Study 

DATE  /  /  Site ID  Patient ID  -

Urine test 4	<input type="checkbox"/>
--------------	--------------------------

### Hospital Visits

3. Has the child attended any hospital visits during the last 3 months? No  Yes

Please include any DNAs (e.g. appointments that were missed).

If 'Yes' please record information relating to any contacts the child has had below:

Please note: Do not include data from the initial recruitment visit if the child was recruited at an A&E Dept.

4. Accident and Emergency department visits during the last 3 months?

No   
Yes  → Total number of visits:

Please record the following for each visit:

	Date of visit	Reason for visit: <i>please state</i>
VISIT 1		
VISIT 2		
VISIT 3		


5. Hospital clinic visits during the last 3 months?

No   
Yes  → Total number of visits:

Please record the following for each visit:

	Date of visit	Type of Clinic (e.g. Paediatrics)	Reason for visit: <i>please state</i>
VISIT 1			
VISIT 2			
VISIT 3			

DUTY: **D**iagnosis of **U**rinary **I**nfections in **Y**oung Children Study

DATE   /   /     Site ID   Patient ID  -     

6. Hospital stays during the last 3 months?

No   
 Yes  → Total number of visits:

Please record the following for each visit:

	Admission Date	Length of stay (# of nights, enter '0' if day case)	Reason for stay: please state
VISIT 1			
VISIT 2			
VISIT 3			

### Hospital Tests

7. Has the child had any hospital tests during the last 3 months?

No   
 Yes  → Total number of tests:

Please record the following for each test:

	Date of test	Type of test*	Results: please state
TEST 1			
TEST 2			
TEST 3			


\* e.g. Ultrasound scan (USS), DMSA scan, Micturating Cysto-urethrogram (MCUG), Indirect MAG3 Cystogram, or any other test

DUTY: **D**iagnosis of **U**rinary **T**ract Infections in **Y**oung Children Study

DATE   /   /

Site ID

Patient ID  -



### Prescribed Medication

8. Please record details of any **prescribed medication** the child received at the index visit and any subsequent visits in the last 3 months.

Please also record any repeat prescriptions.

	Name	Date prescribed	Strength / Dose	Formulation e.g. suspension/ suppository	Quantity and frequency of dose	Length of Dose	Tick if prescribed at recruitment visit
<i>Example</i>	<i>Trimethoprim</i>	<i>01/01/10</i>	<i>50mg/5ml</i>	<i>Oral suspension</i>	<i>5ml twice daily</i>	<i>5 days</i>	
Medicine 1							
Medicine 2							
Medicine 3							
Medicine 4							



## Appendix 7 Laboratory data entry forms

NHS LABORATORY MICROBIOLOGY INFORMATION – entered onto DUTY online database

### *i) Patient Details*

PATIENT UNIQUE ID	
PATIENT D.O.B	dd/mm/yyyy
REGIONAL ID	Regional/study site unique code*
LAB ID	Unique identifier
Date Sample Received at Lab	dd/mm/yyyy
Time Sample Received at Lab	xx.xx (24hr format)
Date Sample Processed	dd/mm/yyyy
Time Sample Processed	xx.xx (24hr format)
Date Sample Reported	dd/mm/yyyy
Time Sample Reported	xx.xx (24hr format)
Able to process urine: Yes/No	If 'no' drop down box:  Leaked sample  Other (please specify)  <i>free text box</i>

### *ii) Urine Results*

#### Microscopy

Microscopy performed: Yes/No			
If Yes indicate method:	Manual	Automated	Not Measured

	Count (mm <sup>3</sup> )					
Microscopy Results	WBC	<10	10-30	30-100	>100	Absolute WBC count
	RBC	0	<5	6-100	>100	
	Squamous Epithelial Cells	0-5	6-75	>75		Not recorded

### Culture

Culture performed: Yes/No

If No: please give reason (*free text*)

### ii) Growth

No Growth	No significant growth		
<b>Growth:</b> (please indicate)	<10 <sup>3</sup> CFU/mL	10 <sup>3</sup> -10 <sup>5</sup> CFU/mL	>10 <sup>5</sup> CFU/mL

<b>Purity:</b> (please indicate)	Predominant	Mixed growth (2 species)	Mixed growth (> 2 species)
<b>Speciation:</b> (please indicate)	For organism 1 (see below)	For organism 1 & 2 (see below)	Not Applicable
<b>Sensitivity:</b> (please indicate)	For organism 1 (see below)	For organism 1 & 2 (see below)	Not Applicable



## iii) Speciation and Sensitivity (for each organism identified from above)

Organism 1 (complete table for each identified)	Speciation	Sensitivity	Results (record for each item tested)		
	Not Identified	Not Tested			
Drop down box:	Coliform  (please indicate species of coliform (e.g. E.coli, Klebsiella etc if available))	Menu of antibiotics tested  (drop down list)	S	I	R
	Proteus**	Menu of antibiotics tested  (drop down list)	S	I	R
	Enterococcus	Menu of antibiotics tested  (drop down list)	S	I	R
	Other organisms:  (please indicate)  Staph saprophiticus  Pseudomonas  Staph aureus  Candida	Not applicable			

	(additional drop down box for alternatives organisms)	
--	---	--

(\*\* Proteus is classified as a coliform, but recorded separately.)

Please record any other information here	Free text box
--	---------------

### **RESEARCH LABORATORY MICROBIOLOGY INFORMATION.**

#### **Data entered onto DUTY online database**

##### *i) Patient Details*

PATIENT UNIQUE ID	
PATIENT D.O.B	dd/mm/yyyy
REGIONAL ID	Regional/study site unique code*
LAB ID	Unique identifier
Date Sample Received at Lab	dd/mm/yyyy
Time Sample Received at Lab	xx.xx (24hr format)
Date Sample Processed	dd/mm/yyyy
Time Sample Processed	xx.xx (24hr format)
Date Sample Reported	dd/mm/yyyy
Time Sample Reported	xx.xx (24hr format)
Able to Process Sample	Yes/No
If 'No' please tick:	Drop down box: Leaked sample

	Other (please specify) <i>free text box</i>
--	---

## ii) Urine Results

Weight Monovette (with urine)	xx.xx (g)
-------------------------------	-----------

**Microscopy:**

Microscopy Method	Manual	Automated	Not Measured
If Microscopy not performed please indicate why: <i>(free text box)</i>			
	Count (mm <sup>3</sup> )		
WBC	Absolute WBC count		
RBC	Absolute RBC count		
Squamous Epithelial Cells	Absolute SEC count		

**Culture:**

Culture Performed	Yes/No
If 'no' please state reason:	<i>free text box</i>

NO GROWTH	NO SIGNIFICANT GROWTH
-----------	-----------------------

Growth: (please indicate)	<u>Enter TOTAL count here (CFU/mL)</u>
------------------------------	--

Purity: (please indicate)	Pure/Predominant	Mixed growth (2 species)	Mixed growth (> 2 species)
Speciation: (please indicate)	For organism 1 (see below)	For organism 1 & 2 (see below)	Not Applicable
Sensitivity: (please indicate)	For organism 1 (see below)	For organism 1 & 2 (see below)	Not Applicable

iii) Speciation and Sensitivity (for each organism identified from above)

Replicate tables generated for each Organism identified

Organism 1	Colony Count (CFU/mL) (enter absolute count)				
Organism 1 (complete table for each identified)	Speciation	Sensitivity (or indicate as 'not tested')	Results (record for each item tested)		
Drop down box:	<i>E.coli</i>	Menu of antibiotics tested (drop down list)	S	I	R
	<i>Enterococci</i>	Menu of antibiotics tested	S	I	R
	<i>Klebsiella-Enterobacter-Serriata</i> group	Menu of antibiotics tested	S	I	R
	<i>Proteus-Morganella-Providencia</i> group	Menu of antibiotics tested	S	I	R
	<i>Pseudomonas spp</i>	Menu of antibiotics tested	S	I	R
	<i>Staphylococci</i>	Menu of antibiotics tested	S	I	R
	Other organisms: (please indicate)  (additional drop down box for alternatives organisms)  <i>free text box for other info</i>	Not applicable			
Please enter any other Information here		Free Text			

Presence of Antimicrobial Substances	Yes/No
--------------------------------------	--------

## Appendix 8 Laboratory methods

**TABLE 86** Typical results recorded for a calibrated loop method

Number of CFU counted using inoculum of	0.3 µl	1 µl	2 µl	5 µl	10 µl
10 <sup>6</sup> CFU/l or 10 <sup>3</sup> CFU/ml	–	–	–	5	10
10 <sup>7</sup> CFU/l or 10 <sup>5</sup> CFU/ml	3	10	20	50	100
10 <sup>8</sup> CFU/l or 10 <sup>6</sup> CFU/ml	30	100	200	500	1000

**TABLE 87** Dilution of urine

Urine opacity	Dilutions made	Dilutions spiral plated
Clear	1:10 <sup>3</sup> (10 µl into 10 ml of water)	Neat urine + 10 <sup>3</sup>
Cloudy	1:10 <sup>3</sup> + 1:10 <sup>6</sup> (10 <sup>3</sup> : 10 µl into 10 ml of water 10 <sup>6</sup> : 10 µl of 10 <sup>3</sup> into 10 ml water)	10 <sup>3</sup> + 10 <sup>6</sup>
Cloudy? (If unsure)	1:10 <sup>3</sup> + 1:10 <sup>6</sup> (10 <sup>3</sup> : 10 µl into 10 ml of water 10 <sup>6</sup> : 10 µl of 10 <sup>3</sup> into 10 ml of water)	Neat urine + 10 <sup>3</sup> + 10 <sup>6</sup>

**TABLE 88** Summary of research laboratory methods

Microscopy:	<b>Automated</b>	
Antimicrobial substance assay:	<b>Phenotypic</b>	
Culture method:	<b>Precise colony counts from spiral plater</b>	
Culture media:	<b>CBA</b>	Chromogenic media
Total count from:	*	
Species specific count from:		*
Culture volume:	<b>50 µl</b>	
Culture plate:	<b>Whole</b>	

TABLE 89 Identification of isolates at the research laboratory

Chromogenic agar	Presumptive ID	Further tests	Confirmatory tests
Large PINK/RED colonies	<i>E. coli</i>	Indole (+)	Phoenix™
Large DARK BLUE/PURPLE colonies	KES ( <i>Klebsiella</i> spp/ <i>Enterobacter</i> spp/ <i>Serratia</i> spp)	N/A	Phoenix™ to differentiate b/w species
Large BROWN/GREEN colonies	<i>Pseudomonas</i> spp	Oxidase (+)	Phoenix™
Large cols with BROWN HALO colonies	<i>Proteus</i> spp/ <i>Providencia</i> spp/ <i>Morganella</i> spp	Indole (+) Oxidase (-) Swarming with <i>Proteus</i>	Phoenix™ to differentiate b/w species
Large opaque colonies	Coliform	Indole (+/-) Oxidase (+/-)	Phoenix™
Small BLUE/GREEN colonies	Enterococci	Catalase (-) Group (D)	Phoenix™
Medium WHITE colonies	Staph/yeast (+ <i>Cryptococcus neoformans</i> )	Check if yeast seen in microscopy and growth on Sabouraud Gram (record result) Catalase (record result) Staphaurex (record result)	Phoenix™ if GPC Auxacolor kit for yeast (Bio-Rad Laboratories Ltd, Hertfordshire, UK)
Small PINK colonies	<i>S. saprophyticus</i>	Staphaurex (+/-) Catalase (+) Novobiocin (R)	Phoenix™
Small opaque colonies	?	Gram (if GPC coagulase and staphaurex)	

## Research laboratory quality control

### Spiral plater

Daily checks on the instrument were carried out according to manufacturers' guidelines. Internal quality assurance was carried out weekly and monthly.

### Bacteriology reagents

Control organisms (with known reliable target results) were tested alongside each ID and susceptibility test performed. If incorrect test results were obtained for any control organism the test was repeated. Control organism test results were recorded in the quality assurance folder.

### Disc susceptibility testing

Control strains *S. aureus* NCTC 6571, *E. faecalis* ATCC 29212, *E. coli* ATCC 25922 and *P. aeruginosa* ATCC 27853 were used on every occasion to quality check the media and disc susceptibility testing conditions. Results were recorded in the DUTY disc-media spreadsheet.

### ***Urine analyser***

Control runs were performed at the start, middle and end of each day to assure quality of testing. Calibration was performed monthly. Results were recorded automatically in the instrument.

### ***Incubator/freezers/fridges***

All temperature-controlled equipment is monitored by an automated system (Comark Instruments Ltd, Norwich, UK). Temperatures are recorded at 15-minute intervals and anomalies e-mailed to managers.

### ***Internal DUTY quality assurance***

An internal quality assurance system was set up for the DUTY study involving randomised urine samples. After initial processing, urines were randomly chosen to be processed again. Urines would be processed as per standard DUTY procedures and the microscopy and culture results compared with original results. Any discrepancies were investigated by further testing.

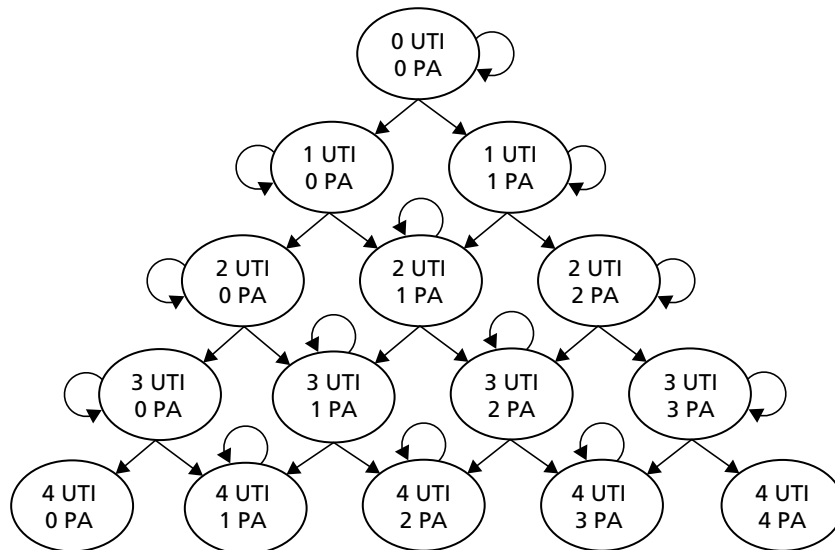
### ***Vertical audit of DUTY data entry***

Each month a processed sample is chosen for vertical audit to assure quality in data entry to the web database. All data from worksheets are compared with the web database entry and any discrepancies recorded.





## Appendix 9 Health economic analysis and modelling of diagnostic strategies



**FIGURE 37** Medium-term model. Figure represents the movement of the patient within the medium-term model for one of the four states corresponding to the VUR status and UTI history of the patient. Patients may move between the untreated VUR and treated VUR if they receive a correct VUR diagnosis. Patients may move from the 'no history of UTI' state to the 'history of UTI state' if they suffer a UTI infection. PA, pyelonephritic attack.

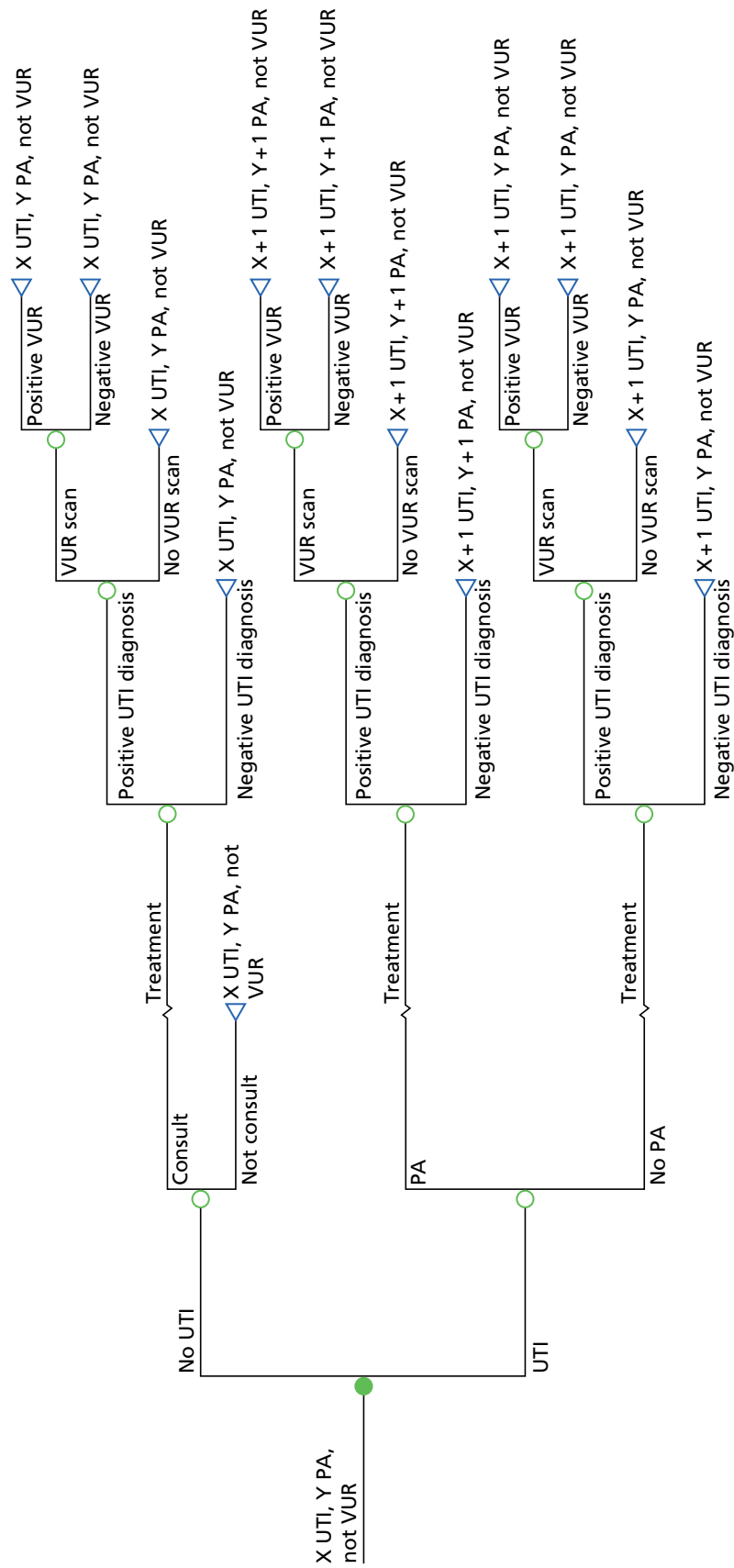


FIGURE 38 Medium-term model, no VUR. PA, pyelonephritic attack.

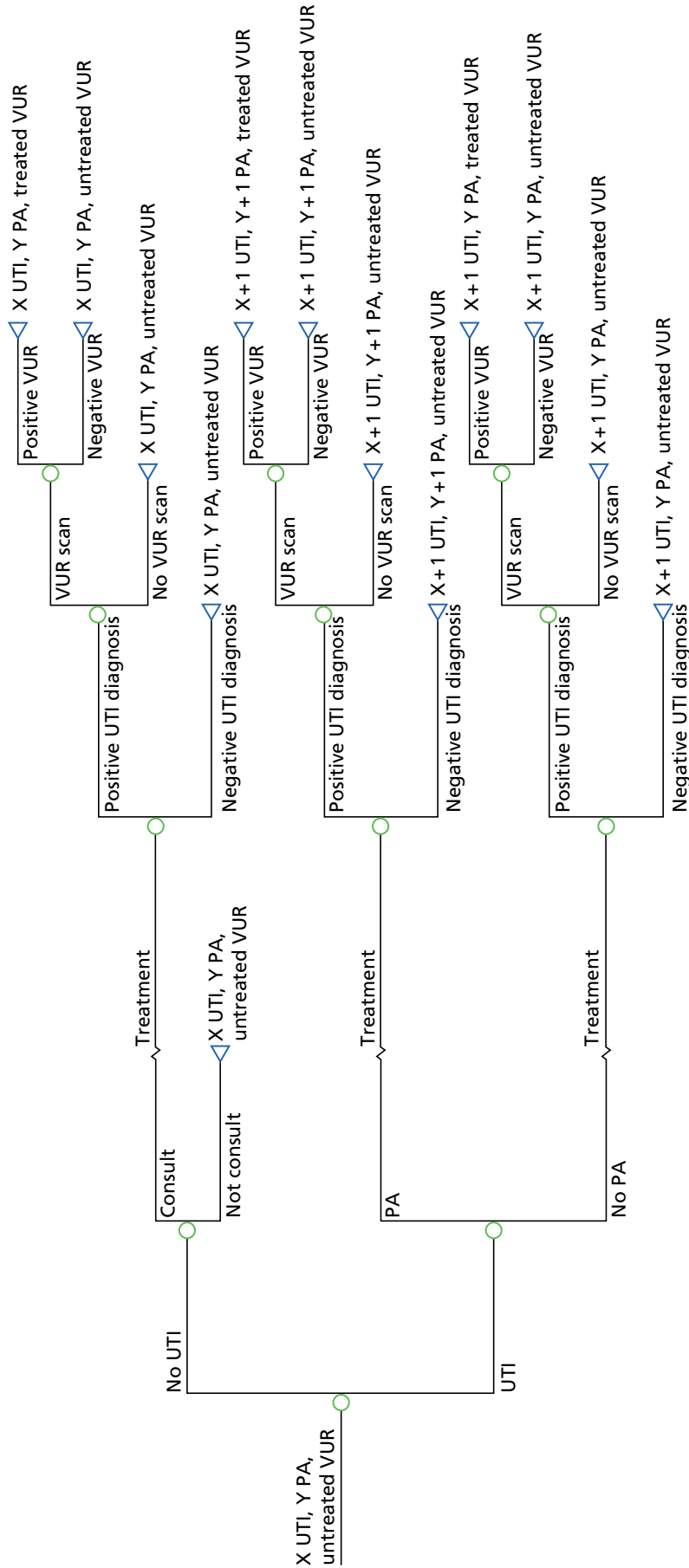


FIGURE 39 Medium-term model, untreated VUR. PA, pyelonephritic attack.

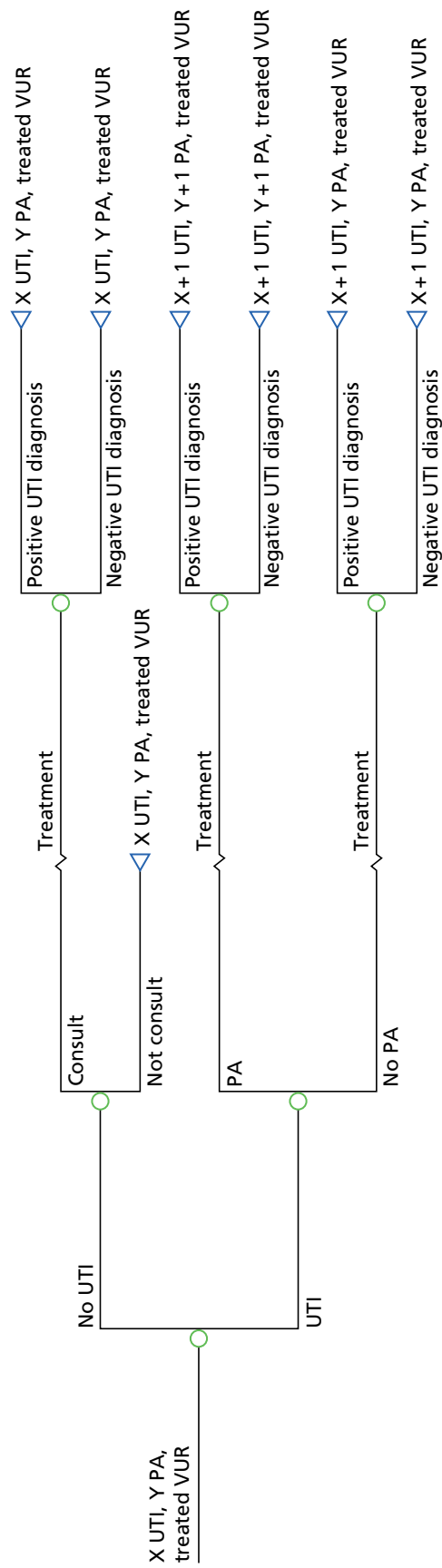
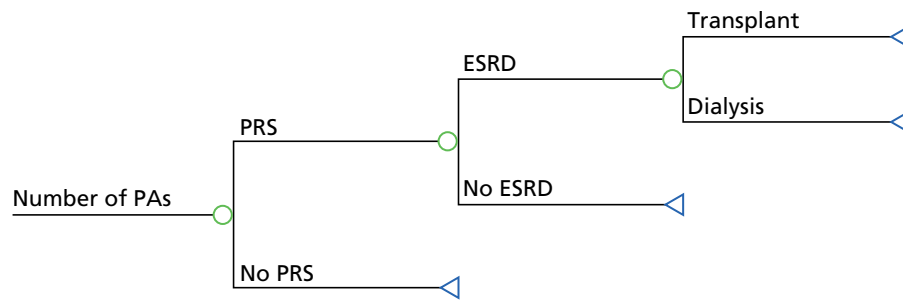
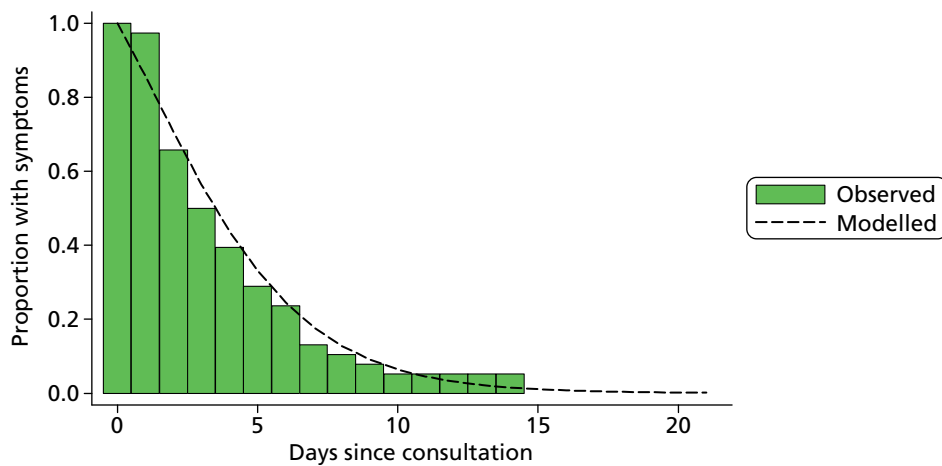


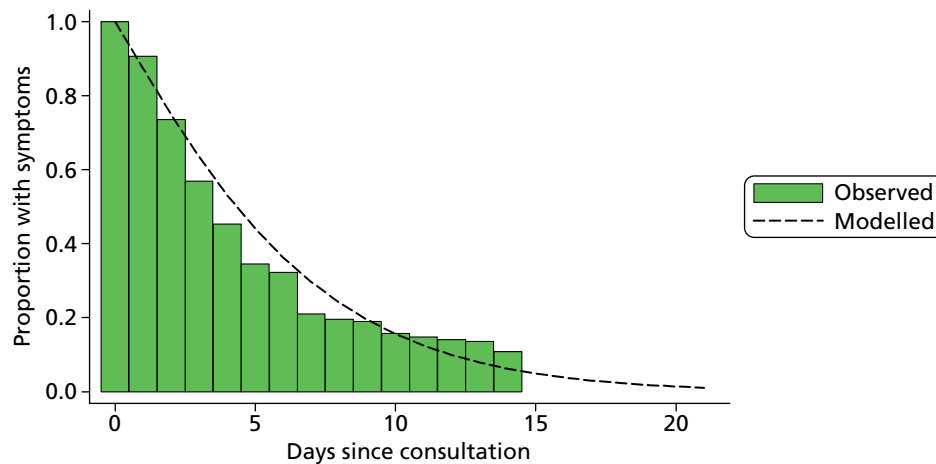
FIGURE 40 Medium-term model, treated VUR. PA, pyelonephritic attack.



**FIGURE 41** Long-term model. PA, pyelonephritic attack.



**FIGURE 42** Comparison of observed and modelled (using a Weibull model: shape = 1.497, scale = 0.091) symptom resolution rates for patients with UTI treated immediately with antibiotics ( $n = 38$ ).



**FIGURE 43** Comparison of observed and modelled (using a Weibull model: shape = 1.246, scale = 0.099) symptom resolution rates for patients without UTI ( $n = 733$ ).

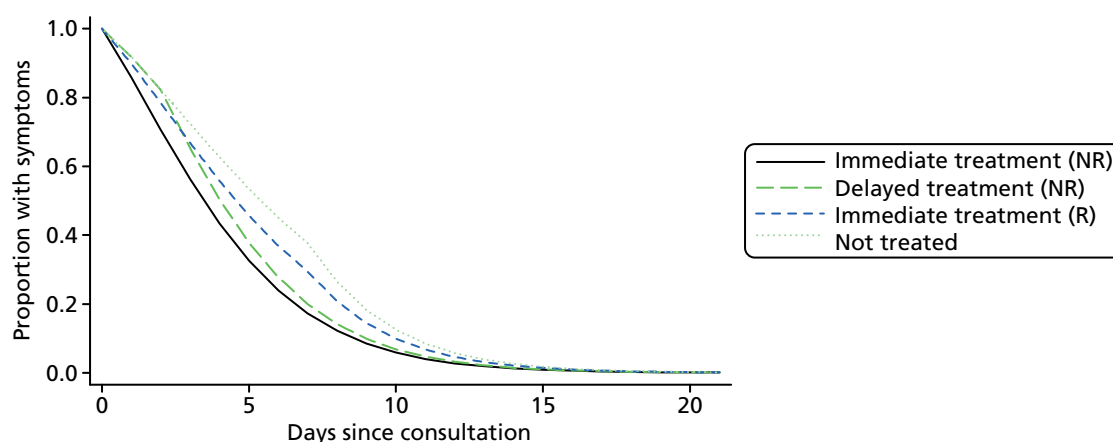


FIGURE 44 Symptom resolution in a selection of treatment states.

## Time and motion methods

We observed two RNs collecting urine samples on six DUTY participants at four sites between February 2012 and April 2012. We collected data using a Samsung mobile device with Timer Pro software (Applied Computer Services Inc., Engelwood, CO, USA), which allowed a range of pre-programmed tasks to be selected and the time taken for each task to be recorded. Timing began when first contact was made with the patient and ended when the urine sample was left for collection by the laboratory. As the study was nested within the DUTY study many tasks, for example gaining informed consent, were undertaken by the nurses, which would not form part of routine NHS care. We excluded these protocol driven tasks from our cost analysis. We defined a usual care pathway for both clean catch and nappy pad collection (*Table 90*). When activities which would normally be undertaken in usual practice were not performed by the nurses (e.g. when completion of laboratory request was undertaken by the GP due to electronic laboratory requesting in the general practice) we imputed the time using the mean other children where this activity was undertaken. We included the cost of time taken for a nurse to give instructions on urine sampling to the parent when a sample was attempted but not successful. We calculated costs by multiplying the number of minutes taken by the cost per minute for a practice nurse (£0.65).<sup>158</sup> The average time taken for sampling was 9.08 minutes (cost: £6.78) for samples with no dipstick test and 12.04 minutes (cost: £7.81) when a dipstick was required.

TABLE 90 Time for activities (minutes)

Activity	Patient number						Average time	Cost, £
	1	2	3	4	5	6		
Nurse gives instructions on urine sampling	NO	NO	NO	0.83	2.23	3.47	2.18	1.41
Nurse administers dipstick test <sup>a</sup>	1.5	0.75	1.53	2.07	1.65	2.00	1.58	1.03
Nurse draws sample	1.00	0.50	0.83	0.88	0.80	1.18	0.87	0.56
Nurse fills in laboratory request form	4.00	4.75	3.90	NO	3.35	NO	4.00	2.60
Nurse prepares laboratory sample	1.67	2.00	1.42	1.60	2.13	3.35	2.03	1.32
Nurse leaves dipstick results for GP <sup>a</sup>	1.17	1.50	1.53	NO	1.91	0.78	1.38	0.89
<b>Total sample obtained without dipstick</b>							<b>9.08</b>	<b>6.78</b>
<b>Total sample obtained with dipstick</b>							<b>12.04</b>	<b>7.81</b>
<b>Total sample not obtained</b>							<b>2.18</b>	<b>1.41</b>

NO, not observed.

<sup>a</sup> Cost only included for diagnostic strategies that include dipstick testing.

## Acute illness costs

Data were collected on the number of clinical contacts, prescription medications, hospital tests, outpatient appointments, inpatient stays and other personal costs in the 14 days following the index consultation.

When a record of a contact was made, but no detailed information was given allowing it to be costed a clean catch separately, the average cost of all contacts within that resource group (e.g. hospital test, outpatient appointment) was used. We mapped drug names, forms and dose to *British National Formulary* (BNF)<sup>185</sup> codes using the prescription cost analysis<sup>160</sup> data set for England in 2011, which is in turn based on BNF data from 2010. When dose was not provided, we mapped to the most commonly prescribed dose for the given drug and formulation. If drug formulation was not provided, we mapped to the average drug chemical cost where this was available, or used the most commonly prescribed formulation within the drug type. We calculated the total number of prescriptions for each patient and used the average cost of a prescription for the formulation or chemical to calculate the total cost of prescribed medication. In a small number of cases, non-specific drug groups were reported (e.g. antibiotic, cough syrup); when this occurred we used the code of the most commonly prescribed formulation within this drug group. When no drug name was given (e.g. unsure) or the drug name could not be matched to any BNF formulation or chemical, we used the mean cost of all prescriptions during the 14 days' follow-up.

Our primary analysis includes only costs borne by the NHS; however, we also collected some personal cost data to enable a secondary analysis taking a societal perspective and cost related to travel, parking, over-the-counter medications and time off work. Personal costs were reported directly; however, for non-prescribed medicines we supplemented this information using similar methods to those applied for prescribed medicines when costs were omitted by parents. We used the Automobile Association's estimate of the average cost per mile<sup>186</sup> to convert miles of travel to a cost. Costs for a pyelonephritic attack were assumed to be 50% higher than a lower UTI.

We also used the costs derived from the 14-day interviews to estimate short-term costs of a UTI infection (*Table 91*). We removed antibiotic and VUR scan costs to avoid double-counting as these are included separately in the model. We calculated the mean cost for infections lasting < 2 days and assumed this to be the fixed cost of a UTI infection, and then we calculated the excess daily cost of infection by dividing the average cost of an infection lasting > 2 days, minus the fixed cost, by the average number of days an infection lasted over 2 days (*Table 92*). A small number of infections had not resolved by the end of the follow-up; when this occurred, we assumed that symptoms had resolved at 21 days.

## Estimation of utility values

In the absence of preference-based measures of health-related quality of life (e.g. EQ-5D) that have been developed and validated in infants, we used TAPQOL questionnaires,<sup>110</sup> completed by parents at 14 days, to evaluate the impact on quality of life. The TAPQOL questionnaire was specifically designed for use in infants and comprises 43 items across eight domains, including sleeping, lung function and anxiety, which aim to describe the health state of the patient. No mapping function<sup>187</sup> has been developed for the TAPQOL; therefore, we cannot use it to directly derive utility scores needed to calculate QALYs. However, it does allow us to compare the health states of children with and without UTI to determine whether or not quality of life is markedly different and, therefore, whether or not utility scores measured in other infant diseases (e.g. rotavirus or respiratory syncytial virus) might provide suitable substitutes. On the basis of these exploratory results we concluded that the impact of UTI on infant quality of life was similar to other diseases commonly presenting in primary care (*Table 93*).

TABLE 91 Fourteen-day resource use

Resource	Not UTI			UTI			Mean cost difference, £ (95% CI) <sup>b</sup>	p-value		
	N <sup>a</sup>	Mean	Unit cost, £	Average cost, £	N	Mean, £			Unit cost, £	Average cost, £
GP primary care face to face	738	0.366	39.72	14.53	58	0.431	39.12	16.86	2.33 (-6.05 to 10.71)	0.586
GP primary care telephone	738	0.157	27.07	4.25	58	0.362	24.67	8.93	4.68 (0.08 to 9.27)	0.046
Nurse primary care face to face	738	0.098	17.43	1.70	58	0.138	15.53	2.14	0.44 (-1.94 to 2.82)	0.717
Nurse primary care telephone	738	0.042	14.66	0.62	58	0.034	8.00	0.28	-0.34 (-0.79 to 0.11)	0.137
Community	738	0.136	23.62	3.20	58	0.086	25.45	2.19	-1.01 (-3.42 to 1.41)	0.415
A&E	740	0.058	103.88	6.04	58	0.069	107.00	7.38	1.34 (-8.25 to 10.94)	0.784
Inpatient days	740	0.032	425.00	13.78	58	0.103	425.00	43.97	30.18 (-25.04 to 85.41)	0.284
Outpatient	740	0.036	105.00	3.83	58	0.069	105.00	7.24	3.41 (-5.16 to 11.98)	0.435
Ambulance	738	0.008	226.06	1.84	58	0.000			-1.84 (-3.60 to -0.08)	0.041
Hospital tests	740	0.023	31.59	0.73	58	0.034	38.40	1.32	0.60 (-1.47 to 2.67)	0.571
Prescription medicine	740	0.824	4.46	3.68	58	1.207	2.52	3.05	-0.63 (-1.93 to 0.67)	0.344
Non-prescription medicine	740	0.936	3.39	3.17	58	0.776	3.24	2.51	-0.66 (-1.41 to 0.10)	0.088
Travel	740	0.331	4.66	1.54	58	0.397	9.64	3.82	2.28 (-1.75 to 6.31)	0.267
Lost earnings	735	0.118	105.84	12.53	58	0.172	138.74	23.92	11.38 (-10.59 to 33.34)	0.310
Other personal costs	735	0.076	70.23	5.35	58	0.086	38.60	3.33	-2.02 (-6.00 to 1.96)	0.319
<b>Total</b>				<b>76.79</b>				<b>126.95</b>		

a Different numbers of observations due to missing data.

b Bootstrapped CIs.



**TABLE 92** Fixed and variable costs of infection excluding antibiotic treatment or imaging scans<sup>a</sup>

Outcome	Not UTI	UTI
Average infection fixed costs (less than 2 days)	19.58	20.46
Average cost of infection given it lasts more than 2 days	58.71	112.22
Average variable cost of infection	39.14	91.77
Average length of infection	7.74	8.60
Average excess length of infection given it lasts for more than 2 days	5.74	6.60
<b>Average cost per excess day</b>	<b>6.82</b>	<b>13.91</b>

a Antibiotic treatment and imaging scan costs are included in the model directly.

**TABLE 93** Comparison of mean TAPQOL domain scores across GP diagnosis<sup>a</sup>

Domain	UTI (n = 58)	URTI (n = 229)	Viral illness (n = 109)	Otitis media (n = 66)	Chest infection (n = 46)	Tonsillitis (n = 29)	Gastroenteritis (n = 26)	Other (n = 112)
Sleeping	0.598	0.569	0.592	0.538	0.511	0.504	0.603	0.611
Appetite	0.603	0.654	0.667	0.667	0.607	0.526	0.564	0.683
Lungs	0.954	0.912	0.939	0.962	0.79	0.848	0.978	0.921
Stomach	0.803	0.870	0.818	0.889	0.855	0.802	0.638	0.804
Skin	0.909	0.875	0.896	0.876	0.909	0.888	0.933	0.878
Motor	0.948	0.939	0.952	0.950	0.924	0.881	0.910	0.942
Social	0.841	0.799	0.817	0.795	0.819	0.795	0.796	0.795
Problem	0.741	0.689	0.706	0.697	0.685	0.653	0.676	0.698
Communication	0.949	0.944	0.971	0.909	0.952	0.847	0.958	0.933
Anxiety	0.845	0.877	0.867	0.889	0.830	0.902	0.859	0.823
Positive	0.739	0.750	0.752	0.750	0.685	0.707	0.699	0.733
Liveliness	0.615	0.702	0.734	0.727	0.645	0.701	0.590	0.687

a Scores are given based on the working diagnosis of the consulting GP, except for UTI, where only those with laboratory-confirmed UTI are included.



## Appendix 10 STROBE checklist

**TABLE 94** STROBE statement: checklist of items that should be included in reports of observational studies

	Item number	Recommendation	Page of DUTY report
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	Title page
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	xviii
Introduction			
Background/ rationale	2	Explain the scientific background and rationale for the investigation being reported	1
Objectives	3	State specific objectives, including any prespecified hypotheses	15
Methods			
Study design	4	Present key elements of study design early in the paper	18
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up and data collection	20, 34
Participants	6	(a) <i>Cohort study</i> : give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	25, 35
		<i>Case-control study</i> : give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	
		<i>Cross-sectional study</i> : give the eligibility criteria, and the sources and methods of selection of participants	
		(b) <i>Cohort study: for matched studies, give matching criteria and number of exposed and unexposed</i>	N/A
		<i>Case-control study: for matched studies, give matching criteria and the number of controls per case</i>	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	33, 103
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	33
Bias	9	Describe any efforts to address potential sources of bias	51
Study size	10	Explain how the study size was arrived at	50
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	N/A

continued

**TABLE 94** STROBE statement: checklist of items that should be included in reports of observational studies (*continued*)

	Item number	Recommendation	Page of DUTY report
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	48, 52, 63, 85, 106, 125
		(b) Describe any methods used to examine subgroups and interactions	N/A
		(c) Explain how missing data were addressed	48, 52, 63, 85, 105, 135
		(d) <i>Cohort study</i> : if applicable, explain how loss to follow-up was addressed	38
		<i>Case-control study</i> : if applicable, explain how matching of cases and controls was addressed	
		<i>Cross-sectional study</i> : if applicable, describe analytical methods taking account of sampling strategy	
	(e) Describe any sensitivity analyses	176, 192	
N/A, not applicable.			

# Appendix 11 STARD checklist

**TABLE 95** STARD checklist for reporting of studies of diagnostic accuracy (version January 2003)

Section and topic	Item #		On page #
Title/abstract/ keywords	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity')	Title page xvii
Introduction	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups	15
Methods			
<i>Participants</i>	3	The study population: the inclusion and exclusion criteria, setting and locations where data were collected	21, 25, 26
	4	Participant recruitment: was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?	26
	5	Participant sampling: was the study population a consecutive series of participants defined by the selection criteria in items 3 and 4? If not, specify how participants were further selected	26, 104
	6	Data collection: was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	16, 33
<i>Test methods</i>	7	The reference standard and its rationale	xvii, xxii, 17, 52, 80, 103, 105, 146, 237
	8	Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard	40, Appendix 8
	9	Definition of and rationale for the units, cut-offs and/or categories of the results of the index tests and the reference standard	34, 103
	10	The number, training and expertise of the persons executing and reading the index tests and the reference standard	41, 55, 105
	11	Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers	52, 80, 105, 145, 146
<i>Statistical methods</i>	12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals)	9, 59, 189, 191, 195, 205, 215
	13	Methods for calculating test reproducibility, if done	102
Results			
<i>Participants</i>	14	When study was performed, including beginning and end dates of recruitment	20
	15	Clinical and demographic characteristics of the study population (at least information on age, gender, spectrum of presenting symptoms)	59, 68, 69
	16	The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to undergo either test (a flow diagram is strongly recommended)	57, 59

continued

**TABLE 95** STARD checklist for reporting of studies of diagnostic accuracy (version January 2003) (*continued*)

Section and topic	Item #		On page #
<i>Test results</i>	17	Time-interval between the index tests and the reference standard, and any treatment administered in between	86, 88, 98, 99, 100, 113, 133, 205, 2113
	18	Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition	219
	19	A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard	109
	20	Any adverse events from performing the index tests or the reference standard	38, 110
<i>Estimates</i>	21	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals)	9, 51, 58, 59, 107, 124, 184, 189, 191, 195, 207, 208, 209
	22	How indeterminate results, missing data and outliers of the index tests were handled	48, 52, 63, 85, 105, 135
	23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centres, if done	112, 117, 118, 120, 122, 123, 124, 126, 134, 135, 137, 138, 140
	24	Estimates of test reproducibility, if done	N/A
Discussion	25	Discuss the clinical applicability of the study findings	242



A decorative graphic consisting of numerous thin, parallel green lines that curve from the left side of the page towards the right, creating a sense of movement and depth.

**EME  
HS&DR  
HTA  
PGfAR  
PHR**

Part of the NIHR Journals Library  
[www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)

*This report presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health*

***Published by the NIHR Journals Library***