NIHR | National Institute for Health and Care Research

Journals Library

Check for updates

Health Technology Assessment

Volume 28 • Issue 77 • November 2024

ISSN 2046-4924

Point-of-care tests for urinary tract infections to reduce antimicrobial resistance: a systematic review and conceptual economic model

Eve Tomlinson, Mary Ward, Chris Cooper, Rachel James, Christina Stokes, Samina Begum, Jessica Watson, Alastair D Hay, Hayley E Jones, Howard Thom and Penny Whiting

DOI 10.3310/PTMV8524

Point-of-care tests for urinary tract infections to reduce antimicrobial resistance: a systematic review and conceptual economic model

Eve Tomlinson,[1] Mary Ward,[1] Chris Cooper,[1] Rachel James,[1] Christina Stokes,[2] Samina Begum,[2] Jessica Watson,[3] Alastair D Hay,[3] Hayley E Jones,[1] Howard Thom,[1] and Penny Whiting[1*]

[1]Bristol TAG, Population Health Sciences, Bristol Medical School, Bristol, UK
[2]Patient representative, UK
[3]Centre for Academic Primary Care, Population Health Sciences, Bristol Medical School, Bristol, UK

[*]Corresponding author

Published November 2024
DOI: 10.3310/PTMV8524

This report should be referenced as follows:

Tomlinson E, Ward M, Cooper C, James R, Stokes C, Begum S, *et al.* Point-of-care tests for urinary tract infections to reduce antimicrobial resistance: a systematic review and conceptual economic model. *Health Technol Assess* 2024;**28**(77). https://doi.org/10.3310/PTMV8524

Health Technology Assessment

NICE TAR and DAR

ISSN 2046-4924 (Online)

Impact factor: 3.6

A list of Journals Library editors can be found on the NIHR Journals Library website

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 3.6 and is ranked 32nd (out of 105 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2022 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), EMBASE (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded™ (Clarivate™, Philadelphia, PA, USA).

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

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta.

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

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

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.

This article

The research reported in this issue of the journal was commissioned and funded by the Evidence Synthesis Programme on behalf of NICE as award number NIHR135710. The protocol was agreed in December 2022. The draft manuscript began editorial review in February 2023 and was accepted for publication in May 2024. 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' manuscript 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 article.

This article presents independent research funded by the National Institute for Health and Care 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, the HTA programme or the Department of Health and Social Care. 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, the HTA programme or the Department of Health and Social Care.

This article was published based on current knowledge at the time and date of publication. NIHR is committed to being inclusive and will continually monitor best practice and guidance in relation to terminology and language to ensure that we remain relevant to our stakeholders.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Newgen Digitalworks Pvt Ltd, Chennai, India (www.newgen.co).

Abstract

Point-of-care tests for urinary tract infections to reduce antimicrobial resistance: a systematic review and conceptual economic model

Eve Tomlinson,[1] Mary Ward,[1] Chris Cooper,[1] Rachel James,[1] Christina Stokes,[2] Samina Begum,[2] Jessica Watson,[3] Alastair D Hay,[3] Hayley E Jones,[1] Howard Thom,[1] and Penny Whiting[1*]

[1]Bristol TAG, Population Health Sciences, Bristol Medical School, Bristol, UK
[2]Patient representative, UK
[3]Centre for Academic Primary Care, Population Health Sciences, Bristol Medical School, Bristol, UK

[*]Corresponding author penny.whiting@bristol.ac.uk

Background: Urinary tract infections are diagnosed by general practitioners based on symptoms, dipstick tests in some and laboratory urine culture. Patients may be given inappropriate antibiotics. Point-of-care tests can diagnose urinary tract infection in near-patient settings quicker than standard culture. Some can identify the causative pathogen or antimicrobial sensitivity.

Objective: To assess whether point-of-care tests for people with suspected urinary tract infection have the potential to be clinically effective and cost-effective to the NHS.

Design: Systematic review and conceptual economic model.

Results: Two randomised controlled trials evaluated Flexicult Human (one against standard care; one against ID Flexicult). One trial found no evidence of a difference between groups in concordant antibiotic use (odds ratio 0.84, 95% confidence interval 0.58 to 1.20), and the other found no difference in appropriate antibiotic prescribing (odds ratio 1.44, 95% confidence interval 1.03 to 1.99). Compared with standard care, Flexicult was associated with reduced antibiotic prescribing at initial consultation (odds ratio 0.56, 95% confidence interval 0.35 to 0.88). No difference was found for other outcomes. Sixteen studies reported test accuracy data. Most were rated as being at unclear or high risk of bias. We identified data on three rapid tests (results < 40 minutes). Lodestar DX (n = 1) had good sensitivity (86%, 95% confidence interval 74% to 99%) and specificity (88%, 95% confidence interval 83% to 94%) for detecting *Escherichia coli*. Uriscreen (n = 4) had modest summary sensitivity (74%, 95% confidence interval 59% to 84%) and specificity (64%, 95% confidence interval 41% to 82%). UTRiPLEX (n = 1) had poor sensitivity (21%) and good specificity (94%). Twelve studies evaluated culture-based tests (results 24 hours). Laboratory-based studies found Dipstreak (n = 2) and Uricult (n = 1) to be highly accurate, but there were limitations with these studies. Uricult Trio (n = 3) had more modest summary sensitivity (73%, 95% confidence interval 63% to 82%) and specificity (70%, 95% confidence interval 52% to 84%). Summary sensitivity for Flexicult Human (n = 4) and ID Flexicult (n = 2) was 79% (95% confidence interval 72% to 85%) and 89% (95% confidence interval 84% to 93%). Summary specificity was 67% (95% confidence interval 30% to 90%) and 70% (95% confidence interval 52% to 84%). Caution is needed in interpreting findings because of heterogeneity and limited data. Five studies evaluated technical performance (Flexicult Human, n = 3; Uricult Trio, n = 2). Limited data suggested that they are easier to use and interpret than standard culture. A conceptual economic model estimated the cost-effectiveness of point-of-care tests for urinary tract infection diagnosis, pathogen identification

and antimicrobial sensitivity testing. Sensitivity and specificity of tests were informed by the clinical effectiveness review. Studies identified by the review were screened for evidence on treatment efficacy, costs and utility data; only two studies provided relevant evidence. A pragmatic search identified eight cost-effectiveness studies that provided further evidence. A decision tree comparing point-of-care tests in a mixed population (Lodestar DX vs. Flexicult Human) and in women with uncomplicated urinary tract infection (Lodestar DX vs. Flexicult Human vs. ID Flexicult) was implemented. The available input data were too limited for the results to be meaningful.

Conclusion and future work: More research is required to determine whether point-of-care tests for urinary tract infection have the potential to be clinically effective and cost-effective to the NHS. Rapid tests such as Astrego PA-100 system and Lodestar DX appear promising, but data are very limited.

Study registration: This study is registered as PROSPERO CRD42022383889.

Funding: This award was funded by the National Institute for Health and Care Research (NIHR) Evidence Synthesis programme (NIHR award ref: NIHR135710) and is published in full in *Health Technology Assessment*; Vol. 28, No. 77. See the NIHR Funding and Awards website for further award information.

Contents

List of tables

List of figures

List of abbreviations

AE	adverse event
AST	antimicrobial sensitivity testing
CE	Conformité Européenne
CENTRAL	Cochrane Central Register of Controlled Trials
CFU	colony-forming unit
CI	confidence interval
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CRD	Centre for Reviews and Dissemination
DTA	diagnostic test accuracy
ESPAUR	English Surveillance Programme for Antimicrobial Utilisation and Resistance
EVA	early value assessment
GP	general practitioner
ICER	incremental cost-effectiveness ratio
ICTRP	International Clinical Trials Registry Platform
LE	leukocyte esterase
NICE	National Institute for Health and Care Excellence
POCT	point-of-care test
PSS	Personal Social Services
QALY	quality-adjusted life-year
RCT	randomised controlled trial
UTI	urinary tract infection
WHO	World Health Organization

Note

This monograph is based on the Diagnostic Assessment Report produced for NICE. The full report contained a considerable number of data that were deemed confidential. The full report was used by the Diagnostic Advisory Committee at NICE in their deliberations. The full report with each piece of confidential data removed and replaced by the statement 'confidential information (or data) removed' is available on the NICE website: www.nice.org.uk.

The present monograph presents as full a version of the report as is possible while retaining readability, but some sections, sentences, tables and figures have been removed. Readers should bear in mind that the discussion, conclusions and implications for practice and research are based on all the data considered in the original full NICE report.

Plain language summary

What is the problem?

Urine infections are very common but can be difficult to diagnose. A GP will diagnose a urine infection based on symptoms, and sometimes they will send a urine sample to the lab. The GP will usually give antibiotics before knowing the lab test results (which can take up to a week). Some people will be given the wrong antibiotics, and some will be given antibiotics unnecessarily.

New 'rapid tests' can be done in the GP surgery or pharmacy and will quickly tell (some in just a few minutes) whether someone has a urine infection. Some tests can also tell which bug is causing the infection and which antibiotics will work best.

What did we do?

We wanted to know whether using 'rapid tests' to diagnose urine infections means that more people are correctly diagnosed, diagnosed more quickly, and treated with the right antibiotics more quickly. We also wanted to know whether these tests are a good use of NHS money. We reviewed existing research and developed an economic (cost) model.

What did we find?

There is very little information available on these 'rapid tests'. Tests were only looked at by a few studies each, and the people studied were different. Rapid tests that can detect a urine infection in under 40 minutes showed promise, but there were not enough data to know whether they are a good use of NHS money. More studies are needed to answer this question and to determine whether results vary across different populations.

Scientific summary

Background

Urinary tract infections (UTIs) are one of the most common causes of infection worldwide. The accurate and timely diagnosis of UTIs is crucial to ensure that appropriate treatment is started to help resolve symptoms, improve quality of life and reduce the risk of complications such as pyelonephritis, kidney failure and sepsis. In the ongoing public health challenge of antibiotic resistance, it is important that antibiotics are prescribed only when necessary and that they target the causative organism of the infection.

However, UTIs can be difficult to diagnose. Currently they are diagnosed by a general practitioner (GP) based on symptoms and laboratory-based urine culture. Dipstick tests can be used to help make a quicker diagnosis in some people, for example children or women aged < 65 years. Dipstick tests involve dipping a specially treated paper or plastic strip into a urine sample to identify the presence of leukocyte esterase, nitrites and blood. However, these tests are not very accurate at diagnosing UTI, and they do not provide any information on the pathogenic cause or on antibiotic resistance. The GP will often prescribe antibiotics before knowing the culture results, which can take up to a week to receive. Some people may therefore be given antibiotics unnecessarily, and some will be given the wrong antibiotics.

Novel point-of-care tests (POCTs) can be conducted in a near-patient setting and can quickly diagnose a UTI. Some can also tell which pathogen is causing the infection and which antibiotic will work best.

Objectives

This project aimed to determine whether POCTs for people with suspected UTI have the potential to be clinically effective and cost-effective to the NHS.

We defined the following objectives to address this overall aim.

- Objective 1: what is the impact on clinical outcomes of using POCTs to diagnose UTI, with or without additional pathogen identification and antimicrobial sensitivity testing (AST)?
- Objective 2: what is the accuracy of POCTs for UTI diagnosis, pathogen identification and AST?
- Objective 3: what is the technical performance (other than accuracy) of POCTs for UTI?
- Objective 4: what are the costs, from a UK NHS and Personal Social Services perspective, of using POCTs for UTI diagnosis, pathogen identification and AST?
- Objective 5: how might a conceptual model be specified in terms of structure and evidence required for parametrisation in order to estimate the cost-effectiveness of POCT for UTI diagnosis, pathogen identification and AST?

Methods

Clinical effectiveness review

A systematic review was conducted in line with published guidance.

Data sources

Four databases and two trial registries were searched. Additional non-bibliographic search methods included searching trial registries, screening reference lists of reviews and study reports, hand-searching relevant websites and reviewing information submitted by test manufacturers.

Study selection and review methods

Studies were eligible for inclusion if they were published during or after the year 2000, enrolled patients with suspected UTI, and evaluated a POCT in scope:

- rapid tests giving results < 40 minutes – Astrego PA-100 system, Lodestar DX, TriVerity, Uriscreen, UTRiPLEX
- culture-based tests giving results in up to 24 hours – Flexicult Human, ID Flexicult, Diaslide, Dipstreak, Chromostreak, Uricult, Uricult Trio, Uricult Plus.

For objective 1, studies had to be randomised controlled trials (RCTs) or non-randomised studies of interventions, set in primary care or the community and use standard care as the reference standard. For objective 2, only diagnostic test accuracy studies were eligible for inclusion. Studies of any design were eligible for objective 3. Studies had to report data on prespecified outcomes to be eligible:

- Objective 1 – any outcome related to antibiotic use/prescription, morbidity, mortality, UTI-associated healthcare resources, health-related quality of life.
- Objective 2 – test accuracy in detecting UTI, identifying pathogens or assessing susceptibility to antimicrobials.
- Objective 3 – test failure rate, ease of use/acceptability, time to results, health-related quality of life, any outcome related to antibiotic use/prescription, UTI-associated healthcare resources, test costs, clinical outcomes. Title and abstract screening was conducted by two reviewers independently. Inclusion assessment, data extraction and risk-of-bias assessment were performed by one reviewer and checked by a second reviewer. Risk of bias was assessed using the RoB 2 tool for RCTs, QUADAS-2 for diagnostic test accuracy studies, and QUADAS-C for comparative accuracy studies.

For each objective, we provided a narrative summary of included study details, risk of bias and results, stratified by POCT. For objective 2, bivariate random-effects meta-analyses were used to pool sensitivity and specificity across studies, separately for each POCT. We presented coupled forest plots of individual study and summary estimates of sensitivity and specificity together with 95% confidence intervals (CIs) to allow visual assessment of results and of heterogeneity across studies. There were not enough studies for formal investigation of heterogeneity, or to stratify analysis based on populations specified in the scope.

Conceptual economic model

We developed a conceptual model to estimate the cost-effectiveness of POCTs for UTI diagnosis, pathogen identification and AST. This represented important short- and long-term costs and quality-of-life impacts on the management of UTIs.

The conceptual model was implemented as a decision tree comparing POCTs with laboratory culture-based tests for UTI. Sensitivity and specificity were informed by the clinical effectiveness review. The decision tree was further informed by screening studies identified by the clinical effectiveness review for any evidence relating to cost-effectiveness or parameters that could inform the conceptual model. This was supplemented by pragmatic searches of Ovid MEDLINE, EMBASE and EconLit for cost-effectiveness studies in UTI. These were supplemented by evidence from National Institute for Health and Care Excellence guidelines, *British National Formulary* costs, and the Personal Social Services Research Unit.

We prioritised tests and populations where evidence was greatest. We also prioritised rapid over culture-based tests and tests that performed AST over those that only identified pathogenic cause and both such tests over those that tested only for UTI.

The decision tree model was implemented in the R statistical programming language (The R Foundation for Statistical Computing, Vienna, Austria).

Results

Clinical effectiveness review

We identified 16 studies for inclusion in the review. All studies were included for objective 2; two were also included for objective 1, while five also provided data for objective 3. Six studies evaluated rapid POCTs (Lodestar DX, *n* = 1; Uriscreen, *n* = 4; UTRiPLEX, *n* = 1) and 12 studies evaluated culture-based POCTs (Flexicult Human, *n* = 4; ID Flexicult, *n* = 2; Uricult Trio, *n* = 3; Uricult, *n* = 1; Dipstreak, *n* = 2). Two studies reported direct comparisons between tests (Flexicult Human and ID Flexicult; Uriscreen and UTRiPLEX). Studies enrolled women, pregnant women, children and people with catheters. There were no data on any other prespecified tests or populations of interest.

Objective 1: clinical outcomes

Two RCTs evaluated the clinical impact of the culture-based test Flexicult Human in women: one compared with standard care (*n* = 653) and the other compared with ID Flexicult (*n* = 376). Both trials were judged as being at low risk of bias. There was no evidence of a difference between intervention groups in the studies' primary outcomes: one evaluated concordant antibiotic use (odds ratio 0.84, 95% CI 0.58 to 1.20) and the other evaluated appropriate antibiotic prescribing (odds ratio 1.44, 95% CI 1.03 to 1.99). Compared with standard care, one study found that the use of Flexicult Human was associated with reduced antibiotic prescribing at initial consultation (odds ratio 0.56, 95% CI 0.35 to 0.88), but no difference was found between groups for other outcomes related to antibiotic use. Neither study reported a difference between intervention groups in duration of symptoms/infection, patient enablement or resource use. There were no data on mortality or health-related quality of life.

Objective 2: diagnostic test accuracy

Sixteen studies reported data on test accuracy. Two studies took place in Wales (*n* = 200 samples; *n* = 144 samples) and one had centres in Wales, England, Spain and the Netherlands (*n* = 289). The other studies were conducted in Israel (two studies; *n* = 795; *n* = 818), Hawaii (one study; *n* = 378), Venezuela (one study; *n* = 150), Mexico (one study; *n* = 108 samples), Philippines (one study; *n* = 200), South Africa (one study; *n* = 374), Republic of Korea (one study; *n* = 151), Argentina (one study; *n* = 2173), Denmark (three studies; *n* = 183 Flexicult Human/*n* = 158 ID Flexicult, *n* = 121 samples, *n* = 117) and Belgium (one study; *n* = 156 Uriscreen/*n* = 292 URiPLEX) (brackets show the number of participants or samples analysed). Twelve studies were conducted in primary or secondary care and four were laboratory-based. Five studies were judged at high risk of bias, eight at unclear risk of bias and three at low risk.

Only three rapid tests were evaluated (six studies). Lodestar DX appeared to be the most promising test. In a laboratory-based study, it had good sensitivity (86%, 95% CI 74% to 99%) and specificity (88%, 95% CI 83% to 94%) for detecting *E. coli*. Uriscreen had modest summary estimates of sensitivity (74%, 95% CI 59% to 84%; four studies) and specificity (64%, 95% CI 41% to 82%). UTRiPLEX had poor sensitivity (21%) but good specificity (94%) in one study recruiting children. Neither Uriscreen or UTRiPLEX provide information on antimicrobial sensitivity or pathogenic cause of infection.

Twelve studies evaluated culture-based tests. Of the culture-based tests evaluated, Dipstreak and Uricult were found to be highly accurate. However, these were assessed by two studies and one study, respectively, and both were conducted in the laboratory and were at high or unclear risk of bias. By contrast, studies of Uricult Trio (an extension of Uricult) in near-patient settings reported more modest

summary sensitivity (73%, 95% CI 63% to 82%) and specificity (70%, 95% CI 52% to 84%). Summary sensitivity for Flexicult Human (three studies) was 79% (95% CI 72% to 85%) and summary specificity was 67% (95% 30% to 90%). For ID Flexicult (two studies), this was 89% (95% CI 84% to 93%) and 70% (95% CI 52% to 84%). Three studies reported data on the accuracy of Flexicult Human in determining antimicrobial sensitivity. Summary sensitivity was 87% (95% CI 83% to 90%), and summary specificity was 93% (95% CI 89% to 95%).

All summary estimates should be interpreted with caution due to heterogeneity across studies.

Objective 3: technical performance

Five studies reported technical performance data. These evaluated culture-based tests only: three on Flexicult Human (n = 653; n = 35; n = 121) and two on Uricult Trio (n = 200; n = 374) Studies reported that POCTs are easier to use and interpret than laboratory tests and produce results more quickly. Clinicians reported that using Flexicult Human had increased their awareness of antibiotic prescribing and positively impacted their prescribing habits. However, they raised concerns regarding limits on when the test can be used, difficulties in result interpretation, limited resources, concerns about prolonging patient discomfort while awaiting test results, and the expense of maintaining a stock of tests. One study reported that Flexicult Human costs £48. (Confidential information has been removed). There were no data on test failure rate or health-related quality of life.

Conceptual economic model

We developed a conceptual model that could be used for a future full economic evaluation of POCTs for UTI and their role in reducing antibiotic resistance. This model identified pathways for benefit from POCTs, namely that they could reduce the use of empiric antibiotics and, by reducing the incidence of UTI complications and improving cure rates, reduce healthcare costs and quality-of-life impacts arising from UTI. Beyond test accuracy, we found only two studies from the clinical effectiveness review with relevant evidence for the economic model. Our pragmatic searches identified only eight cost-effectiveness studies in UTI, none of which modelled POCTs and none of which provided all the evidence needed to inform our economic evaluation. Due to the limited findings on test accuracy, we restricted modelling to a mixed population (Lodestar DX vs. Flexicult Human) and to women with uncomplicated UTI (Lodestar DX vs. Flexicult Human vs. ID Flexicult). Despite our prioritisation of tests and subgroups, broad approach to modelling, and pragmatic approach to searching for evidence, we found that evidence informing our economic model was too weak for results to be meaningful.

Conclusions

Implications for practice

There are few available data concerning the clinical effectiveness and cost-effectiveness of POCTs, particularly rapid POCTs, for people with suspected UTI, making it difficult to determine whether these tests have the potential to be clinically effective and cost-effective to the NHS. There is a clear need for a rapid test that would accurately diagnose a UTI within a short time in GP surgeries or pharmacy settings. Ideally, such tests would also provide information on antimicrobial sensitivity to allow targeted antibiotic use. The only test within scope that meets these criteria is the Astrego PA-100 system. However, there are currently no data available on this test.

Our conceptual model for economic evaluation found potential pathways to benefit from POCTs. They could reduce costs, improve quality of life, reduce antibiotic resistance and reduce complications from UTI. There were insufficient data on test accuracy, targeted versus empiric antibiotic efficacy, or costs and quality-of-life impacts of UTI complications for our model to perform a meaningful comparison.

Strong evidence that POCTs (1) reduce unnecessary antibiotic use, (2) improve symptoms or (3) are cost-effective is needed before such tests are introduced into the NHS.

Recommendations for research

Given the paucity of data on POCTs for diagnosing UTI, further studies are needed to determine whether POCTs for people with suspected UTI have the potential to be clinically effective and cost-effective to the NHS. Ideally, studies would be RCTs with embedded diagnostic test accuracy studies of POCTs and should be conducted in primary care; such studies would provide data on clinical impact and test accuracy. Studies should focus on tests with the greatest potential for clinical impact: the Astrego PA-100 system and Lodestar DX. Either the studies should enrol patients across multiple patient groups of interest (e.g. men, women, pregnant women, children) with results stratified according to patient subgroup, or separate studies should be carried out to determine whether results differ according to subgroups. Studies should also consider the feasibility of introducing rapid POCTs into pharmacy settings.

In addition to further studies on clinical effectiveness, further research on potential cost-effectiveness and impact on antibiotic resistance is needed. This research could build on our conceptual economic model using systematic literature reviews to identify evidence on the efficacy of empiric versus targeted antibiotic treatment of UTI; the efficacy in preventing UTI complications; and both the cost and quality-of-life impacts of these complications.

Study registration

This study is registered as PROSPERO CRD42022383889.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Evidence Synthesis programme (NIHR award ref: NIHR135710) and is published in full in *Health Technology Assessment*; Vol. 28, No. 77. See the NIHR Funding and Awards website for further award information.

Chapter 1 Background

Epidemiology and burden of urinary tract infections

Urinary tract infection (UTI) is one of the most common infections worldwide and is the most commonly seen bacterial infection in general practice.[1] UTI is also the most common hospital-acquired infection in the UK, accounting for almost one in four of all infections, most of which are associated with catheter use.[2] UTI can affect the lower urinary tract when the infection is in the urethra (urethritis) or bladder (cystitis), or the upper urinary tract when the infection is in the kidney (pyelonephritis). The incidence of UTI generally increases with age and is higher in women than in men; a 2019 study reported that around 83% of UTIs in primary care between 2011 and 2015 in England were in women.[3] Lifetime incidence of UTI in women is estimated at approximately 50–60%.[3] Risk factors for recurrent uncomplicated UTIs include frequent intercourse, vulvovaginal atrophy, change of the local bacterial flora, history of UTI, diabetes mellitus and a non-secretor blood type.[1,4]

There are several classifications of UTI, depending on the location and frequency of infection and whether the patient is symptomatic. Classifications of uncomplicated UTI are summarised in *Table 1*. A proportion of patients will suffer from chronic UTI. There is no accepted definition of this, and its prevalence is unclear, but it is generally accepted that these patients will suffer ongoing symptoms with no or little relief between attacks.[5] This is in contrast to recurrent UTI, where symptoms do resolve completely between attacks.

Complications including pyelonephritis, kidney failure and sepsis may arise as a consequence of UTI. Additionally, infections during pregnancy can cause pre-term delivery and low birth weight. Risk factors for complicated UTI include structural or neurological abnormalities, pregnancy, catheterisation, certain infecting organisms and comorbidities such as immunosuppression.[6]

The most common cause of both uncomplicated and complicated UTIs is *Escherichia coli*.[3] A recent UK-based surveillance study found that *E. coli* was isolated from 67% (113/169) of positive urine samples. Other bacteria identified in positive samples included *Klebsiella pneumoniae* (9%), *Citrobacter koseri* (5%), *Enterococcus* spp. (5%) and *Staphylococcus saprophyticus* (3.5%).[7]

TABLE 1 Overview of classification of uncomplicated UTI

Classification	Definition
Uncomplicated UTI	UTI in which there are no relevant functional or anatomical abnormalities in the urinary tract, no relevant kidney function impairment, and no relevant concomitant diseases promoting the UTI or risk of developing serious complications
Acute uncomplicated cystitis	Lower UTI in which the acute symptoms involve only the lower urinary tract, for example urgency, painful voiding (dysuria), pollakiuria, and pain above the symphysis
Acute pyelonephritis	Upper UTI with persistent symptoms including flank pain, flank tenderness or fever (temperature > 38°C)
Asymptomatic bacteriuria	Positive urine culture (> 10^5 colony-forming units/ml) in the absence of urinary symptoms
Recurrent uncomplicated UTI	Recurrent UTI refers to the occurrence of ≥ 2 symptomatic episodes within 6 months or ≥ 3 symptomatic episodes within 12 months

Source: Reproduced from Medina *et al.*[3] © The Author(s), 2019. This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (http://www.creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

Presentation of urinary tract infections

Clinical presentation of UTI varies according to patient group and can be non-specific, making it difficult to identify those who may have a UTI. Symptoms can include dysuria (discomfort/pain/burning with urination), increased daytime frequency, urgency, abdominal/suprapubic pain, haematuria, and changes in urine smell, appearance or consistency.[8,9] In those aged > 65 years, symptoms can be less specific and include delirium, lethargy, a reduced ability to carry out activities of daily living, and anorexia.[6]

Diagnosis

The accurate and timely diagnosis of UTI is important to ensure appropriate treatment to help resolve symptoms and improve quality of life and also to reduce the risk of long-term complications such as pyelonephritis, kidney disease and sepsis.[10]

Urinary tract infections are currently diagnosed using a combination of dipstick tests and laboratory-based urine culture, which usually includes antimicrobial sensitivity testing (AST). Dipstick tests involve dipping a specially treated paper or plastic strip into a urine sample to identify the presence of leukocyte esterase (LE), nitrites and blood. These tests can be used as initial screening for UTI as they can be performed by general practitioners (GPs) and give a result very quickly (within a few minutes), but their accuracy is limited, particularly in certain populations such as men, those aged > 65 years and those who are catheterised, in whom they are not recommended.[11] They are also unable to provide information on the pathogenic cause of the infection or on AST. Thus, even when these tests are used to help diagnose a UTI, follow-up laboratory testing using culture is often needed to confirm the infection and to determine AST. The European Committee on Antimicrobial Susceptibility Testing (EUCAST) provides guidance on AST that includes definitions of susceptibility testing categories with the aim of harmonising breakpoints in Europe.[12]

Culture can take 24–72 hours depending on geographical location and local laboratory facilities, and in some cases, where there are delays in getting urine samples to a laboratory or in processing the test once samples arrive at a laboratory, results can take up to 1 week to be returned to the GP. Public Health England guidance recommends culture in the following groups to help diagnose a UTI:[11]

- men
- people aged > 65 years
- babies aged < 3 months
- children aged < 16 years who do not respond to treatment within 24–48 hours
- pregnant women
- those with suspected complicated UTI (pyelonephritis or sepsis)
- those with failed antibiotic treatment or persistent symptoms
- those who have recurrent UTI
- catheterised patients
- those testing dipstick negative for nitrites but positive LE
- those aged < 3 years, with positive dipstick for nitrite and LE
- those with the following risk factors for resistance:
  - abnormalities of genitourinary tract
  - renal impairment
  - care home resident
  - hospitalised for > 7 days in last 6 months
  - recent travel to country with increased resistance
  - previous resistant UTI.

However, there are also limitations associated with culture and exactly how a UTI should be defined. Culture can be negative even when a UTI is present, particularly in the case of antibiotic-resistant bacteria. Laboratory guidelines differ in how culture result should be interpreted to confirm the presence or absence of UTI[13] and recommend different diagnostic criteria depending on age, symptoms and how urine was collected. Culture has additional limitations in populations such as frail older people in whom long-term colonisation can make diagnosis particularly difficult, and where culture cannot accurately identify those with a UTI.

Treatment of urinary tract infections

An acute uncomplicated UTI generally resolves within around 9 days without treatment,[14] but most patients with UTI will be prescribed antibiotics. Treatment also involves giving advice on self-care such as analgesia and hydration. National Institute for Health and Care Excellence (NICE) guidelines on antimicrobial prescribing for UTI recommend that antibiotics are prescribed immediately in pregnant women, men, and children aged < 16 years.[15] In non-pregnant women, a backup antibiotic (to be taken only if symptoms persist for 48 hours or worsen) or an immediate antibiotic may be prescribed. While dipstick tests and culture are often used to inform the diagnosis and decision on whether to prescribe antibiotics, in some patients antibiotics will be prescribed based on symptoms and examination alone. A recent study of treatment of lower UTI in primary care in England found that the majority of patients (80%) were given empiric antibiotic treatment on the day of diagnosis and that for the majority (83%) no evidence of urine sample collection for laboratory investigation was in their electronic health records.[16] If urine is sent for culture and AST, then the antibiotic choice should be reviewed when the AST results are available. The NICE guidelines contain detailed recommendations of which antibiotic to prescribe as first choice or second choice (if the first choice is not effective or suitable) in different populations. First-choice antibiotics are based on empiric treatment (treatment given based on experience, without exact knowledge of the cause or nature of UTI), usually with nitrofurantoin or trimethoprim. Second-choice antibiotics include pivmecillinam (a penicillin) or fosfomycin in adults and amoxicillin or cefalexin in children.[15] Empiric antibiotics may have side effects, can be less effective than targeted antibiotics (antibiotics targeting the causative pathogen) and increase the risk of antibiotic resistance developing (see *Antibiotic prescribing and resistance*).

An acute recurrent UTI is managed in the same way as acute UTI. NICE guidelines on antimicrobial prescribing for recurrent UTI recommend giving advice on behavioural and personal hygiene measures and self-care treatment to reduce the risk of future UTI. Post-menopausal women with recurrent UTI may be recommended vaginal oestrogen if other measures are not effective. Antibiotic prophylaxis can be considered if none of the other measures is effective. An alternative to this that is being increasingly used is methenamine hippuirate (Hiprex), a non-antibiotic option. This should not be started until the acute UTI has been treated and resolved. Initial prophylaxis should include single-dose antibiotics; if this is not effective, then daily antibiotic prophylaxis can be trialled. This has associated risks of resistance and possible adverse effects.[15]

There are currently no NICE guidelines on the treatment of chronic UTI. Patient organisations suggest that treatment may involve high-dose, extended-course (3–6 months) oral antibiotics or the instillation of antibiotics directly into the bladder.[17] Many patients will also seek relief from alternative therapies for which there is little evidence of effectiveness.[18]

Antibiotic prescribing and resistance

Almost 75% of antibiotic prescribing occurs in primary care,[19] with UTI contributing to a large proportion of this. Antimicrobial resistance, and in particular antibiotic resistance, is one of the greatest public

health challenges faced today. The World Health Organization (WHO) highlights this as one of the current biggest threats to global health, food security and development.[20]

The 2017 English Surveillance Programme for Antimicrobial Utilisation and Resistance (ESPAUR) report says that more than 1 million UTI samples were analysed in NHS laboratories across England in 2016, and that resistance was a 'common' observation. A recent surveillance study, published in June 2020, found that around 30% of *E. coli*, the most common cause of UTI, was resistant to trimethoprim, and around 1% was resistant to nitrofurantoin.[7] This is consistent with data from a study that evaluated the Flexicult test, which reported that around 20% of those with a microbiologically confirmed UTI had an infection that was resistant to any first-line antibiotic (nitrofurantoin, trimethoprim or fosfomycin).[7]

Chapter 2 Decision problem

Population

The population for this scope is people with suspected UTI who:

- would have an initial dipstick test in current practice (population 1)
- would *not* have an initial dipstick test in current practice (population 2).

People with suspected sepsis are not included in the scope. Subgroups of interest include:

- people with suspected acute UTI
- people with suspected recurrent UTI
- people with suspected chronic UTI
- women aged < 65 years
- women aged > 65 years
- men aged < 65 years
- men aged > 65 years
- adults with indwelling urinary catheters
- babies, children and young people aged < 16 years
- children aged < 3 months
- pregnant women
- people who are frail or have dementia
- people who are pre-, peri- or post-menopausal
- people on prophylactic antibiotics for treatment of UTI
- people of different ethnicities
- people with a higher risk of complicated UTI (e.g. people with neurogenic bladder, diabetes, polycystic kidney disease or people who are immunocompromised)
- people with suspected pyelonephritis.

Technologies of interest

Guidance from Public Health England, 'Health matters: antimicrobial resistance',[19] published in 2015, highlights the need for rapid diagnostic tools to help GPs quickly (i.e. within minutes) identify the strain of bacterial infection present and the antibiotics to which the infection is resistant or susceptible. This is also highlighted in the 2021/2 ESPAUR. Tests that give a more accurate, rapid diagnosis of UTI than current dipstick testing, with or without identifying the bacteria or providing information on AST, would have the potential to substantially improve diagnosis of UTI in primary care. Such tests may reduce inappropriate antibiotic prescribing in general, as well as improve appropriate targeting of antibiotics prescribed (see *Antibiotic prescribing and resistance*).[21] They would be particularly useful in those groups in whom dipstick testing is not recommended. Given the high proportion of those presenting with symptoms of UTI who are subsequently found not to have a UTI, novel tests would also have the potential to rule out UTIs, reducing the need for samples to be sent for laboratory testing.

The technologies of interest in this appraisal are novel point-of-care tests (POCTs) that may detect the presence of a UTI and provide information on the strain of bacterial infection present and/or the antibiotics to which the bacteria are susceptible. POCTs are defined as technologies that a healthcare professional can carry out outside a conventional laboratory setting.[22] *Table 2* gives an overview of POCTs for diagnosing UTIs within the scope of this appraisal. These tests were identified as part of the appraisal process and were specified in the NICE scope. These are grouped into rapid tests (those that

TABLE 2 Overview of POCTs for diagnosing UTI within the scope of this assessment

Test name	Test basis	Sample	Antibiotics/ bacteria targeted	Time to detect bacteria	Time to detect pathogenic cause	Time to result AST	Test interpretation	CE-IVD marked
Rapid tests (results < 40 minutes)								
Astrego PA-100 analyser and PA-AST panel U-0501 (Sysmex Astrego)	Microfluidics	Urine	Five commonly used antibiotics (amoxicillin/ clavulanic acid, ciprofloxacin, fosfomycin, nitrofurantoin, trimethoprim)	10–15 minutes	N/A	30–45 minutes for full results	Digital display shows which antibiotics sample is susceptible to	Yes
Lodestar DX (Llusern Scientific)	Molecular diagnostic test	Urine	*E. coli*, *Klebsiella* spp., *Proteus mirabilis*, *S. saprophyticus*, *Enterococcus* spp., *Pseudomonas aeruginosa*	40 minutes	40 minutes	N/A	Digital display – light indicates which bacteria are detected	Expected < 12 months
TriVerity (Inflammatix)	Detects 29 target mRNAs	Blood	Identifies presence, type and severity of infection	30 minutes	N/A	N/A	Unclear	Expected < 12 months
Uriscreen (Savyon Diagnostics Ltd)	Catalase-based test	Urine	Detects catalase activity as indicator of bacteria in somatic cells	2 minutes	N/A	N/A	Visual detection – white foam indicates positive result	Yes
UTRiPLEX (Global Access Diagnostics)	Dipstick for detection of inflammatory biomarkers	Urine	Detects presence of urinary biomarkers MMP8 and HNE	6 minutes	N/A	N/A	Visual reading of dipstick – line indicates UTI	Expected < 12 months
Culture-based tests (results up to 24 hours)								
Flexicult Human, ID Flexicult (SSI Diagnostica)	Culture	Urine	Flexicult Human: five commonly used antibiotics (mecillinam, nitrofurantoin, ampicillin, sulfamethizole and trimethoprim) ID Flexicult gives information on pathogenic cause	16–24 hours	16–24 hours	16–24 hours	Visual assessment of number and type of growths on agar plate	Yes
Diaslide, Dipstreak, Chromostreak (Novamed)	Semi-quantitative culture	Urine	Total bacterial count; presence of Gram-negative bacteria; growth of common UTI-causing bacteria (*E. coli*, *Proteus* and enterococci) – chromastreak only	18–24 hours	18–24 hours	N/A	Number of bacterial colonies is compared with the Colony Density Chart	Yes

TABLE 2 Overview of POCTs for diagnosing UTI within the scope of this assessment (*continued*)

Test name	Test basis	Sample	Antibiotics/ bacteria targeted	Time to detect bacteria	Time to detect pathogenic cause	Time to result AST	Test interpretation	CE-IVD marked
Uricult, Uricult t Trio and Uricult Plus (Aidian; formerly Orion Diagnostica)	Culture	Urine	Uricult identifies presence of Gram-negative bacteria; Uricult Plus also detects enterococci; Uricult Trio also detects Gram-negative, β-glucuroni-dase-producing organisms, e.g. *E. coli*	16–24 hours	16–24 hours	N/A	Visual assessment of growth on agar plate	Yes

HNE, 4-Hydroxynonenal; MMP8, matrix metalloproteinase-8; N/A, not applicable.

provide results in < 40 minutes) and culture-based tests (which take up to 24 hours to give results). The aim of these tests is to provide more accurate, rapid diagnoses of UTIs and improve antibiotic prescribing. The extent to which these POCTs can improve antibiotic prescribing will depend on how quickly they are able to provide results, how accurate they are, whether they provide additional information on the specific pathogen present in the urine, and whether they provide information on AST.

Potential alternative technologies

A number of technologies are currently in development that will be able to rapidly indicate the presence of bacteria, identify the bacteria present and/or provide information on antimicrobial susceptibility, but these do not have a Conformité Européenne or UK Conformity Assessment (UKCA) mark, and are not expected to obtain this in the next 12 months, and so cannot yet be considered for recommendation by NICE.

Comparator

The comparator for this assessment is the current standard of care: (1) urine dipstick followed by confirmatory culture and AST (if necessary; population 1) or (2) urine culture and AST carried out in the laboratory (population 2). This varies according to population. Further details of the treatment pathway are provided in *Current treatment pathway*.

Current treatment pathway

The exact treatment pathway varies according to the population (age, sex and whether catheterised). *Figure 1* provides a general overview of the treatment pathway. A person presents to their GP with symptoms suggestive of UTI. Depending on the patient population, the person may receive dipstick testing. If this test is positive for nitrite and LE, the person will be diagnosed with UTI; in some populations (e.g. women aged < 65 years) a diagnosis can also be made based on a positive nitrite alone or LE, if also positive for blood. A sample may be sent to the laboratory for susceptibility testing. Decisions about whether to prescribe antibiotics, and which antibiotic to prescribe, are often made

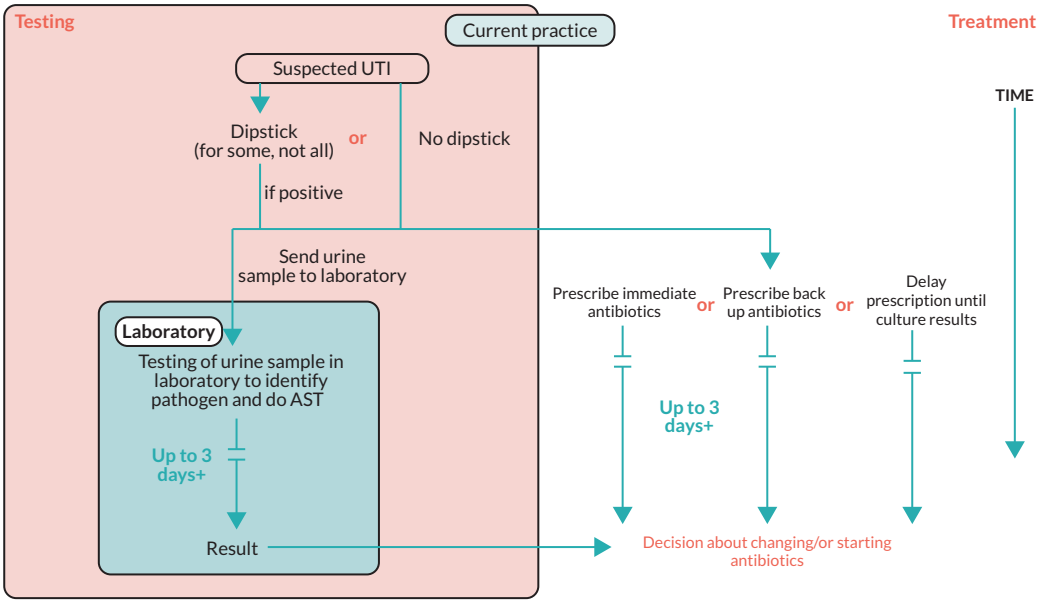

FIGURE 1 Outline of treatment pathway. Reproduced with permission from NICE. © NICE [2023] Point-of-care tests for urinary tract infections to improve antimicrobial prescribing: early value assessment(TS insert superscript 23). Available from www.nice.org.uk/guidance/hte7 All rights reserved. Subject to Notice of rights. NICE guidance is prepared for the National Health Service in England. All NICE guidance is subject to regular review and may be updated or withdrawn. NICE accepts no responsibility for the use of its content in this product/publication.

before culture results are available, particularly if the person has presented with severe symptoms. This means that antibiotics may need to be changed if culture and AST suggest that the person is taking an antibiotic that is not likely to be effective against their infection, or stopped if no infection is detected on culture.

Public Health England has separate pathways for infants/children aged < 16 years, women aged < 65 years, men aged < 65 years, adults who are catheterised and adults aged > 65 years.[11]

The treatment pathways differ in terms of whether an initial dipstick test is done, whether a urine sample should be sent to a laboratory for culture testing and when or if to prescribe antibiotics. *Table 3* provides an overview of recommendations from the treatment pathways for these different groups:

Place of the technology in the treatment pathway

A POCT for suspected UTI would be used as an initial test to diagnose UTI. If its performance is sufficient, then its place in the treatment pathway, as an initial test to diagnose UTI, will be the same in all populations and prespecified subgroups (see *Population*).

A POCT's role in UTI diagnosis will depend on whether it provides additional information on the specific pathogen present in the urine, whether it provides information on AST, and the time taken to produce the result. This will also affect the potential impact of the tests. *Table 4* provides an overview of the potential role and impact of a new POCT based on its features.

TABLE 3 Summary of recommendations for dipstick, culture and antibiotics in different patient groups for lower UTI[11]

Population	Dipstick	Culture	Immediate antibiotics
Children (aged < 16 years)	Yes	If no response to treatment in 24–48 hours or aged < 3 years with positive dipstick for nitrite and LE	Yes (depending on dipstick result)
Men aged < 65 years	Yes – but not to rule out infection	Yes	Yes
Women aged < 65 years	Yes – those without risk factors for complicated UTI Not needed if have two or three key diagnostic signs/symptoms	Dipstick negative for nitrites but positive LE	Delayed prescription may be offered in some patients
Pregnant	Yes	Yes	Yes (depending on dipstick result)
Catheterised	No	Yes	Yes
Men aged > 65 years	No	Yes	Yes
Women aged > 65 years	No	Yes	Yes, or backup antibiotics if symptoms mild

TABLE 4 Overview of the potential role and impact of new POCT based on its features

Test features	Role	Potential impact
Detection of UTI	• Triage – rule out UTI or identify those in whom further testing for AST is required. This includes groups in whom dipstick testing is not currently recommended • Replace dipstick in populations in whom dipstick testing is recommended	• Inform need for antibiotics • Reduce unnecessary antibiotic prescription • Enable quicker access to antibiotics when needed • Reduce need for culture
Detection of UTI plus pathogen identification	• Triage – rule out UTI or identify those in whom further testing for AST is required. This includes groups in whom dipstick testing is not currently recommended • Replace dipstick in populations in whom dipstick testing is recommended	• Inform need for antibiotics • Reduce unnecessary antibiotic prescription • Quicker access to antibiotics when needed • Reduce need for culture • Provide some indication for initial antibiotic prescription based on type of bacteria but not to AST
Detection of UTIs plus AST	• Replace dipstick and laboratory testing	• Inform need for antibiotics • Reduce unnecessary antibiotic prescription • Enable quicker access to antibiotics when needed • Target initial antibiotic prescription to AST • Reduce need for culture and AST

Chapter 3 Objectives

The overall aim of this project is to determine whether POCTs for people with suspected UTI have the potential to be clinically effective and cost-effective to the NHS. We will summarise the available evidence to support the value proposition outlined in the scope and outline where there are evidence gaps.

1. What is the impact on clinical outcomes of using POCTs to diagnose UTI, with or without additional pathogen identification and AST?
2. What is the accuracy of POCTs for UTI diagnosis, pathogen identification and AST?
3. What is the technical performance (other than accuracy) of POCTs for UTI?
4. What are the costs, from a UK NHS and Personal Social Services (PSS) perspective, of using POCTs for UTI diagnosis, pathogen identification and AST?
5. How might a conceptual model be specified in terms of structure and evidence required for parameterisation in order to estimate the cost-effectiveness of POCTs for UTI diagnosis, pathogen identification and AST?

Chapter 4 Methods for assessment of clinical effectiveness

This report contains reference to confidential information provided as part of the NICE Diagnostic Assessment process. This information has been removed from the report and the results, discussions and conclusions of the report do not include the confidential information. These sections are clearly marked in the report.

A systematic review was conducted to summarise the evidence on the accuracy, technical performance and clinical effects of using POCTs in people with suspected UTI. The systematic review followed the principles outlined in the Centre for Reviews and Dissemination (CRD) guidance for undertaking reviews in health care, the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy and the NICE Health Technology Evaluations Manual.[24–26] The review is reported in accordance with the PRISMA 2020 guidance.[27] The review protocol was registered on the PROSPERO database (CRD42022383889).

Inclusion and exclusion criteria

Studies that met the criteria summarised in *Table 5* were eligible for inclusion:

Given the tight timelines for conducting an early value assessment (EVA), it was necessary to restrict the review so that it could be undertaken within the available time. The review was therefore restricted to studies reported (published or unpublished) after 2000. We consider it likely that clinical practice, the spectrum of bacteria causing UTI, and the technical performance of tests evaluated will have changed such that studies published before this date are unlikely to provide useful information to inform this appraisal. Animal studies were excluded.

TABLE 5 Inclusion criteria for objectives 1, 2 and 3

	Objective 1: clinical impact	Objective 2: accuracy	Objective 3: technical performance
Participants	Patients with suspected UTI. Studies in patients with suspected acute, recurrent or chronic UTI will be eligible		
Technology	*Rapid tests*: Astrego PA-100 system, Lodestar DX, TriVerity, Uriscreen, UTRiPLEX *Culture-based tests*: Flexicult Human, ID Flexicult, Diaslide, Dipstreak, Chromostreak, Uricult, Uricult Trio or Uricult Plus		
Comparator/ reference standard	Standard care: dipstick plus culture or culture alone	Culture or other reported reference standard	N/A

Outcome	• Morbidity, including:
	  ○ Recurrence
	  ○ Pyelonephritis
	  ○ Sepsis
	  ○ Adverse effects of antibiotics
	• Any outcome related to antibiotic use or prescription
	• Mortality
	• UTI-associated healthcare resources
	• Health-related quality of life
		Test accuracy in detecting UTI, identifying pathogens or assessing susceptibility to antimicrobials	• Test failure rate
			• Ease of use/acceptability
			• Time to test results
			• Any outcome related to antibiotic use or prescription
			• UTI-associated healthcare resources
			• Health-related quality of life
			• Test costs
			• Any reported data on clinical outcomes, e.g. morbidity/mortality

Setting	Primary care or community setting	Any	Any
Study design	RCT or non-randomised study of interventions	Diagnostic test accuracy study	Any

Study identification

Studies were identified using bibliographic and non-bibliographic search methods following guidance in the NICE technology appraisal manual and recent guidance on searching.[28,29]

Bibliographic searching

The following databases were searched:

- MEDLINE (via Ovid SP)
- EMBASE (via Ovid SP)
- Cochrane Central Register of Controlled Trials (CENTRAL) (via Wiley)
- Cumulative Index to Nursing and Allied Health Literature (CINAHL) (via EBSCO*host*).

We used a sensitive search strategy based on terms for each of the technologies eligible for inclusion and for the manufacturers of these technologies. Full details of the search strategy are available in *Appendix 1*.

Non-bibliographic search methods

Completed and ongoing trials were identified through searches of the following trial registries:

- ClinicalTrials.gov via www.clinicaltrials.gov/
- WHO International Clinical Trials Registry Platform (ICTRP) via www.who.int/clinical-trials-registry-platform

Additional relevant studies were identified by:

- screening reference lists of any reviews (systematic or non-systematic) identified by our searches
- reviewing the reference lists of any study report included at full-text stage
- hand-searching the websites of the manufacturer/or licence holders of each test
- reviewing information submitted by test manufacturers.

Managing the searches

Search results were exported to EndNote 20 [Clarivate Analytics (formerly Thomson Reuters), Philadelphia, PA, USA] for deduplication using the default deduplication settings and manual review of records. Search results were then exported to Microsoft Access (Microsoft Corporation, Redmond, WA, USA) for screening.

Review strategy

Two reviewers independently screened titles and abstracts identified by the searches. Full copies of all reports considered potentially relevant were obtained, and two reviewers independently assessed these for inclusion. Any disagreements were resolved by consensus or discussion with a third reviewer.

Data were extracted using standardised data extraction forms developed in Microsoft Access (objective 2) and Microsoft Word (Microsoft Corporation, Redmond, WA, USA) (objectives 1 and 3). Data extraction forms were piloted on a small sample of papers and adapted as necessary. Data were extracted by one reviewer and checked in detail by a second reviewer. Any disagreements were resolved by consensus or discussion with a third reviewer.

Data were extracted on the following: study design [randomised controlled trial (RCT), diagnostic test accuracy (DTA) or other], objective that study addresses, funding sources (public, industry, mixed), country of study, population, sex, age, inclusion/exclusion criteria, number of participants, rapid-test details (manufacturer, antibiotics targeted, location of test performance, urine sampling methods), comparator or reference standard test(s), and outcomes specified in inclusion criteria

(see *Inclusion and exclusion criteria*). If data were reported on any of the following subgroups of interest, these were extracted separately:

- people with suspected acute UTI
- people with suspected recurrent UTI
- people with suspected chronic UTI
- women aged < 65 years
- women aged > 65 years
- men aged < 65 years
- men aged > 65 years
- adults with indwelling urinary catheters
- babies, children and young people aged < 16 years
- children aged < 3 months
- pregnant women
- people who are frail or have dementia
- people who are pre-, peri- or post-menopausal
- people on prophylactic antibiotics for treatment of UTI
- people of different ethnicities
- people with a higher risk of complicated UTI (e.g. people with neurogenic bladder, diabetes, polycystic kidney disease or people who are immunocompromised)
- people with suspected pyelonephritis.

Dichotomous clinical impact data were extracted as number of patients with events and/or number of events and total number of patients in each treatment arm, where reported. For all types of data, effect estimates (odds ratios, hazard ratios or mean difference), with 95% confidence intervals (CIs) and *p*-values for comparisons between groups, together with details on the methods of analysis and the test statistic, were extracted.

Accuracy data were extracted as 2 × 2 tables comparing the POCT with the reference standard, where available. If measures of accuracy (e.g. sensitivity, specificity, receiver operating characteristic plot) were reported without the information needed to calculated 2 × 2 tables, then these data were extracted. We considered accuracy separately for the following target conditions:

- presence of UTI
- pathogenic cause of UTI
- antimicrobial sensitivity.

Where multiple sets of 2 × 2 data were reported in a single study, for example, for different tests, target conditions, thresholds or subgroups of interest, all data were extracted.

Quality assessment strategy

The methodological quality of included RCTs was assessed using the updated Cochrane Risk of Bias Tool (RoB 2).[30] We had intended to assess the risk of bias in non-randomised studies of interventions using the ROBINS-I tool, but no studies of this design were identified.[31] The methodological quality of DTA studies was assessed using QUADAS-2.[32] We modified the tool slightly in that we did not consider applicability given the broad range of populations and tests for interest defined in the review question. Potential sources of heterogeneity were instead considered in the synthesis. Quality assessment was undertaken by one reviewer and checked by a second reviewer. Any disagreements were resolved by consensus or discussion with a third reviewer.

Synthesis methods

For each of the three systematic review objectives (1–3), a narrative summary of all of the included studies is presented. This includes a summary of the study characteristics, outcomes reported and study

quality. The synthesis was stratified by the test evaluated with tests grouped into rapid tests (produce results in < 40 minutes) and culture-based tests.

For objective 2, coupled forest plots of sensitivity and specificity were used to display results from individual studies to allow a visual assessment of heterogeneity. To create this plot, we selected one set of 2 × 2 data per study/population and test. If multiple index test and culture thresholds were reported in a study, then we selected the same thresholds for index test and culture, where possible. Where results were presented for multiple reference standards, we selected the reference standard considered to be the most likely to give an accurate result (e.g. culture, microscopy and spiral plating was chosen over culture and microscopy alone).

Meta-analysis of sensitivity and specificity was performed separately for each test, producing summary estimates of sensitivity and specificity with 95% CIs. The decision to combine results from studies performed in the laboratory with studies performed in the near-patient setting was made on a test-by-test basis, considering the nature of the test. Meta-analyses assumed binomial likelihoods for numbers of true positives and numbers of true negatives. Where results were pooled across four or more studies, bivariate random-effects meta-analysis was used.[33,34] Where results were pooled across only three or two studies, univariate random-effects or fixed-effect meta-analysis, respectively, was performed, owing to a lack of data for estimating all parameters in a bivariate random-effects model. We did not have sufficient studies for formal investigations of heterogeneity. We had intended to stratify the analysis based on the populations specified in the scope, but there were insufficient data available to do this.

Protocol changes

- We had originally specified that studies would be included for objective 3 only if they evaluated a test that had not been considered as part of objective 1 or 2. However, owing to the very small number of studies that we identified that fulfilled the inclusion criteria for objective 3, we removed this restriction and included studies of any of the technologies of interest.
- In addition to Flexicult Human, we identified a number of studies of ID Flexicult. This test was not specifically in the scope but is included in the review as we consider it possible that ID Flexicult identifies the same information as the control field of Flexicult Human; however, this has not been confirmed by the company.

Chapter 5 Results of the clinical effectiveness review

Search results

The searches of bibliographic databases and trials registries identified 728 unique references after deduplication. After initial screening of titles and abstracts, 38 reports were considered potentially relevant and retrieved for full-paper screening.

In total, 16 studies in 28 reports were included in the review. Two studies in six reports were included for objective 1. Sixteen studies in 20 reports were included for objective 2. Two of these studies were also included in objective 1, and separate reports of DTA substudies provided data for objective 2. Five studies in five reports were included for objective 3. Four of these studies were also included for either objective 1 or objective 2. The final study was a report of a qualitative substudy from one of the studies included for objective 1.

The process of study identification and selection is summarised in *Figure 2*. *Table 6* provides an overview of the number of studies assessing each test for each of our three clinical objectives, stratified by test. There were no data for any of the objectives for the following tests: Astrego PA-100 system, TriVerity, Diaslide, Chromostreak or Uricult Plus. The majority of studies evaluated culture-based tests, which take up to 24 hours to provide results. Uriscreen was the only rapid test to be evaluated in more than one study.

Table 7 provides an overview of the populations defined in the scope and whether data were available for these populations. The majority of populations were not specifically considered in the included studies, although they might have been included in studies that enrolled mixed populations.

We excluded studies published before the year 2000, as outlined in *Chapter 4*. These were excluded after title and abstract screening. *Appendix 2* provides a summary of the 62 studies excluded for this reason, showing which test and objective they potentially evaluated. As these were only screened at title and abstract stage, they were not reviewed at full-text screening stage, and so it is likely that not all of these studies would have been included in the review had the date restriction not been applied. All evaluated culture-based tests: the majority (n = 47) evaluated Uricult, two evaluated Uricult Trio, seven evaluated Uriscreen, one evaluated Diaslide, and it was not possible to tell which test was evaluated in the remaining five.

Objective 1: what is the impact on clinical outcomes of using point-of-care tests to diagnose urinary tract infection, with or without additional pathogen identification and antimicrobial sensitivity testing?

Two individually randomised RCTs evaluated the clinical impact of using Flexicult Human (often referred to in studies as the Flexicult SSI urinary kit): the POCT for urinary tract infection in primary care (POETIC) trial[8] and a Danish trial.[35] Both trials were conducted in primary care and enrolled women aged > 18 years with symptoms suggestive of uncomplicated UTI. In both studies, all participants also had a urine sample sent for laboratory culture, which meant that a diagnostic accuracy substudy could be performed; the results of these two substudies are included for objective 2 (see *Objective 2: what is the accuracy of the point-of-care test for urinary tract infection diagnosis, pathogen identification and antimicrobial sensitivity testing?*).[36,37] Both studies were considered at low risk of bias (see *Appendix 1*).

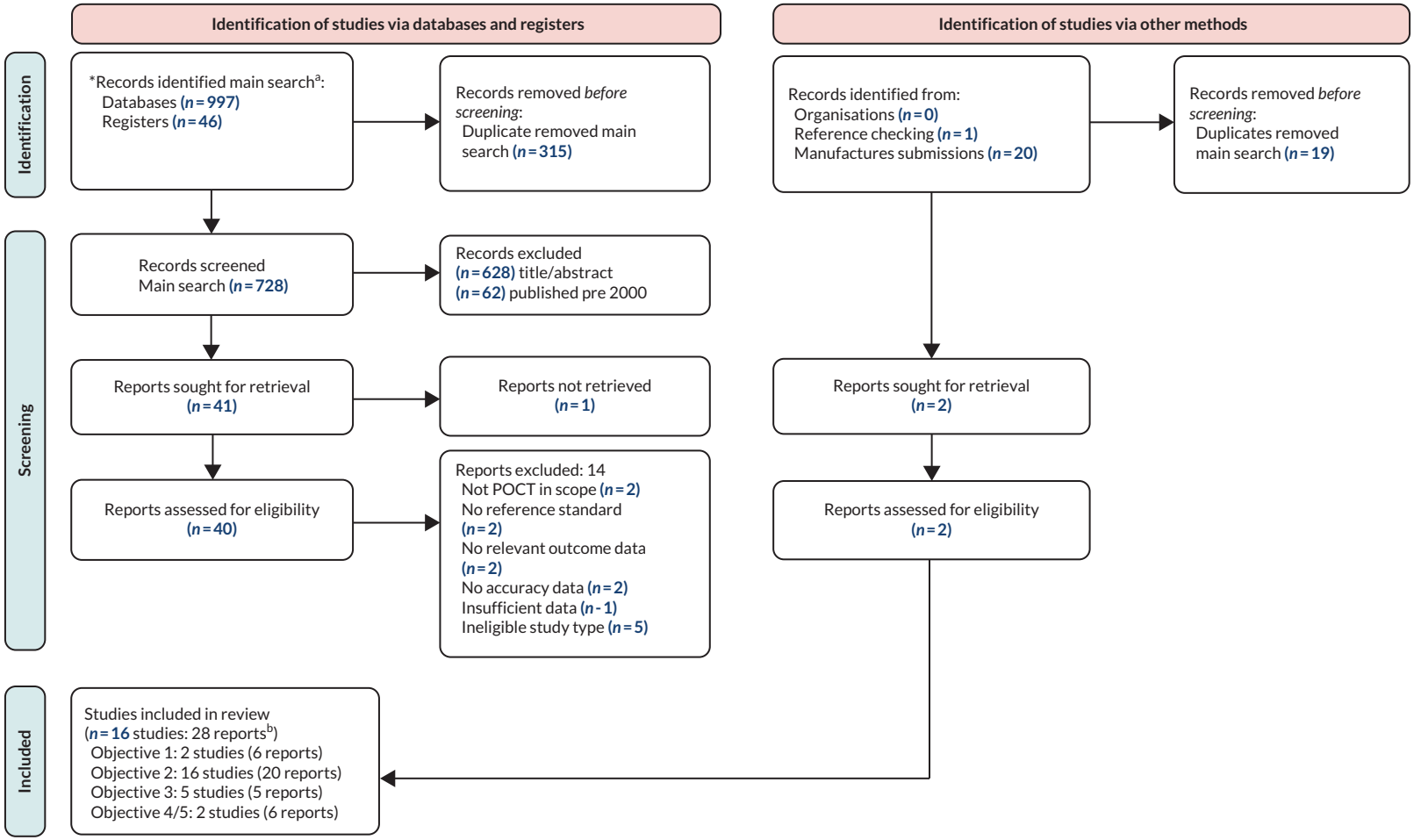

FIGURE 2 Flow of studies through the review process. (a) After the main searches had been completed, as additional test (UTRiPLEX) was added to the scope of the review; (b) studies and study reports contributed to more than one objective.

TABLE 6 Overview of number of studies assessing each test for each of the review objectives

Test	Objective 1	Objective 2	Objective 3
Rapid tests: results < 40 minutes			
Astrego PA-100 system	0	0	0
Lodestar DX	0	1	0
TriVerity	0	0	0
Uriscreen	0	4	0
UTRiPLEX	0	1	0
Culture-based: up to 24 hours for results			
Flexicult Human	2	4	2
ID Flexicult	1	2	0
Diaslide	0	0	0
Dipstreak	0	2	0
Chromostreak	0	0	0
Uricult	0	1	0
Uricult Plus	0	0	0
Uricult Trio	0	3	2

Note
Tests in grey-shaded cells were not evaluated in any included studies. Two studies (one for objective 1 and one for objective 2) evaluated two tests of interest.

TABLE 7 Overview of populations defined in the scope and whether data were available specifically for each population/subgroup of interest

Population	Data available for specific group of interest?
People with suspected acute UTI	Yes
People with suspected recurrent UTI	No
People with suspected chronic UTI	No
Women aged < 65 years	Yes (studies of women only; no age restrictions)
Women aged > 65 years	
Men aged < 65 years	No
Men aged > 65 years	No
Adults with indwelling urinary catheters	Yes
Babies, children and young people aged < 16 years	Yes
Children aged < 3 months	No
Pregnant women	Yes
People who are frail or have dementia	No
People who are pre-, peri- or post-menopausal	No
People on prophylactic antibiotics for treatment of UTI	No
People of different ethnicities	No
People with a higher risk of complicated UTIs	No
People with suspected pyelonephritis	No

Neither study was funded directly by the test manufacturer, although the manufacturer provided the tests in the Danish study.

The POETIC trial was conducted across four countries: England, the Netherlands, Spain and Wales. It randomised 654 participants: 329 to testing with Flexicult Human and treating based on results (England, *n* = 117; Wales, *n* = 109) and 325 to standard care informed by national guidelines (England, *n* = 117; Wales, *n* = 110). One male participant was then excluded, resulting in a sample of 653 women. Flexicult plates specific to the antibiotics most commonly used in each of the four regions were developed. GPs were free to determine how best to use the test. Examples of how it could be used included:

- to determine whether, and what antibiotic class, to prescribe the following day
- to prescribe empirically and to aid in a next-day review of the initial prescribing decision
- to provide delayed antibiotics prescription and to guide the use of delayed prescription.

The Danish trial randomised 376 women to two different Flexicult-based strategies: Flexicult Human (which incorporates susceptibility testing) or ID Flexicult (which does not include susceptibility testing). In both arms, GPs were advised to treat based on test results.

The results of the two trials are summarised in *Table 8*. The POETIC trial reported six different measures of antibiotic use. There was evidence that antibiotic prescribing at the initial consultation had reduced

TABLE 8 Results of trials of clinical impact of the Flexicult Human test

Study	Outcome		Effect measure – estimate (95% CI)
Antibiotic use			
Butler *et al.* (2018)[8] (POETIC trial)	Concordant antibiotic use		OR 0.84 (0.58 to 1.20)
	Antibiotic prescribing at initial consultation		OR 0.56 (0.35 to 0.88)
	Antibiotics prescribed to guidelines at initial consultation		OR 0.99 (0.67 to 1.45)
	Antibiotic consumed day 3		OR 1.24 (0.81 to 1.89)
	Antibiotic consumed (during 2 weeks)		OR 1.38 (0.87 to 2.19)
	New antibiotic prescription (within 2 weeks)		OR 1.11 (0.65 to 1.89)
	Drug type and duration	UTI-specific and 1–3 days	Reference
		UTI-specific and > 3 days	RR 1.15 (0.71 to 1.87)
		Broad spectrum and 1–3 days	N/A (0 events)
		Broad spectrum and > 3 days	RR 1.00 (0.58 to 1.75)
Holm *et al.* (2017)[35] (Danish trial)	Appropriate prescribing		OR 1.44 (1.03 to 1.99)
UTI/symptom incidence or duration			
Butler *et al.* (2018)[8] (POETIC trial)	Microbiologically confirmed UTI (at 2 weeks)		OR 0.94 (0.49 to 1.81)
	Recurrence of UTI within 3-month period		OR 0.72 (0.48 to 1.07)
	Duration of symptoms		HR 1.02 (0.83 to 1.25)
	Duration of moderately bad symptoms		HR 0.98 (0.82 to 1.17)
	Overall urinary symptom burden		MD 0.99 (0.84 to 1.19)
	No significant bacteriuria on day 14		OR 1.15 (0.62 to 2.13)

TABLE 8 Results of trials of clinical impact of the Flexicult Human test (*continued*)

Study	Outcome	Effect measure – estimate (95% CI)
Holm *et al.* (2017)[35] (Danish trial)	Symptom free on day 5	OR 0.91 (0.56 to 1.49)
Enablement		
Butler *et al.* (2018)[8] (POETIC trial)	Patient enablement (measured using Patient Enablement Instrument at day 14 and 3 months[38])	OR 0.99 (0.66 to 1.48)
Resource use		
Butler *et al.* (2018)[8] (POETIC trial)	Re-consultation (within 2 weeks)	OR 0.99 (0.62 to 1.60)
	Hospital stay (within 2 weeks)	Numbers too small

HR, hazard ratio; MD, mean difference; N/A, not applicable; OR, odds ratio; RR, risk ratio.

[odds ratio (OR) 0.56, 95% CI 0.35 to 0.88], but this did not impact on overall antibiotic prescription or on antibiotic use that was concordant with culture results (the primary outcome for the trial). Concordant antibiotic use is defined by Butler *et al.*[8] as 'consumption of an antibiotic on day 3 (or day 1 or day 2 for Fosfomycin), for which a pathogen considered to be causing a UTI isolated in a laboratory was sensitive in vitro; or no antibiotic use by females who did not have a UTI on laboratory culture'. The Danish trial only reported on 'appropriate antibiotic prescribing'; there was some evidence that appropriate prescribing was higher in the control arm than in the Flexicult Human arm (OR 1.11, 95% CI 1.03 to 1.99). Appropriate prescribing is defined by Holm *et al.*[35] as:

1. prescription of a first-line antibiotic to which the infecting pathogen is susceptible if the individual is found to have UTI in the reference
2. prescription of a second-line antibiotic if the individual has UTI but is allergic to the antibiotic or the pathogen is resistant to all first-line antibiotics
3. no antibiotic prescription if the individual is found to not have UTI in the reference.

Both trials also looked at improvement or duration of symptoms and microbiological cure. There was no evidence of any difference between groups for any of these outcomes. The POETIC trial looked at additional outcomes of enablement and resource use (re-consultation or hospital stay within 2 weeks) and found no differences between intervention groups. There were no data for the following outcomes prespecified in our protocol: mortality, health-related quality of life, recurrence, pyelonephritis, sepsis or adverse effects of antibiotics.

Objective 2: what is the accuracy of the point-of-care test for urinary tract infection diagnosis, pathogen identification and antimicrobial sensitivity testing?

Sixteen studies, reported in 20 publications, reported data on test accuracy and were included for this objective.[19,36,37,39–51] Studies were conducted in Denmark (n = 3[37,39,40]), Wales (n = 2[19,51]), Israel (n = 2[49,50]), Hawaii (n = 1[41]), Venezuela (n = 1[42]), Belgium (n = 1[43]), Mexico (n = 1[44]), Philippines (n = 1[45]), South Africa (n = 1[46]), Republic of Korea (n = 1[47]) and Argentina (n = 1[48]), and one study was undertaken in Wales, England, Spain and the Netherlands (n = 1[36]). Most studies were reported in English, with the exception of one in Korean[47] and one in Spanish.[44] These were translated using Google Translate (Google Inc., Mountain View, CA, USA); the Spanish translation was checked by a member of the team whose native language is Spanish and was found to be accurate. One study was included from a manufacturer's submission (submitted in response to a request for information) in the form of a draft manuscript that is academic in confidence.[51] All other studies were published as full reports. *Table 9* provides an overview of the included studies' key characteristics. Full details of each included study are reported in *Appendix 3*.

TABLE 9 Characteristics of the 16 studies reporting the accuracy of POCTs

	Rapid tests (results < 40 minutes)			Culture-based tests (results up to 24 hours)				
	Lodestar DX	Uriscreen	UTRiPLEX	Flexicult Human	ID Flexicult	Uricult Trio	Uricult	Dipstreak
Number of studies[a]	1	4	1	4	2	3	1	2
Reference	51	41–44	43	19,36,37,39	37,40	45–47	48	49,50
Population	1 Mixed	2 Screening – pregnant women 1 Children (aged < 18 years) 1 Catheterised ICU	1 Children (aged < 18 years)	2 Women -uncompli-cated UTI 1 Mixed 1 Mixed	2 Women -uncomplicated UTI	1 Pregnant women 1 Children (aged < 16 years) 1 Aged < 24 months	1 Screening – pregnant women	2 Mixed
Urine sampling	1 NR	1 Mid-stream 1 Mid-stream/ adhesive bags 2 Catheter	1 Mid-stream or adhesive bags	2 Mid-stream 1 Mid-stream/ catheter/ unknown 1 NR	2 Mid-stream	2 Mid-stream 1 Mid-stream/ collection bags	1 Mid-stream	1 Mid-stream 1 NR
Country	1 Wales	1 Hawaii 1 Venezuela 1 Belgium 1 Mexico	1 Belgium	2 Denmark 1 Wales 1 Wales, England, Spain, Netherlands	2 Denmark	1 Philippines 1 South Africa 1 Korea	1 Argentina	2 Israel
Setting	1 Lab	2 Antenatal clinics 1 Primary care 1 ICU	1 Primary care	1 Laboratory 3 Primary care	2 Primary care	2 Secondary care 1 Antenatal clinics	1 Antenatal clinics	2 Laboratory
Funding	1 Industry	2 Non-industry 2 NR	1 NR	3 Non-industry 1 NR	2 Non-industry	2 NR 1 Mixed industry/ non-industry	1 Non-industry	2 NR
Outcome	1 POE	4 POU	1 POU	3 POU + AMS 1 POU	2 POU	2 POU 1 POU + POE	1 POU	1 POU 1 POU + PC
Test location	1 Laboratory	3 Near patient 1 Laboratory	1 Laboratory	1 Laboratory 3 Near patient	2 Near patient	3 Near patient	1 Laboratory	2 Laboratory

AMS, antimicrobial sensitivity; NR, not reported; PC, pathogenic cause; POE, presence of *E. coli*; POU, presence of UTI.

a Two studies reported data on two test comparisons: (1) Flexicult Human and ID Flexicult and (2) Uriscreen and UTRiPLEX. These are counted twice in this table.[37,43]

Note

Mixed: laboratory-based studies using samples from a mixed population, for example hospitalised patients and outpatients (does not refer to whether patients had symptoms or not; this and further details are reported in *Objective 2: what is the accuracy of the POCT for UTI diagnosis, pathogen identification and AST?*).

The majority of studies evaluated culture-based tests that take up to 24 hours to provide results. Four studies evaluated the Flexicult Human test (referred to in all studies as the Flexicult SSI Urinary Kit),[19,36,37,39] three evaluated Uricult Trio,[45–47] two evaluated ID Flexicult,[37,40] two evaluated Dipstreak[49,50] and one evaluated Uricult.[48] The only rapid test to be evaluated in multiple studies was the Uriscreen test, which was evaluated in four studies;[41–44] UTRiPLEX[43] and Lodestar DX[51] were each evaluated in single studies. Two studies evaluated two tests of interest; one evaluated Flexicult Human and ID Flexicult and the other evaluated Uriscreen and UTRiPLEX.[37,43] The manufacturers' submissions highlighted two ongoing studies that will provide data on the accuracy of the Astrego PA-100 AST

test and the Lodestar DX, both rapid tests for UTI. (Confidential information has been removed). The Lodestar submission highlighted TOUCAN, a study evaluating the accuracy of three or four POCTs (details of these not yet available) in up to 800 women who consult their GP with symptoms of UTI. This study was due to complete in October 2023.[52]

Four studies were laboratory-based; three tested fresh urine samples[19,49,50] and one tested both fresh and stored urine samples.[51] The other 12 studies were conducted in primary or secondary care. Most of these studies performed the POCT in a near-patient setting but two performed the test in the laboratory.[43,48]

Four studies recruited pregnant women,[41,42,46,48] three studies recruited women with uncomplicated UTI,[36,37,40] one study enrolled catheterised ICU patients,[44] and three studies recruited children and/or infants aged under 18 years,[43] 16 years[45] and 24 months.[47] Five studies analysed samples from mixed populations: two included people visiting outpatient clinics and hospitalised patients;[49,50] one included symptomatic patients consulting the GP;[39] and two tested samples submitted to the Public Health Wales microbiology laboratory.[19,51] No further information was provided on these mixed populations. Three studies specifically stated that those with recurrent UTI were excluded;[37,44,45] information on whether those with recurrent or chronic UTI were eligible was not reported in the remaining studies.

Seven studies enrolled symptomatic patients[36,37,39,40,43,45,47] and four enrolled asymptomatic patients.[41,42,44,48] One study comprised a mix of asymptomatic and symptomatic patients and stratified results accordingly.[46] The four laboratory-based studies did not specify whether urine samples came from symptomatic patients, but as they tested urine samples that had been referred to the laboratory it seems likely the sample comprised a mix of symptomatic and asymptomatic patients.[19,49–51]

In the 10 studies that enrolled people and then took urine samples to test for UTI,[36,37,40–43,45–48] the number of patients ranged from 117 to 2173 (mean 459 patients). Another study enrolled 57 patients and took multiple samples from each patient, giving a total of 108 samples.[44] In the five studies that tested urine samples rather than enrolling patients,[19,39,49–51] the number of samples ranged from 121 to 955 (mean 578 patients).

One study was funded by the test manufacturer.[51] One study was funded by industry (not the test manufacturer) and non-industry.[45] Seven studies did not report funder details[39,42,44,46,47,49,50] and all other studies were non-industry funded.

All included studies except for one[51] assessed the accuracy of POCTs for detecting the presence of UTI. Three of these studies also reported data on antimicrobial sensitivity,[19,36,39] one reported data on pathogenic cause,[50] and one reported data on presence of *E. coli*.[47] Most studies used culture alone as the reference standard, with the exception of one study that used culture and microscopy, and culture, microscopy and spiral plating.[19] The threshold for culture varied between studies but was often reported as ≥ 10^3 colony-forming unit (CFU), ≥ 10^4 CFU or ≥ 10^5 CFU (see *Appendix 3*).

Risk of bias

Table 10 presents an overview of the risk-of-bias assessment results for the studies included for objective 2; full details are reported in *Appendix 3*. Four studies were judged as being at high risk of bias. In three studies this was because a large proportion of patients had been excluded from the analysis,[46,48,50] and in the remaining study participant selection was unclear and multiple samples were taken from some patients.[44] As interpretation of culture involves some degree of subjectivity, it is important that those interpreting the culture results could not be influenced by knowledge of the POCT results. We considered culture to be an appropriate reference standard (i.e. studies were not judged at risk of bias for using culture), but there are limitations to culture as a reference standard; these are discussed in more detail in *Chapter 7*. Nine studies were judged as being at an unclear risk

TABLE 10 Overview of risk of bias in studies that evaluated the accuracy of POCTs

Study details	Patient selection	Index test	Reference standard	Flow and timing	Overall	Rationale for judgement
Van der Goes (2023)[51] **Test:** Lodestar DX	😊	?	?	😊	?	No information on blinding of interpreter of reference standard
Macias (2002)[44] **Test:** Uriscreen	☹	😊	?	😊	☹	Multiple samples taken from some patients; unclear how patients selected for inclusion
Millar (2000)[41] **Test:** Uriscreen	😊	😊	?	😊	?	No information on blinding of interpreter of reference standard
Teppa (2005)[42] **Test:** Uriscreen	😊	😊	?	😊	?	No information on blinding of interpreter of reference standard
Boon (2022)[a,43,53] **Test:** UTRiPLEX and Uriscreen	😊	😊	😊	😊	😊	No concerns. There was a high amount of exclusion in the Uriscreen vs. culture comparison, but this was due to late introduction of the test
Blom (2002)[39] **Test:** Flexicult Human	?	😊	?	😊	?	No information on blinding of interpreter of reference standard
Bongard (2015)[19] **Test:** Flexicult Human	?	😊	?	😊	?	Unclear if consecutive patients were enrolled. No information on blinding of interpreter of reference standard
Hullegie (2017)[36] **Test:** Flexicult Human	😊	😊	?	😊	?	No information on blinding of interpreter of reference standard
Holm (2017)[a,37] **Test:** Flexicult Human and ID Flexicult	😊	😊	😊	😊	😊	No concerns
Pernille (2019)[40,54] **Test:** ID Flexicult	😊	😊	?	😊	?	No information on blinding of interpreter of reference standard
Colodner (2000)[49] **Test:** Dipstreak	?	😊	?	😊	?	Unclear if consecutive patients were enrolled. No information on blinding of interpreter of reference standard
Yagupsky (2000)[50] **Test:** Dipstreak	?	😊	?	☹	☹	High proportion of patients excluded from analysis
Mignini (2009)[48] **Test:** Uricult	😊	😊	?	☹	☹	High proportion of patients excluded from analysis
Anacleto (2009)[45] **Test:** Uricult Trio	😊	😊	😊	😊	😊	No concerns
Greeff (2002)[46] **Test:** Uricult Trio	😊	😊	?	☹	☹	High proportion of patients excluded from analysis
Lee (2010)[47] **Test:** Uricult Trio	?	😊	?	😊	?	Unclear if consecutive patients were enrolled. No information on blinding of interpreter of reference standard

a QUADAS-C assessments were also conducted for these studies for UTRiPLEX and Uriscreen (Boon 2022) and Flexicult Human and ID Flexicult (Holm 2017). All domains were still rated as low risk.

of bias.[19,36,39–42,47,49,51] The main reason for this was lack of information on blinding of interpreter of the reference standard. Three of these studies had additional concerns outlined in *Table 10*.[19,47,49] Three studies were judged at low risk of bias.[37,43,45] Two of these reported data on test comparisons[31,36] and therefore QUADAS-C assessments were also completed. All domains on QUADAS-C were judged at low risk of bias.

Results

Figure 3 shows paired forest plots of estimates of sensitivity and specificity for the detection of presence of UTI together with 95% CIs, stratified by test. Summary estimates for tests evaluated in at least two studies are shown as diamonds on the plot. Results for each test are discussed below. Where evaluated, data are also presented for the detection of the pathogenic cause of the infection and for the accuracy of the test in detecting antimicrobial sensitivity. *Table 11* provides a summary of whether data were available on diagnosis of UTI, pathogenic cause and antimicrobial sensitivity for each test. Full accuracy results are presented in *Appendix 3*.

Lodestar DX

One study, funded by the test manufacturer, evaluated Lodestar DX.[51] The study was laboratory-based and evaluated the accuracy of Lodestar DX for detecting specific pathogens in fresh urine samples. It did not report the urine sampling method used and was judged as being at unclear risk of bias (see *Table 9*).

Pathogenic cause

Lodestar DX (n = 1) had good sensitivity 86% (95% CI 74% to 99%) and specificity 88% (95% CI 83% to 94%) for detecting *E. coli* in urine samples.

Uriscreen

Four studies evaluated Uriscreen.[41–44] One study analysed 156 children aged < 18 years in primary care in Belgium and conducted the POCT in the laboratory.[43] Three other studies conducted the POCT in a near-patient setting and analysed 378 pregnant women from antenatal clinics in Hawaii,[41] 150 pregnant women from antenatal clinics in Venezuela,[42] and 108 samples from 57 catheterised ICU patients in Mexico.[44] Two studies used catheterised urine samples,[42,44] one used mid-stream sampling[41] and one used mid-stream or adhesive bags.[43] One study was judged as being at low risk of bias,[43] two at unclear risk of bias[41,42] and one at high risk of bias[44] (see *Table 9*).

Presence of urinary tract infection

All four studies reported data on the accuracy of Uriscreen for detecting UTI, using the presence of foam to indicate the presence of UTI. Estimates of sensitivity ranged from 61% to 89% and specificity ranged from 43% to 89%. Summary sensitivity was 74% (95% CI 59% to 84%) and summary specificity was 64% (95% CI 41% to 82%). There were no clear reasons for the observed heterogeneity.

UTRiPLEX

One study evaluated UTRiPLEX.[43] The study analysed 292 children aged < 18 years in primary care in Belgium, although the test was conducted in the laboratory. The study collected urine samples using mid-stream sampling or adhesive bags, as per clinical practice. It was judged at low risk of bias (see *Table 9*).

Presence of urinary tract infection

Using the visualisation of ≥ 2 test lines after 6 minutes as the threshold, sensitivity was low (21%) but specificity was high (94%).

Flexicult Human

Four studies evaluated Flexicult Human.[19,36,37,39] This included test accuracy substudies from the two trials included for objective 1.[36,37] These two studies and one additional study were conducted in primary care settings in Denmark, Wales, and Wales, England, Spain and the Netherlands. The two test accuracy substudies from trials were restricted to women (aged > 18 years) with uncomplicated UTI; one of these analysed 183 women,[37] and one analysed 289 women.[36] One study analysed 121 samples from a mixed population of symptomatic patients in Denmark,[39] and one study was laboratory-based and used 200 fresh urine samples from a mixed population in Wales.[19] Mid-stream urine samples were collected in the two trial substudies.[36,37] The laboratory-based study collected samples using different methods, mid-stream sampling (n = 134) and catheter sampling (n = 7), and for 65 samples the method was unknown.

Test	Population	Setting	Location	Study	TP	FN	TN	FP	Sensitivity (95% CI)	Specificity (95% CI)
Uriscreen	Pregnant (screening)	Antenatal	Near patient	Millar 2000[41]	30	13	150	185	0.70 (0.55 to 0.81)	0.45 (0.40 to 0.50)
	Pregnant (screening)	Antenatal	Near patient	Teppa 2005[42]	17	11	109	13	0.61 (0.42 to 0.76)	0.89 (0.83 to 0.94)
	Catheterised	ICU	Near patient	Macias 2002[44]	55	7	20	26	0.89 (0.78 to 0.94)	0.43 (0.30 to 0.58)
	Children	Primary care	Laboratory	Boon 2022[43]	10	5	97	44	0.67 (0.42 to 0.85)	0.69 (0.61 to 0.76)
									0.74 (0.59 to 0.84)	**0.64 (0.41 to 0.82)**
UTRiPLEX IFU	Children	Primary care	Laboratory	Boon 2022[43]	6	23	248	15	0.21 (0.098 to 0.38)	0.94 (0.91 to 0.97)
Flexicult Human	Women	Primary care	Near patient	Hullegie 2017[36]	140	50	37	62	0.74 (0.67 to 0.79)	0.37 (0.28 to 0.47)
	Women	Primary care	Near patient	Holm 2017[35]	111	18	29	25	0.86 (0.79 to 0.91)	0.54 (0.41 to 0.66)
	Mixed	Primary care	Near patient	Blom 2002[39]	58	17	43	3	0.77 (0.67 to 0.85)	0.93 (0.82 to 0.98)
									0.79 (0.72 to 0.85)	**0.67 (0.30 to 0.90)**
	Mixed	Laboratory	Laboratory	Bongard 2015[19]	50	4	130	16	0.93 (0.82 to 0.97)	0.89 (0.83 to 0.93)
ID Flexicult	Women	Primary care	Near patient	Holm 2017[35]	104	12	24	18	0.90 (0.83 to 0.94)	0.57 (0.42 to 0.71)
	Women	Primary care	Near patient	Pernille 2019[40]	46	6	52	13	0.88 (0.77 to 0.95)	0.80 (0.69 to 0.88)
									0.89 (0.84 to 0.93)	**0.70 (0.52 to 0.84)**
Dipstreak	Mixed	Laboratory	Laboratory	Colodner 2000[49]	167	2	641	8	0.99 (0.96 to 1.00)	0.99 (0.98 to 0.99)
	Mixed	Laboratory	Laboratory	Yagupsky 2000[50]	270	12	509	4	0.96 (0.93 to 0.98)	0.99 (0.98 to 1.00)
									0.97 (0.94 to 0.99)	**0.99 (0.98 to 0.99)**
Uricult	Pregnant (screening)	Antenatal	Laboratory	Mignini 2009[48]	321	8	1836	8	0.98 (0.95 to 0.99)	1.00 (0.99 to 1.00)
Uricult Trio	Pregnant (screening)	Antenatal	Near patient	Greeff 2002[46]	47	11	104	85	0.81 (0.69 to 0.89)	0.55 (0.48 to 0.62)
	Pregnant (symptomatic)	Antenatal	Near patient	Greeff 2002[46]	29	8	44	46	0.78 (0.63 to 0.89)	0.49 (0.39 to 0.59)
	Children	Outpatient	Near patient	Anacleto 2009[45]	70	33	80	17	0.68 (0.58 to 0.76)	0.82 (0.74 to 0.89)
	Children (< 24 months)	Outpatient	Near patient	Lee 2010[47]	19	13	101	18	0.59 (0.42 to 0.74)	0.85 (0.77 to 0.90)
									0.73 (0.63 to 0.82)	**0.70 (0.52 to 0.84)**

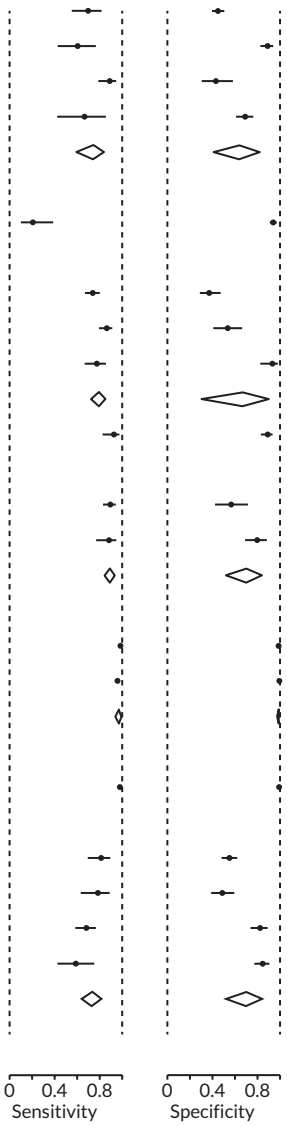

FIGURE 3 Paired forest plots of individual study estimates and summary estimates of sensitivity and specificity for the detection of presence of UTI together with 95% CIs, stratified by test. FN, false negative; FP, false positive; TN, true negative; TP, true positive.

TABLE 11 Summary of whether data were available on diagnosis of UTI, pathogenic cause and antimicrobial sensitivity for each test

Test name	Presence of UTI	Pathogenic cause	Antimicrobial sensitivity
Rapid tests			
Lodestar DX	×	✓	×
Uriscreen	✓	×	×
UTRiPLEX	✓	×	×
Culture-based tests			
Dipstreak	✓	✓	×
Flexicult Human	✓	×	✓
ID Flexicult	✓	×	×
Uricult trio	✓	✓	×
Uricult	✓	×	×

One study did not report how urine samples were collected.[39] Three of the studies were judged to be at unclear risk of bias[19,36,39] and one was judged to be at low risk of bias[37] (see *Table 9*).

Presence of urinary tract infection

All studies provided data on the accuracy of the Flexicult Human test for diagnosing UTI. Three used culture alone as the reference standard.[36,37,39] One study used two reference standards: (1) culture and microscopy and (2) culture, microscopy and spiral plating.[19] Another study used three different reference standard definitions to define a UTI: ≥ 10^4 CFU/ml pure culture of pathogen; ≥ 10^5 CFU/ml mixed growth with one predominant pathogen; *or* ≥ 10^3 CFU/ml of *E. coli* or *S. saprophyticus* (Public Health England/Health Protection Agency), ≥ 10^5 CFU/ml pure culture of uropathogen *or* ≥ 10^5 CFU/ml predominant culture a uropathogen with 3-log difference between the highest and next species (UK laboratory definition) and ≥ 10^3 CFU of uropathogen (European definition).

The Flexicult Human thresholds for defining the presence of UTI varied. Two studies used ≥ 10^3 CFU/ml,[19,37] one used ≥ 10^4 CFU/ml,[39] and one used 10^3 CFU/ml for pure culture of a pathogen and ≥ 10^3 CFU/ml for predominant growth of a pathogen in mixture with normal flora.[36]

Estimates of sensitivity ranged from 74% to 93% and of specificity ranged from 37% to 93%. Estimates were highest in the laboratory-based study of mixed urine samples (93% and 89%).[39] This study used a compound reference standard of culture, microscopy and spiral plating. Estimates were lower when the study used culture and microscopy as the reference standard (87% and 83%) and more similar to the reference standard used in the other studies. The summary estimates of sensitivity and specificity across all three studies in which the Flexicult Human test was conducted in primary care were 79% (95% CI 72% to 85%) and 67% (95% CI 30% to 90%).

Antimicrobial sensitivity

Three studies reported data for antimicrobial sensitivity.[19,36,39] Estimates of sensitivity ranged from 79% to 90% with a summary estimate of 87% (95% CI 83% to 90%). Estimates of specificity ranged from 72% to 94% with a summary estimate of 93% (95% CI 89% to 95%).[19,36,39]

ID Flexicult

Two studies evaluated ID Flexicult.[37,40] Both studies conducted the ID Flexicult test in primary care in Denmark and recruited women with uncomplicated UTI and used mid-stream urine samples. One study analysed 158 people[37] and the other analysed 117. One of these studies also evaluated Flexicult

Human; this was the accuracy study nested within the trial that compared testing and treatment based on Flexicult Human with testing and treatment based on ID Flexicult. One study was judged as being at low risk of bias[37] and one had unclear risk of bias[40] (see *Table 9*).

Presence of urinary tract infection

The test had good sensitivity (90% and 88%), but estimates of specificity were lower, at 56% and 80%. Summary sensitivity was 89% (95% CI 84% to 93%) and summary specificity was 70% (95% CI 52% to 84%). The studies used thresholds of 10^3 CFU/ml (primary pathogens) and 10^4 CFU/ml (secondary pathogens) for the POCT.

Dipstreak

Two studies evaluated Dipstreak.[49,50] Both were conducted in Israel and were laboratory-based studies that tested fresh urine samples from mixed populations. One study analysed 795 mid-stream urine samples;[50] the other analysed 818 samples (urine sampling method not reported). One study was judged at high risk of bias[50] and one was judged at unclear risk of bias[49] (see *Table 9*).

Presence of urinary tract infection

Both studies found Dipstreak to be highly accurate for detecting UTI. Sensitivity was estimated at 96% and 99%; both studies estimated specificity at 99%. Summary sensitivity was 95% (95% CI 94% to 99%) and summary specificity was 99% (95% CI 98% to 99%). One of these studies evaluated two Dipstreak thresholds (10^4 and 10^5 CFU/ml)[49] and found similar results; the other did not report the Dipstreak threshold.[50]

Pathogenic cause of urinary tract infection

Yagupsky *et al.*[50] reported that Dipstreak correctly identified the pathogenic cause of UTI in 211 out of 270 cases (the other 59 were not identified).

Uricult

One study evaluated Uricult.[48] It analysed mid-stream urine samples from 2173 pregnant women from antenatal clinics in Argentina, and performed the test in the laboratory. It was judged at high risk of bias (see *Table 9*).

Presence of urinary tract infection

The study reported very high estimates of sensitivity (98%) and specificity (100%) for Uricult for the detection of the presence of UTI, using a threshold of > 10^5 CFU.

Uricult Trio

Three studies evaluated Uricult Trio.[45–47] Populations varied: one analysed 374 pregnant women in antenatal clinics in South Africa,[46] one analysed 151 infants aged < 24 months from outpatient clinics in Republic of Korea[47] and one analysed 200 children < 16 years from outpatient clinics in the Philippines.[45] The study in pregnant women stratified results according to whether women were symptomatic (n = 127) or asymptomatic (n = 247). All studies used mid-stream urine samples; one also used urine collection bags in infants and another used catheterisation where clean catch was difficult. One study was judged at high risk of bias,[46] one was judged at unclear risk of bias,[47] and one was judged at low risk of bias[45] (see *Table 9*).

Presence of urinary tract infection

Estimates of sensitivity ranged from 59% to 78% and of specificity ranged from 49% to 85%. Summary sensitivity was 73% (95% CI 63% to 82%) and summary specificity was 70% (95% CI 52% to 84%).

Pathogenic cause

One study reported that for detecting the presence of *E. coli* infection, the sensitivity of Uricult Trio was 60% and the specificity was 96%.

Test comparisons

Two studies reported data on two POCTs included in the scope.[37,43] One evaluated both Flexicult Human and ID Flexicult. The other evaluated Uriscreen and UTRiPLEX. Both studies were set in general practice, assessed the accuracy of POCTs for the detection of UTI, and used culture as the reference standard. Both studies were judged to be at low risk of bias when assessed with QUADAS-C.

An accuracy study, nested within a trial, evaluated Flexicult Human and ID Flexicult.[37] The study recruited 341 women in Denmark who had uncomplicated UTI. Patients were randomised to be tested with Flexicult Human or with ID Flexicult. The study reported similar sensitivity and specificity with Flexicult Human (86% and 54%) and ID Flexicult (90% and 56%).

A prospective cross-sectional study evaluated the Uriscreen test and the UTRiPLEX test in children aged < 18 years in Belgium.[43] Three hundred samples were taken systematically and tested. However, far fewer results (156 vs. 292) were available for Uriscreen test than for the UTRiPLEX test because the former was introduced later in the trial, making a comparison of the tests difficult. Sensitivity and specificity were reported as 67% and 69% for Uriscreen and as 21% and 94% for UTRiPLEX.

We are unable to draw comparisons between the tests in other studies due to heterogeneity of population.

Comparison with standard urine dipstick tests

Six studies provided a direct comparison between the POCTs and standard urine dipstick testing for LE or nitrite.[40,41,43,47,48] Four of these defined a positive dipstick test as being positive for either LE or nitrite, one as being positive for both LE and nitrite, and one reported data separately for nitrite and LE dipstick tests. Three studies compared Uriscreen with standard dipstick testing and reported different findings, which may be related to how a positive dipstick test was defined (see *Table 12*). One study also evaluated UTRiPLEX which was found to be less sensitive but more specific than dipstick testing. Three studies compared culture-based POCTs with standard dipstick testing. All found that the POCTs were more sensitive and more specific than standard dipstick tests.

Objective 3: what is the technical performance (other than accuracy) of point-of-care tests for urinary tract infection?

Five publications reported data on technical performance. Three reported data for Flexicult Human[8,55] (two of these reported on the POETIC trial[8,55]) and two reported data for Uricult Trio.[45,46] Of these, one publication was also included for objective 1[8] and three were included for objective 2.[56] A further study[56] appeared relevant to objective 3; however, it was excluded because it was only reported in a trial registry with no data and the trial author did not reply to a request for information. Results are provided in *Appendix 3*. There were no data for the following outcomes prespecified in our protocol: test failure rate; UTI-associated healthcare resources; health-related quality of life; and clinical outcomes, for example morbidity/mortality.

Flexicult Human

The Butler *et al.* trial that compared testing and treating based on the results of Flexicult Human with no treatment reported additional technical performance data on the Flexicult Human test.[8] These data are summarised in *Table 13*. They found that in 63% of participants the management was changed as a result of the test. Estimates of time related to performing the test were 9 minutes to prepare the test, 6 minutes to obtain and record results and 7 minutes to discuss the results with patients. This is in addition to the time that the test takes to perform, which was not reported. The total cost of the intervention, including the cost of the test itself, was estimated at £48.

In a qualitative substudy of the POETIC trial, 35 clinicians were interviewed who used the Flexicult Human test.[55] The study found that 'clinicians overwhelmingly felt that a POCT for UTI management

TABLE 12 Estimates of sensitivity and specificity for standard dipstick tests and POCTs from studies that evaluated both tests

Study	Population	Test	Sensitivity (95% CI)	Specificity (95% CI)
Rapid tests				
Boon (2022)[43]	Children aged < 18 years	UTRiPLEX	21 (8 to 40)	94 (91 to 97)
		Uriscreen	67 (38 to 88)	69 (60 to 76)
		Dipstick (either nitrite or LE positive considered positive)	32 (16 to 52)	86 (82 to 90)
Macias (2002)[44]	Catheterised ICU patients	Uriscreen	66.7	74.1
		Dipstick – nitrite only	66.7	45.2
		Dipstick – LE only	78.9	47.2
Millar (2000)[41]	Pregnant women (screening)	Uriscreen	70 (57 to 84)	45 (40 to 51)
		Dipstick (both nitrite and LE positive considered positive)	81 (69 to 93)	97 (95 to 99.2)
Culture-based tests				
Pernille (2019)[40]	Women – uncomplicated UTI	ID Flexicult	88 (80 to 97)	80 (70 to 90)
		Dipstick (either nitrite or LE positive considered positive)	73 (59 to 84)	75 (63 to 85)
Mignini (2009)[48]	Pregnant women (screening)	Uricult	98 (96 to 99)	99.6 (99.3 to 99.8)
		Dipstick (either nitrite or LE positive considered positive)	53 (48 to 58)	92 (91 to 93)
Lee (2010)[47]	Children aged < 24 months	Uricult Trio	59%	85%
		Dipstick (either nitrite or LE positive considered positive)	50%	76.7%

TABLE 13 Technical performance of the Flexicult Human test

Outcome	Category	Results
Management change as result of Flexicult Human	Overall	63.1%
	Did not start antibiotic	7.4%
	Stopped taking antibiotic	5.3%
	Started taking antibiotic	15.3%
	Continued with antibiotic	33.2%
	New antibiotic prescribed	38.9%
Time to perform test	Prepare test	9 minutes
	Obtain and record result	6 minutes
	Discuss result with patient	7 minutes
Cost	Cost per person, including POCT cost in UK	£48

would be useful'. It reported that most clinicians agreed that the Flexicult Human test gave quicker results than laboratory tests (24 hours vs. 3–4 days), reassured patients and had a positive impact on clinician confidence in diagnosing UTI. There was an even split between those who thought it would have no impact on prescribing and those who stated that it had increased their awareness about antibiotic prescribing and, therefore, they were more cautious about prescribing. However, they noted difficulties in interpreting test results, limitations in when the test can be used, limited resources to undertake testing and concerns about prolonging a patient's discomfort while waiting for test results and about the potential expense of maintaining a regular stock of tests. They highlighted that an ideal POCT for UTI would give fast results; ease of use, accuracy and reliability were mentioned much less.

A further study conducted in primary care reported that GPs considered Flexicult Human to be easy to handle and read.[39]

Uricult Trio

One study reported that Uricult Trio was convenient to use and easy to interpret.[45] Another study[46] agreed that results could be obtained quicker and easier with Uricult Trio than with a laboratory test and stated that this would impact the cost of hospitalisation. It reported fewer lost specimens with Uricult Trio than with laboratory tests that require transportation (0 vs. 79 lost). However, it also reported that 'the Uricult Trio did not add anything in terms of managing the patient more efficiently' and said it 'is not useful for screening asymptomatic bacteriuria or for diagnosing UTIs in women with symptoms suggestive of an infection'.

Chapter 6 Objectives 4 and 5: assessment of cost-effectiveness

In this chapter, we describe the methods and findings of our assessment of cost-effectiveness of POCTs for UTI to reduce antimicrobial resistance. This comprises a conceptual model for POCTs in UTI and summary of identified evidence, and a potential implementation of the conceptual model using the available evidence. The implemented model is described in *Evaluating costs, quality of life and cost-effectiveness* and was coded in the R programming language.[57] Results of the implemented model are not presented as the evidence was too limited for findings to be meaningful.

Conceptual modelling of costs, quality of life and cost-effectiveness

A decision-analytic model was conceptualised to estimate the incremental costs and quality-adjusted life-years (QALYs) for POCTs for UTI in comparison with culture with or without dipstick tests. The model described below is for all possible comparators and populations/subgroups described in *Population*. Separate models would be required for each population/subgroup.

In *Review of evidence on cost-effectiveness*, we review the clinical evidence identified in *Chapter 5*, and evidence identified by pragmatic searches, to narrow the focus on tests and populations where evidence and impact are greatest.

Testing strategies

The POCTs considered were those included in the scope outlined in *Table 2*. These include rapid tests (results < 40 minutes) that perform AST (e.g. Astrego PA-100), rapid tests that only identify pathogenic cause (e.g. Lodestar DX), culture-based tests (results up to 24 hours) that perform AST (e.g. Flexicult), and culture-based tests that only identify pathogenic cause (e.g. Dipstreak).

As described in *Comparator*, the comparator was diagnosis based on clinical features plus dipstick tests with laboratory culture-based confirmation (in population 1) or diagnosis based on clinical features plus laboratory culture-based without dipstick test (in population 2).

In the case of this comparator, where results can take several days, and culture-based tests where results take up to 24 hours, it was assumed that some patients would be prescribed and begin antibiotics without knowing whether they had a UTI, pathogenic cause, or antimicrobial sensitivity status.

Subgroups of interest

As per *Population*, the population in scope is those with suspected UTI, but subgroups of interest include the following.

Patient subgroups identified by Public Health England guidance:

A. women aged < 65 years
B. women aged > 65 years
C. men aged < 65 years
D. men aged > 65 years
E. adults with indwelling urinary catheters
F. babies, children and young people aged < 16 years

Other patient subgroups:

G. people with suspected acute UTI
H. people with suspected recurrent UTI
I. people with suspected chronic UTI
J. children aged < 3 months
K. pregnant women
L. people who are frail or have dementia
M. people who are pre-, peri- or post-menopausal
N. people on prophylactic antibiotics for treatment of UTI
O. people of different ethnicities
P. people with a higher risk of complicated UTI (e.g. people with neurogenic bladder, diabetes, polycystic kidney disease or people who are immunocompromised)
Q. people with suspected pyelonephritis.

Conceptual model

Our conceptual model is illustrated in *Figure 4*. Arrows indicate the influence of components on the rest of the model.

Our conceptualisation was divided into short-term and long-term components. In the short term, the important elements to consider were the symptoms of complicated and uncomplicated UTI, characteristics and consequences of antibiotics, expected efficacy of antibiotics, and any response to ineffectiveness of antibiotics. In the long term, the model links to a generic model for UTI and covers the key complications of sepsis, pyelonephritis and kidney failure. Furthermore, the development or continuation of chronic or recurrent UTI was considered, and it was recognised that this would be particularly common in patients with risk factors such as catheters.

Costs were assumed to be from an NHS and PSS perspective and include all elements from the short-term or long-term components. The tests to compare are those detailed in *Table 2*, as described in *Testing strategies*.

Our conceptual model reflects the influence on the costs, health outcomes and model structures of the choice of populations and subgroups. UTIs themselves are categorised in acute, recurrent and chronic. Furthermore, UTIs divide into those that are uncomplicated and complicated at GP presentation, while our model reflects that patients with either uncomplicated or complicated UTI can still suffer complicated UTI at the end of testing and treatment.

Rates of complicated UTI, and the costs and health outcomes of the model, also depend on the subgroup under investigation. We conceptualised these to be broad and include the subgroups identified in *Population*.

Review of evidence on cost-effectiveness

In this section, we review the relevant evidence on cost-effectiveness identified by the clinical effectiveness review and separate pragmatic literature searches. We use this as a basis for narrowing the tests and subpopulations to only those that are feasible for modelling.

Relevant evidence from clinical effectiveness review

The search for the clinical effectiveness review (see *Chapter 4*) was not limited by study design or publication type search filters and therefore identified economic evidence. The process of study identification and selection is summarised in *Figure 2*. We identified two relevant studies from this, discussed below.[8,35]

Butler 2018 (POETIC)

The Butler *et al.*[8] POETIC study, described in *Chapter 5*, was an RCT that assessed the clinical effectiveness and cost-effectiveness of Flexicult Human compared with standard care in adult women with a clinical diagnosis of uncomplicated UTI. Cost-effectiveness was measured by total cost per unit increase in concordant antibiotic prescribing, and, on this basis, Flexicult testing was not cost-effective. The study found that clinicians generally prescribed broad/empiric antibiotics rather than waiting for the Flexicult results, and they seldom withdrew antibiotic treatment in response to test results (see *Table 13*). In both treatment arms, the duration of all UTI symptoms was reported as 8 (range 5–14) days, and the duration of moderately bad symptoms was 4 (range 2–6) days.

Holm 2017

The Holm *et al.*[35] study, discussed under objective 1 in *Objective 1: what is the impact on clinical outcomes of using POCT to diagnose UTI, with or without additional pathogen identification and AST?*, was a RCT comparing Flexicult Human and ID Flexicult in women with suspected uncomplicated UTI. The primary outcome was appropriate antibiotic prescribing, as described in *Objective 1: what is the impact on clinical outcomes of using POCT to diagnose UTI, with or without additional pathogen identification and AST?*. The study found that including POCT AST did not improve antibiotic prescribing in general practice. As summarised in *Table 8*, the study reported results on appropriate prescribing and on patient enablement (measured using the Patient Enablement Instrument at day 14 and 3 months). However, this cannot be used for modelling because neither outcome matches sufficiently to any outcome in the conceptual economic model.

Additional pragmatic searches for cost-effectiveness evidence

We conducted pragmatic searches of MEDLINE (via Ovid), EMBASE (via Ovid) and EconLit (via EBSCO*host*) databases using search terms listed in *Table 14*. There were 24 studies identified after removal of duplicates. Thirteen were identified at title/abstract screening as potentially having useful information, although two of these were conference abstracts related to two full-text records. Two were studies related to the POETIC trial that had already been identified in the clinical effectiveness review. One study was a potentially relevant cost-effectiveness evaluation of trimethoprim-sulfamethoxazole and amoxicillin in UTI, but it was inaccessible and published in 1987, and so it was not considered further.[58] The remaining eight records were evaluated at full-text stage.

Wang 2021

Wang *et al.*[59] reported a US-based decision tree model that considered both empiric antibiotics and culture-directed antibiotics; the latter aligns with our treatment strategy of targeted antibiotics. The focus of their analysis was the impact of antibiotic resistance on cost-effectiveness of treatment strategies. The authors found that empiric antibiotics were the most cost-effective strategy if resistance was < 6%, while symptomatic treatment was most cost-effective if resistance was > 80%. However, at most levels of resistance, the study found that empiric antibiotics, with simultaneous urine culture and later targeting of antibiotics, was the most cost-effective strategy. This aligns with our assumed standard of care: laboratory culture-based testing with empiric/broad antibiotics. This study reported quality-adjusted life-days for UTI cured, UTI, and pyelonephritis, presented in *Table 15*.

TABLE 14 Details and results of the additional pragmatic searches for cost-effectiveness evidence

Database (date range)	Search term	Results
Ovid MEDLINE 1946 to present	("urinary tract infection" and "cost-effectiveness").ti.	20
EMBASE 1974 to present	("urinary tract infection" and "cost-effectiveness").ti.	24[a]
EconLit	("urinary tract infection" and "cost-effectiveness").ti.	0

a These hits included all studies identified by Ovid MEDLINE.

TABLE 15 Summary of input parameters that could be used in the cost-effectiveness model

Input	Name used for code/ equations	Value(s), random distribution	Source of value(s)	Comments
Probability of having a (true) UTI	p_uti	*0.6* Beta (α = 2.212762, β = 1.475174)	Wang and LaSala 2021,[60] Schmiemann *et al.* 2010[61]	p_uti is different for each patient subgroup Diagnosis of UTI given symptoms was 0.6 (0–1) Wang and LaSala 2021[60] and Schmiemann *et al.* 2010[61]
Probability of correctly detecting a UTI (sensitivity or true-positive rate)	p_uti_tp	See *Table 19*	See *Table 19*	p_uti_tp is different for each test
Probability of incorrectly diagnosing a non-UTI patient as having UTI and then giving them antibiotics (false-positive rate)	p_uti_fp	See *Table 19*	See *Table 19*	p_uti_fp is different for each test
Probability of identifying specific antibiotic for targeted treatment, given that a UTI was detected using POCT with AST	p_targ	See *Table 19*	See *Table 19*	p_targ is different for each POCT with AST test
Probability of becoming 'healthy' on targeted treatment, i.e. had a true UTI but are now healthy with no complications from UTI, but at risk of recurrence	p_healthy_targ			Estimate probabilities of complications first, then calculate p_healthy_targ = 1 – p_sepsis_targ – p_kidney_failure_targ – p_pyelonephritis
Probability of sepsis on targeted treatment	p_sepsis_targ	No data	No data	No data
Probability of kidney failure on targeted treatment	p_kidney_ failure_targ	No data	No data	No data
Probability of pyelonephritis on targeted treatment	p_ pyelonephritis_ targ	No data	No data	Probability of pyelonephritis in treated pregnant women identified from Smaill and Vazquez 2015[62] and NICE NG109,[15] but this did not distinguish between targeted and empiric treatment
Probability of becoming 'healthy' on empiric treatment, i.e. had a true UTI but are now healthy with no complications from UTI, but at risk of recurrence	p_healthy_emp	Women: 61.8% (complete resolution NICE NG109, Falagas 2009) Mixed: assumed same as in women		p_healthy_emp is different for each test as some non-AST tests can still detect bacteria Estimate probabilities of complications first, then calculate p_healthy_emp = 1 – p_sepsis_emp – p_kidney_failure_emp – p_pyelonephritis_emp Older people: 61% (bacteriological cure NICE NG109, Zalmanovici-Trestioreanue 2015)
Probability of sepsis on empiric treatment	p_sepsis_emp	No data	No data	No data
Probability of kidney failure on empiric treatment	p_kidney_ failure_emp	No data	No data	No data

TABLE 15 Summary of input parameters that could be used in the cost-effectiveness model (*continued*)

Input	Name used for code/ equations	Value(s), random distribution	Source of value(s)	Comments
Probability of pyelonephritis on empiric treatment	p_pyelonephritis_emp	Women: 5.6% (NICE NG109,[15] Smaill and Vazquez 2015[62]) Mixed: assume same as in women		(0–0.02) in Wang and LaSala 2021,[60] Ferry *et al.* 2007,[63] Christiaens *et al.* 2002[64] 0.04 in Sadler *et al.* 2017,[65] for risk of pyelonephritis if clinical cure not achieved, Little *et al.* 2009[66] We use Smaill and Vazquez 2015[62] and NICE NG109[15] as divided into treated and untreated although it relates to pregnant women and does not distinguish between targeted and empiric treatment
Probability of becoming 'healthy' on 'no treatment', i.e. had a true UTI but are now healthy with no complications from UTI, but at risk of recurrence	p_healthy_no_treatment	Non-pregnant women: 25.7% (complete resolution NICE NG109, Falagas 2009) Mixed: assume average of non-pregnant women and older people: 21.35%		p_healthy_no_treatment is different for each of the three types of test as for culture testing patients may be given antibiotics while awaiting test results Estimate probabilities of complications first, then calculate p_healthy_no_treatment = 1 – p_sepsis_no_treatment – p_kidney_failure_no_treatment – p_pyelonephritis_no_treatment Older people: 17% (bacteriological cure NICE NG109, Zalmanovici-Trestioreanue 2015)
Probability of sepsis on 'no treatment'	p_sepsis_no_treatment	No data	No data	No data
Probability of kidney failure on 'no treatment'	p_kidney_failure_no_treatment	No data	No data	No data
Probability of pyelonephritis on 'no treatment'	p_pyelonephritis_no_treatment	Women: 66.3%		This was for pregnant women (NICE NG109,[15] Smaill and Vazquez 2015[62])
Probability of needing more than one course of antibiotics	p_multiple_courses	No data	No data	No data
Proportion of patients who are given antibiotics despite test not detecting a UTI	prop_emp_when_no_detected_uti	No data	No data	No data
Probability of side effects on antibiotics	p_side_effects_antibiotics	10% (5–30%) Log-normal (meanlog = −2.303, sdlog = 0.457)		Used in Fenwick et al. 2000[59] but from Norrby 1990[67]
Duration side effects from antibiotics		3 days (2–4 days) Normal (mean 3, SD 0.5)		Used in Fenwick *et al.* 2000[59] but from Carlson and Mulley 1985[68]

continued

TABLE 15 Summary of input parameters that could be used in the cost-effectiveness model (*continued*)

Input	Name used for code/ equations	Value(s), random distribution	Source of value(s)	Comments
Overall cost of test	cost_test	(Confidential information has been removed) Flexicult: £48 ID Flexicult: unavailable use (confidential information has been removed)	Flexicult: Butler *et al.* 2018[8] Lodestar: manufacturer's submission	This should include the actual cost of the test from the manufacturer and the cost of processing the test, as different tests take different lengths of time and therefore may need more laboratory time and a follow-up appointment/ attention to prescribe chosen antibiotic Flexicult is total cost per person of the intervention, including the cost of the POCT and in text authors say nearly 90% (£43.90) are distribution cost (Confidential information has been removed). We add £43.90 (from Flexicult) distribution costs to Lodestar costs Manufacturers did not provide prices for ID Flexicult. We assume the highest cost estimated for Lodestar
Cost of follow-up appointment/attention if required to prescribe chosen antibiotic at a later point due to length of wait for results	cost_fol-lowup_appt	£42	*Unit Costs of Health and Social Care 2022* (PSSRU and Centre for Health Economics)	GP appointment cost is £42, including direct care staff costs (nurses)
Overall cost per course of antibiotics	mapped_treat-ment_costs	See *Table 16*	NICE guidelines and BNF	This is different for each antibiotic and also varies with dosage and course length according to patient group
Cost of treating sepsis	cost_sepsis	No data	No data	This is likely to be complex to calculate and will include costs of additional GP appointments and hospital admissions
Cost of treating kidney failure	cost_kidney_ failure	No data	No data	This is likely to be complex to calculate and will include costs of additional GP appointments and hospital admissions
Cost of treating pyelonephritis	cost_pyelone-phritis	£1221.26 (2022 price, inflated from the 2016 price of £986.40)	Sadler *et al.* 2017[65]	Hospitalisation cost of pyelonephritis (2016 price): £3992 Days of hospitalisation for pyelonephri-tis: 2 Outpatient visit cost of pyelonephritis (2016 price): £94 Risk of hospitalisation if pyelonephritis: 0.20
QALY loss from uncom-plicated UTI	qaly_loss_uti		Wang and LaSala 2021[60] and Bermingham and Ashe 2012[69]	0.68 (0.56–0.72) was QALDs for UTI
Additional QALY loss from sepsis in the short-term model	qaly_loss_sepsis	No data	No data	No data
Additional QALY loss from kidney failure in the short-term model	qaly_loss_ kidney_failure	No data	No data	No data

TABLE 15 Summary of input parameters that could be used in the cost-effectiveness model (*continued*)

Input	Name used for code/ equations	Value(s), random distribution	Source of value(s)	Comments
Additional QALY loss from pyelonephritis in the short-term model	qaly_loss_pyelonephritis		Wang and LaSala 2021[60] and Bermingham and Ashe 2012[69]	0.59 (0.48–0.64) QALDs for pyelonephritis Duration of treated pyelonephritic attack was 10 days and untreated was 14 days in Whiting *et al.* 2006[70] and Barry *et al.* 1997.[71] Decrements were 0.010225 and 0.014315 in treated and untreated, respectively
QALY loss from antibiotic AE	qaly_loss_antibiotic_ae	No data	No data	No data
Utility for healthy		0.82 (0.58, 0.92)	Wang and LaSala 2021[60] and Bermingham and Ashe 2012[69]	0.82 (0.58–0.92) QALDs for UTI cured

BNF, British National Formulary; PSSRU, Personal Social Services Research Unit; QALD, quality-adjusted life-day.

Sadler 2017

Sadler *et al.*[65] reported a UK-based decision tree economic model that compared the cost-effectiveness of four antibiotics (fosfomycin, nitrofurantoin, pivmecillinam and trimethoprim) for adult women with signs and symptoms of uncomplicated UTI in primary care. Results were stratified by resistance to trimethoprim. Trimethoprim was most cost-effective if resistance was < 35%, fosfomycin was most cost-effective if resistance was between 30% and 35%, and either fosfomycin or nitrofurantoin was most effective at > 35%.

Fenwick 2000

Fenwick *et al.*[59] used a decision tree model to compare the cost-effectiveness of management strategies for UTI. The model included branches for symptoms disappearing, symptoms persisting and antibiotics working. The authors found that empiric antibiotic treatment based on symptoms was largely cost-effective compared with no treatment, empiric using culture-based testing, and empiric using dipstick with/without culture-based testing. Antibiotics included NICE recommended amoxycillin, cefalexin, amoxicillin-clavulanic acid and trimethoprim, as well as the no-longer-recommended cephradine (see *Table 16*). We therefore used the probability and duration of side effects from this study (see *Table 15*).

Whiting 2006

Whiting *et al.*[70] reported a systematic review and economic model of effectiveness and cost-effectiveness of tests for the diagnosis and investigation of UTI in children. The review resulted in an algorithm for the diagnosis of UTI in children under the age of 5 years.

Only one prior economic evaluation was identified by the systematic review: a US-based cost-effectiveness decision tree model comparing diagnosis and management strategies for UTI in children aged 2 months to 2 years.[72] The model compared diagnostic strategies for children presenting with symptoms suggestive of UTI, with eight subgroups of age and gender considered. This used a decision tree using combinations of dipstick, microscopy and laboratory culture-based tests to diagnose patients with UTI and vesicoureteral reflux. A long-term model was used to model the consequences of pyelonephritis and the possibility and consequences of end-stage renal disease. At lower willingness-to-pay thresholds, treating all children without any prior diagnostic test was most cost-effective. At higher

TABLE 16 Assumptions and sources for costing courses of antibiotic treatment for UTI

Antibiotic name	Patient group it is recommended for	Empiric or targeted	Recommended dosage and course length for patient group	Source of recommendation	Unit cost from BNF	Cost per course of antibiotics (£)
Nitrofurantoin	Non-pregnant women aged ≥ 16 years with a lower UTI (and eGFR ≥ 45 ml/minute)	Empiric and targeted	100 mg modified-release twice a day for 3 days	'UTI (lower): antimicrobial prescribing' NICE guidelines May 2022 (www.nice.org.uk/guidance/ng109/resources/visual-summary-pdf-6544021069)	Macrobid 100 mg modified-release capsules: £9.50 per 14 capsules	4.07
Nitrofurantoin	Children aged ≥ 3 months with a lower UTI (and eGFR ≥ 45 ml/minute)	Empiric and targeted	3 months to 11 years, 750 µg/kg four times a day for 3 days; 12–15 years, 50 mg four times a day or 100 mg modified-release twice a day for 3 days	'UTI (lower): antimicrobial prescribing' NICE guidelines May 2022 (www.nice.org.uk/guidance/ng109/resources/visual-summary-pdf-6544021069)	Macrobid 100 mg modified-release capsules: £9.50 per 14 capsules Nitrofurantoin 50 mg tablets: £3.43 per 28 tablets Nitrofurantoin 50 mg capsules: £5.30 per 30 capsules	
Nitrofurantoin	Pregnant women aged ≥ 12 years with a lower UTI and (and eGFR ≥ 45 ml/minute)	Empiric and targeted	100 mg modified-release twice a day for 7 days	'UTI (lower): antimicrobial prescribing' NICE guidelines May 2022 (www.nice.org.uk/guidance/ng109/resources/visual-summary-pdf-6544021069)	Macrobid 100 mg modified-release capsules: £9.50 per 14 capsules	9.50
Nitrofurantoin	Men aged ≥ 16 years with a lower UTI and (and eGFR ≥ 45 ml/minute)	Empiric and targeted	100 mg modified-release twice a day for 7 days	'UTI (lower): antimicrobial prescribing' NICE guidelines May 2022 (www.nice.org.uk/guidance/ng109/resources/visual-summary-pdf-6544021069)	Macrobid 100 mg modified-release capsules: £9.50 per 14 capsules	9.50
Nitrofurantoin	Non-pregnant women and men aged ≥ 16 years with a catheter (and eGFR ≥ 45 ml/minute)	Empiric and targeted	100 mg modified-release twice a day for 7 days	'UTI (catheter): antimicrobial prescribing' NICE guidelines September 2019 (www.nice.org.uk/guidance/ng113/resources/visual-summary-pdf-6599495053)	Macrobid 100 mg modified-release capsules: £9.50 per 14 capsules	9.50

TABLE 16 Assumptions and sources for costing courses of antibiotic treatment for UTI (*continued*)

Antibiotic name	Patient group it is recommended for	Empiric or targeted	Recommended dosage and course length for patient group	Source of recommendation	Unit cost from BNF	Cost per course of antibiotics (£)
Cefalexin	Pregnant women aged ≥ 12 years with a catheter	Empiric	500 mg twice or three times a day for 7–10 days	'UTI (catheter): antimicrobial prescribing' NICE guidelines September 2019 (www.nice.org.uk/guidance/ng113/resources/visual-summary-pdf-6599495053)	Cefalexin 500 mg tablets: £2.70 per 21 tablets Cefalexin 500 mg capsules: £2.42 per 21 capsules	1.61–3.86
Cefalexin	Non-pregnant women and men aged ≥ 16 years with acute pyelonephritis	Empiric	500 mg twice or three times a day for 7–10 days	'Pyelonephritis (acute): antimicrobial prescribing' NICE guidelines September 2019 (www.nice.org.uk/guidance/ng111/resources/visual-summary-pdf-6544161037)	Cefalexin 500 mg tablets: £2.70 per 21 tablets Cefalexin 500 mg capsules: £2.42 per 21 capsules	1.61–3.86
Cefalexin	Pregnant women and men aged ≥ 12 years with acute pyelonephritis	Empiric	500 mg twice or three times a day for 7–10 days	'Pyelonephritis (acute): antimicrobial prescribing' NICE guidelines September 2019 (www.nice.org.uk/guidance/ng111/resources/visual-summary-pdf-6544161037)	Cefalexin 500 mg tablets: £2.70 per 21 talets Cefalexin 500 mg capsules: £2.42 per 21 capsules	1.61–3.86
Fosfomycin	Adults with acute uncomplicated lower UTI	Targeted	3 g per one dose (granules)	Dosage from BNF	Fosfomycin 3 g granules sachets: £4.86 per sachet	
Trimethoprim	Women aged ≥ 16 years with lower UTI	Targeted	200 mg twice daily for 3 days	Dosage from BNF	Trimethoprim 200 mg tablets: £1.76 per 14 tablets	0.75
Trimethoprim	Men aged ≥ 16 years with lower UTI	Targeted	200 mg twice daily for 7 days	Dosage from BNF	Trimethoprim 200 mg tablets: £1.76 per 14 tablets	1.76
Trimethoprim	Children	Targeted	Dosage depends on age and weight	Dosage from BNF		
Pivmecillinam hydrochloride	Children with UTI	Targeted	5–10 mg/kg every 6 hours	Dosage from BNF		
Ampicillin	Adults aged ≥ 18 years with UTI	Targeted	0.5–1 g every 6 hours	Dosage from BNF	Ampicillin 500 mg capsules: £47.96 per 28 capsules	
Ampicillin	Children with UTI	Targeted	Dosage depends on age	Dosage from BNF		

BNF, British National Formulary.

thresholds, including the £20,000–30,000 per QALY commonly used by NICE, nitrite and leucocyte esterase followed by micturating cystourethrography was most cost-effective. Nitrate or laboratory leucocyte esterase/culture-based testing followed by micturating cystourethrography was also cost-effective. These have limited relevance to our evaluation, as POCTs were not considered and guidelines on UTI treatment have been updated in the past 18 years. The population was also children only, so there was limited generalisability across our subgroups.

Utility data came from Barry *et al.*,[71] a US-based cost–utility analysis of evaluation strategies for UTI in ambulatory women. Although this source is outdated, the authors reported duration of treated pyelonephritic attack as 10 days and untreated as 14 days, and utility decrements of 0.010225 and 0.014315 in treated and untreated, respectively. We use the durations in our model (see *Table 15*).

Gaither 2020

Gaither *et al.*[73] developed a decision tree model to estimate the cost-effectiveness of routine screening renal bladder ultrasound in children aged 2–24 months after a first febrile UTI. The study's main outcomes were the incremental cost-effectiveness ratio (ICER) and the recurrent UTI rate, where a recurrent UTI was defined to be a second UTI occurring within 1 year. They used a US health system perspective with a willingness-to-pay threshold of US$100,000 per QALY. The authors found that screening renal bladder ultrasound after a first febrile UTI was not cost-effective when compared with their control arm of screening after a second UTI. Using data from the Careful Urinary Tract Infection Evaluation (CUTIE) trial, they estimated the recurrent UTI rate to be 0.19 in patients without genitourinary anomalies or vesicoureteral reflux and with index UTI occurring between the ages of 2 and 72 months.[74]

Sanyal 2019

Sanyal *et al.*[75] used a decision tree model to compare the cost-effectiveness and budget impact of the management of uncomplicated UTI in women when initiated by community pharmacists compared with family physicians or emergency physicians. Costs were based on data from Canada. The authors concluded that, from the perspective of the Canadian public healthcare system, community pharmacist-initiated management would likely be a cost-effective strategy for uncomplicated UTI. In their model, 88.6% of patients were cured of UTI in the pharmacist-initiated group and 90% of patients were cured of UTI in the family and emergency physician-initiated groups, although it is not clear which tests were used to assess UTI. They used quality-adjusted life-months to model health outcomes, but did not explicitly report the quality-adjusted life-months used for different health states. Instead, they reported the utilities at the start and end of the 28-day assessment period. Ernst *et al.*[76] (their source) provided more detailed data from which a curve could be fitted to estimate the quality-adjusted life-months.

Kassabian 2022

Kassabian *et al.*[77] used a decision tree model to perform a cost-effectiveness analysis comparing fosfomycin with nitrofurantoin and trimethoprim-sulfamethoxazole (TMP-SMX) as treatment for uncomplicated UTI from a US perspective. They concluded that fosfomycin may be considered cost-effective, especially if antibiotic stewardship is taken into account. In their model, the probability of UTI resolution following an initial course of antibiotics was 88.17% for fosfomycin, 85.94% for nitrofurantoin and 81.78% for trimethoprim-sulfamethoxazole. These estimates were derived using estimates of bacterial susceptibility and the proportions of UTI caused by different bacteria.

Implications for cost-effectiveness modelling

The available evidence drove our selection of tests and subgroups to model. We prioritised the modelling of rapid tests, with results in < 40 minutes, over culture-based tests, with results in < 24 hours, due to their greater potential impact on clinical practice. We also prioritised tests that performed AST (i.e. Astrego PA-100, Flexicult Human) over those that only identified pathogenic cause (i.e. Lodestar DX, ID Flexicult, Chromostreak, Uricult Plus). Both of these types of tests were prioritised over those that only detected UTI.

As summarised under objective 2 and in *Table 6*, the only rapid tests with accuracy data were Lodestar DX, Uriscreen and ULTRiPLEX. None of these can perform AST and only Lodestar DX can detect pathogenic cause. We therefore selected only Lodestar DX for modelling. However, data on Lodestar DX were only available on accuracy of identifying specific bacteria (*E. coli*) and not on accuracy of detecting UTI itself. The only culture-based tests with accuracy data that performed AST were Flexicult Human and ID Flexicult, while the one that identified pathogenic cause alone was Uricult Trio. Dipstreak provides some information on pathogenic cause by detecting the presence of Gram-negative bacteria. However, only laboratory-based studies were found for Dipstreak. We therefore excluded it from modelling. Only one study, at high risk of bias, provided accuracy data on the Uricult tests (see *Table 10*), and this test also only identifies the presence of Gram-negative bacteria, and was in a laboratory setting, so Uricult was also not selected for modelling.

Therefore, we included Lodestar DX, Flexicult Human and ID Flexicult in modelling. Astrego-PA was the test with highest potential impact (AST in < 40 minutes), but there were no accuracy data, so it could not be meaningfully modelled. The final selection of tests is summarised in *Table 17*.

The populations of interest evaluated in the included studies are summarised in *Table 17*. Lodestar DX was only evaluated in a mixed population. Due to its importance as the only rapid test with accuracy data, we assume that this estimate can be generalised to other populations. Flexicult Human and ID Flexicult were only evaluated in a mixed population and/or women with uncomplicated UTI. We thus focused evaluations on two populations in which we could model up to three tests each.

Mixed population

- Lodestar DX
- Flexicult Human

Women with uncomplicated UTI:

- Lodestar DX (assuming same accuracy as in mixed population)
- Flexicult Human
- ID Flexicult

TABLE 17 Final selection of tests, and summary of evidence, for modelling based on data availability and potential impact

Test	Rapid or culture based	AST or only identifies bacteria	Bias in accuracy data, other comments	Cost data	Populations (number of studies)
Included					
Lodestar DX	Rapid	Identifies bacteria	No UTI detection accuracy data One study at unclear risk of bias	Yes	Mixed (*n* = 1)
Flexicult Human	Culture based	AST	3 at unclear, 1 at low	Yes	Women – uncomplicated UTI (*n* = 2); mixed (*n* = 1)
ID Flexicult	Culture based	AST	1 at low risk, 1 at unclear	No	Women – uncomplicated UTI (*n* = 2)
Could be modelled but no comparator in available populations					
Uricult Trio	Culture based	Identifies bacteria	1 at high risk, 1 at unclear risk, 1 at low risk	No	Pregnant women (*n* = 2); children aged < 16 years (*n* = 1); children aged < 24 months (*n* = 1)

Evaluating costs, quality of life and cost-effectiveness

Using the conceptual model in *Conceptual modelling of costs, quality of life and cost-effectiveness* and evidence sources summarised in *Review of evidence on cost-effectiveness*, we developed a structure and identified the necessary evidence to evaluate the costs, quality of life and cost-effectiveness of POCTs for UTI. Our model also assesses the reduction in use of empiric/broad-spectrum antibiotics, and therefore antibiotic use overall, as POCTs with AST can yield targeted treatment and POCTs without AST can indicate when no UTI is present. An NHS and PSS perspective was taken with a lifetime horizon where costs and QALYs were discounted at an annual rate of 3.5%.

Our conceptual model could be extended to a full model with systematic literature reviews and other evidence-gathering exercises; the analyses described below are therefore what should be done if a full timescale for this work were ever to be made available, rather than the truncated timing of an EVA. However, a simple coded model for the tests and subgroups identified in *Implications for cost-effectiveness modelling* has been implemented in the R programming language. The results are not presented from this model as the evidence identified is too limited for the results to be meaningful, even for the subset of tests and populations evaluated.

Model structure

Our model structure comprises a decision tree over which the costs and consequences of testing for UTI would play out. Decision tree was the only type of model we identified as having been used previously in the UTI literature.[59,60,65,70,72] The key model assumptions are presented in *Table 18*.

TABLE 18 Key structural and parameter assumptions of the cost-effectiveness model

Assumptions of the cost-effectiveness model	
(i)	The underlying probability of UTI (p_uti) is the same regardless of the test used, but varies according to patient subgroup
(ii)	Test accuracy does not vary by subgroup. A notable exception is that manufacturers' submissions note that Astrego can be used only in women
(iii)	Probability of antibiotic cure and side effects varies by population
(iv)	The probability of 'healthy' on targeted treatment (p_healthy_targ) is the same for each targeted antibiotic
(v)	As some tests can identify pathogenic cause or type of infection, despite not performing AST, the probability of 'healthy' on empiric treatment (p_healthy_emp) depends on the type of test used but not on which empiric antibiotic was prescribed
(vi)	Costs and health impacts of pyelonephritis, sepsis and kidney failure can be modelled as once-off costs and disutilities
(vii)	The probability of requiring more than one course of antibiotics is higher if we prescribe empiric antibiotics as there is a higher probability of the first course not targeting the correct bacteria
(viii)	Not modelling long-term impact of unnecessary antibiotic prescription. Instead modelling extent of empiric antibiotic treatment used for suspected UTI
(ix)	AST for patients without UTI will not detect specific antibiotic sensitivity, so these patients can only be falsely given broad-spectrum/empiric antibiotics
(x)	The UTI is eventually cured by targeted or empiric courses of antibiotics, although patients may suffer complications and remain at risk of recurrent/chronic UTI
(xi)	Antibiotic treatment may be given while awaiting culture-based testing results. All patients with a detected UTI will eventually be treated with antibiotics, but some may be treated only after the culture-based testing results have been received
(xii)	Patients started on antibiotics while awaiting culture-based testing will complete their course of antibiotics, even if the culture-based test eventually comes back negative

TABLE 18 Key structural and parameter assumptions of the cost-effectiveness model (*continued*)

Assumptions of the cost-effectiveness model	
(xiii)	Patients without UTI may benefit from POCT or culture-based testing as underlying cause of symptoms may have specific antibiotic sensitivity
(xiv)	Patients suspected of UTI but with (true or false) negative test results may be given no further treatment or non-specific empiric/broad-spectrum antibiotics
(xv)	Costs and QALYs of complications do not vary by subgroups

Pyelonephritis, kidney failure and sepsis can be modelled as a once-off cost and quality-of-life decrement. We furthermore did not need to model a later recurrence of UTI. Such a repeat UTI would already be modelled by the decision tree model, as the tree does not distinguish between first and repeat UTI. We therefore did not adopt a long-term model, such as a cohort Markov model, for the long-term outcomes in *Figure 4*.

Our decision tree is illustrated in *Figure 5*. This structure is for rapid POCTs that perform AST or identify pathogenic cause (e.g. Astrego PA-100, Lodestar DX), POCTs with culture-based testing (e.g. Flexicult Human, ID Flexicult), and laboratory culture-based testing (with or without dipstick). The model could be extended to include no testing, as is often the strategy for women with uncomplicated UTI and typical symptoms.[78] Patients are assumed to have either a true UTI or no underlying UTI. Our conceptualisation is that the POCTs with AST or pathogenic cause identification would identify a patient as having UTI and a specific antibiotic to which the patient is susceptible, identify a patient as having UTI but not identify a specific antibiotic to which the patient is susceptible, or identify a patient as not having UTI. It is assumed that the POCTs with AST may not always detect the antibiotic to which the UTI is susceptible as they do not detect all possible bacteria. Laboratory culture-based testing can initially assign patients to broad-spectrum/empiric antibiotics before targeted treatment is enabled by the results of the test. Under all strategies, if no UTI is detected, the patient is assumed to be assigned to no further treatment. False positives (i.e. patients without UTI but diagnosed with UTI) are assumed to always receive broad-spectrum/empiric antibiotics.

The probabilities of detecting UTI and, when with AST, detecting antibiotic susceptibility would differ between POCTs, as per the analyses in *Results*.

Treatment with broad-spectrum/specific antibiotics is modelled to include multiple courses of antibiotics. It also includes switching from one antibiotic to a targeted antibiotic in response to the results of a POCT or laboratory culture-based testing.

The decision tree then assumes that antibiotics would be assigned accordingly (e.g. targeted if specific susceptibility is known, empiric/broad-spectrum if unknown, and not given if known not to be a UTI). Empiric antibiotics are assumed to be potentially followed by targeted treatment if the initial antibiotic is unsuccessful and the results of culture-based tests become available. Treatment can be successful and leave a patient healthy without complications, or unsuccessful with complications from UTI. Our model assumes that all patients are eventually cured of UTI but may suffer complications, in line with Wang and LaSala,[60] Sadler *et al.*[65] and all previous economic models we identified. Patients who recover without complications are 'healthy' but at risk of recurrent or chronic UTI. Dipstick with laboratory culture-based testing, or culture-based testing alone, is assumed to initially lead to broad/empiric antibiotic treatment as specific susceptibility is unknown.

The 'UTI but not specific, empiric/broad treatment' arms include additional costs and QALY losses from further testing being required to identify and prescribe an effective antibiotic. Recurrence of UTI takes place after the decision tree and may include chronic UTI.

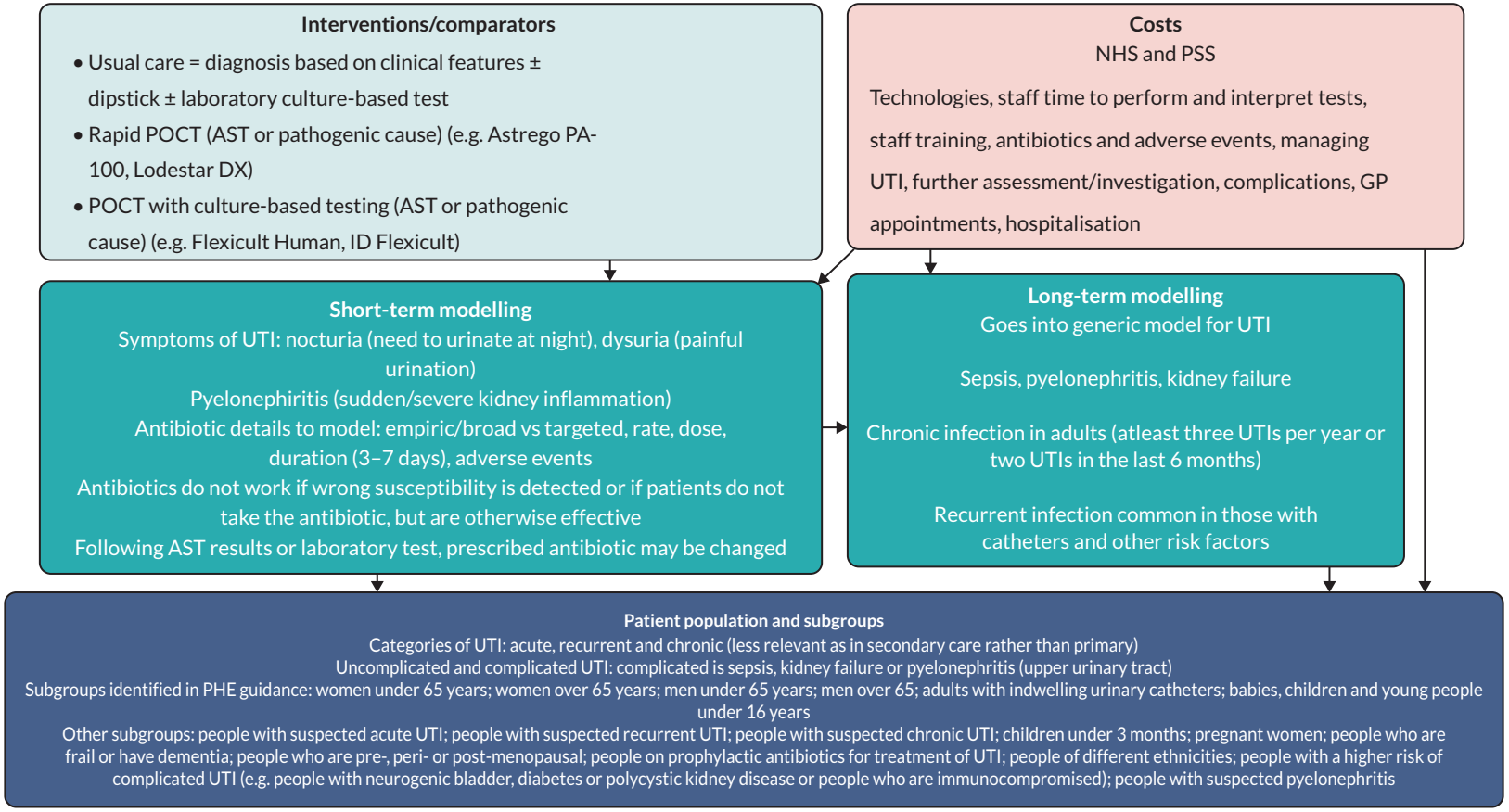

FIGURE 4 Conceptual model for POCTs in UTI. Boxes illustrate important elements to consider. Arrows illustrate influence. AS, antibiotic susceptibility; PHE, Public Health England.

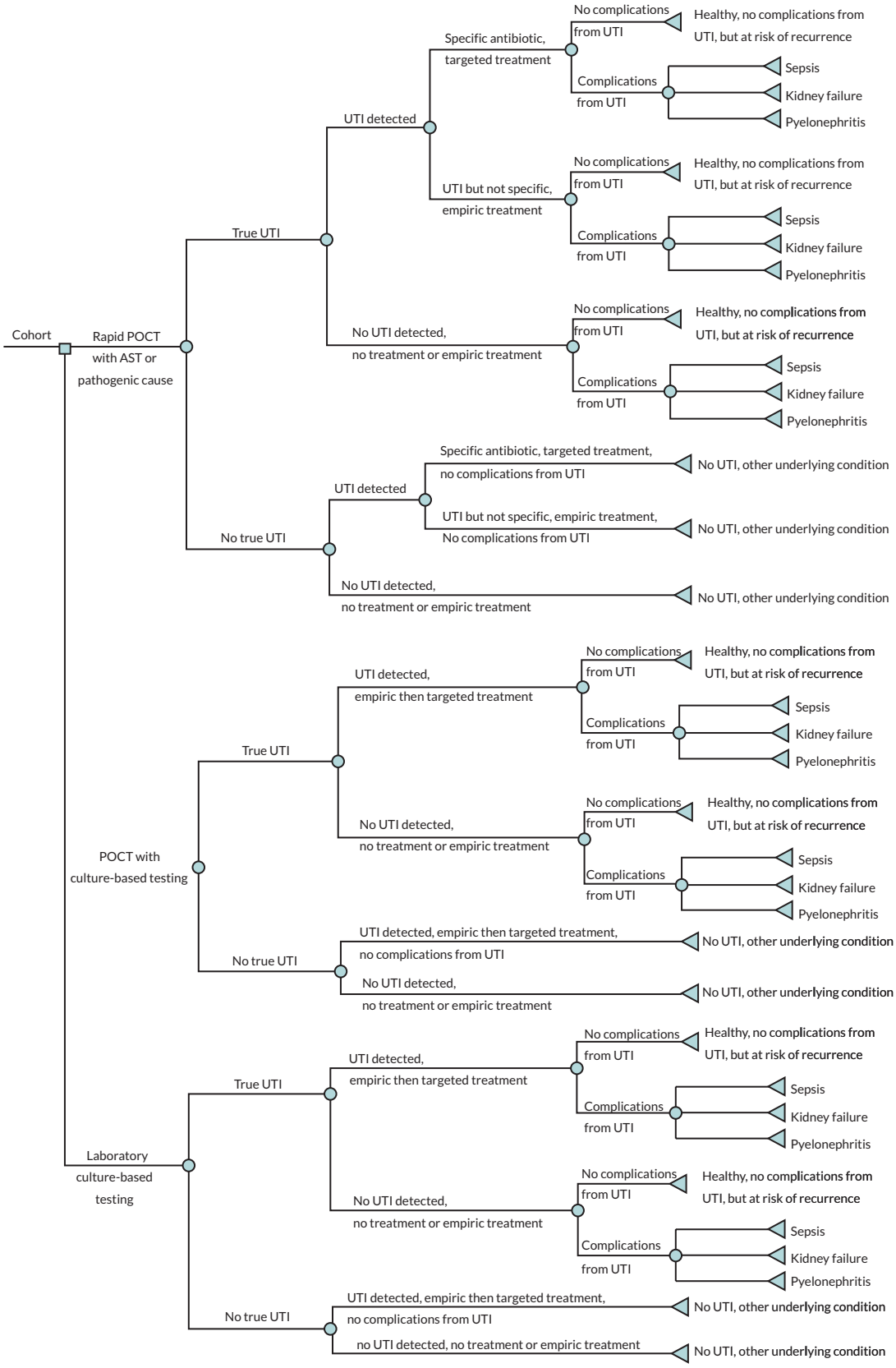

FIGURE 5 Decision tree structure for short-term modelling. 'False positives' do not incur further costs or consequences. We assume that these branches only incur the cost/disutility of treatment, and that they have no benefit from the POCT.

Model inputs

Where possible, model inputs were derived from the clinical review, from our additional searches in *Additional pragmatic searches for cost-effectiveness evidence*, or from expert opinion. We would recommend further systematic literature reviews and expert elicitation in a full-scale evaluation.

Test accuracy parameters

Test accuracy data are summarised in *Table 19* but were derived from *Objective 2: what is the accuracy of the point-of-care test for urinary tract infection diagnosis, pathogen identification and antimicrobial sensitivity testing?*. Although estimates of sensitivity and specificity for detecting UTI were identified, few reliable data were identified on the probability of identifying antibiotic susceptibility or pathogenic cause to direct targeted treatment. Sensitivity and specificity for detecting *E. coli* were identified for Lodestar DX (see *Table 19*).

Other model input parameters

Values, distributions and evidence sources for other model input parameters are summarised in *Table 15*.

Health outcomes

In the decision tree, we need to quantify the quality of life with a complicated or an uncomplicated UTI, and the impact on quality of life of testing and of the 3- to 7-day course of antibiotics, including their adverse events (AEs). We did this using utilities, disutilities and QALYs over defined time periods, for example a disutility for antibiotic AEs along with a proportion of the cohort expected to experience these AEs; and the QALYs accrued by patients with complicated or uncomplicated UTI over the period of the short-term model. These are summarised in *Table 15*.

The utility and QALY estimates could then be used to generate total QALYs over the time horizon of the overall model for each strategy.

The model is designed to additionally estimate the proportion of patients assigned to empiric antibiotic treatment under each treatment pathway. This aimed to assess the impact on antibiotic resistance.

Costs

Costs of testing technologies, staff time to perform the tests, GP appointments, antibiotics courses, managing complicated/uncomplicated UTI and managing each complication were gathered from

TABLE 19 Accuracy parameters of tests that could be included in the cost-effectiveness model

Test name	Type of test (POCT with/without AST)	Probability of correctly detecting UTI (sensitivity or true-positive rate)	Probability of incorrectly diagnosing a non-UTI patient as having UTI and then giving them antibiotics (specificity or false-positive rate)	Probability of identifying specific antibiotic for targeted treatment	Source of values
Flexicult Human (SSI Diagnostica)	Culture-based, with AST	Sensitivity 0.79 (0.72 to 0.85)	Specificity 0.67 (0.30 to 0.90)	No reliable data identified by systematic review	Meta-analysis of women and mixed populations in primary care near patients (see *Figure 3*)
Lodestar DX (Llusern Scientific)	Rapid, identifies pathogenic cause	Sensitivity 0.86 (0.74 to 0.99)	Specificity 0.88 (0.83 to 0.94)	No reliable data identified by systematic review	Published Lodestar abstract[51]
ID Flexicult	Culture-based, with AST	Sensitivity 0.89 (0.84 to 0.93)	Specificity 0.70 (0.52 to 0.84)	No reliable data identified by systematic review	Meta-analysis of women in primary care near patients (see *Figure 3*)

evidence sources described in *Review of evidence on cost-effectiveness*. These were supplemented by routine NHS sources (NHS reference costs, Personal Social Services Research Unit, *British National Formulary*) and discussions with clinical advisors. The costs of antibiotic treatment are summarised in *Table 16*, while other costs are summarised in *Table 15*.

Costs of training staff to use innovative tests were considered, but these are a budget impact rather than a cost to be included in the cost-effectiveness analysis as they relate to cost of setup rather than routine use.

Analyses

Probabilistic analysis where parameter uncertainty is captured with probability distributions and simulation would be used to estimate ICERs and expected net benefits at commonly used NICE willingness-to-pay thresholds. Uncertainty should be presented using cost-effectiveness acceptability curves and cost-effectiveness planes.

Scenario and subgroup analyses

As explained in *Implications for cost-effectiveness modelling*, we model only two populations with three available tests in each.

Mixed population:

- Lodestar DX
- Flexicult Human.

Women with uncomplicated UTI:

- Lodestar DX (assuming same accuracy as in mixed population)
- Flexicult Human
- ID Flexicult.

In a full economic evaluation, other subgroup and scenario analyses would be conducted.

One-way sensitivity analyses would be recommended for all key parameters in a full evaluation, including all parameters based on expert opinion.

Summary of evaluation of cost-effectiveness

In *Conceptual modelling of costs, quality of life and cost-effectiveness*, we developed a conceptual model that could be used for a future full economic evaluation of POCTs for UTI and the role of these tests in reducing antibiotic resistance. Our evaluation of the identified evidence (see *Review of evidence on cost-effectiveness*) and attempt to inform a decision tree implementation of the conceptual model (see *Evaluating costs, quality of life and cost-effectiveness*) reveal that the evidence is too limited for results to be meaningful. This is despite the restriction to a narrow set of tests and subgroups in *Implications for cost-effectiveness modelling*. We summarise below the areas where our evidence is most limited. However, these do not constitute formal gaps in the evidence. The clinical effectiveness systematic review of *Chapters 4* and *5* was restricted to addressing objectives 1–3 of *Chapter 3*, which relate to the clinical efficacy, accuracy and technical performance of POCTs. Systematic literature reviews were not conducted on, for example, quality of life with UTI, efficacy of antibiotics for treating UTI, or costs related to complications of UTI. Our pragmatic search of *Additional pragmatic searches for cost-effectiveness evidence* identified eight previous economic models in UTI, but it was not a systematic search as no formal inclusion criteria was specified, the search terms were potentially insensitive, and screening was performed by only one analyst.

Evidence on test accuracy that could be used in our cost-effectiveness model is summarised in *Table 19*. The sensitivity and specificity of detecting UTI estimates for Flexicult Human and ID Flexicult were identified by the clinical effectiveness systematic review, but no reliable data were identified on the accuracy of detecting specific antibiotic sensitivity. For Lodestar DX, the sensitivity and specificity were identified for detecting *E. coli* estimates but not for detecting UTI overall.

There were more substantial evidence limitations in the other model parameters summarised in *Table 15*. No evidence was identified on the probabilities of sepsis and kidney failure resulting from UTI on targeted antibiotics, empiric antibiotics or no treatment. The probability of pyelonephritis on treatment was identified using NICE guideline NG109, but this guideline did not distinguish between targeted and empiric treatment and related to pregnant women. We would need to assume that this applies to non-pregnant women and the mixed population, which is questionable.

Full costing was possible for single courses of antibiotics to treat UTI (see *Table 16*). However, no evidence was identified on the probability of needing more than one course of antibiotics. There was also no evidence on the proportion of patients given antibiotics if their initial test did not detect UTI. Cost data on POCTs themselves were limited. The total cost per person of the Flexicult test was estimated in Butler *et al.*,[8] which included administration and interpretation costs, but similar estimates were not available for Lodestar DX or ID Flexicult. The manufacturer of Lodestar DX provided only the price of the test, plus an estimate of the distribution cost. The price per test of ID Flexicult was not provided by the manufacturer.

Evidence on the costs and QALY impacts of sepsis and kidney failure in UTI was not identified.

These substantial weaknesses in our evidence base limit the utility of our model results for decision-making. Further systematic reviews and expert elicitation would be required to fully inform the model and use it in a full economic evaluation.

Chapter 7 Discussion

Statement of principal findings

There were limited data on the clinical effectiveness of POCTs for UTI. The majority of the included studies evaluated culture-based tests that take up to 24 hours to give a result: Flexicult Human (four studies), Uricult Trio (three studies), Dipstreak (two studies) and ID Flexicult (two studies). The rapid test Uriscreen was evaluated in four studies, with Lodestar DX and UTRiPLEX evaluated in single studies. We did not identify any relevant data on the rapid tests Astrego PA-100 system or TriVerity. The Astrego PA-100 system has the potential to be the most useful of the tests included in the scope for this appraisal as it is able to determine AST within 40 minutes. There were also no data on Chromostreak or Diaslide, but these are linked to the Dipstreak test or Uricult Plus, the latter of which is linked to the Uricult and Uricult Trio tests. These limited clinical effectiveness data also limited the feasibility of an economic evaluation.

Included studies assessed only the following specific populations defined in the scope: women (four studies, not stratified on age), pregnant women (four studies), children (four studies) and those with catheters (one study). There were no data on any of the other prespecified populations of interest. This further limited the scope of economic evaluation to these populations. However, those studies that enrolled a mixed population will most likely have included patients from these populations, but they did not report data separately for the different included populations.

There was very little evidence on the impact of using POCTs for UTI on clinical outcomes. We identified only two trials, both of which evaluated Flexicult Human: one compared with standard care and the other compared with testing with ID Flexicult (which can tell only if UTI is present and does not give information on antibiotic sensitivity). Both trials were judged at low risk of bias. Neither trial reported evidence of a difference in the primary outcome (concordant antibiotic use and appropriate antibiotic prescribing) between intervention groups. Although the study that compared Flexicult Human with standard care found that antibiotic prescribing was reduced at the initial consultation, it did not find a difference between groups for any other outcome related to antibiotic use. Neither study reported a difference between intervention groups for other outcomes: duration of symptoms/infection, patient enablement and resource use. There were no data on mortality or health-related quality of life. The lack of evidence on the impact of antibiotics prescribing also limited the feasibility of economic evaluation.

There were also limited data on the accuracy of POCTs for diagnosing UTI, detecting the pathogenic cause of the infection or detecting antimicrobial sensitivity. Although 16 studies were included for this objective, individual POCTs were each assessed in a maximum of four studies. Where there were data from multiple studies for a single test, studies were heterogeneous in terms of setting, population and where the POCT was performed (near patient or in a laboratory). The limited data suggested that performing the POCT in the laboratory may overestimate accuracy compared with performing the test in a near-patient setting, particularly for culture-based tests. Using stored rather than fresh urine samples was also found to overestimate accuracy in one study that used both types of sample. Some studies were judged to be at risk of bias, and so results should be interpreted with caution. Five were judged at high risk of bias because they had a large proportion of missing data (three studies), included multiple samples from the same patients (one study) and had selected enrolment of patients (one study). Blinding of the person interpreting the reference standard (usually culture) was often not reported and so this may have introduced bias in these studies. There were only two studies that reported direct comparisons between tests. Extreme caution should therefore be applied to the summary estimates and what these mean for the relative accuracy of the tests.

The Lodestar DX tests showed the greatest clinical value potential of the three rapid tests for which data were available. The study of Lodestar DX showed promising results for detecting the presence of *E. coli*, with 86% sensitivity (95% CI 74% to 99%) and 88% specificity (95% CI 83% to 94%). However, this study was conducted in a laboratory setting using fresh urine samples. It did not provide information on the accuracy of the test for detecting other pathogens using fresh urine samples (the test runs six panels each for different pathogens). Despite the importance of Lodestar DX for economic evaluation, this reliance on laboratory-based testing indicated that economic results could be biased in favour of Lodestar DX. Further data from a clinical setting are required to confirm the accuracy of this test.

Uriscreen was the most commonly evaluated rapid test. This simple test, which involves adding a test reagent powder that enables catalase detection followed by hydrogen peroxide to the urine sample, shaking the collection tube and then observing whether a foam ring has formed, is able to tell whether a UTI is present in a few minutes. However, it does not provide any information on antimicrobial sensitivity or on the pathogenic cause of the infection. The results suggested that both sensitivity and specificity were modest, with summary estimates of 74% (95% CI 59% to 84%) for sensitivity and 70% (95% CI 52% to 84%) for specificity. A single study of UTRiPLEX in children in primary care found very poor sensitivity (21%) but very good specificity (94%). This test, which uses a dipstick to detect inflammatory markers and provides data only on whether a UTI is present, is less likely to be of value given the poor sensitivity suggested by this study.

There were more data on culture-based POCTs, but these tests are less likely to be of value in a primary care setting because of the time they take to provide results (up to 24 hours), although they do provide results more quickly than standard laboratory-based culture. As demonstrated by the POETIC trial, the delay in providing results means that clinicians often start antibiotic treatment during the 24-hour wait for the result (reducing the tests' value in avoiding unnecessary antibiotics). The limited data suggested that Dipstreak (two studies) and Uricult (one study) were highly accurate tests, but studies were at high or unclear risk of bias. Both studies of Dipstreak were performed in the laboratory and assessed urine samples from mixed populations (outpatient clinics and hospitalised patients), not all of whom would have presented with symptoms of a UTI. Further studies in a primary care setting are therefore needed to confirm these findings. Uricult was assessed by one study at high risk of bias using samples from secondary care and tested in the laboratory and reported very high sensitivity and specificity of 98% and 100%, respectively. However, studies of Uricult Trio, an extension of Uricult that provides additional information on whether Gram-negative, β-glucuronidase-producing organisms (e.g. *E. coli*) are present, reported more modest accuracy, with summary sensitivity and specificity estimated at 73% (95% CI 63% to 82%) and 70% (95% CI 52% to 84%), respectively. These studies were conducted in near-patient settings and so were likely to produce more reliable estimates for the use of this test in practice. Flexicult Human (four studies) and ID Flexicult (two studies) were found to be modestly accurate in the detection of UTI, with summary sensitivity of 79% (95% CI 72% to 85%) and 89% (95% CI 84% to 93%) and summary specificity of 67% (95% CI 30% to 90%) and 70% (95% CI 52% to 84%), although these data should be interpreted with caution because of the substantial variation across studies. All studies included in the meta-analysis were conducted and interpreted in primary care; one laboratory-based study of Flexicult Human reported higher estimates of sensitivity and specificity (this study was not included in the meta-analysis for this reason). Flexicult Human was shown to have good accuracy for AST, with summary sensitivity of 87% (95% CI 83% to 90%) and summary specificity of 93% (95% CI 89% to 95%). Two studies of culture-based tests provided information on the tests' accuracy in correctly identifying the pathogenic cause. One study of Dipstreak reported sensitivity of 78% with no bacteria incorrectly identified (i.e. where bacteria were detected, all were correctly identified). A study of Uricult Trio looked only at the detection of the *E. coli* infection and reported sensitivity of 60% and specificity of 96%.

There were also very few data on the technical performance of the tests. We did not find any studies that reported only data on technical performance; all data for this objective came from five studies included for either objective 1 or objective 2 and relate to culture-based tests. Three studies evaluated Flexicult Human and two evaluated Uricult Trio. Technical performance data suggested that POCTs

are easier to use and interpret than laboratory tests and produce results more quickly. The study of Uricult Trio reported fewer lost specimens using this POCT than with laboratory tests that need to be transported. The POETIC study included for objective 1 provided additional data on outcomes in the Flexicult Human arm only. These showed that it was quick to perform the test and obtain and record results and to discuss these with patients, although data on time between taking the sample and obtaining a test were not reported. A qualitative substudy of the POETIC trial suggested that around half of clinicians considered that Flexicult had increased their awareness about antibiotic prescribing and had positively impacted their prescribing habits. However, there were barriers to implementation, including limits on when the test can be used, difficulties in test result interpretation, limited resources, concerns about prolonging patient discomfort while awaiting test results, and the expense of maintaining a regular stock of tests. Only one study reported data on cost; Flexicult Human was reported to cost £48. (Confidential information has been removed). There were no other data on costs, and no data on test failure rate or health-related quality of life.

New POCTs would need to be more accurate and cheaper than standard dipstick tests or provide additional information to inform treatment. Although these tests give results within a few minutes, they are able only to suggest whether or not a UTI is present; they do not provide any information on pathogenic cause or on antimicrobial sensitivity. Six studies provided a direct comparison of POCTs with standard dipstick tests. These showed that culture-based tests were both more sensitive and more specific than standard dipstick tests. Results were more variable for the studies that compared rapid tests with standard dipstick tests.

We developed a conceptual model that could be used for a future full economic evaluation of POCTs for UTI and their role in reducing antibiotic resistance. This model identified pathways for benefit of POCTs, namely that they could reduce the use of empiric antibiotics and, by reducing the incidence of UTI complications and improving cure rates, reduce the healthcare costs and quality-of-life impacts arising from UTI.

The above limitations of the clinical evidence were compounded by limited findings of our further pragmatic searches for economic models. We found only eight previous economic models in UTI management, which provided limited evidence on rates of complications, treatment effects, quality of life and costs. We further explored NICE guidelines on antibiotics for UTI treatment, but these also yielded estimates of efficacy in a small range of subgroups and in broad ‘treated’ or ‘untreated’ groups. This made it impossible to show benefit of targeted versus empiric antibiotic treatment. Given the limitations in the clinical evidence, we restricted our potential implementation of the economic model to a mixed population (Lodestar DX vs. Flexicult Human) and in women with uncomplicated UTI (Lodestar DX vs. Flexicult Human vs. ID Flexicult). Even with this narrow comparison, it was decided that the results of our economic model would not be meaningful, and our findings are limited to the conceptual level.

Strengths and limitations of the assessment

Our systematic review followed published guidance on conducting systematic reviews of DTA studies[26] and is reported in accordance with PRISMA 2020 guidance[27] and PRISMA-DTA guidance, making our review processes transparent and robust. The protocol was pre-registered in the PROSPERO database (CRD42022383889). The only changes that we made to the protocol were to broaden our inclusion criteria such that objective 3 was not restricted to studies of tests that had not been evaluated for objective 1 or 2 and to include studies of ID Flexicult in addition to those of Flexicult Human.

We conducted extensive literature searches designed to maximise the retrieval of relevant studies and did not apply any language or date restrictions to these searches. However, the review was restricted to studies published after the year 2000 so that it could be completed within the tight timescales of an EVA. We documented those studies considered potentially eligible but excluded due to their

publication date; 62 studies were excluded for this reason. All evaluated culture-based POCTs; the majority evaluated Uricult/Uricult Trio, with a small number evaluating Uriscreen and Diaslide. We did not exclude studies based on the language of the report and included two non-English studies: one in Spanish and one in Korean. We used Google Translate to enable us to include these studies. The Spanish translation was checked by a member of the team whose first language is Spanish and this was found to be accurate; we were unable to verify the accuracy of the Korean translation. We conducted a formal assessment of the risk of bias of included studies using the RoB 2 tool for RCTs[30] and the QUADAS-2 tool for DTA studies[32] and its extension QUADAS-C[79] to assess the two comparative accuracy studies included in the review. We modified QUADAS-2 to exclude the assessment of applicability. This is because our review question was broad with multiple populations and tests of interest. Instead of a formal assessment of applicability, we extracted information that could result in variation across studies and considered this in our synthesis of results. These data included population, setting, location of test performance, POCT and culture threshold, and reference standard. However, due to the small number of eligible studies that evaluated each individual test, it was not possible to draw strong conclusions regarding the impact of these features on test performance. Our synthesis included a meta-analysis where more than one study evaluated the same test. Included studies only were conducted in women, pregnant women, children and those with catheters. We calculated summary estimates of sensitivity and specificity across patient subgroups. This assumes that accuracy would not vary by subgroup, but this may not be the case; there were insufficient data to investigate whether accuracy varied across different populations. Estimates from these should be interpreted with caution due to clinical and statistical heterogeneity across studies. Estimates should not be applied to populations beyond those from which the estimates were drawn.

We did not include a formal assessment of publication bias due to the small number of included studies and the difficulties in assessing publication bias for DTA studies for which there is no clear threshold for 'significance'.[26]

We prespecified clearly defined, objective inclusion criteria. These specified that studies should be conducted in a population with suspected UTI. We interpreted this broadly such that studies in which pregnant women were screened for UTI and those in which mixed samples sent to the laboratory for testing were also included. However, we excluded studies that only assessed the technical validity of the tests, where control samples with known pathogens were tested using the POCT. These studies do not reflect how the test will perform in practice; they are an initial stage evaluation to determine whether the test can, in principle, be used to process patient urine. Such studies are likely to overestimate test performance. The submission from Astrego highlighted two technical performance studies of the Astrego PA-100 system, a test for which we did not identify any studies that fulfilled the inclusion criteria.[80,81] These studies showed that the test can, in principle, detect the presence of UTI and correctly identify antimicrobial sensitivity. This is potentially a very promising test as it can provide information on the presence of UTI and on antibiotic resistance within 10–15 minutes, but further data on the accuracy and clinical impact are needed. (Confidential information has been removed).

A potential limitation of the evidence base is exactly how UTI should be defined. The gold-standard test for UTI is culture, and the concept of significant bacteriuria, usually defined as > 10^5 CFU/ml, was established in the 1960s by Kass from a study of 415 women attending a prenatal clinic who were screened for bacteriuria, of whom only 35 were culture positive.[82] However, there are limitations to culture as a reference standard. Culture can be negative even when a UTI is present, particularly in the case of antibiotic-resistant bacteria. Laboratory guidelines differ in how culture results should be interpreted to confirm the presence or absence of UTI[13] and recommend different diagnostic criteria depending on age, symptoms and how urine was collected. All but one of the studies included in our review used culture alone as the reference standard, with thresholds to define the presence of UTI ranging from ≥ 10^3 CFU to ≥ 10^5 CFU. In some studies, this was based only on the presence of a single organism, while others had different thresholds for mixed growth, for example 'Single organism 10^4 CFU

or two organisms when colony count of one > 10^5 CFU'. One study used a compound reference standard consisting of culture, microscopy and spiral plating, which is likely to have given a more accurate determination of whether or not a UTI was present.

A further problem is potential contamination of urine samples or asymptomatic bacteriuria.[83] Culture does not distinguish between pathogenic and non-pathogenic bacteria, so bacteria growing on culture will not necessarily indicate the presence of a UTI, particularly in asymptomatic patients. The accuracy of all tests for UTI will depend on how the urine sample was collected and the potential risk of contamination. Where method of collection was reported, most studies included in this review used mid-stream urine samples or urine collection bags for children. Although such methods of urine collection do have a greater risk of contamination than other methods such as suprapubic aspiration or catheterisation, these methods reflect how urine is likely to be collected in practice and so were appropriate.

The available accuracy evidence drove our selection of the tests and subgroups to include in the economic evaluation. We took a pragmatic approach to prioritising the modelling of tests with the greatest potential for impact. This led us to focus on modelling rapid tests over culture-based tests and to prioritise tests that performed AST over those that only identified pathogenic cause, and both over those that only detected UTI. The only rapid tests with accuracy data were Lodestar DX, Uriscreen and UTRiPLEX; none of these can perform AST, and only Lodestar DX can detect pathogenic cause. The only culture-based tests with accuracy data that performed AST were Flexicult Human and ID Flexicult. We therefore aimed to model only Lodestar DX, Flexicult Human and ID Flexicult.

The limited evidence on accuracy further drove our selection of populations to include in the economic evaluation. Lodestar was only evaluated in a mixed population, while Flexicult Human and ID Flexicult were only evaluated in mixed and/or women with uncomplicated UTI. We therefore restricted modelling to a mixed population (Lodestar DX vs. Flexicult Human) and in women with uncomplicated UTI (Lodestar DX vs. Flexicult Human vs. ID Flexicult).

Sensitivity and specificity of detecting UTI estimates for Flexicult Human and ID Flexicult were identified by the clinical effectiveness systematic review, but no reliable data were identified on the accuracy of detecting specific antibiotic sensitivity. Sensitivity and specificity of Lodestar DX were identified for detecting *E. coli* estimates but not for detecting UTI overall.

We utilised cost-effectiveness evidence identified by the clinical systematic review, but this was limited to only two studies. We took a pragmatic approach to searching for additional cost-effectiveness evidence, with searches of Ovid MEDLINE, EMBASE and EconLit. We did not restrict to models and, by not specifying a PICOS, were able to flexibly include any study with potentially useful evidence. However, we found only eight studies, none of which modelled POCTs and none of which provided all of the evidence needed to inform our economic evaluation.

We used a broad conceptual model to reflect the influence of the choice of populations and subgroups on the costs, health outcomes and model structures. This covered all costs, outcomes, tests and populations specified in the scope. We furthermore designed a decision tree to reflect the short-term aspects of our conceptual model. Despite our prioritisation of tests and subgroups, broad approach to modelling and pragmatic approach to searching for evidence, we found that the evidence informing our economic model was too weak for results to be meaningful.

Uncertainties

Given the limited data available for this appraisal, a number of uncertainties remain. These include the accuracy of rapid tests for diagnosing UTI in primary care settings, the comparative accuracy of

tests, whether accuracy varies according to population, how test interpretation varies between the laboratory and near-patient settings, the impact of recurrent or chronic UTI on test performance, and economic modelling.

We identified only a small body of evidence, with evidence particularly lacking for the more novel rapid POCT. There were insufficient data to investigate whether test performance differed across the different populations defined in the scope, or to consider how having recurrent or chronic UTI could impact on test performance.

Although the POCTs were designed to be carried out in a near-patient setting, nine studies performed a POCT in a laboratory setting. Six of these used samples sent to the laboratory, and the others collected the samples in antenatal clinics or primary care and then sent the samples to the laboratory for testing. Studies in which tests were performed in laboratories tended to overestimate accuracy compared with those carried out in near-patient settings. The only primary care settings in which studies were conducted were GP practices and antenatal clinics. There were no data on pharmacy settings. Further data are needed on how these tests perform in a near-patient setting.

The limitations of the clinical effectiveness evidence also limited the scope of the economic evaluation. Despite prioritising those tests and subgroups for which evidence and potential for impact were greatest, it was still decided that results of the economic model would not be meaningful for decision-making.

Although limited sensitivity and specificity data were identified for our prioritised tests (Lodestar DX, Flexicult Human and ID Flexicult), few reliable data were identified on the probability of identifying antibiotic susceptibility or pathogenic cause to direct targeted treatment. Sensitivity and specificity of Lodestar DX were identified for detecting *E. coli* estimates but not for detecting UTI overall.

There were more substantial evidence limitations of the other model parameters summarised in *Table 15*. No evidence was identified on probabilities of sepsis and kidney failure resulting from UTI on targeted antibiotics, empiric antibiotics or no treatment. The probability of pyelonephritis on treatment was identified using NICE guideline NG109, but the guideline did not distinguish between targeted and empiric treatment and related to pregnant women. No evidence was identified on the probability of needing more than one course of antibiotics. There was no evidence on the proportion of patients given antibiotics if their initial test did not detect UTI. We also need better ways of determining the long-term impact at both individual and societal levels of using point of care diagnostics to better target antibiotics.

Cost data on POCTs themselves were limited. The total cost per person of the Flexicult test was estimated in Butler *et al.*,[8] which included administration and interpretation costs, but similar estimates were not available for Lodestar DX or ID Flexicult. The manufacturer of Lodestar DX provided only the price of the test, plus an estimate of the distribution cost. The price per test of ID Flexicult was not provided by the manufacturer. Evidence on costs and QALY impacts of sepsis and kidney failure in UTI was not identified.

In addition to this evidence weakness, the structure of the model was subject to limitations. All assumptions in *Table 18* could be questioned. In particular, the assumption that accuracy does not vary by subgroup could be challenged by *Figure 3*, for example pregnant women versus catheterised people for Uriscreen, specificity of Flexicult Human in mixed population versus women, or pregnant women versus children for Uricult Trio. As further evidence that test accuracy can vary by population, the manufacturers' submissions note that Astrego can only be used in women.

Our choice of a decision tree (see *Figure 5*) to represent the conceptual model (see *Figure 4*) is a substantial structural uncertainty. All economic models in UTI that we identified used decision trees, but these were largely restricted to modelling pyelonephritis as a complication of UTI. Kidney failure,

sepsis, recurrent UTI and chronic UTI are all potential long-term consequences of poor management of UTI. A Markov model, as illustrated in *Figure 6*, could be used to model the long-term consequences of complication branches of our decision tree.

Equality, diversity and inclusion

Our research was based on existing literature and so we had no control over the participants enrolled. We were broad in our inclusion criteria such that studies from any country and in any language of publication were eligible. We had intended to investigate how the accuracy of included tests varied across different populations, but there were insufficient data to allow us to do this.

Our team included researchers with a broad range of experience and expertise. The lead authors are junior researchers within Bristol TAG, who were given the opportunity to lead on the writing of this report to help develop their research skills and portfolio. They were supported by the two senior authors, who provided advice and mentorship to the junior researchers leading on the reviews and health economic modelling. The team included those with expertise in systematic reviews, health economics and medical statistics.

Patient and public involvement

We involved two patient representatives in this project who have lived experience of UTI. They attended meetings with the clinical effectiveness team (one at the beginning of the project and one closer to the end of the project), gave feedback on the plain language summary for the protocol and main report and wrote the following section about the difference that POCTs may make to patients with UTI. The involvement of patients had a positive impact on this project, particularly in highlighting the importance of not having to wait for test results. Discussions around this topic led us to stratify our results section into rapid tests and culture-based tests.

Impact on patients

The first and most important impact for patients is that a test that can be given immediately in the GP's surgery, particularly if it suggests the appropriate antibiotic for treatment, can relieve symptoms much more quickly and effectively with less impact on antimicrobial resistance.

Urinary tract infections can be extremely painful and uncomfortable. They make leaving the house and being away from a toilet very difficult and they therefore impact on the ability of people to manage

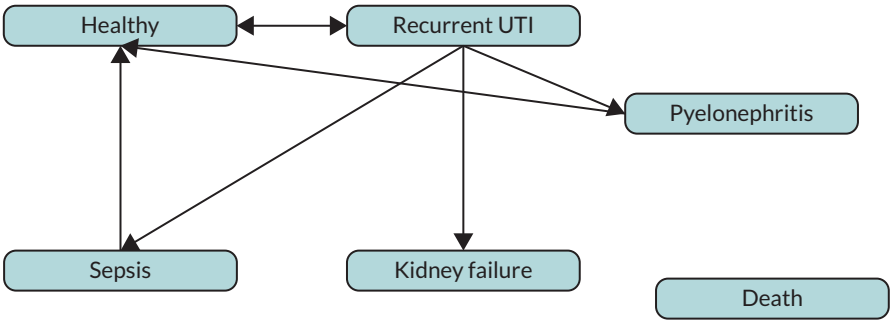

FIGURE 6 Possible long-term Markov model from decision tree. Hospitalisation is a factor for each of the complication states. Death is possible from any state.

their everyday lives. For this reason, anything that can make treatment quicker and more effective is immensely valuable to patients. It also means that patients are less likely to attend accident and emergency departments, relieving pressure on those services and reducing a patient's likelihood of coming into contact with other communicable diseases or spending long and painful hours waiting for treatment.

The benefit of being able to be diagnosed in a person's local GP surgery in one visit would have a major impact on those with busy lives and would make life much better for those who find it difficult to get to the surgery. It would also reduce the number of appointments being booked, freeing up appointments for others to use.

In fact, these tests could be carried out at community pharmacies. It has been shown that during the COVID-19 pandemic more people sought advice and accessed pharmacies and trusted the advice provided. The fact that pharmacies are in the community and accessible, with longer opening hours, including at weekends, benefits patients.

If those GP-based tests can also suggest the most appropriate antibiotic or show immediately that the patient is unlikely to have a UTI, this will lead to less use of antibiotics overall, which must help to reduce antimicrobial resistance. This is positive for the future treatment of infections. This is also likely to cost less in antibiotic prescribing, which would be positive for the NHS.

Chapter 8 Conclusions

Implications for practice

There is a clear need for a rapid test that can accurately diagnose UTI within a short time period in primary care settings, including GP surgeries and pharmacies. Ideally, such tests would also provide information on antimicrobial sensitivity, which would allow appropriate targeted antibiotic use, meaning that patients would be treated appropriately more quickly and the total burden of antibiotic prescriptions would be reduced. The only test within scope that meets these criteria is the Astrego PA-100 system. However, there are currently no data available on this test. Tests such as Lodestar DX that are able to rapidly identify pathogenic cause would also be of value as while these would provide direct information about which antibiotic the causative organism is susceptible to, they would help guide treatment as different pathogens are known to respond differently to certain antibiotics.

Flexicult Human, like the Astrego PA-100 system, is able to provide information on whether a patient has a UTI and on antimicrobial sensitivity. However, this test takes up to 24 hours to produce a result, and this is likely to be longer for samples that are taken on a Friday as the result would not be available until Monday. This makes it more difficult to implement in a primary care setting. Evidence from two trials suggested that using Flexicult had little impact on antibiotic prescribing or on other outcomes such as symptom duration or resource use. Accuracy of the test was found to be modest. Other culture-based tests had similar accuracy when conducted in near-patient settings.

Our conceptual model for economic evaluation found potential pathways to benefit of POCTs. They could reduce costs, improve quality of life and reduce antibiotic resistance by better targeting antibiotic use and reducing complications from UTI. However, we did not have sufficient evidence on test accuracy, targeted versus empiric antibiotic efficacy, or costs and quality-of-life impacts of UTI complications for our model to perform a meaningful comparison. A full evaluation is needed before any recommendation can be made regarding the cost-effectiveness of POCTs or their ability to impact antibiotic resistance.

Strong evidence that POCTs (1) reduce unnecessary antibiotic use, (2) improve symptoms or (3) are cost-effective is needed before such tests are introduced into the NHS.

Suggested research priorities

Given the paucity of data on POCTs for diagnosing UTI, further studies are needed to determine whether POCTs for people with suspected UTI have the potential to be clinically effective and cost-effective to the NHS. Future studies should prioritise those tests with the greatest potential to improve patient outcomes and reduce inappropriate antibiotic prescribing. The most promising tests, of those in scope, are the rapid POCT Astrego PA-100 system (which provides information on antibiotic susceptibility) and Lodestar DX (which provides information on pathogenic cause). Studies should also investigate the feasibility of introducing testing within a pharmacy setting, which could take pressure off GP practices and ensure quicker access to appropriate treatments in the current climate where it can be difficult to access GP appointments. Future research should also encourage the continued development of new diagnostic technologies.

The ideal study would use a similar design to the POETIC study: it would be conducted in primary care (GP surgery and/or pharmacy) and would randomise GP practices/pharmacies to either ‘test and treat appropriately’ or standard practice. Outcomes such as those from a recently published core outcome set[84] (to improve standardisation and comparability between trials), for example symptom duration and

AEs, would then be compared between intervention arms. Ideally, studies would also include a nested diagnostic accuracy study to provide additional information on the accuracy of the test. Either studies should enrol patients across multiple patient groups of interest (e.g. men, women, pregnant women, children), with results stratified according to patient subgroup, or separate studies should be carried out to determine whether results differ according to subgroups. Before such studies are conducted, it may be appropriate to conduct efficacy studies to demonstrate that the technology can work under ideal conditions, in which patient recovery is closely monitored, which cannot be done in a pragmatic RCT as described above.

In addition to further studies on the clinical effectiveness of POCTs, further research on potential cost-effectiveness and impact on antibiotic resistance is needed. This research could build on our conceptual economic model using systematic literature reviews to identify evidence. Such reviews should focus on the efficacy of empiric versus targeted antibiotic treatment of UTI, efficacy in preventing UTI complications, and both the cost and quality-of-life impacts of these complications.

Additional information

Contributions of authors

Eve Tomlinson (https://orcid.org/0000-0002-0969-602X) (Research Associate, Evidence Synthesis, Systematic Reviews) contributed to and checked data extraction and quality assessment for objective 2, extracted data for objective 3, drafted the results sections for objectives 2 and 3 and assisted in drafting clinical effectiveness sections of the discussion.

Mary Ward (https://orcid.org/0000-0002-2321-2546) (Senior Research Associate, Health Economic Modelling, Health Economics) helped design the cost-effectiveness model, gathered input parameters, implemented the model in R and ran analyses.

Chris Cooper (https://orcid.org/0000-0003-0864-5607) (Research Fellow, Health Technology Assessment and Information Science, Systematic Reviews) designed and undertook the literature searches, contributed to the reporting of the systematic review, reviewed the company submissions and worked on the review of cost-effectiveness.

Rachel James (https://orcid.org/0000-0001-9299-9461) (Research Associate, Evidence Synthesis, Systematic Reviews) screened and extracted studies and reviewed the company submissions.

Christina Stokes (Patient Representative) provided a patient perspective on the project, edited the plain language summary and wrote the section of the report on impact on patients.

Samina Begum (Patient Representative) provided a patient perspective on the project, edited the plain language summary and wrote the section of the report on impact on patients.

Jessica Watson (https://orcid.org/0000-0002-8177-6438) (NIHR Clinical Lecturer, General Practice, Primary Care and Test Evaluation) provided clinical advice for the project.

Alastair D Hay (https://orcid.org/0000-0003-3012-375X) (Professor, Primary Care and Test Evaluation) provided clinical advice for the project.

Hayley E Jones (https://orcid.org/0000-0002-4265-2854) (Associate Professor, Medical Statistics and Test Evaluation) provided statistical advice and carried out the meta-analyses of diagnostic accuracy data.

Howard Thom (https://orcid.org/0000-0001-8576-5552) (Senior Lecturer, Health Economics) drafted the cost-effectiveness section of the report, designed the cost-effectiveness model and led the cost-effectiveness assessment.

Penny Whiting (https://orcid.org/0000-0003-1138-5682) (Professor, Clinical Epidemiology, Systematic Reviews and Test Evaluation) drafted the clinical effectiveness sections of the protocol and led the clinical systematic review.

All authors were involved in commenting on the final report. Penny Whiting is the senior author and guarantor.

Acknowledgements

We would like to thank Gus Hamilton for microbiological advice relating to this project, and Charlene Wisdom-Trew, Bristol TAG, for providing administrative support.

Data-sharing statement

All data extracted for the systematic review and the results of the risk of bias assessments are provided in full in the appendices to this report. The R code for the cost-effectiveness model is provided as a publicly accessible repository on GitHub.

Ethics statement

The research included in this report is secondary research and as such did not require ethical approval.

Information governance statement

There were no personal data involved in the production of this report.

Disclosure of interests

Full disclosure of interests: Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at https://doi.org/10.3310/PTMV8524.

Primary conflicts of interest: Alastair D Hay has been a member of the National Institute for Health and Care Research Efficacy and Mechanism Evaluation Funding Committee since 2021 and National Institute for Health and Care Excellence common infections guidelines since 2020.

References

1. Tandogdu Z, Wagenlehner FM. Global epidemiology of urinary tract infections. *Curr Opin Infect Dis* 2016;**29**:73–9. https://doi.org/10.1097/QCO.0000000000000228
2. National Institute for Health and Care Excellence (NICE). *Urinary Tract Infection (Lower) – Men: How Common Is It?* URL: https://cks.nice.org.uk/topics/urinary-tract-infection-lower-men/background-information/prevalence/ (accessed 15 October 2022).
3. Medina M, Castillo-Pino E. An introduction to the epidemiology and burden of urinary tract infections. *Ther Adv Urol* 2019;**11**:1983217. https://doi.org/10.1177/1756287219832172
4. Storme O, Tiran Saucedo J, Garcia-Mora A, Dehesa-Davila M, Naber KG. Risk factors and predisposing conditions for urinary tract infection. *Ther Adv Urol* 2019;**11**:814382. https://doi.org/10.1177/1756287218814382
5. Chronic Urinary Tract Infection Campaign. *What Is Chronic UTI?* 2021. URL: https://cutic.co.uk/what-is-chronic-uti/ (accessed 29 November 2022).
6. National Institute for Health and Care Excellence (NICE). *Urinary Tract Infection (Lower): Women*. 2022. URL: https://cks.nice.org.uk/topics/urinary-tract-infection-lower-women/ (accessed 29 November 2022).
7. Watts V, Brown B, Ahmed M, Charlett A, Chew-Graham C, Cleary P, *et al.* Routine laboratory surveillance of antimicrobial resistance in community-acquired urinary tract infections adequately informs prescribing policy in England. *JAC Antimicrob Resist* 2020;**2**:dlaa022. https://doi.org/10.1093/jacamr/dlaa022
8. Butler CC, Francis NA, Thomas-Jones E, Longo M, Wootton M, Llor C, *et al.* Point-of-care urine culture for managing urinary tract infection in primary care: a randomised controlled trial of clinical and cost-effectiveness. *Br J Gen Pract* 2018;**68**:e268–78.
9. Giesen LG, Cousins G, Dimitrov BD, van de Laar FA, Fahey T. Predicting acute uncomplicated urinary tract infection in women: a systematic review of the diagnostic accuracy of symptoms and signs. *BMC Fam Pract* 2010;**11**:78. https://doi.org/10.1186/1471-2296-11-78
10. National Institute for Health and Care Excellence (NICE). *Pyelonephritis: Acute*. 2021. URL: https://cks.nice.org.uk/topics/pyelonephritis-acute/ (accessed 15 October 2022).
11. Public Health England. *Diagnosis of Urinary Tract Infections: Quick Reference Tool for Primary Care for Consultation and Local Adaptation*. London: Public Health England; 2020.
12. Kahlmeter G, Brown DF, Goldstein FW, MacGowan AP, Mouton JW, Odenholt I, *et al.* European Committee on Antimicrobial Susceptibility Testing (EUCAST) Technical Notes on antimicrobial susceptibility testing. *Clin Microbiol Infect* 2006;**12**:501–3. https://doi.org/10.1111/j.1469-0691.2006.01454.x
13. Grabe M, Bishop M, Bjerklund-Johansen T, Botto H, Çek M, Lobel B, *et al. Guidelines on Urological Infections*. Arnhem: European Association of Urology; 2010.
14. Hoffmann T, Peiris R, Mar CD, Cleo G, Glasziou P. Natural history of uncomplicated urinary tract infection without antibiotics: a systematic review. *Br J Gen Pract* 2020;**70**:e714–22. https://doi.org/10.3399/bjgp20X712781
15. National Institute for Health and Care Excellence (NICE). *Urinary Tract Infection (Lower): Antimicrobial Prescribing*. 2018 (accessed 15 October 2022).
16. Pujades-Rodriguez M, West RM, Wilcox MH, Sandoe J. Lower urinary tract infections: management, outcomes and risk factors for antibiotic re-prescription in primary care. *EClinicalMedicine* 2019;**14**:23–31. https://doi.org/10.1016/j.eclinm.2019.07.012

17. Chronic UTI Info. *Antibiotics*. 2022. URL: www.chronicutiinfo.com/treatment/conventional-medicine/antibiotics/ (accessed 29 November 2022).
18. Chronic UTI Info. *Complementary Medicine*. 2022. URL: www.chronicutiinfo.com/treatment/complementary-medicine/ (accessed 30 November 2022).
19. Bongard E, Frimodt-Moller N, Gal M, Wootton M, Howe R, Francis N, *et al.* Analytic laboratory performance of a point of care urine culture kit for diagnosis and antibiotic susceptibility testing. *Eur J Clin Microbiol Infect Dis* 2015;**34**:2111–19.
20. World Health Organization. *Antibiotic Resistance*. 2020. URL: www.who.int/news-room/fact-sheets/detail/antibiotic-resistance (accessed 29 November 2022).
21. UK Health Security Agency. *English Surveillance Programme for Antimicrobial Utilisation and Resistance (ESPAUR): Report 2021 to 2022*. London: UK Health Security Agency; 2022.
22. Medicines and Healthcare products Regulatory Agency. *Management and Use of IVD Point of Care Test Devices*. 2021. URL: www.gov.uk/government/publications/in-vitro-diagnostic-point-of-care-test-devices/management-and-use-of-ivd-point-of-care-test-devices (accessed 29 November 2022).
23. National Institute for Health and Care Excellence (NICE). *Point-of-care tests for urinary tract infections to improve antimicrobial prescribing: early value assessment*. 2023. URL: www.nice.org.uk/guidance/hte7 (accessed 21 October 2024).
24. Centre for Reviews and Dissemination. *CRD's Guidance for Undertaking Reviews in Health Care*. York: University of York; 2009.
25. National Institute for Health and Care Excellence (NICE). *NICE Health Technology Evaluations: The Manual*. 2022. URL: www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation (accessed 29 November 2022).
26. Deeks J, Bossuyt PM, Leeflang M, Takwoingi Y, Flemyng E. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*. 2022. URL: https://training.cochrane.org/handbook-diagnostic-test-accuracy (accessed 29 November 2022).
27. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *J Clin Epidemiol* 2021;**134**:178–89. https://doi.org/10.1016/j.jclinepi.2021.03.001
28. Cooper C, Booth A, Britten N, Garside R. A comparison of results of empirical studies of supplementary search techniques and recommendations in review methodology handbooks: a methodological review. *Syst Rev* 2017;**6**:234. https://doi.org/10.1186/s13643-017-0625-1
29. National Institute for Health and Care Excellence (NICE). *Diagnostic Assessment Programme Manual*. 2011. URL: www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-diagnostics-guidance/Diagnostics-assessment-programme-manual.pdf (accessed 29 November 2022).
30. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, *et al.* RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;**366**:l4898. https://doi.org/10.1136/bmj.l4898
31. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, *et al.* ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016 **355**:i4919. https://doi.org/10.1136/bmj.i4919
32. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, *et al.*; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;**155**:529–36. https://doi.org/10.7326/0003-4819-155-8-201110180-00009

33. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;**58**:982–90. https://doi.org/10.1016/j.jclinepi.2005.02.022

34. Chu HT, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *J Clin Epidemiol* 2006;**59**:1331–2. https://doi.org/10.1016/j.jclinepi.2006.06.011

35. Holm A, Cordoba G, Moller Sorensen T, Rem Jessen L, Frimodt-Moller N, Siersma V, Bjerrum L. Effect of point-of-care susceptibility testing in general practice on appropriate prescription of antibiotics for patients with uncomplicated urinary tract infection: a diagnostic randomised controlled trial. *BMJ Open* 2017;**7**:e018028.

36. Hullegie S, Wootton M, Verheij TJM, Thomas-Jones E, Bates J, Hood K, *et al.* Clinicians' interpretations of point of care urine culture versus laboratory culture results: analysis from the four-country POETIC trial of diagnosis of uncomplicated urinary tract infection in primary care. *Fam Pract* 2017;**34**:392–9.

37. Holm A, Cordoba G, Sorensen TM, Jessen LR, Frimodt-Moller N, Siersma V, Bjerrum L. Clinical accuracy of point-of-care urine culture in general practice. *Scand J Prim Health Care* 2017;**35**:170–7.

38. Howie JG, Heaney DJ, Maxwell M, Walker JJ. A comparison of a Patient Enablement Instrument (PEI) against two established satisfaction scales as an outcome measure of primary care consultations. *Fam Pract* 1998;**15**:165–71. https://doi.org/10.1093/fampra/15.2.165

39. Blom M, Sorensen TL, Espersen F, Frimodt-Moller N. Validation of FLEXICULT SSI-Urinary Kit for use in the primary health care setting. *Scand J Infect Dis* 2002;**34**:430–5.

40. Pernille H, Lars B, Marjukka M, Volkert S, Anne H. Sampling of urine for diagnosing urinary tract infection in general practice: first-void or mid-stream urine? *Scand J Prim Health Care* 2019;**37**:113–19. https://doi.org/10.1080/02813432.2019.1568708

41. Millar L, DeBuque L, Leialoha C, Grandinetti A, Killeen J. Rapid enzymatic urine screening test to detect bacteriuria in pregnancy. *Obstetr Gynecol* 2000;**95**:601–4.

42. Teppa RJ, Roberts JM. The uriscreen test to detect significant asymptomatic bacteriuria during pregnancy. *J Soc Gynecol Investig* 2005;**12**:50–3.

43. Boon HA, De Burghgraeve T, Verbakel JY, Van Den Bruel A. Point-of-care tests for pediatric urinary tract infections in general practice: a diagnostic accuracy study. *Fam Pract* 2022;**39**:616–22.

44. Macias AE, Arreguin V, Nieto MA, Munoz JM, Medina H. [Catalase test (Uriscreen) in the detection of bacteriuria-candiduria in hospitalized adults with Foley catheter.] *Rev Invest Clin Organ Hosp Enfermed Nutr* 2002;**54**:521–6.

45. Anacleto FE, Resontoc LP, Padilla GH. Bedside diagnosis of outpatient childhood urinary tract infection using three-media Dipslide culture test. *Pediatr Nephrol* 2009;**24**:1539–43.

46. Greeff A, Jeffery B, Pattinson RC. Uricult trio as a screening test for bacteriuria in pregnancy. *South Afr J Obstetr Gynaecol* 2002;**8**:61–4.

47. Lee J, Kim EJ, Lee TJ, Chang JK, Cha SH. Evaluation of Uricult Trio test as a rapid screening of UTI in children with fever [Korean]. *Korean J Pediatr Infect Dis* 2010;**17**:74–82.

48. Mignini L, Carroli G, Abalos E, Widmer M, Amigot S, Nardin JM, *et al.*; World Health Organization Asymptomatic Bacteriuria Trial Group. Accuracy of diagnostic tests to detect asymptomatic bacteriuria during pregnancy. *Obstet Gynecol* 2009;**113**:346–52. https://doi.org/10.1097/AOG.0b013e318194f109

49. Colodner R, Keness Y. Evaluation of DipStreak containing CNA-MacConkey agar: a new bedside urine culture device. *Israel Med Assoc J* 2000;**2**:563–5.

50. Yagupsky P, Rider M, Peled N. Clinical evaluation of a novel chromogenic agar dipslide for diagnosis of urinary tract infections. *Eur J Clin Microbiol Infect Dis* 2000;**19**:694–8.

51. van der Goes A, Diggle J, Nieuwland J, Roula A, Hayhurst E. P32 Performance of a novel molecular test designed for point-of-care UTI diagnosis. *JAC Antimicrob Resist* 2023;**5**:dlad077-036.

52. Hayward G, Butler C. *TOUCAN*. 2023. URL: www.spcr.nihr.ac.uk/research/projects/toucan-platform-for-uti-diagnostic-evaluation (accessed 15 January 2023).

53. KU Leuven. *ERNIE4: Urine and CRP Point-of-Care Test in Acutely Ill Children. NCT03835104*. 2021. URL: www.cochranelibrary.com/central/doi/10.1002/central/CN-01965319/full (accessed 15 November 2022).

54. University of Copenhagen. *Diagnostic Accuracy of Point of Care Test of First Voided Urine Compared to Midstream Voided Urine in Primary Care. NCT02585115*. 2015. URL: www.clinicaltrials.gov/ct2/show/NCT02585115?term=Flexicult&draw=2&rank=2 (accessed 15 November 2022).

55. Brookes-Howell L, Thomas-Jones E, Bates J, Bekkers MJ, Brugman C, Coulman E, *et al.* Challenges in managing urinary tract infection and the potential of a point-of-care test guided care in primary care: an international qualitative study. *BJGP Open* 2019;**3**:bjgpopen18X101630.

56. DRKS00017273. *Management of UTI in German Primary Care: Feasibility of FLEXICULT™ (MAFL)*. 2019. URL: www.drks.de/DRKS00017273 (accessed 15 October 2022).

57. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation; 2019. URL: www.R-project.org/ (accessed 15 December 2022).

58. Carlson KJ. Cost-effectiveness analysis of single-dose therapy of urinary tract infection compared to conventional treatment. *Eur Urol* 1987;**13**:45–7. https://doi.org/10.1159/000472859

59. Fenwick EA, Briggs AH, Hawke CI. Management of urinary tract infection in general practice: a cost-effectiveness analysis. *Br J Gen Pract* 2000;**50**:635–9.

60. Wang R, LaSala C. Role of antibiotic resistance in urinary tract infection management: a cost-effectiveness analysis. *Am J Obstet Gynecol* 2021;**225**:550.e1–10. https://doi.org/10.1016/j.ajog.2021.08.014

61. Schmiemann G, Kniehl E, Gebhardt K, Matejczyk MM, Hummers-Pradier E. The diagnosis of urinary tract infection: a systematic review. *Dtsch Arztebl Int* 2010;**107**:361–7. https://doi.org/10.3238/arztebl.2010.0361

62. Smaill FM, Vazquez JC. Antibiotics for asymptomatic bacteriuria in pregnancy. *Cochrane Database Syst Rev* 2015;**2015**:CD000490. https://doi.org/10.1002/14651858.CD000490.pub3

63. 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. https://doi.org/10.1080/02813430601183074

64. 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.

65. Sadler S, Holmes M, Ren S, Holden S, Jha S, Thokala P. Cost-effectiveness of antibiotic treatment of uncomplicated urinary tract infection in women: a comparison of four antibiotics. *BJGP Open* 2017;**1**:bjgpopen17X101097. https://doi.org/10.3399/bjgpopen17X101097

66. Little P, Turner S, Rumsby K, Warner G, Moore M, Lowes JA, *et al.* Dipsticks and diagnostic algorithms in urinary tract infection: development and validation, randomised trial, economic analysis, observational cohort and qualitative study. *Health Technol Assess* 2009;**13**(19). https://doi.org/10.3310/hta13190

67. Norrby SR. Short-term treatment of uncomplicated lower urinary tract infections in women. *Rev Infect Dis* 1990;**12**:458–67. https://doi.org/10.1093/clinids/12.3.458

68. Carlson KJ, Mulley AG. Management of acute dysuria. A decision-analysis model of alternative strategies. *Ann Intern Med* 1985;**102**:244–9.

69. Bermingham SL, Ashe JF. Systematic review of the impact of urinary tract infections on health-related quality of life. *BJU Int* 2012;**110**:E830–6. https://doi.org/10.1111/j.1464-410X.2012.11337.x

70. 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). https://doi.org/10.3310/hta10360

71. 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.

72. 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. https://doi.org/10.1542/peds.103.4.e54

73. Gaither TW, Selekman R, Kazi DS, Copp HL. Cost-effectiveness of screening ultrasound after a first, febrile urinary tract infection in children age 2-24 months. *J Pediatr* 2020;**216**:73–81.

74. Keren R, Shaikh N, Pohl H, Gravens-Mueller L, Ivanova A, Zaoutis L, *et al.* Risk factors for recurrent urinary tract infection and renal scarring. *Pediatrics* 2015;**136**:e13–21. https://doi.org/10.1542/peds.2015-0409

75. Sanyal C, Husereau DR, Beahm NP, Smyth D, Tsuyuki RT. Cost-effectiveness and budget impact of the management of uncomplicated urinary tract infection by community pharmacists. *BMC Health Serv Res* 2019;**19**:1–3

76. Ernst EJ, Ernst ME, Hoehns JD, Bergus GR. Women's quality of life is decreased by acute cystitis and antibiotic adverse effects associated with treatment. *Health Qual Life Outcomes* 2005;**3**:45. https://doi.org/10.1186/1477-7525-3-45

77. Kassabian M, Calderwood MS, Ohsfeldt R. A cost-effectiveness analysis of fosfomycin: a single-dose antibiotic therapy for treatment of uncomplicated urinary tract infection. *Health Serv Insights* 2022;**15**:11786329221126340

78. Public Health England. *Diagnosis of Urinary Tract Infections: Quick Reference Tool for Primary Care for Consultation and Local Adaptation*. London: Public Health England; 2020.

79. Yang B, Mallett S, Takwoingi Y, Davenport CF, Hyde CJ, Whiting PF, *et al.*; QUADAS-C Group. QUADAS-C: a tool for assessing risk of bias in comparative diagnostic accuracy studies. *Ann Intern Med* 2021;**174**:1592–9. https://doi.org/10.7326/m21-2234

80. Baltekin Ö, Hammar P, Kovachev P, Myzithra M, Wistrand-Yuen E. *Reproducibility of Fully Automated AST for Direct Near Patient Testing*. Abstract no. 1059, 23–26 April 2022. ECCMID 2022, Lisbon.

81. Baltekin O, Boucharin A, Tano E, Andersson DI, Elf J. Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging. *Proc Nat Acad Sci USA* 2017;**114**:9170–5. https://doi.org/10.1073/pnas.1708558114

82. Cohen SN, Kass EH. A simple method for quantitative urine culture. *N Engl J Med* 1967;**277**:176–80. https://doi.org/10.1056/nejm196707272770403

83. Cormican M, Murphy AW, Vellinga A. Interpreting asymptomatic bacteriuria. *BMJ* 2011;**343**:d4780. https://doi.org/10.1136/bmj.d4780

84. Beecher C, Duane S, Vellinga A, Murphy AW, Cormican M, Smyth A, *et al.* COSUTI: a core outcome set (COS) for interventions for the treatment of uncomplicated urinary tract infection (UTI) in adults. *Antibiotics* 2022;**11**:1846. https://doi.org/10.3390/antibiotics11121846

85. POETIC. *POETIC Trial Registration: Point of Care Testing for Urinary Tract Infection in Primary Care. ISRCTN65200697*. 2013. URL: www.isrctn.com/ISRCTN65200697 (accessed 16 November 2022).

86. Bates J, Thomas-Jones E, Pickles T, Kirby N, Gal M, Bongard E, *et al.* Point of care testing for urinary tract infection in primary care (POETIC): protocol for a randomised controlled trial of the clinical and cost effectiveness of FLEXICULT TM informed management of uncomplicated UTI in primary care. *BMC Fam Pract* 2014;**15**:187.

87. Holm A, Cordoba G, Sørensen TM, Jessen LR, Siersma V, Bjerrum L. Point of care susceptibility testing in primary care – does it lead to a more appropriate prescription of antibiotics in patients with uncomplicated urinary tract infections? Protocol for a randomized controlled trial. *BMC Fam Pract* 2015;**16**:106. https://doi.org/10.1186/s12875-015-0322-x

88. Holm A. *Point of Care Susceptibility Testing in Primary Care. NCT02323087*. 2014. URL: https://clinicaltrials.gov/ct2/show/NCT02323087 (accessed 23 November 2022).

89. Wikler MA. *Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically: Approved Standard. Abstract No. 1069*. URL: https://clsi.org/media/1928/m07ed11_sample.pdf (accessed 15 October 2022)

Appendix 1 Literature search strategies

We used one search to inform the clinical review and the review of cost-effectiveness. This was possible because our searches were not limited by study design, date of publication or language.

Resource	Hits
MEDLINE (MEDALL)	526
EMBASE	416
Cochrane	33
CINHAL	12
ClinicalTrials.gov	29
ICTRP	17
Total (prior to deduplication)	1035
– duplicates	−304
N to screen	731

Database: MEDLINE (MEDALL)

Host: Ovid

Data parameters: 1946 to 2 December 2022

Date of search: 5 December 2022

#	Search	Results
1	(Astrego* or (“PA-100” and (urin* or infect*))).ti,ab,kw,kf.	4
2	“Sysmex Astrego”.ab,in.	0
3	flexicult*.ti,ab,kw,kf.	12
4	(“SSI Diagnostica” or “Statens Serum Institut” or “Statens Serum Institute”).ab.	162
5	Lodestar*.ti,ab,kw,kf.	22
6	“Llusern Scientific”.ab,in.	0
7	TriVerity*.ti,ab,kw,kf.	0
8	Inflammatix.ab,in.	40
9	“Uriscreen*”.ti,ab,kw,kf.	16
10	“Savyon Diagnostics”.ab,in.	25
11	(Diaslide* or Dipstreak* or Chromostreak*).ti,ab,kw,kf.	6
12	Novamed.ab,in.	51
13	Uricult*.ti,ab,kw,kf.	66
14	(Aidian or Orion Diagnostic*).ab,in.	145
15	(NCT02323087 or ISRCTN65200697 or NCT02585115 or NCT03835104 or NCT02368847).af.	6

continued

#	Search	Results
16	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15	544
17	exp animals/ not humans.sh.	5,070,893
18	16 not 17	526

Database: EMBASE

Host: Ovid

Data parameters: 1974 to 2 December 2022

Date of search: 5 December 2022

#	Search	Results
1	(Astrego* or ("PA-100" and (urin* or infect*))).ti,ab,kw,kf.	12
2	"Sysmex Astrego".ab,in.	0
3	flexicult*.ti,ab,kw,kf.	12
4	("SSI Diagnostica" or "Statens Serum Institut" or "Statens Serum Institute").ab.	262
5	Lodestar*.ti,ab,kw,kf.	26
6	"Llusern Scientific".ab,in.	0
7	TriVerity*.ti,ab,kw,kf.	0
8	Inflammatix.ab,in.	58
9	"Uriscreen*".ti,ab,kw,kf.	17
10	"Savyon Diagnostics".ab,in.	47
11	(Diaslide* or Dipstreak* or Chromostreak*).ti,ab,kw,kf.	8
12	Novamed.ab,in.	81
13	Uricult*.ti,ab,kw,kf.	70
14	(Aidian or Orion Diagnostic*).ab,in.	229
15	(NCT02323087 or ISRCTN65200697 or NCT02585115 or NCT03835104 or NCT02368847).af.	6
16	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15	817
17	(Animal/ or Nonhuman/) not Human/	6,258,009
18	16 not 17	742
19	limit 18 to EMBASE	416

Database: Cochrane (CENTRAL and CDSR)

Host: Wiley

Data parameters: Issue 12 of 12, December 2022

Date of search: 5 December 2022

#	Search	Results
1	(astrego OR ("PA-100" AND (urin* OR infect*)) OR flexicult OR "SSI diagnostica" OR lodestar OR "Llusern scientific" OR uriscreen OR "savyon diagnostics" OR triverity OR inflammatix OR diaslide OR dipstreak OR chromostreak OR novamed OR uricult OR aidian OR "orion diagnostica")	32
2	(NCT02323087 OR ISRCTN65200697 OR NCT02585115 OR NCT03835104 OR NCT02368847)	5
3	#1 or #2	35

Database: Cumulative Index to Nursing and Allied Health Literature (CINAHL)

Host: EBSCO*host*

Data parameters: 1981–current

Date of search: 5 December 2022

#	Search	Results
S2	TI ((astrego* or ("PA-100" and (urin* or infect*)) or flexicult* or "SSI diagnostica*" or lodestar* or "Llusern scientific*" or uriscreen* or "savyon diagnostics*" or triverity* or inflammatix* or diaslide* or dipstreak* or chromostreak* or novamed or uricult* or aidian* or "orion diagnostica*")) OR AB ((astrego* or ("PA-100" and (urin* or infect*)) or flexicult* or "SSI diagnostica*" or lodestar* or "Llusern scientific*" or uriscreen* or "savyon diagnostics*" or triverity* or inflammatix* or diaslide* or dipstreak* or chromostreak* or novamed or uricult* or aidian* or "orion diagnostica*"))	12
S1	TI ((astrego* or ("PA-100" and (urin* or infect*)) or flexicult* or "SSI diagnostica*" or lodestar* or "Llusern scientific*" or uriscreen* or "savyon diagnostics*" or triverity* or inflammatix* or diaslide* or dipstreak* or chromostreak* or novamed or uricult* or aidian* or "orion diagnostica*")) OR AB ((astrego* or ("PA-100" and (urin* or infect*)) or flexicult* or "SSI diagnostica*" or lodestar* or "Llusern scientific*" or uriscreen* or "savyon diagnostics*" or triverity* or inflammatix* or diaslide* or dipstreak* or chromostreak* or novamed or uricult* or aidian* or "orion diagnostica*"))	31

Note
A server-side deduplication was undertaken at S2 to remove studies included in the MEDLINE database.

Trials registry resources

Clinical Trials.gov

www.clinicaltrials.gov/ct2/results/refine?show_xprt=Y

5 December 2022

#	Search
1	(astrego OR ("PA-100" AND (urine OR urinary OR infection)) OR flexicult OR "SSI diagnostica" OR lodestar OR "Llusern scientific" OR uriscreen OR "savyon diagnostics" OR triverity OR inflammatix OR diaslide OR dipstreak OR chromostreak OR novamed OR uricult OR aidian OR "orion diagnostica")
2	(NCT02323087 OR ISRCTN65200697 OR NCT02585115 OR NCT03835104 OR NCT02368847)
3	1 or 2

WHO International Clinical Trials Registry Platform (ICTRP)

https://trialsearch.who.int/

5 December 2022

#	Search
1	(astrego OR ("PA-100" AND (urine OR urinary OR infection)) OR flexicult OR "SSI diagnostica" OR lodestar OR "Llusern scientific" OR uriscreen OR "savyon diagnostics" OR triverity OR inflammatix OR diaslide OR dipstreak OR chromostreak OR novamed OR uricult OR aidian OR "orion diagnostica")
2	(NCT02323087 OR ISRCTN65200697 OR NCT02585115 OR NCT03835104 OR NCT02368847)
3	1 or 2

A new test (UTRiPLEX) was added by NICE to the scope of this review after the original searches were undertaken. The searches for UTRiPLEX followed the same methods and procedure as for the original searches.

Resource	*N*
MEDLINE	1
EMBASE	3
Cochrane	0
CINAHL	1
Clinical Trials.gov	0
ICTRP	0
Total (prior to deduplication)	5
- duplicates	−2
N to screen	3

Database: MEDLINE (MEDALL)

Host: Ovid

Data parameters: 1946 to present

Date of search: 12 December 2022

#	Search	Results
1	UTRiPLEX*.ti,ab,kw,kf.	1
2	Global Access Diagnostics.ab,in.	0
3	1 or 2	1

Database: EMBASE

Host: Ovid

Data parameters: 1974 to 9 December 2022

Date of search: 12 December 2022

#	Search	Results
1	UTRiPLEX*.ti,ab,kw,kf.	1
2	Global Access Diagnostics.ab,in.	2
3	1 or 2	3

Database: The Cochrane Library (CENTRAL and CDSR)

Host: Wiley

Data parameters: Issue 12 of 12, December 2022

Date of search: 12 December 2022

#	Search	Results
1	(UTRiPLEX* or "Global Access Diagnostics"):ti,ab,kw	0

Database: Cumulative Index to Nursing and Allied Health Literature (CINAHL)

Host: EBSCO*host*

Data parameters: 1981–current

Date of search: 12 December 2022

#	Search	Results
1	TI ((UTRiPLEX* or "Global Access Diagnostics")) OR AB ((UTRiPLEX* or "Global Access Diagnostics"))	1

Trials registry resources

ClinicalTrials.gov

12 December 2022

(UTRiPLEX* or "Global Access Diagnostics")

ICTRP

12 December 2022

(UTRiPLEX* or "Global Access Diagnostics")

Web searching

Searcher: Christopher Cooper

Searcher location: London, UK

Date of search: 6 December 2022

Test name	Manufacturer	Website URL	Search approach	Results (checked/ included)
Astrego PA-100 system and PA-AST panel	Sysmex Astrego website	https://astrego.se/products/	Hand-search of the website followed by Google overlay search: PA-100 site: https://astrego.se/	0/0
Flexicult Human	SSI Diagnostica website	https://ssidiagnostica.com/international/solutions/flexicult/human/	Hand-search of the website followed by Google overlay search: Flexicult Human site: https://ssidiagnostica.com/	1/0
Lodestar DX	Llusern Scientific website	https://llusern.co.uk	Hand-search of the website followed by Google overlay search: Lodestar DX site: https://llusern.co.uk/	0/0
TriVerity	Inflammatix website	https://inflammatix.com/?creative=538983415339&keyword=inflammatix&matchtype=b&network=g&device=c	Hand-search of the website. Followed by manual review of the TriVerity publications tab	37/0
Uriscreen	Savyon Diagnostics Ltd	www.savyondiagnostics.com/product/uriscreen/	Hand-search of the website	2/0
Diaslide, Dipstreak, Chromostreak	Novamed	www.novamed.co.il/culture-device	Hand-search of the website	0/0
Uricult, Uricult Trio and Uricult Plus	Aidian; formerly Orion Diagnostica	www.aidian.eu/microbiology/uricult/uricult-tests#generally	Hand-search of the website	0/0
UTRiPLEX	Global Access Diagnostics	www.globalaccessdx.com/	Hand-search of the website followed by Google overlay search: Flexicult Human site: https://ssidiagnostica.com/	0/0

Appendix 2 List of excluded studies, with rationale

Pre-2000 studies

The table below provides an overview of the studies identified as potentially relevant during title and abstract screening that were excluded because they were published before the year 2000.

Study details	Test evaluated	Objective assessed
Rosenberg M, Berger SA, Barki M, Goldberg S, Fink A, Miskin A. Initial testing of a novel urine culture device. *J Clin Microbiol* 1992;**30**:2686–91	Diaslide	Unclear
Edwards B, White RH, Maxted H, Deverill I, White PA. Screening methods for covert bacteriuria in schoolgirls. *Br Med J* 1975;**2**:463–7	Unclear	Unclear
Van Dorsten JP, Bannister ER. Office diagnosis of asymptomatic bacteriuria in pregnant women. *Am J Obstet Gynecol* 1986;**155**:777–80	Unclear	Unclear
Carroll KC, Hale DC, Von Boerum DH, Reich GC, Hamilton LT, Matsen JM. Laboratory evaluation of urinary tract infections in an ambulatory clinic. *Am J Clin Pathol* 1994;**101**:100–3	Unclear	Unclear
Deguchi K, Yokota *N*, Koguchi M, Suzuki Y, Fukayama S, Ishihara R, *et al.* [Detection of bacteria in urine using dip-slides (1). Possible occurrence of false-negative results when dip-slides are used for urine containing antibacterial agents.] *Jpn J Antibiot* 1995;**48**:155–62	Unclear	Unclear
Roca A, Diez O, Puncernau M, Sanz R, Vinamata B, Carbonell JM. Semiquantitative tests in the diagnosis of urinary infection in pediatric primary care. [Catalan]. *Pediatr Catal* 1998;**58**:147–50	Unclear	Unclear
Zoller L, Tobler L. [Comparison of culture count determination with the Uricult pour-plate.] *Med Lab* 1969;**22**:214–17	Uricult	Unclear
Breitfellner G. [Experiences with Uricult, a new method for the quantitative determination of bacteria in urine.] *Wiener Med Wochen* 1970;**120**:235–43	Uricult	Unclear
Haahr J, Bohn L. [Uricult. A simple method of semiquantitative urine culture.] *Ugeskrift Laeger* 1970;**132**:1360–2	Uricult	Unclear
Orellana M, Linde J, Schmidt V. [Significant bacteriuria. Assessment of a new diagnostic method (Uricult) and presentation of a simple quantitative pipetter dilution method.] *Ugeskrift Laeger* 1970;**132**:1966–70	Uricult	Unclear
Schmid I, Pletscher E. [Uricult, a simple procedure for the determination of bacterial count in urine.] *Med Lab* 1970;**23**:254–6	Uricult	Unclear
Fuchs T, Gutensohn G. [Comparative studies on the value of Uricult-procedure in the diagnosis of urinary tract infections.] *Medizinische Welt* 1971;**18**:735–40	Uricult	Unclear
Bruhl P, Adams E, Straube W. [Results and experiences in the diagnosis of bacteriuria with Uricult.] *Urologe* 1971;**10**:14–17	Uricult	Unclear
Haahr J, Bohn J. Uricult. A simple method of semi-quantitative culture from urine. *Acta Paediatr Scand* 1971;**60**:245–6	Uricult	Unclear
Bailey MJ, Neary JT, Notelovitz M. The Uricult dip-slide in significant bacteriuria. *S Afr Med J* 1972;**46**:1323–6	Uricult	Unclear
Buchanan *N*. Uricult dip-slide in significant bacteriuria. *S Afr Med J* 1972;**46**:1654	Uricult	Unclear
Dayer JM, Humair L. [Bacteriuria: importance and value of the semi-quantitative method of Uricult. Comparative study.] *Schweizer Rundsc Med Praxis* 1972;**61**:384–8	Uricult	Unclear
Hellwig I. [Demonstrations of urinary tract infections using Uricult.] *Deutsch Med Wochensc* 1972;**97**:1687–9	Uricult	Unclear

continued

Study details	Test evaluated	Objective assessed
Mongeau JG, Robillard JE, Brousseau Y. Screening for bacteriuria in children: comparison of two dip-tests. *Can Med Assoc J* 1972;**107**:227–9	Uricult	Unclear
Maugeri TL, Cefali M, Galletti G. [Determination of bacteriuria using Uricult, a new formula.] *Quad Sclavo Diagnost Clin Lab* 1973;**9**:950–63	Uricult	Unclear
Bailey MJ, Notelovitz M. Appraisal of the Uricult dip-slide method in the diagnosis of urinary infections. *S Afr Med J* 1973;**47**:1135	Uricult	Unclear
Finlayson MH, Coates JK, Brede HD, Mitchell P. An appraisal of the uricult dip-slide method in the diagnosis of urinary infections. *S Afr Med J* 1973;**47**:725–7	Uricult	Unclear
Jackaman FR, Darrell JH, Shackman R. The dip-slide in urology. *Br Med J* 1973;**1**:207–8	Uricult	2
Simplaceanu L, Mosora N, Munteanu E. The Uricult test compared with quantitative bacteriuria in diabetics (Rumanian). [Romanian]. *Bacteriol Virusol Parazitol Epidemiol* 1974;**19**:405–10	Uricult	2
Steiner PO, Gerber A, Sigrist W. Independent bacteriologic urine examination with the new Enterotube in a regional hospital. [German]. *Schweiz Med Wochenschr* 1974;**104**:1091–3	Uricult	2
Narbutowicz B, Kostrzewska K, Krawczynski J. [Detection of bacteriuria by means of the Uricult test.] *Pediatr Pol* 1974;**49**:1387–91	Uricult	Unclear
Mackinnon AE, Strachan CJL, Sleigh JD, Burns MM. Screening for bacteriuria with a dip stick test for urinary glucose. *Br J Urol* 1974;**46**:101–5	Uricult	2
Joffe BI, Seftel HC, Distiller LA. Asymptomatic bacteriuria in diabetes mellitus. *S Afr Med J* 1974;**48**:1306–8	Uricult	Unclear
Christen JP, Zawodnik S, Girardet P. Infection and the search for a radiologic anomaly of the urinary tract in a pediatric outpatient practice. [French]. *Schweiz Med Wochenschr* 1974;**104**:430–4	Uricult	Unclear
Anonymous. [New drugs: object culture carrier for the determination of urinary pathogeons (Merckognost Bakteriurie, Uricult, Urifekt resp. CLED-Urifekt, Urotube Roche).] *Urologe (Ausg A)* 1974;**13**:51	Uricult	Unclear
Berbik I, Lampe L, Orosz Toth M. Diagnostic use of the URICULT test in urinary tract infection infections pregnancy (Hungarian). [Hungarian]. *Orvosi Hetilap* 1975;**116**:1403–6	Uricult	Unclear
Havlik I. [Screening of asymptomatic bacteriuria in pregnant women by means of Uricult (author's transl).] *Ceskoslov Gynekol* 1975;**40**:581–3	Uricult	Unclear
Ellner PD, Papachristos T. Detection of bacteriuria by dip-slide. Routine use in a large general hospital. *Am J Clin Pathol* 1975;**63**:516–21	Uricult	2
Wencel J, Dzierzanowska D. Correlation of results of quantitative urine analysis by the method of Hoeprich and by the dip method, using the Uricult set (Polish). [Polish]. *Pol Tyg Lek* 1975;**30**:107–8	Uricult	2
Novakova M, Petracek E. [Personal experience with Uricult.] *Zdravotn Pracovn* 1975;**25**:651–3	Uricult	Unclear
Berbik I, Lampe L, Orosz TM. [The Uricult test in the diagnosis of urinary tract infections in pregnancy.] *Orvosi Hetilap* 1975;**116**	Uricult	Unclear
Cvoric A, Zecevic B, Nikolic V, Markovic M. [Determination of bacteriuria by means of Uricult method.] *Srp Arh Celok Lek* 1976;**104**:145–9	Uricult	Unclear
Tepavcevic P, Burka E, Jeremic D, Fele D, Beric M. [Comparative studies on the value of the Uricult technic in the estimation of the number of bacteria in urine.] *Med Pregled* 1976;**29**:513–17	Uricult	Unclear
Duerden BI, Moyes A. Comparison of laboratory methods in the diagnosis of urinary tract infection. *J Clin Pathol* 1976;**29**:286–91	Uricult	Unclear
Adamczewska K. Applicability of the 'uricult' test in evaluation of significant bacteriuria in pregnant women, especially in cases of EPH toxemia. [Polish]. *Ginek Pol* 1977;**48**:961–6	Uricult	Unclear
Golebiowska M, Chlebna-Sokol D, Kostenko D. Uricult test in urinary tract screening of children aged 6–36 months. [Polish]. *Pediatr Pol* 1977;**52**:1219–22	Uricult	Unclear
Jojart G, Eder I. [Comparative study of urinary nitrite content and Uricult reactions.] *Orvosi Hetilap* 1977;**118**:1975–8	Uricult	Unclear

Study details	Test evaluated	Objective assessed
Bordt J, Beller FK. Is examination of urinary sediment in prenatal check-up still up-to-date?. [German]. *Diagnostik* 1979;**12**:148–9	Uricult	Unclear
Dornbusch K, Lindeberg B, Nord CE, Thunell S. Bacteriuria diagnosis and antibiotic susceptibility testing in a group practice by Dipslide techniques. *Chemotherapy* 1979;**25**:227–32	Uricult	2
Emans SJ, Grace E, Masland Jr RP. Asymptomatic bacteriuria in adolescent girls: II – screening methods. *Pediatrics* 1979;**64**:438–41	Uricult	Unclear
Kjaerulff E, Dybkjaer L, Granlie K, Magnusson B. The diagnosis of urinary infections in general practice. A comparative investigation with Microstix and Uricult. [Danish]. *Ugeskrift Laeger* 1979;**141**:1477–80	Uricult	Unclear
Sebbesen O, Nielsen E. Demonstration of bacteriuria with transport agar. Comparison between Uricult and Urotube. [Danish]. *Ugeskrift Laeger* 1979;**141**:375–6	Uricult	Unclear
Winn WC Jr, Gillenwater JY. Evaluation of Uricult dip slide in two hospital populations. *Urology* 1980;**15**:44–6	Uricult	2
Arbus GS, McCuaig CC, Yeung C, Leers WD. Comparison of the Ontario Ministry of Health dipspoon with Uricult and Microstix-3 as methods of screening for bacteriuria. *Can Med Assoc J* 1981;**124**:48–50	Uricult	Unclear
Ferry S, Burman LG, Holm SE. Uricult and Sensicult dipslides for diagnosis of bacteriuria and prediction of drug resistance in primary health care. *Scand J Prim Health Care* 1989;**7**:123–8	Uricult	Unclear
Lorentzon S, Hovelius B, Miorner H, Tendler M, Aberg A. The diagnosis of bacteriuria during pregnancy. *Scand J Prim Health Care* 1990;**8**:81–3	Uricult	2
Cid E, Fernandez Seara MJ, Buznego R, Pavon P, Rodrigo E, Castro-Gago M. Comparative study between Uricult and urine culture for the diagnosis of urinary infections in infants. [Spanish]. *Revi Espanola Pediatr* 1992;**48**:23–5	Uricult	2
Villanustre Ordonez C, Buznego Sanchez R, Rodicio Garcia M, Rodrigo Saez E, Fernandez Seara MJ, Pavon Belinchon P, *et al.* Comparative study of semiquantitative methods (leukocytes, nitrite test and uricult) with urine culture for the diagnosis of urinary tract infection during infancy. [Spanish]. *Anal Espanoles Pediatr* 1994;**41**:325–8	Uricult	Unclear
Dalet F, Segovia T. Evaluation of a new agar in Uricult-Trio for rapid detection of *Escherichia coli* in urine. *J Clin Microbiol* 1995;**33**:1395–8	Uricult trio	Unclear
Larinkari U, Rautio M. Evaluation of a new dipslide with a selective medium for the rapid detection of beta-glucuronidase-positive *Escherichia coli*. *Eur J Clin Microbiol Infect Dis* 1995;**14**:606–9	Uricult trio	Unclear
Andreu A, Xairo D. [Evaluation of a new method for urine screening based on the study of catalase.] *Enfermed Infec Microbiol Clin* 1991;**9**:162–4	Uriscreen	Unclear
Pezzlo MT, Amsterdam D, Anhalt JP, Lawrence T, Stratton NJ, Vetter EA, *et al.* Detection of bacteriuria and pyuria by URISCREEN a rapid enzymatic screening test. *J Clin Microbiol* 1992;**30**:680–4	Uriscreen	Unclear
Dalton MT, Comeau S, Rainnie B, Lambert K, Forward KR. A comparison of the API Uriscreen with the Vitek Urine Identification-3 and the leukocyte esterase or nitrite strip as a screening test for bacteriuria. *Diagnost Microbiol Infect Dis* 1993;**16**:93–7	Uriscreen	Unclear
Nauschuetz WF, Harrison LS, Trevino SB, Becker GR, Benton J. Two rapid urine screens for detection of bacteriuria: an evaluation. *Curr Microbiol* 1993;**26**:43–5	Uriscreen	Unclear
Hagay Z, Levy R, Miskin A, Milman D, Sharabi H, Insler V. Uriscreen, a rapid enzymatic urine screening test: useful predictor of significant bacteriuria in pregnancy. *Obstet Gynecol* 1996;**87**:410–13	Uriscreen	Unclear
Palmer LS, Richards I, Kaplan WE. Clinical evaluation of a rapid diagnostic screen (URISCREEN) for bacteriuria in children. *J Urol* 1997;**157**:654–7	Uriscreen	Unclear
Waisman Y, Zerem E, Amir L, Mimouni M. The validity of the uriscreen test for early detection of urinary tract infection in children. *Pediatrics* 1999;**104**:e41	Uriscreen	2

Studies excluded after full-text assessment

Study details	Test	Reason for exclusion
Aspevall O, Kjerstadius T, Lindberg L, Hallander H. Performance of Uricult Trio assessed by a comparison method and external control panels in primary healthcare. *Scand J Clin Lab Invest* 2000;**60** Aspevall O, Forsum U, Kjerstadius T, Hallander H. Evaluation of two methods for improving quality of diagnosis of bacteriuria by culture in primary healthcare. *Scand J Clin Lab Invest* 2000;**60**	Uricult Trio	Technical performance; data not reported on relevant outcomes
Cordoba G, Holm A, Hansen F, Hammerum AM, Bjerrum L. Prevalence of antimicrobial resistant *Escherichia coli* from patients with suspected urinary tract infection in primary care, Denmark. *BMC Infect Dis* 2017;**17**	N/A	Did not evaluate POCT of interest
Dilek AR, Dereci S, Ozkasap S, Sahin K. Validity of urine and blood tests for detection of urinary tract infections in children. *Cocuk Enfeksiyon Dergisi* 2014;**8**	N/A	Did not evaluate POCT of interest
DRKS00017273. *Management of UTI in German Primary Care: Feasibility of FLEXICULT™ (MAFL).* 2019. URL: www.drks.de/DRKS00017273	Flexicult	Feasibility study; single-arm study
Espinoza J, Michelli E, De Donato M. Frequency and antibiotic susceptibility of enterobacteria isolated from urocultures in communities of Sucre State during 2005–2006. [Spanish]. *Salus* 2009;**13**	Uricult	Prevalence study – not evaluation of test
Frimodt-Moller N, Espersen F. Evaluation of calibrated 1 and 10 microl loops and dipslide as compared to pipettes for detection of low count bacteriuria in vitro. *APMIS* 2000;**108**	Uricult	Analytical validity
Jameson M, Edmunds Otter M, Williams C, Modha D, Lim F, Conroy SP. Which near-patient tests might improve the diagnosis of UTI in older people in urgent care settings? A mapping review and consensus process. *Eur Geriatr Med* 2019;**10**	N/A	Not a primary study (mapping review) References were checked to identify[47]
Kollerup I, Aagaard Thomsen AK, Kornum JB, Paulsen KI, Bjerrum L, Hansen MP. Use and quality of point-of-care microscopy, urine culture and susceptibility testing for urinalysis in general practice. *Scand J Prim Health Care* 2022;**40**	Flexicult SSI	Analytical validity
KU Leuven. 2015. *Urinary Tract Infections in Older Persons Admitted to a Psychogeriatric Ward*. NCT02368847. 2015. URL: http://clinicaltrials.gov/show/NCT02368847 (accessed November 2022)	Uricult	Trial record only: insufficient data for analysis following author contact
Olsen BE, Hinderaker SG, Lie RT, Gasheka P, Baerheim A, Bergsjo P, *et al.* The diagnosis of urinary tract infections among pregnant women in rural Tanzania; prevalences and correspondence between different diagnostic methods. *Acta Obstet Gynecol Scand* 2000;**79**	Uricult	Agreement with dipstick tests – no reference standard and no other outcomes
Scarparo C, Piccoli P, Ricordi P, Scagnelli M. Evaluation of the DipStreak, a new device with an original streaking mechanism for detection, counting, and presumptive identification of urinary tract pathogens. *J Clin Microbiol* 2002;**40** Schaeffer AJ. Evaluation of the DipStreak, a new device with an original streaking mechanism for detection, counting, and presumptive identification of urinary tract pathogens. *J Urol* 2003;**169**	Dipstreak	No reference standard for evaluation of accuracy
Wigton RS. The Uriscreen test was not better than standard urinalysis and dipstick tests for detecting urinary tract infection in children. *Evid Based Med* 2000;**5**	Uriscreen	Not a primary study – secondary report of existing study that was excluded due to publication date of 1999

Studies included in manufacturer's submission that did not meet inclusion criteria

Study details	Document type	Manufacturer	Test evaluated	Reason for exclusion
Baltekin Ö, Boucharin A, Tano E, Andersson DI, Elf J. Antibiotic susceptibility testing in <30 minute using direct single-cell imaging. *Proc Nat Acad Sci* 2017;**114**	Journal article – including supporting information	Astrego	PA-100 AST System	Exclude – population; analytical validity based on known samples
Baltekin Ö, Hammar P, Kovachev P, Myzithra M, Wistrand-Yuen E. *Reproducibility of Fully Automated AST for Direct Near Patient Testing*. Poster presentation, ECCMID, 23–26 April 2022, Lisbon, Portugal	Poster	Astrego	PA-100 AST System	Exclude – population; analytical validity based on known samples
Sysmex Europe SE. *How to Perform Real-time Antimicrobial Susceptibility Testing (AST).* 2022. URL: www.sysmex-europe.com/fileadmin/media/f100/Academy/Documents/Whitepaper/Nanofluidics_Whitepaper_EN_01.pdf (accessed October 2022)	Web page	Astrego	AST testing	General discussion page
Llusern Scientific. *UTI Test Kit: Instructions For Use* [test insert]. (accessed January 2023)	Test package information	LLusern	Lodestar DX analyser and Llusern UTI test kit	Package insert for the test
Safarika A, Wacker JW, Katsaros K, Solomonidi N, Giannikopoulos G, Kotsaki A, *et al.* A 29-mRNA host response test from blood accurately distinguishes bacterial and viral infections among emergency department patients. *Intensive Care Med Exp* 2021;**9**	Journal article	Triverity	Inflammatix Classifier (InSep)	Population – not UTI
Bauer W, Kappert K, Galtung N, Lehmann D, Wacker J, Cheng HK, *et al.* A novel 29-messenger RNA host-response assay from whole blood accurately identifies bacterial and viral infections in patients presenting to the emergency department with suspected infections: a prospective observational study. *Crit Care Med* 2021;**49**	Journal article	Triverity	Inflammatix Classifier (InSep)	Population – not UTI
Galtung N, Diehl-Wiesenecker E, Lehmann D, Markmann N, Bergström WH, Wacker J, *et al.* Prospective validation of a transcriptomic severity classifier among patients with suspected acute infection and sepsis in the emergency department. *Eur J Emerg Med* 2022;**29**	Journal article	Triverity	Inflammatix Classifier (InSep)	Population – not UTI
Kostaki A, Wacker JW, Safarika A, Solomonidi N, Katsaros K, Giannikopoulos G, *et al.* A 29-mrna host response whole-blood signature improves prediction of 28-day mortality and 7-day intensive care unit care in adults presenting to the emergency department with suspected acute infection and/or sepsis. *Shock* 2022;**58**	Journal article	Triverity	Inflammatix Classifier (InSep)	Population – not UTI
Brakenridge SC, Starostik P, Ghita G, Midic U, Darden D, Fenner B, *et al.* A transcriptomic severity metric that predicts clinical outcomes in critically ill surgical sepsis patients. *Crit Care Explor* 2021;**3**	Journal article	Triverity	Inflammatix Classifier (InSep)	Population – not UTI
Brakenridge SC, Chen U, Loftus T, Ungaro R, Dirain M, Kerr A, *et al.* Evaluation of a multivalent transcriptomic metric for diagnosing surgical sepsis and estimating mortality among critically ill patients. *JAMA Netw Open* 2022;**5**	Journal article	Triverity	Inflammatix Classifier (InSep)	Population – not UTI

continued

Study details	Document type	Manufacturer	Test evaluated	Reason for exclusion
Moore AR, Roque J, Shaller BT, Asuni T, Remmel M, Rawling D, *et al.* Prospective validation of an 11-gene mRNA host response score for mortality risk stratification in the intensive care unit. *Sci Rep* 2021;**11**	Journal article	Triverity	Inflammatix Classifier (InSep)	Population – not UTI
He YD, Wohlford EM, Uhle F, Buturovic L, Liesenfeld O, Sweeney TE. The optimization and biological significance of a 29-host-immune-mRNA panel for the diagnosis of acute infections and sepsis. *J Person Med* 2021;**11**	Journal article	Triverity	Inflammatix Classifier (InSep)	General discussion paper on optimisation
Schneider JE, Romanowsky J, Schuetz P, Stojanovic I, Cheng HK, Liesenfeld O, *et al.* Cost impact model of a novel multi-mRNA host response assay for diagnosis and risk assessment of acute respiratory tract infections and sepsis in the emergency department. *J Health Econ Outcome Res* 2020;**7**	Journal article	Triverity	Inflammatix Classifier (InSep)	Cost impact model
Mayhew MB, Midic U, Choi K, Khatri P, Buturovic LJ, Sweeney TE, editors. Towards equitable patient subgroup performance by gene-expression-based diagnostic classifiers of acute infection. *medRxiv* 2022	Preprint	Triverity	Inflammatix Classifier (InSep)	General discussion paper: not a primary evaluation of tests
Uricult. *Test Package Information*. 2019 (accessed January 2023)	Test package information	Uricult	Uricult	Package insert for the test
Uricult. *Test Package Information*. 2019 (accessed January 2023)	Test package information	Uricult	Uricult Plus	Package insert for the test
Uricult. *Test Package Information*. 2022 (accessed January 2023)	Test package information	Uricult	Uricult Trio	Package insert for the test
UTRiPLEX. *Rapid Urine Test for Urinary Tract Infection. Instructions for Use*. September 2023 (accessed January 2023)	Test package information	Utriplex	UTRiPLEX test assay	Package insert for the test

Appendix 3 Data extraction tables

Objective 1

Baseline details

Study details	Participants	POCT details	Group 1	Control
Author (year): Butler (2018)[8,85,86] **Study name:** POETIC trial **Country:** England, Netherlands, Spain and Wales **Study design:** RCT (individual randomised) **Recruitment:** July 2013–August 2014 **Funding:** European Commission Seventh Framework Programme **Setting:** Primary care	**Population:** Women aged ≥ 18 years – uncomplicated UTI **Inclusion criteria:** Presenting to primary care with any of the following symptoms: dysuria, urgency or frequency with clinical diagnosis of uncomplicated UTI **Exclusion criteria:** Suspected pyelonephritis; long-term antibiotic treatment; antibiotics for UTI in preceding 4 weeks; significant genitourinary tract abnormalities; terminal illness **Number of eligible patients (randomised):** 654 (653) **Age:** 47.6 years (SD 27.6) **Sex:** All female	Flexicult SSI Urinary Kit (SSI Diagnostica, Denmark) Urine poured onto agar plate and incubated overnight in desktop incubator in GP practice. Results reviewed after 18–24 hours Flexicult plates specific for antibiotics most commonly used in three participating regions **Sample collection:** Urine samples collected using Peezy midstream urine collection kit. Flexicult group, urine sample split – portion kept for intervention test; rest sent for culture	Flexicult SSI Urinary Kit (SSI Diagnostica, Denmark) to guide management GPs could decide how best to use the test. Examples of how it could be used include: • Determine whether, and what antibiotic class, to prescribe the following day • Prescribe empirically and use the test to aid in a next-day review of initial prescribing decision • Provide delayed antibiotics prescription and use the test to guide use of delayed prescription	Care informed by national guidelines; clinicians received summary of relevant national treatment guidelines

Study details	Participants	POCT test details	Group 1	Group 2
Author (year): Holm (2017)[35,87,88] **Study name:** N/A **Country:** Denmark **Study design:** RCT (individual randomised) **Recruitment:** March 2015–May 2016 **Funding** a. 2016, the University of Copenhagen b. Læge Sofus Carl Emil Friis og Hustru Olga Doris Friis' legat c. SSI Diagnostika (materials) **Setting:** Primary care	**Population:** Women aged ≥ 18 years – uncomplicated UTI **Inclusion criteria:** Presenting to GP with dysuria, frequency or urgency, for ≤ 7 days for which the GP suspected uncomplicated UTI, including elderly patients >65 years, patients with recurrent UTI and patients with orally treated diabetes without complications **Exclusion criteria:** Negative dipstick analysis on both leucocytes and nitrites, serious comorbidities, former participation in the study and patients presenting on a Friday (as point-of-care culture is read the following day) **Number of eligible patients (randomised):** Unclear (376) **Age:** Not reported **Sex:** All female	Flexicult SSI – intervention group including susceptibility testing All patients had to wait until following day for result of POCT before starting treatment Urine sample split – portion kept for POCT; rest sent for culture	POCT culture plus susceptibility testing – Flexicult SSI Urinary Kit (SSI Diagnostica, Denmark) Treatment based on test results	POCT culture alone – ID Flexicult (SSI Diagnostica, Denmark) Treatment based on test results

ResuLTS

Study	Outcome	Definition	Group 1		Group 2		Effect measure – estimate (95% CI)
			n	%	*n*	%	
Butler (2018)[8,87,88]	Concordant antibiotic use	Consumption of antibiotic on day 3 (or day 2 for fosfomycin) that pathogen considered to be causing UTI was sensitive to OR no antibiotic use if did not have UTI	153	60.7	137	55.9	OR 0.84 (0.58 to 1.20)
	Antibiotic prescribing at initial consultation		267	82.4	282	88.4	OR 0.56 (0.35 to 0.88)
	Antibiotics prescribed to guidelines at initial consultation		156	58.9	166	59.5	OR 0.99 (0.67 to 1.45)
	Patient enablement	Measured using Patient Enablement Instrument at day 14 and 3 months[37]	171	70.1	177	69.7	OR 0.99 (0.66 to 1.48)
	Antibiotic consumed day 3	NR	217	79.2	200	76.6	OR 1.24 (0.81 to 1.89)
	Antibiotic consumed (during 2 weeks)	NR	234	85.1	217	81.6	OR 1.38 (0.87 to 2.19)
	New antibiotic prescription (within 2 weeks)	NR	33	10.3	30	9.7	OR 1.11 (0.65 to 1.89)
	Re-consultation (within 2 weeks)	NR	41	12.9	41	13.2	OR 0.99 (0.62 to 1.60)
	Hospital stay (within 2 weeks)	NR	3	0.9	4	1.3	Numbers too small
	Microbiologically confirmed UTI (at 2 weeks)	NR	20	8.7	20	9.2	OR 0.94 (0.49 to 1.81)
	Recurrence of UTI within 3 month period	NR	54	17	69	22.3	OR 0.72 (0.48 to 1.07)
	Duration of symptoms	NR	N/A	N/A	N/A	N/A	HR 1.02 (0.83 to 1.25)
	Duration of moderately bad symptoms	NR	N/A	N/A	N/A	N/A	HR 0.98 (0.82 to 1.17)
	Overall urinary symptom burden	NR	N/A	N/A	N/A	N/A	MD 0.99 (0.84 to 1.19)
	Management changed as result of Flexicult	NR	190	63.1	N/A	N/A	N/A
	Change of management	Did not start antibiotic	14	7.4	N/A	N/A	N/A
		Stopped taking antibiotic	10	5.3	N/A	N/A	N/A
		Started taking antibiotic	29	15.3	N/A	N/A	N/A
		Continued with antibiotic	63	33.2	N/A	N/A	N/A
		New antibiotic prescribed	74	38.9	N/A	N/A	N/A
	Time to perform test	Prepare test	N/A	N/A	N/A	N/A	9 minutes
		Obtain and record result	N/A	N/A	N/A	N/A	6 minutes
		Discuss result with patient	N/A	N/A	N/A	N/A	7 minutes
	Cost	Cost per person, including POCT cost in UK	N/A	N/A	N/A	N/A	£48

Study	Outcome	Definition	Group 1		Group 2		Effect measure – estimate (95% CI)
			n	%	*n*	%	
Holm (2017)[35,87,88]	Appropriate prescribing	Prescription of a first-line antibiotic to which the infecting pathogen was susceptible, if the individual was found to have UTI in the reference Prescription of a second-line antibiotic, if the individual had UTI but was allergic to the antibiotic or the pathogen was resistant to all first-line antibiotics No antibiotic prescription if the individual was found to not have UTI in the reference	120	67	121	75	OR 1.44 (1.03 to 1.99)
	Symptom free on day 5	NR	NR	NR	NR	NR	OR 0.91 (0.56 to 1.49)
	No significant bacteriuria on day 14	NR	NR	NR	NR	NR	OR 1.15 (0.62 to 2.13)

Risk of bias

Domain	Concerns	Rationale
Identify the trial you are examining: POETIC: Butler et al. (2018)[8,85,86]		
Risk of bias arising from the randomisation process	*Low concerns*	Online central randomisation with allocation concealed – allocation sent electronically once randomisation details entered. Groups comparable at baseline
Risk of bias due to deviations from the intended interventions	*Low concerns*	Pragmatic trial. Blinding not possible due to nature of the intervention; the clinician and patient need to be aware whether they are in the Flexicult arm so that they can act on the Flexicult result. No evidence of deviations from intended interventions, and this would be very difficult given nature of the intervention. Both per-protocol and intention-to-treat analyses reported (as sensitivity analysis)
Risk of bias due to missing outcome data	Low concerns	Large proportion of missing data; proportion similar between groups, no evidence of difference between those with and without missing data and ITT analysis confirmed conclusions Baseline data available on 324/329 randomised in intervention group and 319/325 randomised in control group. Data for primary outcome required each participant to have 2-week diary and urinalysis data available 252/329 in intervention group were included in analysis for primary outcome 245/325 in control group were included in analysis for primary outcome
Risk of bias in measurement of the outcome	Low concerns	Outcome assessors were not blinded. However, outcome is based on antibiotic use, which is objective and not likely to be influenced by outcome assessor
Risk of bias in selection of the reported result	Low concerns	Protocol available; outcomes specified in protocol reported in results
Overall	**Low concerns**	**No concerns identified for any domain**
Identify the trial you are examining: Holm (2017)[35,87,88]		
Risk of bias arising from the randomisation process	*Low concerns*	The randomisation code was produced by an online random number generator as permuted block randomisation in blocks of 10 by the investigators. The allocation of each included patient was placed in an opaque, sequentially numbered, sealed envelope, which was opened in general practice after inclusion of the patient

continued

Domain	Concerns	Rationale
Risk of bias due to deviations from the intended interventions	*Low concerns*	Pragmatic trial. Blinding not possible due to nature of the intervention; the clinician and patient need to be aware whether they are in the Flexicult arm so that they can act on the Flexicult result. Six patients in the culture-only group had the wrong test performed (culture and susceptibility testing). Both per-protocol and intention-to-treat analysis reported (as sensitivity analysis)
Risk of bias due to missing outcome data	Low concerns	Small proportion of missing data; proportion similar between groups, no evidence of difference between those with and without missing data 13 patients excluded from the analysis: 8 in intervention group and 5 in control. Reasons for exclusion included consent withdrawn (n = 2), did not fulfil inclusion criteria (n = 7), other (n = 4)
Risk of bias in measurement of the outcome	Low concerns	Outcome assessors were not blinded. However, outcome is based on antibiotic use, which is objective and not likely to be influenced by outcome assessor
Risk of bias in selection of the reported result	Low concerns	Protocol available; outcomes specified in protocol reported in results
Overall	**Low concerns**	**No concerns identified for any domain**

Objective 2

Baseline details

Study details	Participants	POCT details	Reference standard
Anacleto (2009)[45] **Country:** Philippines **Language:** English **Funding:** Institute of Child Health and Human Development of the National Institutes of Health, Manila, Philippines, the Philippine Society of Nephrology, Inc., and Pediatric Associates, Inc	**Setting and population:** Secondary care; uncomplicated UTI; age < 16 years **Inclusion criteria:** Infants and children aged 0–7 years with symptoms suggestive of UTI and positive LE or nitrite dipstick test **Exclusion criteria:** Poor intake of antibiotics; obstructive uropathy; congenital anomalies of kidneys urinary tract; midline defects; failure to thrive; concomitant infections; recurrent UTI; asymptomatic bacteriuria; other comorbid conditions **Number included (number analysed):** 200 (200) **Age:** 4 months to 7 years **% Female:** 43	**Urine sampling method:** Samples were obtained from clean-voided midstream urine, supervised by a trained physician. In subjects from whom clean catch was difficult, urethral catheterisation was performed **Target condition:** Presence of UTI **Location of test performance:** Outpatient department **POCT**: Uricult Trio – Dipslide unscrewed from the tube without being allowed to touch the agar surfaces. Holding Uricult Trio by the cap, the operator dipped the slide into the urine sample so that the agar surfaces were totally immersed. Excess urine allowed to drain from the slide. The last drops were blotted on absorbent paper. The slide was screwed tightly back into the tube and placed upright in an incubator (36 ± 2°C) for 24 hours **Threshold:** ≥ 10^4 CFU	**Reference standard:** Culture – standard laboratory culture **Threshold:** ≥ 10^4 CFU
Blom (2002)[39] **Country:** Denmark **Language:** English **Funding:** Not reported	**Setting and population:** Primary care – mixed symptomatic patients **Inclusion criteria:** 19 GPs were asked to use Flexicult in addition to standard diagnostic procedures in patients with symptoms of UTI **Exclusion criteria:** Not reported **Number included (number analysed):** 121 **Age:** NR **% Female:** NR	**Urine sampling method:** Not reported **Target condition:** Presence of UTI; antimicrobial resistance **Location of test performance:** GP surgery – *field trial* **POCT:** Flexicult SSI Urinary Kit – suspensions of bacteria diluted in 50 ml of sterile urine to various concentrations. Each suspension was poured into a Flexicult SSI Urinary Kit for 1–2 seconds and then incubated overnight at 35°C **Threshold:** > 10^5 for UTI diagnosis; growth on kit for antimicrobial resistance	**Reference standard:** Culture; bacteria growing on the Flexicult SSI Urinary Kit had their MIC values for trimethoprim, sulfamethoxazole, ampicillin, nitrofurantoin and mecillinam determined according to NCCLS guidelines using standard procedures[89] **Threshold:** > 10^5 for UTI diagnosis; MIC concentration

Study details	Participants	POCT details	Reference standard
Bongard (2015)[19] **Country:** Wales **Language:** English **Funding:** Medical Research Council, Cardiff University, European Community's Seventh Framework Programme, R-GNOSIS consortium	**Setting and population:** Laboratory based; mixed **Inclusion criteria:** Fresh urine samples (within ≈9 hours) submitted from primary and secondary care in course of routine patient care. 124 (62%) from outpatients, 72 (36%) from inpatients and 4 (2%) unknown **Exclusion criteria:** Urine samples collected in boric acid (as this may interfere with the antibiotic sections of Flexicult) and urines < 5 ml volume after routine processing	**Urine sampling method:** Urine sampling MSU (134), catheter (7), unknown (65) *(numbers do not add up)* **Target condition:** Presence of UTI; antimicrobial resistance **Location of test performance:** Laboratory at University Hospital Wales **POCT:** Flexicult SSI Urinary Kit – urine poured to cover all compartments. After ≈5 seconds, excess urine poured off and test was inverted and incubated aerobically overnight at 36 ± 1°C **Threshold:** Antibiotic resistance profile was read if ≥ 10^3 CFU/ml of a clinically significant UTI organism alone or in a predominant quantity.	**Reference standard:** Culture and microscopy and spiral plating in false-positive results only; antimicrobial susceptibility testing performed on significant isolates using the appropriate urine antimicrobial disc set and standard disc diffusion method **Threshold:** If positive on microscopy then culture to confirm. Criteria for positive microscopy: ≥ 5 bacteria, ≥ 100 white blood cells (WBC), ≥ 20,000 ASP (any small particles),
	Number included (number analysed): 211 (200) **Age:** < 18 years to > 65 years – no further details **% Female:** 70	If growth in one antibiotic compartment much lower than in the quantification compartment – or if there is no growth at all – bacterium considered susceptible to the antibiotic	≥ 50 WBC+≥ 2000 ASP, ≥ 50 WBC+≥ 1000 ASP+≥ 3 bacteria, ≥ 3 WBC+≥ 6000 ASP Culture: > 10^5 CFU/ml pure or predominant growth (× 1000) of a clinically significant UTI pathogen
Boon (2022)[43,53] **Country:** Belgium; ERNIE4 study **Language:** English **Funding:** Research Foundation Flanders and by a KU Leuven starting grant	**Setting and population:** Primary care; uncomplicated UTI; age < 18 years **Inclusion criteria:** Age 3 months–18 years; acute illness of maximum 10 days' duration **Exclusion criteria:** Urinary catheter, trauma as main presenting problem, needed referral to hospital at presentation, critically unstable or had taken immunosuppressant medication in previous 30 days or antibiotics in previous 7 days excluded **Number included (number analysed):** 834 (300) **Age:** 5 to 18 years **% Female:** 46	**Urine sampling method:** Mid-stream, clean catch, or adhesive bags as per clinical practice **Target condition:** Presence of UTI **Location of test performance:** One central clinical laboratory (Algemeen Medisch Laboratorium Antwerp) **POCT:** Uriscreen POCT (Savyon Diagnostics Ltd., Ashdod, Israel) – measures bacteria and somatic cells (pyuria, haematuria) in urine by detecting catalase activity **Threshold:** Visual assessment of presence of foam 1–2 minutes after addition of 4 drops of hydrogen peroxide to urine **POCT Test:** Utriplex test (Investigational use, Mologic Ltd, Bedfordshire, UK) – measures three inflammatory markers – HNE, MMP8 and Cystatin C **Threshold:** Visualisation of ≥ 2 test lines after 6 minutes indicates UTI	**Reference standard:** Culture **Threshold:** ≥ 10^5 CFU/ml of a single pathogen

continued

Study details	Participants	POCT details	Reference standard
Colodner (2000)[49] **Country:** Israel **Language:** English **Funding:** Not reported	**Setting and population:** Laboratory based; mixed **Inclusion criteria:** Fresh urine samples from outpatient clinics (74%) and hospitalised patients (26%) **Exclusion criteria:** NR **Number included (number analysed):** 1000 (1000) **Age**: NR **% Female:** NR	**Urine sampling method:** NR **Target condition:** Presence of UTI **Location of test performance:** Microbiology laboratory, Central Emek Medical Center, Afula, Israel **POCT:** Dipstreak – urine culture device (closed system) for isolating and enumerating bacteria in urine. Study used MacConkey agar/CNA combination. Device results in series of streaks of decreasing inoculum concentration that permit isolation of single colonies and then incubated overnight for culture evaluation the next day **Threshold:** Evaluated according to manufacturer's chart. Two thresholds evaluated – 10^4 and 10^5 CFU	**Reference standard:** Culture – standard culture plates – MacConkey Agar, CAN and SBA **Threshold:** Single organism 10^4 CFU or two organisms when colony count of one > 10^5 CFU. Mixed (contaminated) growth of two organisms with counts between 10^4 and 10^5 or three or more different organisms
Greeff (2002)[46] **Country:** South Africa **Language:** English **Funding:** Not reported	**Setting and population:** Antenatal clinics; screening; pregnant women **Inclusion criteria:** Two populations of patients from the Pretoria region were involved: (1) asymptomatic pregnant women attending the antenatal clinic for the first time or presenting in labour; and (2) pregnant women with symptoms suggestive of UTI **Exclusion criteria:** NR **Number included (number analysed):** 453 (374) **Age:** NR **% Female:** 100	**Urine sampling method:** Self-collected mid-stream urine **Target condition:** Presence of UTI **Location of test performance:** Antenatal clinic **POCT:** Uricult Trio – dipped into urine and placed directly in the incubator and incubated for 16–23 hours **Threshold:** > 10^3 CFU/ml	**Reference standard:** Culture – standard laboratory culture **Threshold:** > 10^5 CFU/ml
Holm (2017)[37] **Country**: Denmark; DTA study nested in Danish RCT[35] **Language:** English **Funding:** 2016, (a) University of Copenhagen, (b) Læge Sofus Carl Emil Friis og Hustru Olga Doris Friis' Legat and (c) SSI Diagnostika (materials)	**Setting and population:** Primary care; uncomplicated UTI; women **Inclusion criteria:** Age ≥ 18 years, female, non-pregnant women with symptoms of UTI (dysuria, frequency or urgency) **Exclusion criteria:** Negative dipstick analysis on leucocytes and nitrites, complicated UTI (except uncomplicated diabetes, elderly patients and recurrent UTI), previous participation in the study and patients presenting on a Friday (POCT is read the following day) **Number included (number analysed):** 376 (341) **Age:** 48.5 years **% Female**: 100	**Urine sampling method:** Mid-stream urine sample **Target condition:** Presence of UTI **Location of test performance:** General practice **POCT:** Flexicult SSI Urinary Kit; agar dish consisting of one big well containing agar material and five small wells containing agar with one of five antibiotics GPs registered the index test as 'significant growth of uropathogens', 'no significant growth of uropathogens' or 'inconclusive' **Threshold:** Significant growth prespecified as ≥ 10^3 CFU/ml for any uropathogen. 'Inconclusive' labelled as negative **POCT:** ID Flexicult; chromogenic agar allowing identification and quantification of six types of bacteria **Threshold:** ≥ 10^3 CFU/ml for *E. coli* and *S. saprophyticus*, 10^4 CFU/ml for other typical uropathogens in accordance with European consensus	**Reference standard:** Culture; urine samples sent to reference lab for culture **Threshold:** ≥ 10^3 CFU/ml for *E. coli* and *S. saprophyticus*, ≥ 10^4 CFU/ml for other typical uropathogens, ≥ 10^5 for possible uropathogens in accordance with European consensus

Study details	Participants	POCT details	Reference standard
Hullegie (2017)[36] **Country:** Wales, England, Spain and Netherlands. DTA substudy from POETIC study[8] **Language:** English **Funding:** European Community's Seventh Framework Programme and R-GNOSIS consortium	**Setting and population:** Primary care; uncomplicated UTI; women **Inclusion criteria:** Women randomised to Flexicult arm of POETIC trial; aged ≥ 18 years with symptoms of UTI (dysuria, urgency or frequency) **Exclusion criteria:** Women who were terminally ill, were receiving treatment for life-threatening cancer, were having severe systemic symptoms or had received antibiotics for UTI within the past 4 weeks **Number included (number analysed):** 325 (312) **Age:** 49 years **% Female:** 100	**Urine sampling method:** Mid-stream urine samples collected using urine collection device (Peezy Midstream, Forte Medical) **Target condition:** Presence of UTI; antimicrobial resistance **Location of test performance:** Primary care **POCT:** Flexicult SSI urinary kit **Threshold:** Presence of UTI: 10^3 CFU/ml, pure culture of a urinary tract pathogen ≥ 10^3 CFU/ml, predominant growth of urinary tract pathogen in mixture with normal flora Recorded bacterial growth as none, pure or mixed organism (if mixed then presence of predominant growth). Bacterial quantification assessed the number of colonies (< 15, 15–20, i.e. at or < $10e^3$ CFU/ml, ≥ 20, i.e. 10e3-1035 CFU/ml, semi confluent/confluent, i.e. ≥ $10e^5$ CFU/ml). If bacterial growth ≥ 10^3 CFU/ml of pure/predominant organism, then clinicians were asked to record antibiotic susceptibility	**Reference standard:** Culture **Threshold: Three thresholds evaluated:** (1) *PHE/HPA definition*: ≥ 10^4 CFU/ml pure culture of pathogen; ≥ 10^5 CFU/ml mixed growth with one predominant pathogen; OR ≥ 10^3 CFU/ml of *E. coli* or *S. saprophyticus*; (2) *UK laboratory definition*: ≥ 10^5 CFU/ml pure culture of uropathogen OR ≥ 10^5 CFU/ml predominant culture a uropathogen with 3-log difference between highest and next species; (3) *European definition*: ≥ 10^3 CFU of uropathogen
Lee (2010)[47] **Country:** Republic of Korea **Language:** Korean – extracted using Google Translate **Funding:** Not reported	**Setting and population:** Secondary care; uncomplicated UTI; age < 24 months **Inclusion criteria:** Febrile infants aged< 24 months who attended outpatient department **Exclusion criteria:** Last dose of antibiotics < 48 hours **Number included (number analysed):** 158 **Age:** 15 months **% Female**: 46	**Urine sampling method:** Mid-stream urine or urine collection bags **Target condition:** Presence of UTI; presence of UTI – caused by *E. coli* **Location of test performance:** Outpatient setting **POCT:** Uricult Trio – composed of green CLED medium, reddish-brown MacConkey medium, and colourless *E. coli* medium. Compared against colony density chart for interpretation. Read at next outpatient clinic **Threshold:** > 10^5 CFU	**Reference standard:** Culture **Threshold:** ≥ 10^5 CFU single bacterium; ≥ 10^4 CFU/ml in patients with symptoms
Macias (2002)[44] **Country:** Mexico **Language:** Spanish **Funding:** NR	**Setting and population:** ICU; indwelling catheter **Inclusion criteria:** Hospitalised adults; indwelling catheter **Exclusion criteria:** Recognised history of recent or recurrent UTI. Severe immunosuppression **Number included (number analysed):** 57 patients, 108 samples **Age:** NR **% Female:** NR	**Urine sampling method:** From catheter – took 3–5 ml per puncture of the probe. Samples taken every 72 hours **Target condition:** Presence of UTI **Location of test performance:** Not reported but likely in hospital **POCT:** Uriscreen – 2 ml of urine placed in tube with catalyst, to which four drops of H20 added. After mixing gently for 5 seconds, formation of foam observed on surface of mixture **Threshold:** Formation of foam according to manufacturer's specifications, in addition to this classification: 1. +foam ring on surface with clear centre 2. ++ foam band < 1mm covering the entire surface 3. +++ foam band > 1mm	**Reference standard:** Culture **Threshold:** 10^3 CFU/ml

continued

Study details	Participants	POCT details	Reference standard
Mignini (2009)[48] **Country:** Argentina **Language:** English **Funding:** Supported by UNDP/UNFPA/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction	**Setting and population:** Antenatal clinics; screening Pregnant women **Inclusion criteria:** All women attending antenatal clinics who presented with live fetuses at gestational weeks 12–35 **Exclusion criteria:** Underlying disease that required continuous steroid or antibiotic treatment; use of antibiotics before assessment; treatment for UTI at any time during pregnancy; history of nitrofurantoin hypersensitivity; symptoms suggesting symptomatic UTI; previous negative urine culture or culture positive with organism resistant to nitrofurantoin **Number included (number analysed):** 3048 (3047) **Age:** NR **% Female:** 100%	**Urine sampling method:** Clean catch mid-stream urine sample in sterile container. Sample divided into three aliquots for testing with index test(s) and reference standard **Target condition:** Presence of UTI **Location of test performance:** Central Laboratory (Department of Public Health of the Municipality of Rosario) **POCT:** Uricult – Dipslides inoculated by dipping the agar-coated slides into the urine and incubated at 37°C for 24 hours. Results were determined by comparison of the microbial density on the slide with a model chart provided by the manufacturer **Threshold:** ≥ 10^5 CFU/ml or higher of a single microorganism or when two different colonies were present but one was ≥10^5 CFU/ml	**Reference standard:** Culture Classic quantitative culturing in the microbiology laboratory **Threshold:** ≥ 10^5 CFU/ml of a single potential uropathogen or of two organisms not consistent with kin flora were isolated
Millar (2000)[41] **Country:** USA (Hawaii) **Language:** English **Funding:** Supported by a Research Centers in Minority Institutions award, from the National Center for Research Resources, National Institutes of Health	**Setting and population:** Antenatal clinics; screening Pregnant women **Inclusion criteria:** Pregnant women screened for bacteriuria at initial prenatal visits **Exclusion criteria:** NR **Number included (number analysed):** 383 (378) **Age:** NR **% Female:** 100	**Urine sampling method:** Clean catch mid-stream urine **Target condition:** Presence of UTI **Location of test performance:** Antenatal clinic **POCT:** Uriscreen – 2 ml of urine poured into a test tube containing Uriscreen reagent powder. Four drops of Uriscreen 10% hydrogen peroxide solution were added to each test tube and mixed gently for 5 seconds. The specimen was monitored for 2 minutes for foam formation **Threshold:** Test was considered positive if foam was generated and formed a continuous ring along the test tube wall or layer on the surface of the liquid. Test was considered negative if no foam was generated or the ring of foam was incomplete at the end of 2 minutes	**Reference standard:** Culture – standard laboratory culture **Threshold:** ≥ 10^4 CFU/ml of single potential uropathogen. Cultures were considered negative if < 10^4 CFU/ml of a single pathogen or any non-uropathogenic bacteria were isolated Cultures were considered contaminated if multiple organisms were identified with at least one potential uropathogen
Pernille (2019)[40,54] **Country:** Denmark **Language** English **Funding:** University of Copenhagen, 2016 funds, and The PLU fond (Praktiserende Laegers Undervisningsfond)	**Setting and population:** Primary care; uncomplicated UTI Women **Inclusion criteria:** Women aged ≥ 18 years; presenting with one or more symptoms of UTI (dysuria, frequency or urge) **Exclusion criteria:** Pregnant; recent bladder surgery; urinary tract abnormality **Number included (number analysed):** 122 (117) **Age:** Sample include age < 30 years to > 61 years **% Female**: 100	**Urine sampling method:** First void urine sample in one cup and mid-stream urine sample in second cup. Results reported for mid-stream urine analysis **Target condition:** Presence of UTI **Location of test performance:** Primary care **POCT:** ID Flexicult **Threshold:** > 5 colonies (corresponds to 10^3 CFU/ml) of a primary uropathogen or > 50 colonies (corresponds to 10^4 CFU/ml) of a secondary uropathogens	**Reference standard:** Culture – standard laboratory culture **Threshold:** ≥ 10^3 CFU/ml for *E. coli* and *S. saprophyticus*, ≥ 10^4 CFU/ml for other typical uropathogens and ≥ 10^5 CFU/ml for possible uropathogens. Growth of more than two different colonies (mixed cultures) considered as non-significant growth

Study details	Participants	POCT details	Reference standard
Teppa (2005)[42] **Country:** Venezuela **Language:** English **Funding:** Not reported	**Setting and population:** Antenatal clinics; screening; pregnant women **Inclusion criteria:** Pregnant women who had routine prenatal screening for asymptomatic bacteriuria **Exclusion criteria:** Patients with urinary symptoms, active vaginal bleeding, or previously on antibiotics therapy were excluded from the study **Number included (number analysed):** 150 (150) **Age:** 27.3 **% Female:** 100	**Urine sampling method:** Catheterised urine samples – first morning urine samples **Target condition:** Presence of UTI **Location of test performance:** Maternal-Fetal Unit of the Department of Obstetrics and Gynaecology **POCT Test:** Uriscreen – 2 ml of urine poured into test tube containing Uriscreen reagent powder. Four drops of Uriscreen 10% hydrogen peroxide solution were added to each test tube and mixed gently for 5 seconds. The specimen was monitored for 2 minutes for foam formation **Threshold:** Considered positive if foam was generated and formed a continuous ring along the test tube wall or layer on the surface of the liquid. The test was considered negative if no foam was generated or the ring of foam was incomplete at the end of 2 minutes	**Reference standard:** Culture – standard laboratory culture **Threshold:** ≥ 10^5 CFU/ml of single pathogen or any non-uropathogneic bacteria. Contaminated if multiple organisms identified
Van der Goes (2023)[51] **Country:** Wales **Language:** English **Funding:** Llusern	**Setting and population:** Laboratory; mixed **Inclusion criteria:** Fresh samples collected by Public Health Wales **Exclusion criteria:** NR **Number included (number analysed):** 144 fresh urine samples **Age:** NR **% Female:** NR	**Urine sampling method:** NR – < 2 days **Target condition:** Presence of *E. coli* **Location of test performance:** Laboratory **POCT:** Lodestar DX (Llusern Scientific) UTI test kit containing: (1) assay panel with individual LAMP reactions for six common uropathogens. Each LAMP reaction consists of a proprietary mix of isothermal mastermix, primers and an intercalating dye; (2) novel real-time LAMP analyser (Lodestar DX). Amplification detected as a clear and steep increase of fluorescence < 40 minutes. An embedded algorithm was used to call a sample positive, negative or inconclusive **Threshold:** N/A	**Reference standard:** Culture **Threshold:** NR
Yagupsky (2000)[50] **Country:** Israel **Language:** English **Funding:** Not reported	**Setting and population:** Laboratory based; uncomplicated UTI **Inclusion criteria:** Fresh urine samples from 251 hospitalised patients and 819 outpatients **Exclusion criteria:** NR **Number included (number analysed):** 1070 (1070) **Age:** NR **% Female:** NR	**Urine sampling method:** Mid-stream urine samples **Target condition:** Presence of UTI; pathogenic cause **Location of test performance:** Laboratory **POCT:** Dipstreak – performed using the Uriselect three blood agar configuration, following the manufacturer's instructions. If no growth was observed or the colony count was < 10 CFU, plates and Dipstreak devices were reincubated for 24 hours to exclude false-negative results caused by insufficient incubation **Threshold:** NR; may have been same as reference standard but not clear	**Reference standard:** Culture Standard laboratory culture **Threshold:** ≥ 10^5 CF/ml of single organism or a mixed culture of 10^5 CFU/ml of one uropathogen and < 10^3 CFU/ml of other organisms accompanied by non-significant growth of other bacteria. Growth of 10^4–10^5 CFU/ml of one or two organisms indicated the need for a repeat culture

HNE, 4-Hydroxynonenal; MMP8, matrix metalloproteinase-8.

Results

Study details	Population and setting	POCT	Reference standard	Target condition	TP	FP	FN	TN	Sensitivity	Specificity	Missing samples/notes
Blom (2002)[39]	**Population:** Mixed symptomatic **Location of test performance:** Near-patient setting (field trial)	Flexicult SSI Urinary Kit	Culture	Antimicrobial resistance	54	17	6	257	NR	NR	Data relate to 67 samples – each sample tested five times (once for each antibiotic)
				Presence of UTI	58	3	17	43	NR	NR	None
Bongard (2015)[19]	**Population:** Mixed **Location of test performance:** Laboratory	Flexicult SSI Urinary Kit	Culture and Microscopy	Presence of UTI	39	27	6	128	87	83	None
			Culture and microscopy and spiral plating	Presence of UTI	50	16	4	130	NR	NR	None
			Culture	Antimicrobial resistance	84	2	22	33	NR	NR	2 × 2 data obtained by summing across all antibiotics
Boon (2022)[43]	**Population:** Children (aged < 18 years) **Location of test performance:** Laboratory	Uriscreen	Culture	Presence of UTI	10	44	5	97	67	69	Results available for 156/300 samples (test introduced at late stage of trial)
		UTRiPLEX IFU			6	15	23	248	21	94	Results available for 292/300 samples obtained
Colodner (2000)[49]	**Population:** Mixed – fresh urine samples **Location of test performance:** Laboratory	Dipstreak: 10^5 threshold	Culture	Presence of UTI	121	5	1	691	99	99	180 contaminated on Dipstreak; 178 on conventional culture; 176 on both
		Dipstreak: 10^4 threshold			167	8	2	641	99	99	
Greeff (2002)[46]	**Population:** Symptomatic pregnant women; screening pregnant women **Location of test performance:** Near-patient setting Symptomatic	Uricult trio	Culture	Presence of UTI	29	46	8	44	78	49	79 samples did not reach the lab and were excluded.
	Asymptomatic				47	85	11	104	81	55	

Study details	Population and setting	POCT	Reference standard	Target condition	TP	FP	FN	TN	Sensitivity	Specificity	Missing samples/notes
Holm (2017)[37]	**Population:** Women – uncomplicated UTI **Location of test performance:** Near-patient setting	Flexicult SSI Urinary Kit	Culture	Presence of UTI	111	25	18	29	86	54	No missing index test results; 22 had no reference standard result across the total sample
		ID Flexicult	Culture		104	18	12	24	90	56	
Hullegie (2017)[36]	**Population:** Women – uncomplicated UTI **Location of test performance:** Laboratory	Flexicult SSI Urinary Kit	Culture **Threshold:** PHE/HPA definition	Presence of UTI	108	94	29	58	79	38	Result for 289/306. 17 missing results (7 missing reference standard data; 10 missing Flexicult data)
			Threshold: UK laboratory definition		74	128	20	67	79	34	
			Threshold: European definition		140	62	50	37	74	37	
			Culture	Antimicrobial resistance	203	5	23	13	NR	NR	Results summed across all antibiotics
Lee (2010)[47]	**Population:** Children (aged < 16 years) **Location of test performance:** Near-patient setting	Uricult Trio	Culture	Presence of UTI	19	18	13	101	59	85	Seven missing samples – two patients failed to collect sample, three only had urine culture tests performed and two patients only performed index test
				Presence of *E. coli*	12	5	8	126	60	96	
Macias (2002)[44]	**Population:** Catheterised ICU patients **Location of test performance:** Near-patient setting	Uriscreen	Culture	Presence of UTI – any	55	26	7	20	89	43	No missing samples reported
				Presence of UTI – +++, foam band > 1 mm	35	14	27	32	57	70	
Mignini (2009)[48]	**Population:** Screening – pregnant women **Location of test performance:** Laboratory	Uricult	Culture	Presence of UTI	321	8	8	1836	98	100	830 samples excluded due to contamination
Millar (2000)[41]	**Population:** Screening – pregnant women **Location of test performance:** Near-patient setting	Uriscreen	Culture	Presence of UTI	30	185	13	150	70	45	5/383 samples contaminated and excluded Inter-rater reliability: 28/30 samples interpreted consistently

continued

Study details	Population and setting	POCT	Reference standard	Target condition	TP	FP	FN	TN	Sensitivity	Specificity	Missing samples/notes
Pernille (2019)[40]	**Population:** Women – uncomplicated UTI **Location of test performance:** Near-patient setting	ID Flexicult	Culture	Presence of UTI – mid-stream urine samples analysed immediately	46	13	6	52	88	80	Results also presented for first void samples and analysed after 1- and 4-hour delay. Test was more accurate for mid-stream urine; little impact of delay in analysis
Teppa (2005)[42]	**Population:** Screening – pregnant women **Location of test performance:** Near-patient setting	Uriscreen	Culture	Presence of UTI	17	13	11	109	61	89	10/150 samples contaminated – repeat culture indicated negative results in all cases, included in analysis as negative culture
Van der Goes (2023)[51]	**Population:** Mixed – fresh urine sample Laboratory	Lodestar DX – 40-minute run time; 1 µl of urine		Presence of *E. coli*	25	14	4	106	57.9	96.1	149 samples
Yagupsky (2000)[50]	**Population:** Mixed – fresh urine samples **Location of test performance:** Laboratory	Dipstreak	Culture	Presence of UTI	270	4	12	509	96	99	275 excluded due to contamination
				Pathogenic cause	211	N/A	59	N/A	NR	NR	211/270 correctly identified. None incorrectly identified but 59 were not identified

PHE, Public Health England.

Risk of bias

Study details	Anacleto (2009)[45]
Index test	Uricult Trio

Domain 1: patient selection	
Consecutive patients; had to have tested positive on LE or nitrite so applicability issues but low risk of bias	
Was a consecutive or random sample of patients enrolled?	Yes
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	**Low**

DOMAIN 2: INDEX TEST	
Prespecified, standard threshold. No information on blinding but likely that test was interpreted before the ref standard	
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**

Domain 3: reference standard	
Standard culture. The routine plates were read independently by one bacteriologist	
Was an appropriate reference standard used?	Yes
Were the reference results interpreted without knowledge of the results of the index test?	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias?	**Low**

DOMAIN 4: FLOW AND TIMING	
No missing data. Same sample used for index test and reference standard	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the selection of patients have introduced bias?	**Low**

OVERALL RISK OF BIAS	**Low**
Rationale for judgement: no concerns	

Study details	**Blom (2002)[39]**
Index test:	Flexicult Human

Domain 1: patient selection	
Field trial – patients recruited by GPs, no further details	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	**Unclear**

DOMAIN 2: INDEX TEST	
Flexicult – no information on interpretation but appears unlikely that would have been aware of result as likely to have been interpreted first. Prespecified standard threshold	
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**

Domain 3: reference standard	
Culture. No information on blinding	
Was an appropriate reference standard used?	Yes
Were the reference results interpreted without knowledge of the results of the index test?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	**Unclear**

DOMAIN 4: FLOW AND TIMING	
One patient missing data for susceptibility testing on ref standard. Same urine sample	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the selection of patients have introduced bias?	**Low**

OVERALL RISK OF BIAS	Unclear
Rationale for judgement: no information on blinding of interpreter of reference standard	

Study details	Bongard (2015)[19]
Index test:	Flexicult Human

Domain 1: patient selection	
Convenience sample of urines available in the laboratory	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	**Unclear**

DOMAIN 2: INDEX TEST	
Flexicult performed on existing laboratory samples. Performed on same day as routine urine sample testing. No information on blinding	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**

Domain 3: reference standard	
Culture and microscopy with additional check using spiral plating. No information on interpretation of test result	
Was an appropriate reference standard used?	Yes
Were the reference results interpreted without knowledge of the results of the index test?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	**Unclear**

DOMAIN 4: FLOW AND TIMING	
None for accuracy; only subsample assessed for antimicrobial sensitivity – high risk of bias for this analysis; tests performed on the same day using the same urine sample	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the selection of patients have introduced bias?	**Low**

OVERALL RISK OF BIAS	Unclear
Rationale for judgement: unclear if consecutive patients were enrolled; no information on blinding of interpreter of reference standard	

Study details	Boon (2022)[43,53]

Comparative review question

Patients	300 children aged < 18 years
Index test A	UTRiPLEX IFU
Index test B	Uriscreen
Reference standard and target condition	Culture; presence of UTI

Domain 1: patient selection	
Children aged < 18 years enrolled consecutively	
Was a consecutive or random sample of patients enrolled?	Yes
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	**Low**
Comparative accuracy (QUADAS-C)	Yes
Was the risk of bias for each index test judged 'low' for this domain?	Yes
Was a fully paired or randomised design used?	Yes
Was the allocation sequence random?	Not applicable
Was the allocation sequence concealed until patients were enrolled and assigned to index tests?	Not applicable
Could the selection of patients have introduced bias in the comparison?	**Low**

DOMAIN 2: INDEX TEST	
Flexicult performed on existing laboratory samples. Performed on same day as routine urine sample testing. No information on blinding	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**
Comparative accuracy (QUADAS-C)	
Was the risk of bias for each index test judged 'low' for this domain?	Yes
Were the index test results interpreted without knowledge of the results of the other index test(s)?	Unclear
Is undergoing one index test unlikely to affect the performance of the other index test(s)?	Yes
Were the index tests conducted and interpreted without advantaging one of the tests?	Yes
Could the conduct or interpretation of the index tests have introduced bias in the comparison?	**Low**

Domain 3: reference standard	
Culture. 'Laboratory staff performing the reference standard were unaware of patient characteristics and treating physicians were blinded for all urine test results conducted as part of the study'	
Was an appropriate reference standard used	Yes
Were the reference results interpreted without knowledge of the results of the index test?	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias?	**Low**
Comparative accuracy (QUADAS-C)	
Was the risk of bias for each index test judged 'low' for this domain?	Yes
Did the reference standard avoid incorporating any of the index tests?	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias in the comparison?	**Low**

DOMAIN 4: FLOW AND TIMING	
834 eligible; 643 sample receive; 354 sample analysed at central laboratory; 292 sample with UTRiPLEX test; 156 sample with Uriscreen test; same urine sample	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the selection of patients have introduced bias?	**Low**
Comparative accuracy (QUADAS-C)	
Was the risk of bias for each index test judged 'low' for this domain?	Yes
Was there an appropriate interval between the index tests?	Yes
Was the same reference standard used for all index tests?	Yes
Are the proportions and reasons for missing data similar across index tests?	No
Could the patient flow have introduced bias in the comparison?	Low

OVERALL RISK OF BIAS	Low
Rationale for judgement: no concerns. There was a high amount of exclusion in the Uriscreen vs. culture comparison but this was due to late introduction of the test	

Study details	Colodner (2000)[49]
Index test	Dipstreak

Domain 1: patient selection	
Laboratory-based study – very few details on samples provided	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	**Unclear**

DOMAIN 2: INDEX TEST	
Dipstreak performed on existing laboratory samples. No information on blinding	
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**

Domain 3: reference standard	
Culture. No information on blinding	
Was an appropriate reference standard used	Yes
Were the reference results interpreted without knowledge of the results of the index test?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	**Unclear**

DOMAIN 4: FLOW AND TIMING	
Results available for all 1000 urine samples – large number of contaminated results but these are reported in detail; same urine sample	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the selection of patients have introduced bias?	**Low**

OVERALL RISK OF BIAS	Unclear
Rationale for judgement: unclear if consecutive patients were enrolled; no information on blinding of interpreter of reference standard	

Study details	Greeff (2002)[46]
Index test	Uricult Trio

Domain 1: patient selection	
Women attending antenatal clinic – appears to be screening but unclear. Unclear if all patients (i.e. consecutive patients) enrolled	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	No
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	**Low**

DOMAIN 2: INDEX TEST	
No information on blinding but likely that test was interpreted before the reference standard	
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**

Domain 3: reference standard	
Culture. No information on blinding	
Was an appropriate reference standard used	Yes
Were the reference results interpreted without knowledge of the results of the index test?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	**Unclear**

DOMAIN 4: FLOW AND TIMING	
79 urine specimens lost; same urine sample	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the selection of patients have introduced bias?	**High**

OVERALL RISK OF BIAS	High
Rationale for judgement: high proportion of patients excluded from analysis	

Study details	Holm (2017)[37]

Comparative review question

Patients	**376 women with uncomplicated UTI**
Index test A	Flexicult SSI kit
Index test B	ID Flexicult
Reference standard and target condition	Culture; presence of UTI

Domain 1: patient selection	
Consecutive women with suspected UTI	
Was a consecutive or random sample of patients enrolled?	Yes
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	**Low**
Comparative accuracy (QUADAS-C)	
Was the risk of bias for each index test judged 'low' for this domain?	Yes
Was a fully paired or randomized design used?	Yes
Was the allocation sequence random?†	Yes
Was the allocation sequence concealed until patients were enrolled and assigned to index tests?†	Yes
Could the selection of patients have introduced bias in the comparison?	**Low**

DOMAIN 2: INDEX TEST	
Flexicult – standard threshold interpreted blind to lab culture (as was interpreted before – explicitly reported in paper)	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**
Comparative accuracy (QUADAS-C)	
Was the risk of bias for each index test judged 'low' for this domain?	Yes
Were the index test results interpreted without knowledge of the results of the other index test(s)?	N/A
Is undergoing one index test unlikely to affect the performance of the other index test(s)?	N/A
Were the index tests conducted and interpreted without advantaging one of the tests?	Yes
Could the conduct or interpretation of the index tests have introduced bias in the comparison?	**Low**

DOMAIN 3: REFERENCE STANDARD	
Culture, reported blind to POCT	
Was an appropriate reference standard used	Yes
Were the reference results interpreted without knowledge of the results of the index test?	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias?	**Low**
Comparative accuracy (QUADAS-C)	
Was the risk of bias for each index test judged 'low' for this domain?	Yes
Did the reference standard avoid incorporating any of the index tests?	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias in the comparison?	Low

DOMAIN 4: FLOW AND TIMING	
35/376 excluded from analysis: 22 patients had missing laboratory data, 2 withdrew consent, 7 did not fulfil inclusion criteria, 4 for other reasons; same sample	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the selection of patients have introduced bias?	**Low**
Comparative accuracy (QUADAS-C)	
Was the risk of bias for each index test judged 'low' for this domain?	Yes
Was there an appropriate interval between the index tests?	Yes
Was the same reference standard used for all index tests?	Yes
Are the proportions and reasons for missing data similar across index tests?	Unclear
Could the patient flow have introduced bias in the comparison?	**Low**

OVERALL RISK OF BIAS	Low
Rationale for judgement: no concerns	

Study details	Hullegie (2017)[36]
Index test	Flexicult Human

Domain 1: patient selection	
DTA study nested in trial	
Was a consecutive or random sample of patients enrolled?	Yes
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	**Low**

DOMAIN 2: INDEX TEST	
Flexicult – standard threshold most likely interpreted blind to laboratory culture (as was interpreted before)	
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**

Domain 3: reference standard	
Culture, no information on blinding	
Was an appropriate reference standard used	Yes
Were the reference results interpreted without knowledge of the results of the index test?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	**Unclear**

DOMAIN 4: FLOW AND TIMING	
6/312 cultures were not available. 13/325 Flexicult missing – in 10 cases clinician did not complete CRF, in three cases test was not performed	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the selection of patients have introduced bias?	**Low**

OVERALL RISK OF BIAS	Unclear
Rationale for judgement: no information on blinding of interpreter of reference standard	

Study details	Lee (2010)[47]
Index test	Uricult Trio

Domain 1: patient selection	
Children presenting to outpatient department	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	**Unclear**

DOMAIN 2: INDEX TEST	
Prespecified, standard threshold. No information on blinding but likely that test was interpreted before the reference standard	
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**

Domain 3: reference standard	
Culture. No information on blinding	
Was an appropriate reference standard used?	Unclear
Were the reference results interpreted without knowledge of the results of the index test?	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias?	**Unclear**

DOMAIN 4: FLOW AND TIMING	
3/158 patients failed to collect urine sample; 2 patients only had culture tests and 2 patients only had Uricult Trio test; same urine sample	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes

DOMAIN 4: FLOW AND TIMING	
Were all patients included in the analysis?	No
Could the selection of patients have introduced bias?	**Low**

OVERALL RISK OF BIAS	Unclear
Rationale for judgement: unclear if consecutive patients were enrolled; no information on blinding of interpreter of reference standard	

Study details	Macias (2002)[44]
Index test	Uriscreen

Domain 1: patient selection	
ICU patients – no details of how selected. Multiple samples taken for each patient	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	**High**

Domain 2: index test	
Threshold clearly defined and prespecified. No information on blinding but test performed before reference standard results would be available	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**

Domain 3: reference standard	
Culture. No information on blinding	
Was an appropriate reference standard used	Yes
Were the reference results interpreted without knowledge of the results of the index test?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	**Unclear**

DOMAIN 4: FLOW AND TIMING	
Results reported for all included patients; tests performed on same urine sample	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the selection of patients have introduced bias?	**Low**

OVERALL RISK OF BIAS	High
Rationale for judgement: multiple samples taken from some patients; unclear how patients selected for inclusion	

Study details	Mignini (2009)[48]
Index test	Uricult

Domain 1: patient selection	
Consecutive pregnant women. Exclusion for multiple reasons, which may have restricted study sample	
Was a consecutive or random sample of patients enrolled?	Yes
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	**Low**

Domain 2: index test	
Uricult. Standard threshold used. Appears likely that index test interpreted before refence standard results available as POCT	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**

Domain 3: reference standard	
Standard laboratory-based culture. No information on blinding of interpreter	
Was an appropriate reference standard used	Yes
Were the reference results interpreted without knowledge of the results of the index test?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	**Unclear**

Domain 4: flow and timing	
Large proportion of samples excluded due to contamination; test performed on same urine samples	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the selection of patients have introduced bias?	**High**

OVERALL RISK OF BIAS	High
Rationale for judgement: high proportion of patients excluded from analysis	

Study details	Millar (2000)[41]
Index test	Uriscreen

Domain 1: patient selection	
Consecutive women	
Was a consecutive or random sample of patients enrolled?	Yes
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	**Low**

DOMAIN 2: INDEX TEST	
Standard threshold; interpreted before reference standard results available	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**

Domain 3: reference standard	
Standard laboratory-based culture. No information on blinding of interpreter	
Was an appropriate reference standard used	Yes
Were the reference results interpreted without knowledge of the results of the index test?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	**Unclear**

DOMAIN 4: FLOW AND TIMING	
5/383 samples were contaminated and were excluded from analysis; same sample	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the selection of patients have introduced bias?	**Low**

OVERALL RISK OF BIAS	Unclear
Rationale for judgement: no information on blinding of interpreter of reference standard	

Study details	Pernille (2019)[40,54]
Index test	ID Flexicult

Domain 1: patient selection	
Women presenting to primary care with symptoms of UTI	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	**Low**

DOMAIN 2: INDEX TEST	
Interpreters were blind to culture result. Standard threshold used	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**

Domain 3: reference standard	
Culture. No information on whether culture was interpreted blind to POCT	
Was an appropriate reference standard used?	Yes
Were the reference results interpreted without knowledge of the results of the index test?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	**Unclear**

DOMAIN 4: FLOW AND TIMING	
Five women excluded – two unable to deliver sufficient urine; three had already participated; same urine samples	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the selection of patients have introduced bias?	**Low**

OVERALL RISK OF BIAS	Unclear
Rationale for judgement: no information on blinding of interpreter of reference standard	

Study details	Teppa (2005)[42]
Index test	Uriscreen

Domain 1: patient selection	
Pregnant women – unclear if consecutive sample	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	**Low**

DOMAIN 2: INDEX TEST	
Standard threshold; interpreted before reference standard results available	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**

Domain 3: reference standard	
Culture. No information on whether culture was interpreted blind to POCT	
Was an appropriate reference standard used	Yes
Were the reference results interpreted without knowledge of the results of the index test?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	**Unclear**

DOMAIN 4: FLOW AND TIMING	
All patients included in 2 × 2 table; same sample	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the selection of patients have introduced bias?	**Low**

OVERALL RISK OF BIAS	Unclear
Rationale for judgement: no information on blinding of interpreter of reference standard	

Study details	Van der Goes (2023)[51]
Index test	Lodestar DX

Domain 1: patient selection	
Stored urine samples and fresh urine samples – mixture of cloudy and non-cloudy urine	
Was a consecutive or random sample of patients enrolled?	Yess
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	**Unclear**

DOMAIN 2: INDEX TEST	
Lodestar – threshold specified. No information on how test was interpreted. Both tests performed in the same laboratory so potential for bias in interpretation	
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Unclear**

Domain 3: reference standard	
Stored samples: 'Samples were also cultured on UTI Chromoselect Agar to confirm bacterial growth' *Fresh samples: 'standard PHW methods including culture'. No information on blinding*	
Was an appropriate reference standard used?	Yes
Were the reference results interpreted without knowledge of the results of the index test?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	**Unclear**

DOMAIN 4: FLOW AND TIMING	
Results available for all samples; index test and reference standard performed on same urine sample	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the selection of patients have introduced bias?	**Low**

OVERALL RISK OF BIAS	Unclear
Rationale for judgement: no information on blinding of interpreter of reference standard	

Study details	Yagupsky (2000)[50]
Index test	Dipstreak

Domain 1: patient selection	
Unclear how samples were collected – whether convenience sample	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	**Unclear**

DOMAIN 2: INDEX TEST	
Dipstreak performed in laboratory setting – no information on blinding and both tests performed in same laboratory so potential for unblinding	
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**

Domain 3: reference standard	
Culture – no information on blinding and both tests performed in same laboratory so potential for unblinding	
Was an appropriate reference standard used	Yes
Were the reference results interpreted without knowledge of the results of the index test?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	**Unclear**

DOMAIN 4: FLOW AND TIMING	
275/1000 excluded due to contamination/need for repeat culture	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes

DOMAIN 4: FLOW AND TIMING	
Were all patients included in the analysis?	No
Could the selection of patients have introduced bias?	**High**

OVERALL RISK OF BIAS	High
Rationale for judgement: high proportion of patients excluded from analysis	

Objective 3

Study details[a]	Participants and test	Results
Anacleto (2009)[45] **Country:** Philippines **Language:** English **Funding:** Institute of Child Health and Human Development of the National Institutes of Health, Manila, Philippines, the Philippine Society of Nephrology, Inc., and Pediatric Associates, Inc.	**Setting and population:** Secondary care; uncomplicated UTI; aged < 16 years **Inclusion criteria:** Infants and children aged 0–7 years with symptoms suggestive of UTI and positive LE or nitrite dipstick test **Exclusion criteria:** Poor intake of antibiotics; obstructive uropathy; congenital anomalies of kidneys seven urinary tract; midline defects; failure to thrive; concomitant infections; recurrent UTI; asymptomatic bacteriuria; other comorbid conditions **Number included (number analysed):** 200 (200) **Age:** 4 months to 7 years **% Female:** 43 **Test:** Uricult Trio	• 'Uricult trio method was convenient to use and easy to interpret'
Blom (2002)[39] **Country:** Denmark **Language:** English **Funding:** Not reported	**Setting and population:** Primary care – mixed symptomatic patients **Inclusion criteria:** 19 GPs asked to use Flexicult in addition to standard diagnostic procedures in patients with symptoms of UTI **Exclusion criteria:** Not reported **Number included (number analysed):** 121 **Age:** NR **% Female:** NR **Test:** Flexicult SSI Urinary Kit	• **Ease of use/acceptability –** 'the participating GPs considered the kit to be easy to handle and read'
Brooks-Howell (2019)[55] **Country:** Wales, England, Spain, Netherlands **Language:** English **Funding:** EU funding as part of the R-GNOSIS programme	**Setting and population:** Telephone interviews; primary care clinicians and health professionals **Inclusion criteria:** Participation in POETIC trial **Number included (number analysed):** 35 **Age:** NR **% Female:** 77 **Test:** Flexicult SSI Urinary Kit	• Overall reaction to POCT positive, perceived impact of Flexicult use on antibiotic prescribing even split between 'no change' and 'more awareness and therefore more cautious prescribing habits' • 'Clinicians overwhelmingly felt that a POCT for UTI management would be useful. When describing the "ideal" test, the key component seemed to be fast results, while ease of use and accuracy and reliability were mentioned far less. Many described the Flexicult POCT as the ideal test but some felt that it would be better if it gave faster results' • **Ease of use/acceptability –** Increased confidence in diagnosing UTI with POCT but difficulties reported in interpretation of results and limitations on when POCT can be used • **Time to test results –** Quicker results than laboratory test (targeted treatment within 24 hours instead of 3–4 days) but some concern about possible patient discomfort while waiting to obtain results rather than prescribing straight away

Study details[a]	Participants and test	Results
		• **Any outcome related to antibiotic use or prescription –** Positive impact on awareness of health professionals regarding antibiotic prescribing • **UTI-associated healthcare resources –** Concerns testing all patients would strain care delivery due to staffing issues and limited capacity to conduct and follow up on test • **Health-related quality of life –** Clinicians felt the use of POCT reassured patients, but were concerned that waiting for test results before prescribing would prolong patient discomfort • **Test costs –** Concerns about potential expense of maintaining regular stock of tests
Butler (2018)[8] **Country:** England, Netherlands, Spain and Wales **Language:** English **Funding:** European Commission Seventh Framework Programme	**Setting and population:** Primary care; women aged ≥ 18 years – uncomplicated UTI **Inclusion criteria:** Presenting to primary care with any of the following symptoms: dysuria, urgency or frequency with clinical diagnosis of uncomplicated UTI **Exclusion criteria:** Suspected pyelonephritis; long-term antibiotic treatment; antibiotics for UTI in preceding 4 weeks; significant genitourinary tract abnormalities; terminal illness **Number of eligible patients (randomised):** 654 (653) **Age:** 47.6 years (SD 27.6 years) **Sex:** All female **Test:** Flexicult SSI Urinary Kit	• **Time to perform test –** 9 minutes to prepare test, 6 minutes to obtain and record result, 7 minutes to discuss result with patient • **Cost –** Cost per person including POCT cost in UK is £48 • **Management change as a result of test –** o Overall 63.1% o Did not start antibiotic 7.4% o Stopped taking antibiotic 5.3% o Started taking antibiotic 15.3% o Continued with antibiotic 33.2% o New antibiotic prescribed 38.9%
Greeff (2002)[46] **Country:** South Africa **Language:** English **Funding:** Not reported	**Setting and population:** Antenatal clinics; screening Pregnant women **Inclusion criteria:** Two populations of patients from the Pretoria region were involved: (1) asymptomatic pregnant women attending the antenatal clinic for the first time or presenting in labour; and (2) pregnant women with symptoms suggestive of UTI **Exclusion criteria:** NR **Number included (number analysed):** 453 (374) **Age:** NR **% Female:** 100 **Test:** Uricult Trio	• **Ease of use/acceptability –** o 'the Uricult Trio did not add anything in terms of managing the patient more efficiently' and 'is not useful for screening asymptomatic bacteriuria or for diagnosing UTIs in women with symptoms suggestive of an infection' o 'the advantage of this on-site test is that none of the Uricult Trio specimens got lost, as opposed to 79 laboratory specimens in this study. This highlights the value of an on-site test. Another advantage of the Uricult Trio is that one can potentially obtain the result sooner and more easily than a conventional laboratory culture, which would also have a great impact on the cost of hospitalisation'

a All studies were accuracy studies with the exception of Brooks-Howell, which employed a qualitative thematic analysis of semistructured interviews.

EME
HSDR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library