NICE National Institute for Health and Care Excellence

Diagnostics Assessment Programme

Point of care tests for urinary tract infections to reduce antimicrobial resistance - early value assessment

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Point of care tests for urinary tract infections to reduce antimicrobial resistance - early value assessment

Diagnostics Assessment Programme

Contents:

Any information supplied to NICE which has been marked as confidential has been redacted. All personal information has also been redacted.

Technology Assessment Report commissioned by the NIHR Evidence Synthesis Programme on behalf of the National Institute for Health and Care Excellence

Title of Project: Point of care tests for urinary tract infections (UTI) to reduce antimicrobial resistance: a systematic review and conceptual economic model to inform Early Value Assessment (EVA) (DAP 69)

Produced by: Bristol Technology Assessment Group

Authors: Eve Tomlinson, Research Associate in Evidence Synthesis, Bristol TAG, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol *(Joint first author)*

Mary Ward, Senior Research Associate in Health Economic Modelling, Bristol TAG, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol *(Joint first author)*

Chris Cooper, Research Fellow in Health Technology Assessment and Information Science, Bristol TAG, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol

Rachel James, Research Associate in Evidence Synthesis, Bristol TAG, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol

Christina Stokes, Patient Representative

Samina Begum, Patient Representative

Jessica Watson, NIHR Clinical Lecturer in General Practice, Centre for Academic Primary Care, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol

Alastair D Hay, Professor of Primary Care, Centre for Academic Primary Care, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol

Hayley E Jones, Associate Professor in Medical Statistics, Bristol TAG, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol

Howard Thom, Senior Lecturer in Health Economics, Bristol TAG, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol *(Joint last author)*

Penny Whiting, Professor of Clinical Epidemiology, Bristol TAG, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol *(Joint last author)*

Source of funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR135710.

Declared competing interests of the authors

None of the authors have any conflicts of interest.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

Commercial in confidence (CiC) data are highlighted in blue throughout the report.

Academic in confidence (AiC) data are highlighted in yellow throughout the report.

Abstract

Background

Urinary tract infections (UTIs) are diagnosed by GPs based on symptoms, dipstick tests in some, and laboratory urine culture. Patients may be given inappropriate antibiotics. Point of care tests (POCT) can diagnose UTI in near-patient settings quicker than standard culture. Some can identify the causative pathogen or antimicrobial sensitivity.

Objective

To assess whether POCT for people with suspected UTI have the potential to be clinically and cost-effective to the NHS.

Design

Systematic review and conceptual economic model.

Results

Two RCTs evaluated Flexicult Human (one against standard care; one against ID Flexicult). Both trials found no difference between groups in concordant antibiotic use (OR 0.84 95% CI 0.58, 1.20) or appropriate antibiotic prescribing (OR 1.44 95% CI 1.03, 1.99). Compared to standard care, Flexicult was associated with reduced antibiotic prescribing at initial consultation (OR 0.56 95% CI 0.35, 0.88). No difference was found for other measures of antibiotic use, symptom duration, patient enablement, or resource use.

Sixteen studies reported test accuracy data. Most were at unclear or high risk of bias. We identified data on three rapid tests (results <40min).

[redacted]Uriscreen (4 studies) had modest summary sensitivity 74% (95% CI 59, 84) and specificity 70% (95% CI 52, 84). UTRiPLEX (1 study) had poor sensitivity (21%) and good specificity (94%).

Twelve studies evaluated culture-based tests (results 24hr). Laboratory based studies found Dipstreak (n=2) and Uricult (n=1) to be highly accurate. Uricult Trio (n=3) had more modest summary sensitivity 73% (95% CI 63, 82) and specificity 70% (95% CI 52, 84). Summary sensitivity for Flexicult Human (n=4) and ID Flexicult (n=2) were 79% (95% CI 72, 85) and 89% (95% CI 84, 93). Summary specificity was 67% (95% 30, 90) and 70% (95% CI 52, 84). Caution is needed in interpreting findings due to heterogeneity and limited data.

Five studies evaluated technical performance (3 Flexicult Human; 2 Uricult Trio). Limited data suggested they are easier to use and interpret than standard culture. Clinicians reported some positives of using Flexicult Human and barriers to implementation.

A conceptual economic model estimate the cost-effectiveness of POCT for UTI diagnosis, pathogen identification and AST. Sensitivity and specificity of tests was informed by the

clinical effectiveness review. Studies identified by the review were screened for evidence on treatment efficacy, costs, and utility data; only 2 studies provided relevant evidence. A pragmatic search identified 8 cost-effectiveness studies that provided further evidence. A decision tree comparing POCTs in a mixed population (Lodestar DX vs Flexicult Human) and in women with uncomplicated UTI (Lodestar DX vs Flexicult Human vs ID Flexicult) was implemented. The available input data were too limited for results to be meaningful.

Conclusion

More research is required to determine whether POCT for UTI have potential to be clinically and cost effective to the NHS. Rapid tests such as Astrego PA-100 system or Lodestar DX appear promising, but data is very limited.

Word count: 496 words

Scientific Summary

Background

Urinary tract infections (UTI) are one of the most common causes of infection worldwide. Accurate and timely diagnosis of UTI is crucial to ensure 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 only prescribed when necessary and that they target the causative organism of the infection.

However, UTIs can be difficult to diagnose. Currently they are diagnosed by the 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 (LE), 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 antibiotic.

Novel point of care tests (POCT) 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 POCT for people with suspected UTI have the potential to be clinically 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 POCT to diagnose UTI, with or without additional pathogen identification and AST?
- Objective 2: What is the accuracy of the POCT for UTI diagnosis, pathogen identification and AST?
- Objective 3: What is the technical performance (other than accuracy) of POCT for UTI?
- Objective 4: What are the costs, from a UK NHS and Personal Social Services (PSS) perspective, of using POCT for UTI diagnosis, pathogen identification and AST?
- Objective 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 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 searches of 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 <40min:* Astrego PA-100 system, Lodestar DX, TriVerity, Uriscreen, UTRiPLEX.
- *Culture-based tests giving results in up to 24hr*: 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 pre-specified outcomes to be eligible.

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 allow 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 POCT for UTI diagnosis, pathogen identification and AST. This represented important short and long-term costs and quality of life impacts in the management of UTIs.

The conceptual model was implemented as a decision tree comparing POCTs to 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 NICE guidelines, British National Formulary (BNF) costs, and the Personal Social Services Research Unit (PRSSU).

We prioritised tests and populations where evidence was greatest. We also prioritised rapid over culture-based tests and tests that perform AST over those than 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.

Results

Clinical effectiveness review

We identified 16 studies for inclusion in the review. All studies were included for objective 2 – 2 were also included for objective 1, while 5 also provided data for objective 3. Six studies evaluated rapid POCT (1 Lodestar DX, 4 Uriscreen, 1 UTRiPLEX) and 12 studies evaluated culture-based POCT (4 Flexicult Human, 2 ID Flexicult, 3 Uricult Trio, 1 Uricult, 2 Dipstreak). 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 pre-specified tests or populations of interest.

Objective 1: Clinical outcomes

Two RCTs evaluated the clinical impact of Flexicult Human in women - one compared to standard care the other to ID Flexicult. Both trials were at low risk of bias. There was no difference between intervention groups in the primary outcomes: concordant antibiotic use (OR 0.84 95% CI 0.58, 1.20) and appropriate antibiotic prescribing (OR 1.44 95% CI 1.03, 1.99). Compared to standard care, the use of Flexicult Human was associated with reduced antibiotic prescribing at initial consultation (OR 0.56 95% CI 0.35, 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. Three of these studies took place in Wales and one had centres in England and Wales (as well as Spain and the Netherlands). The other studies were conducted in Denmark, Israel, Hawaii, Venezuela, Belgium, Mexico,

Philippines, South Africa, Korea, Argentina. Twelve studies were conducted in primary or secondary care and four were laboratory-based. Five studies were judged at high risk of bias, eight unclear risk of bias and three were low risk.

Only three rapid tests were evaluated. Lodestar DX appeared to be the most promising test. [redacted] Uriscreen had modest summary estimates of sensitivity 74% (95% CI 59, 84; 4 studies) and specificity 70% (95% CI 52, 84). 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.

Of the culture-based tests evaluated, Dipstreak and Uricult were found to be highly accurate. However, these were assessed by 2 studies and 1 study respectively, both were conducted in the laboratory and were at high or unclear risk of bias. In contrast, studies of Uricult Trio (an extension of Uricult) in near-patient settings reported more modest summary sensitivity 73% (95% CI 63, 82) and specificity 70% (95% CI 52, 84). Summary sensitivity for Flexicult Human (4 studies) was 79% (95% CI 72, 85) and summary specificity was 67% (95% 30, 90). For ID Flexicult (2 studies), this was 89% (95% CI 84, 93) and 70% (95% CI 52, 84). Three studies reported data on the accuracy Flexicult human in determining antimicrobial sensitivity. Summary sensitivity was 87% (95% CI 83, 90), and summary specificity was of 93% (95% CI 89,95).

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

Objective 3: Technical performance

Five studies, reported technical performance data. These studies evaluated culture-based tests only (3 Flexicult Human; 2 Uricult Trio). Studies reported that POCT 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 whilst awaiting test results, and the expense of maintaining stock of tests. One study reported that Flexicult Human costs £48. [redacted] 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 of POCTs, namely that they could reduce the use of empiric antibiotics and by, by reducing the incidence of UTI complications and improving cure rates, reduce healthcare costs quality of life impacts arising from UTIs. Beyond test accuracy, we found only 2 studies from the clinical effectiveness review with relevant evidence for the economic model. Our pragmatic searches identified only 8 cost-effectiveness studies in UTI, none of which modelled POCTs and none of which provided all 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 in 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 is too weak for results to be meaningful.

Conclusions

Implications for practice

There is little available data concerning the clinical and cost-effectiveness of POCT for people with suspected UTI, particularly for rapid POCT, making it difficult to determine whether these tests have the potential to be clinically 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 of 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 of the 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 vs empiric antibiotic efficacy, or costs and quality of life impacts of UTI complications for our model to perform a meaningful comparison.

Strong evidence that POCT (i) reduce unnecessary antibiotic use; (ii) improve symptoms or (iii) are cost-effective, is needed before such tests are introduced to the NHS.

Recommendations for research

Given the paucity of data on POCT test for diagnosing UTI, further studies are needed to determine whether POCT for people with suspected UTI have the potential to be clinically and cost effective to the NHS. Ideally studies would be randomised controlled trials with embedded diagnostic test accuracy studies of POCT and should be conducted in primary care – such studies would provide data on clinical impact and on test accuracy. Studies should focus on tests with the greatest potential for clinical impact – the Astrego PA-100

system and Lodestar DX. They should either 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 POCT in 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 vs targeted antibiotic treatment of UTI; efficacy in preventing UTI complications; and both the cost and quality of life impacts of these complications.

Study registration

The review was registered at PROSPERO (CRD42022383889).

Funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR135710.

Word count: 2224

Plain English Summary

What is the problem?

Urine infections are very common but can be difficult to diagnose. The 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 antibiotic and some will have 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 you have a urine infection. Some can also tell which bug is causing the infection and which antibiotic will work best.

What did we do?

We wanted to know whether using “rapid tests” to diagnose urine infections means more people are: correctly diagnosed, diagnosed more quickly, and treated with the right antibiotic 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 differed a lot. 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.

Word count: 250 words

Table of contents

List of Tables

List of Figures

Definition of Terms and List of Abbreviations

Abbreviation	Definition
AE	Adverse Event
AiC	Academic in Confidence
AST	Antimicrobial sensitivity testing
BNF	British National Formulary
CE	Conformité Européenne
CEAC	Cost Effectiveness Acceptability Curves
CEAF	Cost Effectiveness Acceptability Frontiers
CENTRAL	Cochrane Central Register of Controlled Trials
CFU	Colony Forming Unit
CI	Confidence Interval
CiC	Commercial in Confidence
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CRD	Centre for Reviews and Dissemination
DAR	Diagnostics Assessment Report
DPD	Depersonalised Data
DTA	Diagnostic Test Accuracy
ESPAUR	English Surveillance Programme for Antimicrobial Utilisation and Resistance
EUCAST	The European Committee on Antimicrobial Susceptibility Testing
EVA	Early Value Assessment
GP	General practitioner
HNE	4-Hydroynonenal
ICER	Incremental Cost-Effectiveness Ratio
ICTRP	International Clinical Trials Registry Platform
LE	Leukocyte Esterase
MM	Markov Model
MMP8	Matrix metalloproteinase-8
NB	Net Benefits
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Insititute for Health Research
NR	Not reported
NRSI	Non-Randomised Study of Interventions
PC	Pathogenic Cause
PHE	Public Health England
POCT	Point of care test
POE	Presence of E.Coli

Abbreviation	Definition
POU	Presence of UTI
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
QALDS	Quality Adjusted Life Days
QALMs	Quality Adjusted Life Months
QALYs	Quality-Adjusted Life Years
RBUS	Renal Bladder Ultrasound
RCT	Randomised Controlled Trial
ROB	Risk of Bias
ROC	Receiver Operating Characteristic
TMP-SMX	Trimethoprim-Sulfamethoxazole
UK	United Kingdom
UKCA	UK Conformity Assessment
UTI	Urinary Tract Infection
WHO	World Health Organisation

1 Background

1.1 Epidemiology and burden of UTI

Urinary tract infections (UTI) are one of the most common causes of infection worldwide, and are the most commonly seen bacterial infections in general practice.[1] UTI is also the most common hospital acquired infection in the UK, accounting for almost 1 in 4 of all infections, most of which are associated with catheter use.[2] UTIs 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). 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 UTIs, 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 for uncomplicated UTI are summarised in Table 1. A proportion of patients will suffer from chronic UTI. There is no accepted definition of this and the 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.

Table 1 Overview of classification of uncomplicated UTI, reproduced from Medina et al. (2019)[3]

Classification	Definition
Uncomplicated UTI	UTI where 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 (>38°C)
Asymptomatic bacteriuria	Positive urine culture (>105 colony-forming units/ml) in the absence of urinary symptoms
Recurrent uncomplicated UTIs	Recurrent UTI refers to the occurrence of ≥2 symptomatic episodes within 6 months or ≥3 symptomatic episodes within 12months

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, catheterization, certain infecting organisms and co-morbidities such as immunosuppression.[6]

The most common cause of UTI is *Escherichia coli* (*E. coli*) in both uncomplicated and complicated UTIs.[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]

1.2 Presentation of UTI

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), frequency, urgency, abdominal/suprapubic pain, haematuria, and changes in urine smell, appearance or consistency.[8,9,6] In those aged over 65 years symptoms can be less specific and include delirium, lethargy, reduced ability to carry out activities of daily living and anorexia. [6]

1.3 Diagnosis

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

UTIs 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 can be used as an initial screening test 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 over 65 years or in those who are catheterised, and so they are not recommended in these groups.[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 which includes definitions of susceptibility testing categories with the aim of harmonising breakpoints in Europe.[12]

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

- Suspected UTI in men
- Age > 65 years
- Babies <3 months

- Children <16 years who do not respond to treatment within 24-48 hours
- Pregnant women
- Suspected complicated UTI (pyelonephritis or sepsis)
- Failed antibiotic treatment or persistent symptoms
- Recurrent UTI
- Catheterised patients
- Dipstick negative for nitrites but positive LE
- Age <3 years, positive dipstick for nitrite and LE
- Risk factor for resistance:
    - Abnormalities of genitourinary tract
    - Renal impairment
    - Care home resident
    - Hospitalisation for >7 days in last 6 months
    - Recent travel to country with increased resistance
    - Previous resistant UTI

1.4 Treatment of UTI

Acute uncomplicated UTI generally resolves within around 9 days without treatment,[13] but most UTIs will be prescribed antibiotics. Treatment also involves giving advice on self-care such as analgesia and hydration. NICE guidelines on antimicrobial prescribing for UTI recommends that antibiotics are prescribed immediately in pregnant women, men and children under 16 years.[14] In non-pregnant women, a back-up antibiotic (to be taken only if symptoms persist for 48 hours or worsen) or immediate antibiotic may be prescribed. Whilst 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 empirical antibiotic treatment on the day of diagnosis and that the majority (83%) had no evidence of urine sample collection for laboratory investigation in their electronic health records.[15] If urine is sent for culture and AST then the antibiotic choice should be reviewed when results of AST are available. The NICE guideline contains detailed recommendations on which antibiotic to prescribe as first or second choice (if first choice is not effective or suitable) in different populations. First choice antibiotics are based on empirical 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.[14] 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 section 1.5).

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. Postmenopausal 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 are effective. An alternative to this which 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.[14]

There are currently no NICE guidelines on treatment of chronic UTI. Patient organisations suggest that treatment may involve high-dose, extended course (3-6 months) oral antibiotics or instillation of antibiotics directly into the bladder. [16] Many patients will also seek relief from alternative therapies with little evidence of effectiveness.[17]

1.5 Antibiotic prescribing and resistance

Almost 75% of antibiotic prescribing occurs in primary care,[18] with UTIs contributing to a large proportion of this use. Antimicrobial resistance, and in particular antibiotic resistance, is one of the greatest public health challenges faced today. The World Health Organisation (WHO) highlight this as one of the biggest threats to global health, food security and development today.[19]

The 'English Surveillance Programme for Antimicrobial Utilisation and Resistance' (ESPAUR) report from 2017 says 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]

2 Decision Problem

2.1 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 under 65

- Women over 65
- Men under 65
- Men over 65
- Adults with indwelling urinary catheters
- Babies, children and young people under 16
- Children under 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 UTIs (for example people with neurogenic bladder, diabetes, polycystic kidney disease or people who are immunocompromised)
- People with suspected pyelonephritis

2.2 Technologies of interest

Guidance from Public Health England on 'Health matters: antimicrobial resistance' [18] published in 2015, highlights the need for rapid diagnostic tools to help GPs quickly (within minutes) identify the strain of bacterial infection present and the antibiotics to which it is resistant or susceptible. This is also highlighted in the 2021/2022 English surveillance programme for utilisation resistance (ESPAUR). Tests that are able to 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 section 1.5).[20] 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 UTI reducing the need for samples to be sent for laboratory testing.

The technologies of interest for this appraisal are novel point of care tests (POCT) that may detect the presence of a UTI, provide information on the strain of bacterial infection present and/or the antibiotic(s) to which the bacteria is susceptible. POCT are defined as technologies that can be done by a healthcare professional outside a conventional laboratory setting [21]. Table 2 gives an overview of POCT for diagnosing UTI within the scope of this appraisal. These are grouped into rapid tests (those that provide results in <40 minutes) and culture-based tests that take up to 24 hours to give results. The aim of these tests is to provide a more accurate, rapid diagnosis of UTI and improve antibiotic prescribing. The extent to which these POCT 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.

Table 2 Overview of POCT 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 mins)								
Astrego PA-100 analyser and PA-AST panel U-0501 (Sysmex Astrego)	Microfluidics	Urine	5 commonly used antibiotics (amoxicillin-Clavulanic acid, ciprofloxacin, fosfomycin, nitrofurantoin, trimethoprim)	10-15 minutes	NA	30 to 45 minutes for full results	Digital display shows which antibiotics sample is susceptible to	Yes
Lodestar DX (Llusern Scientific)	Molecular diagnostic test	Urine	Escherichia coli (E-coli), Klebsiella spp, Proteus mirabilis, Staphylococcus saprophyticus, Enterococcus spp, Pseudomonas aeruginosa	40 minutes	40 minutes	NA	Digital display – light indicates which bacteria is detected	Expected <12 months
TriVerity (Inflammatix)	Detects 29 target mRNAs	Blood	Identifies presence, type and severity of infection.	30 minutes	NA	NA	Unclear	Expected <12 months
Uriscreen (Savyon Diagnostics Ltd)	Catalase based test	Urine	Detects catalase activity as indicator of bacteria in somatic cells	2 minutes	NA	NA	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	NA	NA	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: 5 commonly used antibiotics (mecillinam, nitrofurantoin, ampicillin, sulfamethizol and trimethoprim).	16-24 hours	16 to 24 hours	16 to 24 hours	Visual assessment of number & type of growths on agar plate.	Yes

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
			ID Flexicult gives information on pathogenic cause					
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	NA	number of bacterial colonies is compared with the Colony Density Chart	Yes
Uricult, Uricult 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-glucuronidase-producing organisms e.g. E. coli	16-24 hours	16-24 hours	NA	Visual assessment of growth on agar plate.	Yes

NA: Not applicable. MMP-8: Matrix metalloproteinase-8. HNE: 4-Hydroynonenal.

2.3 Potential alternative technologies

There are a number of technologies currently in development that are able to provide a rapid indication of the presence of bacteria, identify the bacteria present and/or provide information on antimicrobial susceptibility, but these do not have a Conformité Européenne (CE) 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.

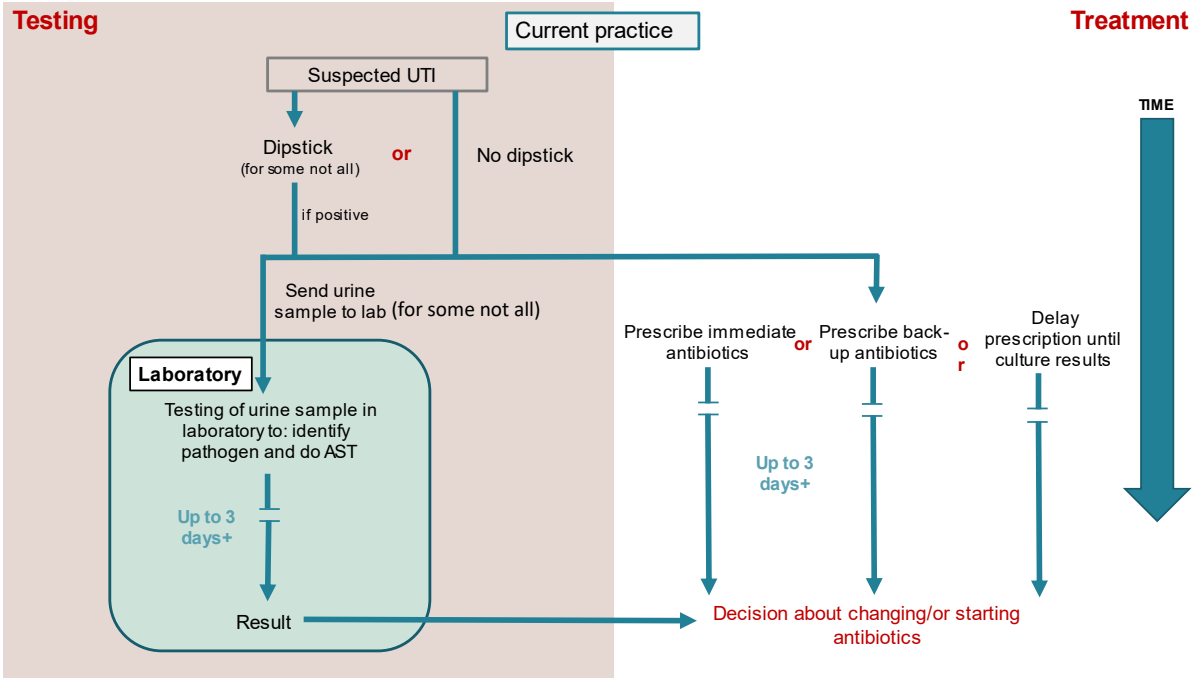
2.4 Comparator

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 done in the laboratory (population 2). This varies according to population. Further details on the treatment pathway are provided in section 2.5.

2.5 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. People present to their GP with symptoms suggestive of UTI. Depending on the patient population, they may receive dipstick testing. If this is positive for nitrite and LE they 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 before culture results are available, particularly if the patient is presenting with severe symptoms. This means that antibiotics may need to be changed if culture and AST suggest that the patient is taking an antibiotic that is not likely to be effective against their infection, or stopped if no infection is detected on culture.

Figure 1 Outline treatment pathway



Public Health England has separate pathways for infants/children under 16 years, women under 65 years, men under 65 years, adults who are catheterised, and adults over 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:

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

Population	Dipstick	Culture	Immediate antibiotics
Children (age <16 years)	Yes	If do not respond to treatment in 24-48 hours or age <3 years & positive dipstick for nitrite and LE	Yes (depending on dipstick result)
Men age <65	Yes – but not to rule out infection	Yes	Yes
Women age<65	Yes – those without risk factors for complicated UTI. Not needed if have 2 or 3 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 age >65	No	Yes	Yes
Women age>65	No	Yes	Yes, or backup antibiotics if symptoms mild

2.6 Place of the technology in the treatment pathway

POCT for suspected UTIs would be used as an initial test to diagnose UTI. If performance is sufficient then the place of the test in the treatment pathway, as an initial test to diagnose UTI, will be the same in all populations and pre-specified subgroups (section 2.1).

The role of POCT tests for UTI will depend on whether they provide additional information on the specific pathogen present in the urine, whether they provide information on AST and the time it takes to produce. This will also affect the potential impact of the tests. Table 4 provides an overview of the potential role and impact of the new POCT based on the features of the test.

Table 4 Overview of the potential role and impact of the new POCT based on the features of the test

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. • Replacement of dipstick in populations where dipstick testing is recommended	• Inform need for antibiotics • Reduce unnecessary antibiotic prescription • 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. • Replacement of dipstick in populations where 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 UTI plus AST	• Replacement of dipstick & laboratory testing	• Inform need for antibiotics • Reduce unnecessary antibiotic prescription • Quicker access to antibiotics when needed • Target initial antibiotic prescription to AST • Reduce need for culture & AST

3 Objectives

The overall aim of this project is to determine whether POCT for people with suspected UTI have the potential to be clinically 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 POCT to diagnose UTI, with or without additional pathogen identification and AST?
2. What is the accuracy of the POCT for UTI diagnosis, pathogen identification and AST?
3. What is the technical performance (other than accuracy) of POCT for UTI?
4. What are the costs, from a UK NHS and Personal Social Services (PSS) perspective, of using POCT 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 POCT for UTI diagnosis, pathogen identification and AST?

4 Methods for assessment of clinical effectiveness

A systematic review was conducted to summarise the evidence on the accuracy, technical performance and clinical effects of using POCT for 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.[22-24] The review is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) 2020 guidance[25] – see Appendix 4: PRISMA 2020 Checklist. The review protocol was registered on the PROSPERO database (CRD42022383889).

4.1 Inclusion and exclusion criteria

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

Table 5 Inclusion Criteria for Objectives 1, 2 and 3

	Obj 1: Clinical Impact	**Obj 2: Accuracy**	**Obj 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	NA
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 (NRSI)	Diagnostic test accuracy (DTA) study	Any

Given the tight timelines to conduct 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 before will have changed such that studies published before this date are unlikely to provide useful information to inform this appraisal. Animal studies were excluded.

4.2 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.[26, 27]

4.2.1 Bibliographic searching

The following databases were searched:

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

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: Literature search strategies.*

4.2.2 Non-bibliographic search methods

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

- ClinicalTrials.gov via https://www.clinicaltrials.gov/
- WHO International Clinical Trials Registry Platform (ICTRP) via https://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
- Hand searching the websites of the manufacturer/or licence holders for each test
- Reviewing information submitted by test manufacturers

4.2.2.1 Managing the searches

Search results were exported to EndNote 20 for deduplication using the default deduplication settings and manual review of records. Search results were then exported to Microsoft Access for screening.

4.2.3 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 (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 (section 4.1). 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 under 65
- Women over 65
- Men under 65
- Men under 65
- Adults with indwelling urinary catheters
- Babies, children and young people under 16
- Children under 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 UTIs (for example 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) together with 95% confidence Intervals (CI) 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 2x2 tables comparing the POCT with the reference standard, where available. If measures of accuracy (e.g. sensitivity, specific, Receiver Operating Characteristic (ROC) plot) were reported without providing the information needed to calculated 2x2 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 2x2 data were reported in a single study, for example for different tests, target conditions, thresholds, or subgroups of interest, all data were extracted.

4.2.4 Quality assessment strategy

The methodological quality of included RCTs was assessed using the updated Cochrane Risk of Bias Tool (RoB 2).[28] We had intended to assess the risk of bias in NRSI using the ROBINS-I tool, but no studies of this design were identified.[29] DTA studies were assessed for methodological quality using QUADAS-2.[30] 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.

4.2.5 Synthesis methods

For each of the three systematic review objectives (1 to 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) culture based tests.

For Objective 2, coupled forest plots of sensitivity and specificity were used to display results from individual studies, to allow visual assessment of heterogeneity. For this plot we selected one set of 2x2 data per study/population and test. If multiple index test and culture thresholds were reported 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% confidence intervals (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.[31, 32] Where results were pooled across only three or two studies, univariate random effects or fixed effect meta-analyses respectively was performed, due to lack of data to estimate 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.

4.3 Protocol changes

- We had originally specified that studies would only be included for objective 3 if they evaluated a test that had not been considered as part of objectives 1 or 2. However, due 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 be confirmed by the company.

5 Results of the clinical effectiveness review

5.1 Search Results

The searches of bibliographic databases and trials registries identified 728 unique references after de-duplication. 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, separate reports of diagnostic test accuracy sub-studies 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 2. The final study was a report of a qualitative sub-study from the POETIC trial.

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 3 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 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 mins			
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

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

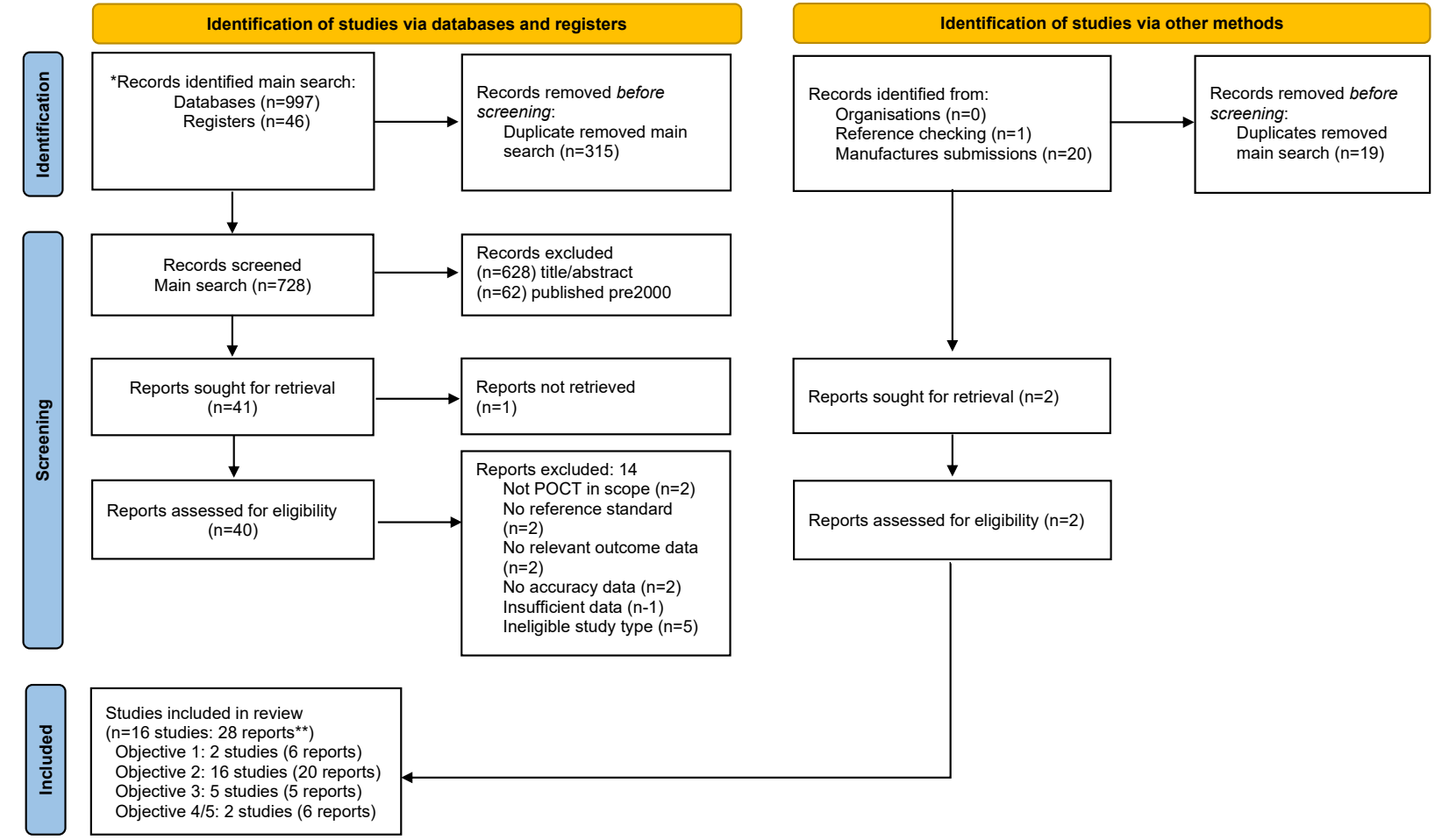
Table 7 provides an overview of the populations defined in the scope and whether data were available for these populations. The majority of populations were not specifically considered in the included studies, although may have been included in studies that enrolled mixed populations.

Table 7 Overview of populations defined the scope and whether data were available specifically for each population of interest

Population	Data available for specific groups of interest?
People with suspected acute UTI	Yes
People with suspected recurrent UTI	No
People with suspected chronic UTI	No
Women under 65	Yes (studies of women only; no age restrictions)
Women over 65	
Men under 65	No
Men over 65	No
Adults with indwelling urinary catheters	Yes
Babies, children and young people under 16	Yes
Children under 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

We excluded studies published before the year 2000, as outlined in the Methods section. These were excluded after title and abstract screening. Appendix 2.1 Pre-2000 studies 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 5.

Figure 2 Flow of studies through the review process



* after the main searches had been completed, as additional test (UTRiPLEX) was added to the scope of the review.
** Studies and study reports contributed to more than one objective.

5.2 Objective 1: What is the impact on clinical outcomes of using POCT to diagnose UTI, with or without additional pathogen identification and AST?

Two individually randomised RCTs evaluated the clinical impact of using Flexicult Human (often referred to in studies as the Flexicult™ SSI urinary kit) – the point of care testing for urinary tract infection in primary care (POETIC) trial[8] and a Danish trial.[33] Both trials were conducted in primary care and enrolled women aged over 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 sub-study could be performed – results for these two sub-studies are included for objective 2 (section 0).[34, 35] Both studies were considered at low risk of bias (Appendix 1.1). 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, 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 was then excluded, resulting in a sample of 653 women. Flexicult plates specific for the antibiotics most commonly used in each of the three regions were developed. GPs were free to determine how best to use the test. Examples of how it could be used included:

- 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

The Danish trial randomised 376 women to two different Flexicult based strategies – Flexicult Human (which incorporate 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 was reduced at the initial consultation (odds ratio (OR) 0.56, 95% confidence interval (CI) 0.35, 0.88) but this did not impact on overall antibiotic prescription or on antibiotic use that was concordant with culture results (the primary outcome). Concordant antibiotic use was defined 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 a suggestion that appropriate prescribing was higher in the control arm rather than the Flexicult Human arm (OR 1.11, 95% CI 1.03, 1.99). Appropriate prescribing was defined as:

(1) if the patient had UTI in the reference: to prescribe a first-line antibiotic to which the infecting pathogen was susceptible

(2) if the patient had UTI but was allergic to the antibiotic or the pathogen was resistant to all first-line antibiotics: to prescribe a second-line antibiotic
(3) if the patient did not have UTI in the reference: not to prescribe an antibiotic

Both trials also looked at improvement or duration of symptoms and microbiological cure. There was no evidence for 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.

Table 8 Results of trials of clinical impact of the Flexicult human test

Study	Outcome		Effect measure – estimate (95% CI)
Antibiotic use			
Butler (2018)[8] *(POETIC Trial)*	Concordant antibiotic use		OR = 0.84 (0.58, 1.20)
	Antibiotic prescribing at initial consultation		OR = 0.56 (0.35, 0.88)
	Antibiotics prescribed to guidelines at initial consultation		OR = 0.99 (0.67, 1.45)
	Antibiotic consumed day 3		OR = 1.24 (0.81, 1.89)
	Antibiotic consumed (during 2 weeks)		OR = 1.38 (0.87, 2.19)
	New antibiotic prescription (within 2 weeks)		OR = 1.11 (0.65, 1.89)
	Drug type and duration	UTI-specific and 1–3 days	Reference
		UTI-specific and >3 days	RR = 1.15 (0.71, 1.87)
		Broad spectrum and 1–3 days	NA (0 events)
		Broad spectrum and >3 days	RR = 1.00 (0.58, 1.75)
Holm (2017)[33] *(Danish Trial)*	Appropriate prescribing		OR = 1.44 (1.03,1.99)
UTI/symptom incidence or duration			
Butler (2018)[8] *(POETIC Trial)*	Microbiologically confirmed UTI (at 2 weeks)		OR = 0.94 (0.49, 1.81)
	Recurrence of UTI within 3 month period		OR = 0.72 (0.48, 1.07)
	Duration of symptoms		HR = 1.02 (0.83, 1.25)
	Duration of moderately bad symptoms		HR = 0.98 (0.82, 1.17)
	Overall urinary symptom burden		MD = 0.99 (0.84, 1.19)
	No significant bacteriuria on day 14		OR = 1.15 (0.62, 2.13)
Holm (2017)[33] *(Danish Trial)*	Symptom free on day 5		OR = 0.91 (0.56, 1.49)
Enablement			
Butler (2018)[8] *(POETIC Trial)*	Patient enablement (measured using Patient Enablement Instrument at day 14 and 3 months[36])		OR = 0.99 (0.66, 1.48)
Resource use			
Butler (2018)[8] *(POETIC Trial)*	Re-consultation (within 2 weeks)		OR = 0.99 (0.62, 1.60)
	Hospital stay (within 2 weeks)		Numbers too small

5.3 Objective 2: What is the accuracy of the POCT for UTI diagnosis, pathogen identification and AST?

Sixteen studies, reported in 20 publications, reported data on test accuracy and were included for this objective.[18, 34, 35, 37-49] Studies were conducted in Denmark (3),[35, 37, 38] Wales (2),[18, 49] Israel (2),[47, 48] Hawaii (1),[39] Venezuela (1),[40] Belgium (1),[41] Mexico (1),[42] Philippines (1),[43] South Africa (1),[44] Korea (1),[45] Argentina (1),[46] and one study was undertaken in Wales, England, Spain, and the Netherlands (1).[34] Most studies were reported in English with the exception of one in Korean [45] and one in Spanish.[42] These were translated using Google translate. 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.[49] All other studies were published as full reports. Table 9 provides an overview of the included studies' key characteristics. Full study details of each included study are reported in Appendix 3.2: Objective 2.

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 as the Flexicult™ SSI urinary kit in all studies);[18, 34, 35, 37] three Uricult trio; [43-45] two ID Flexicult;[35, 38] two Dipstreak,[47, 48] and one evaluated Uricult.[46] The only rapid test to be evaluated in multiple studies was the Uriscreen test, which was evaluated in four studies;[39-42] UTRiPLEX[41] and Lodestar DX[49] 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.[35, 41] The manufacturer's submissions highlighted two ongoing studies that will provide data on the accuracy of the Astrego PA-100 AST test and the Lodestar DX, rapid tests for UTI.

[redacted] The Lodestar submission highlighted the TOUCAN study, a study evaluating the accuracy of three or four POCT (details of these not yet available) in up to 800 women who consult their GP with symptoms of UTI. This study is due to complete in October 2023.[50]

Four studies were laboratory-based; three tested fresh urine samples [18, 47, 48] and [redacted].[49] 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.[41, 46]

Four studies recruited pregnant women,[39, 40, 44, 46] three recruited women with uncomplicated UTI, [34, 35, 38] one enrolled catheterised ICU patients,[42] and three studies recruited children and/or infants aged under: 18 years,[41] 16 years,[43] and 24 months.[45] Five studies analysed samples from mixed populations: people visiting outpatient clinics and hospitalized patients;[47, 48] symptomatic patients consulting the GP;[37]

[redacted].[18, 49] No further information was provided on these mixed populations. Three studies specifically stated that those with recurrent UTI were excluded;[35, 42, 43] information on whether those with recurrent or chronic UTI were eligible was not reported in the remaining studies.

Seven studies enrolled symptomatic patients[34, 35, 37, 38, 41, 43, 45] and four enrolled asymptomatic patients.[39, 40, 42, 46] One study included a mixture of asymptomatic and symptomatic patients and stratified results accordingly.[44] The four laboratory-based studies did not specify whether urine samples came from symptomatic patients, [redacted].[18, 47-49]

In the 10 studies that enrolled people and then took urine samples to test for UTI,[34, 35, 38-41, 43-46] the number of patients ranged from 117 to 2173 (mean 459). Another study enrolled 57 patients and took multiple samples from each patients giving a total of 108 samples.[42] [redacted].

One study was funded by the test manufacturer.[49] One study was funded by industry (not test manufacturer) and non-industry.[43] Seven studies did not report funder details [37, 40, 42, 44, 45, 47, 48] and all other studies were non-industry funded.

All included studies except for one[49] assessed the accuracy of POCT for the detection of the presence of UTI. Three of these studies also reported data on antimicrobial sensitivity,[18, 34, 37] one reported data on pathogenic cause,[48] and one reported data on presence of E.Coli.[45] [redacted] 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.[18] The threshold for culture varied between studies but was often reported at $\geq 10^3$ Colony Forming Unit (CFU), $\geq 10^4$ CFU, or $\geq 10^5$ CFU (see Appendix 3.2: Objective 2).

Table 9 Characteristics of the 16 studies reporting on the accuracy of POCT

	Rapid tests (results <40 minutes)			Culture based tests (results up to 24 hours)				
	[redacted]	**Uriscreen**	**UTRiPLEX**	**Flexicult Human**	**ID Flexicult**	**Uricult trio**	**Uricult**	**Dipstreak**
# studies*	[redacted]	4	1	4	2	3	1	2
Reference	[redacted]	[39-42]	[41]	[18, 34, 35, 37]	[35, 38]	[43-45]	[46]	[47, 48]
Population	[redacted]	**2** Screening - pregnant women **1** Children (<18 yr) **1** Catheterised ICU	**1** Children (<18yr)	**2** Women - uncomplicated UTI **1** Mixed **1** Mixed	**2** Women - uncomplicated UTI	**1** Pregnant women **1** Children (<16yr) **1** Aged <24 months	**1** Screening - pregnant women	**2** Mixed
Urine sampling	[redacted]	**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	[redacted]	**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	[redacted]	**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	[redacted]	**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	[redacted]	**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	[redacted]	**3** Near patient **1** Laboratory	**1** Laboratory	**1** Laboratory **3** Near patient	**2** Near patient	**3** Near patient	**1** Laboratory	**2** Laboratory

*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[35, 41]

**[redacted]

Note: POU: Presence of UTI. AMS: Antimicrobial sensitivity. PC: Pathogenic cause. POE: Presence of E.Coli. GP: General practice. NR: Not reported. Mixed: Laboratory-based studies using samples from a mixed population e.g. hospitalised patients and outpatients (does not refer to whether patients had symptoms or not – this and further detail is reported in section 5.3).

5.3.1 Risk of bias

Table 10 presents an overview of the results of the risk of bias assessment for the studies included for objective 2; full details are reported in Appendix 3: Data extraction tables. Five studies were judged as being at high risk of bias. In three studies this was due to the exclusion of a large proportion of patients from the analysis,[44, 46, 48] in one study participant selection was unclear and multiple samples were taken from some patients,[42] and [redacted][49] 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 results of the POCT. 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 with culture as a reference standard – this is discussed in more detail in the discussion section. Eight studies were judged as being at an unclear risk of bias.[18, 34, 37-40, 45, 47] 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.[18, 45, 47] Three studies were judged as low risk of bias.[35, 41, 43] Two of these reported data on test comparisons[30, 35] therefore QUADAS-C assessments were also completed. All domains on QUADAS-C were judged at low risk of bias.

Table 10 Overview of risk of bias in studies that evaluated the accuracy of POCT tests

Study Details	Patient Selection	Index test	Reference standard	Flow & Timing	Overall	Rationale for Judgement
[redacted] **Test:** Lodestar DX	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Macias(2002)[42] **Test:** Uriscreen	☹	☺	?	☺	☹	Multiple samples taken from some patients; unclear how patients selected for inclusion.
Millar(2000)[39] **Test:** Uriscreen	☺	☺	?	☺	?	No information on blinding of interpreter of reference standard
Teppa(2005)[40] **Test:** Uriscreen	☺	☺	?	☺	?	No information on blinding of interpreter of reference standard
Boon(2022)*[41, 51] **Test:** UTRiPLEX & Uriscreen	☺	☺	☺	☺	☺	No concerns. There was a high amount of exclusion in the Uriscreen v culture comparison but this was due to late introduction of the test.
Blom (2002)[37] **Test:** Flexicult Human	?	☺	?	☺	?	No information on blinding of interpreter of reference standard
Bongard(2015)[18]	?	☺	?	☺	?	Unclear if consecutive patients were enrolled; No information

Study Details	Patient Selection	Index test	Reference standard	Flow & Timing	Overall	Rationale for Judgement
Test: Flexicult Human						on blinding of interpreter of reference standard
Hullegie(2017)[34] **Test:** Flexicult Human	☺	☺	?	☺	?	No information on blinding of interpreter of reference standard
Holm(2017)*[35] **Test:** Flexicult Human & ID Flexicult	☺	☺	☺	☺	☺	No concerns
Pernille(2019)[38, 52] **Test:** ID Flexicult	☺	☺	?	☺	?	No information on blinding of interpreter of reference standard
Colodner(2000)[47] **Test:** Dipstreak	?	☺	?	☺	?	Unclear if consecutive patients were enrolled; No information on blinding of interpreter of reference standard
Yagupsky(2000)[48] **Test:** Dipstreak	?	☺	?	☹	☹	High proportion of patients excluded from analysis
Mignini(2009)[46] **Test:** Uricult	☺	☺	?	☹	☹	High proportion of patients excluded from analysis
Anacleto(2009)[43] **Test:** Uricult Trio	☺	☺	☺	☺	☺	No concerns
Greeff(2002)[44] **Test:** Uricult Trio	☺	☺	?	☹	☹	High proportion of patients excluded from analysis
Lee(2010)[45] **Test:** Uricult Trio	?	☺	?	☺	?	Unclear if consecutive patients were enrolled; No information on blinding of interpreter of reference standard

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

5.3.2 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 is 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 3.2: Objective 2.

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

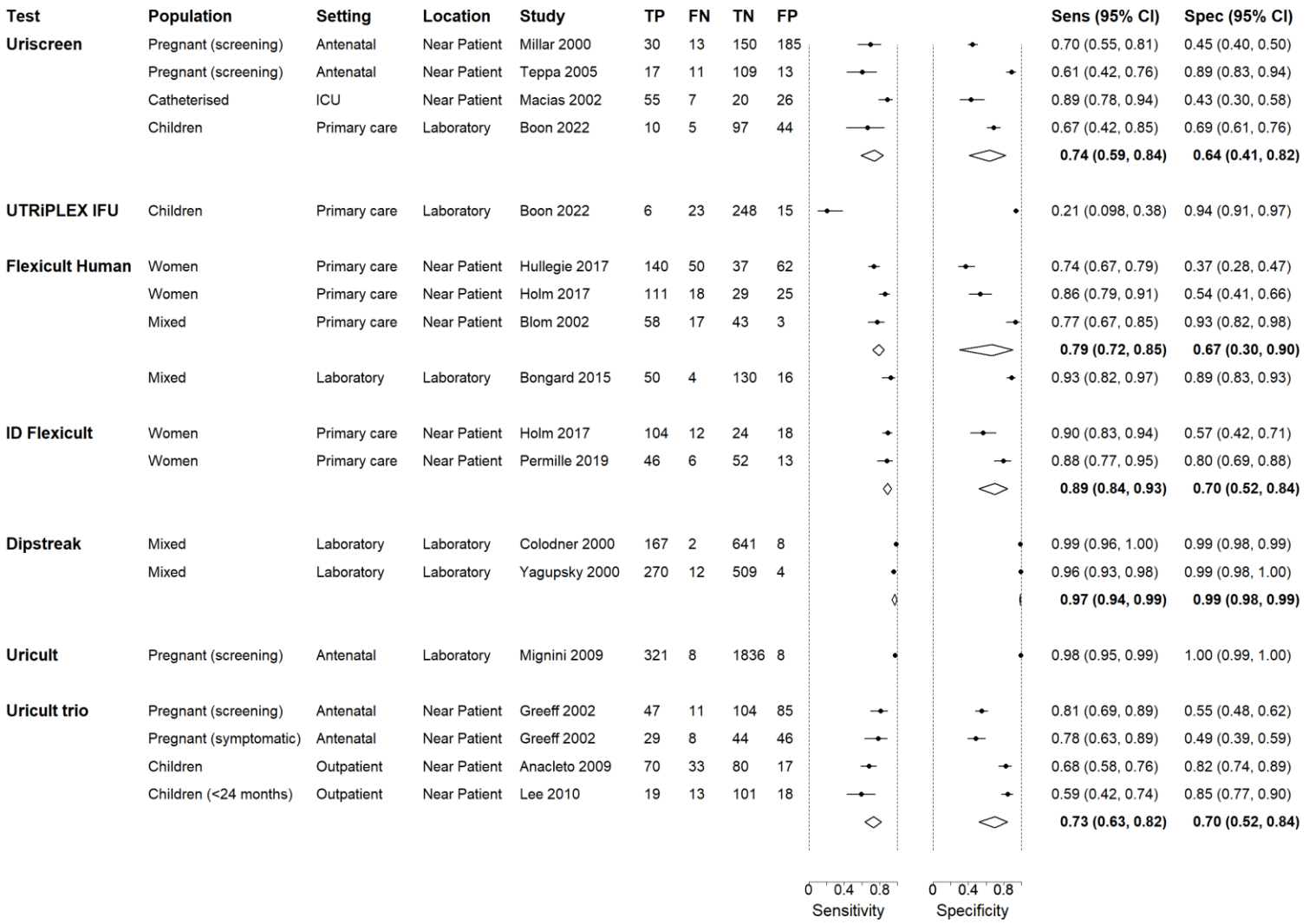


Table 11 Summary of whether data were available on diagnosis of UTI, pathogenic cause and antimicrobial sensitivity for each test

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	✓	✗	✗

5.3.2.1 Lodestar DX

[redacted]

Pathogenic cause

[redacted]

[redacted]

[redacted]

Table 12[redacted]

Table 12 Accuracy of Lodestar DX for detecting pathogens using stored urine samples

Target condition	Sens	Spec
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

5.3.2.2 Uriscreen

Four studies evaluated Uriscreen.[39-42] One study analysed 156 children aged <18 years in primary care in Belgium and conducted the POCT in the laboratory.[41] Three other studies conducted the POCT in a near-patient setting and analysed 378 pregnant women from antenatal clinics in Hawaii,[39] 150 pregnant women from antenatal clinics in Venezuela,[40] and 108 samples from 57 catheterised ICU patients in Mexico.[42] Two studies used catheterised urine samples,[40, 42], one used mid-stream sampling,[39] and one used mid-stream or adhesive bags.[41] One study was judged as being at low risk of bias,[41] two at unclear risk of bias,[39, 40] and one at high risk of bias[42] (see Table 9).

Presence of UTI

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, 84) and summary specificity was 64% (95% CI 41, 82). There were no clear reasons for the observed heterogeneity.

5.3.2.3 UTRiPLEX

One study evaluated UTRiPLEX .[41] 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 via mid-stream sampling or the use of adhesive bags, as per clinical practice. It was judged at low risk of bias (see Table 9).

Presence of UTI

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

5.3.2.4 Flexicult Human

Four studies evaluated Flexicult Human.[18, 34, 35, 37] This included test accuracy sub-studies from the two trials included for objective 1.[34, 35] 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 sub-studies from trials were restricted to women (over 18 years) with uncomplicated UTI – one of these analysed 183 women,[35] and one analysed 289 women[34]. One study analysed 121 samples from a mixed population of symptomatic patients in Denmark,[37] and one study was laboratory-based using 200 fresh urine samples from a mixed population in Wales.[18] Mid stream urine samples were collected in the two trial sub-studies[34, 35]. The laboratory-based study collected samples using different methods: mid-stream sampling (n=134), catheter sampling (n=7), and for 65 samples the method was unknown. One study did not report how urine samples were collected.[37] Three of the studies were judged to be at unclear risk of bias[18, 34, 37] and one was at low risk of bias[35] (Table 9).

Presence of UTI

All studies provided data on the accuracy of the Flexicult Human test for diagnosing a UTI. Three used culture alone as the reference standard.[34, 35, 37] One study used two reference standards: 1) culture and microscopy and 2) culture, microscopy and spiral plating.[18] Another study used three different reference standard definitions to define a UTI: $\geq 10^4$ CFU/ml pure culture of pathogen; $\geq 10^5$ CFU/ml mixed growth with one predominant pathogen; OR $\geq 10^3$ CFU/ml of E.coli or S. saprophyticus (Public Health England/ Health Protection Agency), $\geq 10^5$ CFU/ml pure culture of uropathogen OR $\geq 10^5$ CFU/mL predominant culture a uropathogen with 3 log difference between highest and next species (UK lab definition) and $\geq 10^3$ CFU of uropathogen (European definition).

The Flexicult Human thresholds used to define the presence of UTI varied. Two studies used $\geq 10^3$ CFU/ml,[18, 35] one used $\geq 10^4$ CFU/ml,[37] and one used 10^3 CFU/ml for pure culture of a pathogen and $\geq 10^3$ CFU/ml for predominant growth of a pathogen in mixture with normal flora.[34]

Estimates of sensitivity ranged from 74% to 93% and specificity ranged from 37% to 93%. Estimates were highest in the laboratory based study of mixed urine sample (93% and

89%).[37] 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%), 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, 85) and 67% (95% CI 30, 90).

Antimicrobial sensitivity

Three studies reported data for antimicrobial sensitivity.[18, 34, 37] Estimates of sensitivity ranged from 79% to 90% with a summary estimate of 87% (95% CI 83, 90). Estimates of specificity ranged from 72% to 94% with a summary estimate of 93% (95% CI 89,95).[18, 34, 37]

5.3.2.5 ID Flexicult

Two studies evaluated ID Flexicult.[35, 38] 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[35]; 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 [35] and one had unclear risk of bias[38] (see Table 9).

Presence of UTI

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, 93) and summary specificity was 70% (95% CI 52, 84). The studies used thresholds of 10^3 CFU/mL (primary pathogens) and 10^4 CFU/mL (secondary pathogens) for the POCT.

5.3.2.6 Dipstreak

Two studies evaluated Dipstreak.[47, 48] 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;[48] the other analysed 818 samples (urine sampling method not reported). One study was judged at high risk of bias [48] and one was at an unclear risk of bias[47] (see Table 9).

Presence of UTI

Both studies found Dipstreak to be highly accurate in detecting UTI. Sensitivity was estimated at 96% and 99%; both studies estimated specificity at 99%. Summary sensitivity was 95% (95% CI 94, 99) and summary specificity was 99% (95% CI 98, 99). One of these studies evaluated two Dipstreak thresholds (10^4 & 10^5 CFU/ml)[47] and found similar results; the other did not report the Dipstreak threshold.[48]

Pathogenic cause of UTI

Yagupsky (2000) reported that Dipstreak correctly identified the pathogenic cause of UTI in 211/270 cases (the other 59 were not identified).

5.3.2.7 Uricult

One study evaluated Uricult.[46] 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 UTI

The Uricult test had excellent sensitivity (98%) and specificity (100%) in detecting the presence of UTI, using a threshold of $>10^5$ CFU.

5.3.2.8 Uricult trio

Three studies evaluated Uricult Trio.[43-45] Populations varied: one analysed 374 pregnant women in antenatal clinics in South Africa[44], one analysed 151 infants aged <24 months from outpatient clinics in Korea[45] and one analysed 200 children <16 years from outpatient clinics in the Philippines.[43] 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,[44] one at unclear risk of bias,[45] and one at low risk of bias[43](see Table 9).

Presence of UTI

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

Pathogenic cause

One study reported that the sensitivity of Uricult Trio for detection of the presence of E.Coli infection was 60% and specificity was 96%.

5.3.2.9 Test comparisons

Two studies reported data on two POCT tests included in the scope.[35, 41] 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 POCT for the detection of the presence 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.[35] The study recruited 341 women with uncomplicated UTI in Denmark. Patients were randomized 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 under 18 years in Belgium.[41] Three hundred samples were taken systematically and tested. However, much fewer results (156 vs 292) were available for Uriscreen test than the UTRiPLEX test because it was introduced later in the trial, making it difficult to compare the tests. Sensitivity and specificity was reported at 67% and 69% for Uriscreen and 21% and 94% for UTRiPLEX.

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

5.3.2.10 Comparison with standard urine dipstick tests

Six studies provided a direct comparison between the POCT tests and standard urine dipstick testing for LE or nitrite.[38, 39, 41, 45, 46] Four of these defined a positive dipstick tests 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 to standard dipstick testing, with different findings which may be related to how a positive dipstick test was defined (Table 13). One study also evaluated UTRiPLEX which was found to be less sensitive but more specific than dipstick testing. Three studies compared culture based POCT to standard dipstick testing. All found that the POCT culture were more sensitive and more specific than standard dipstick tests.

Table 13 Estimates of sensitivity and specificity for standard dipstick tests and POCT tests from studies that evaluated both tests

Study	Population	Test	Sensitivity (95% CI)	Specificity (95% CI)
Rapid tests				
Boon(2022)[41]	Children <18 years	UTRiPLEX	21 (8, 40)	94 (91,97)
		Uriscreen	67 (38, 88)	69 (60, 76)
		Dipstick (either nitrite or LE positive considered positive)	32 (16, 52)	86 (82,90)
Macias (2002)[42]	Catheterised ICU patients	Uriscreen	66.7	74.1
		Dipstick – nitrite only	66.7	45.2
		Dipstick – LE only	78.9	47.2
Millar (2000)[39]	Pregnant women (screening)	Uriscreen	70 (57, 84)	45 (40,51)
		Dipstick (both nitrite and LE positive considered positive)	81 (69, 93)	97 (95,99.2)
Culture based tests				
Pernille (2019)[38]	Women – uncomplicated UTI	ID Flexicult	88 (80,97)	80 (70, 90)
		Dipstick (either nitrite or LE positive considered positive)	73 (59,84)	75 (63,85)

Mignini(2009)[46]	Pregnant women (screening)	Uricult	98 (96,99)	99.6 (99.3, 99.8)
		Dipstick (either nitrite or LE positive considered positive)	53 (48,58)	92 (91,93)
Lee (2010)[45]	Children <24 months	Uricult Trio	59%	85%
		Dipstick (either nitrite or LE positive considered positive)	50%	76.7%

5.4 Objective 3: What is the technical performance (other than accuracy) of POCT for UTI?

Data on technical performance data were reported in five publications. Three reported data for Flexicult Human[8, 53] (two of these reported on the POETIC trial[8, 53]) and two reported data for Uricult Trio.[43, 44] Of these, one publication was also included for objective 1[8] and three were included for objective 2.[54] A further study[54] 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 3.3: Objective 3.

5.4.1 Flexicult Human

The Butler (2018) trial that compared testing and treating based on results of Flexicult Human with no treatment reported additional technical performance data on the Flexicult Human test.[8] These data are summarised in Table 14. They found that in 63% of participants the management was changed as a result of the test. Estimates of time to perform the test were 9 mins 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.

Table 14 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 mins
	Obtain and record result	6 mins
	Discuss result with patient	7 mins
Cost	Cost per person, including POCT cost in UK	£48

A qualitative sub-study of the POETIC trial, interviewed 35 clinicians who used the Flexicult Human test in the POETIC trial.[53] The study found that "clinicians overwhelmingly felt that a POCT for UTI management would be useful." It reported that most clinicians agreed that the Flexicult Human test gave quicker results than lab tests (24hr vs 3-4 days), reassured patients, and had a positive impact on clinician confidence in diagnosing UTI. There was an even split between those that thought it would have no impact on prescribing and those who stated that it had increased their awareness about antibiotic prescribing and they therefore had more cautious prescribing habits. However, they noted difficulties in test result interpretation, limitations on when it can be used, limited resources to undertake testing, and concerns about prolonging patient discomfort whilst waiting for test results and

about the potential expense of maintaining regular stock of tests. They highlighted that an ideal POCT test 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 general practitioners considered Flexicult Human to be easy to handle and read.[37]

5.4.2 Uricult Trio

One study reported that Uricult Trio was convenient to use and easy to interpret.[43] Another study[44] agreed 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".

6 Assessment of cost-effectiveness

In this section 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 Section 0 and was coded in the R programming language.[55] Results of the implemented model are not presented as evidence was too limited for findings to be meaningful.

6.1 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 POCT for UTI in comparison to culture with or without dipstick tests. The model described below is for all possible comparators and populations/subgroups described in Section 2.1. Separate models would be required for each population/subgroup.

In Section 6.2 we review the clinical evidence identified in Section 5, and evidence identified by pragmatic searches, to narrow the focus on tests and populations where evidence and impact are greatest.

6.1.1 Testing strategies

The POCT considered were those included in the scope outlined in Table 2. These include rapid tests (results <40 mins) 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 Section 2.4, 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.

6.1.2 Subgroups of interest

As per Section 2.1 the populations in scope are those with suspected UTI, but subgroups of interest include:

Patient subgroups identified by Public Health Enland guidance:

A. Women under 65
B. Women over 65

C. Men under 65
D. Men over 65
E. Adults with indwelling urinary catheters
F. Babies, children and young people under 16

Other patient subgroups:

G. People with suspected acute UTI
H. People with suspected recurrent UTI
I. People with suspected chronic UTI
J. Children under 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 UTIs (for example people with neurogenic bladder, diabetes, polycystic kidney disease or people who are immunocompromised)
Q. People with suspected pyelonephritis

6.1.3 Conceptual model

Our conceptual model is illustrated in Figure 5. Arrows indicate the influence of components on the rest of the model.

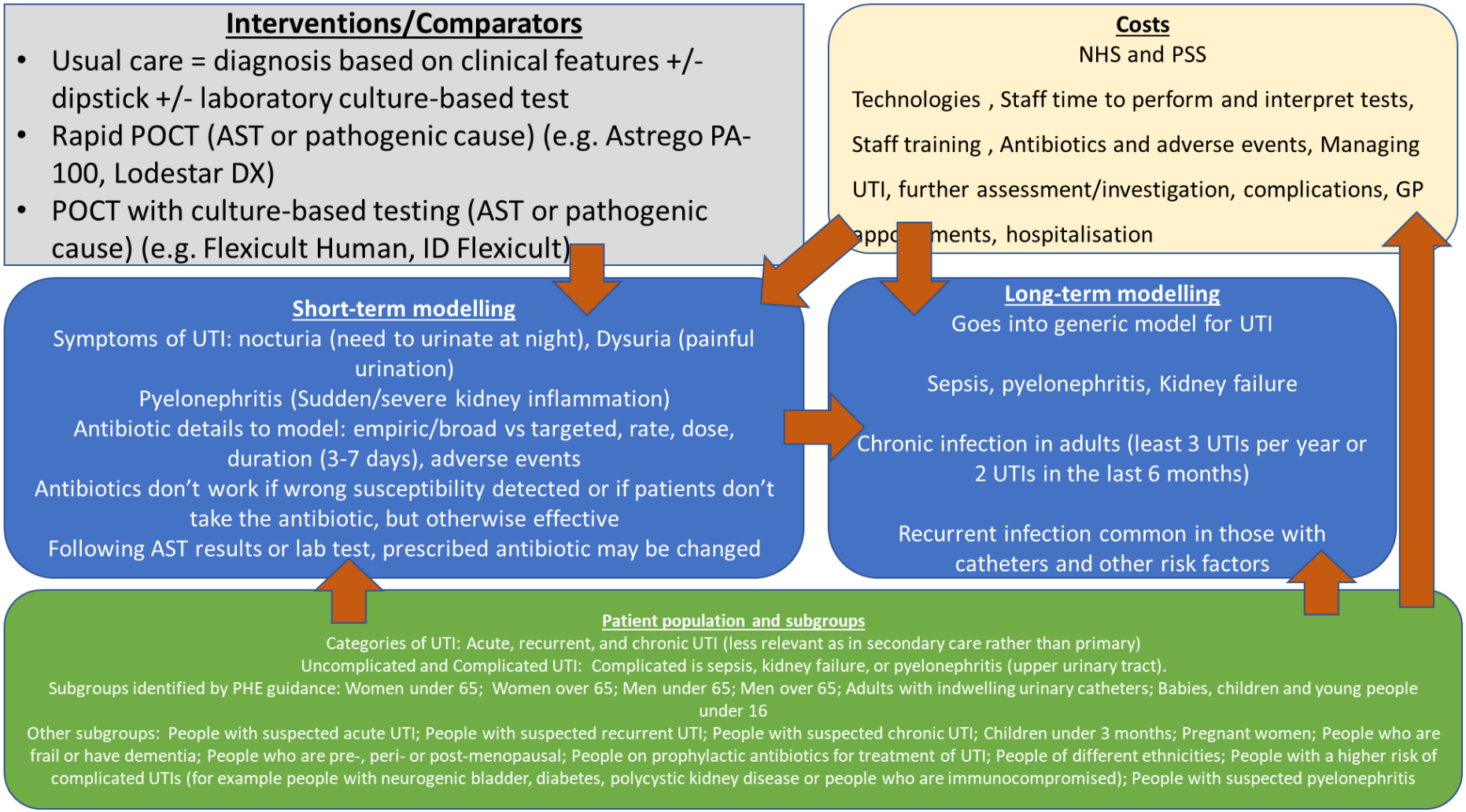
Our conceptualisation was divided into short-term and long-term components. In the short-term, the important elements to consider were the symptoms of complicated and uncomplicated UTI, characteristics and consequences of antibiotics, expected efficacy of antibiotics, and any response to ineffectiveness 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 Section 6.1.1.

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 Section 2.1.

Figure 5 Conceptual model for point of care tests in UTI*



*Boxes illustrate important elements to consider. Arrows illustrate influence. AS=antibiotic susceptibility; AST=antibiotic susceptibility test; GP=General Practice; NHS=National Health Service; PHE=Public Health England; POCT=Point of Care Test; PSS=Personal Social Services

6.2 Review of evidence on cost-effectiveness

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

6.2.1 Relevant evidence from clinical effectiveness review

The search for the clinical effectiveness review (Section 4) was not limited by study design or publication type search filters, and therefore would also identify any relevant economic evidence. Our screening process is summarised in Figure 2 Flow of studies through the review process. We identified two relevant studies, these are discussed below.[8refs, 33]

6.2.1.1 Butler 2018 (POETIC)[8]

The Butler 2018 POETIC study was an RCT assessing the clinical and cost-effectiveness of Flexicult Human compared to standard of care in adult women who already had a clinical diagnosis of uncomplicated UTI. This study is described in Section 5. Cost-effectiveness was measured by total cost per unit increase in concordant antibiotic prescribing, but on this basis Flexicult testing was not cost-effective. They found that clinicians generally prescribed broad/empiric antibiotics, rather than waiting for the Flexicult results, and seldom withdrew antibiotic treatment in response to test results (Table 14). They also reported that duration of all UTI symptoms in both arms was 8 (range 5-14) days, and of moderately bad symptoms was 4 (2-6) days.

6.2.1.2 Holm 2017[33]

This study has been discussed under Objective 1 in Section 5.2. It was an RCT in women with suspected uncomplicated UTI comparing Flexicult Human and ID Flexicult. The primary outcome was appropriate antibiotic prescribing, as described in Section 5.2. Their overall finding was that including POCT AST did not improve antibiotic prescribing in general practice. As summarised in Table 8, they reported results on appropriate prescribing and on patient enablement (measured using Patient Enablement Instrument at day 14 and 3 month). However, neither outcome matches sufficiently to any in the conceptual economic model to be useable.

6.2.2 Additional pragmatic searches for cost-effectiveness evidence

We conducted pragmatic searches of MEDLINE (Ovid), Embase (Ovid) and Econlit (EbscoHost) databases using search terms described in Table 15. There were 24 studies identified after removal of duplicates. Thirteen were identified at title/abstract screening as potentially having useful information, though two of these were conference abstracts related to two full-text records. Two were studies related to the POETIC trial and already identified by the clinical effectiveness review. One study was a potentially relevant cost-effectiveness evaluation of trimethoprim-sulfamethoxazole and amoxicillin in UTI, but was inaccessible and published in 1987, and therefore not further considered.[56] The remaining 8 records were evaluated at full-text.

Table 15 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*
Econlit	("urinary tract infection" and "cost-effectiveness").ti.	0

*These hits included all studies identified by Ovid MEDLINE.

6.2.2.1 Wang 2021

Wang 2001 was a US based decision tree model considered both empirical antibiotics and culture-directed antibiotics, the latter of which aligns with our treatment strategy of targeted antibiotics.[55] The focus of their analysis was the impact of antibiotic resistance on cost-effectiveness of treatment strategies. They found that empirical antibiotics were the most cost-effective strategy if resistance was below 6%, while symptomatic treatment was most cost-effective if resistance was above 80%. However, at most levels of resistance empirical antibiotics, with simultaneous urine culture and later targeting of antibiotics, was the most cost-effective strategy. This aligns with our assumed standard of care of laboratory culture-based testing with empiric/broad antibiotics. This study reported Quality Adjusted Life Days (QALDs) for UTI cured, UTI, and pyelonephritis, which were extracted to Table 19.

6.2.2.2 Sadler 2017

Sadler 2017 was a UK based decision tree economic model 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.[57] Results were stratified by resistance to trimethoprim. Trimethropin 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 were most effective at over 35%.

6.2.2.3 Fenwick 2000

Fenwick 2000 used a decision tree model to compare cost-effectiveness of management strategies for UTI.[58] The model include branches for symptoms disappearing, persisting, and antibiotics working. They found that empiric antibiotic treatment based on symptoms was largely cost-effective compared to 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 (Table 20). We therefore use the probability and duration of side effects from this study (Table 19).

6.2.2.4 Whiting 2006

Whiting 2006 was a systematic review and economic model of effectiveness and cost-effectiveness of tests for the diagnosis and investigation of UTI in children.[59] The ultimate result of the overall systematic review was an algorithm for diagnosis of UTI in children under the age of 5.

Only one prior economic evaluation was identified by the systematic review, which was a US based cost-effectiveness decision tree model comparing diagnosis and management strategies for UTI in children aged 2 months to 2 years.[60] The Whiting 2006 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 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 thresholds, including the £20,000-£30,000/QALY commonly used by NICE, nitrite and leucocyte esterase followed by micturating cystourethrography was most cost-effective, as was nitrate or laboratory leucocyte esterase/culture-based testing, followed by micturating cystourethrography. These have limited relevance for our evaluation as POCT were not considered and guidelines on UTI treatment have been updated in the past 18 years. The population was also children only, so limited generality across our subgroups.

Utility data came from Barry 1997, which was a US-based cost-utility analysis of evaluation strategies for UTI in ambulatory women.[61] Although this source is out-dated, they 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 (Table 19).

6.2.2.5 Gaither 2020

Gaither 2020 developed a decision tree model to estimate the cost-effectiveness of routine, screening renal bladder ultrasound (RBUS) in children aged 2-24 months after a first febrile UTI. Their 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 a year. They used at US health system perspective with a willingness-to-pay threshold of $100,000 per QALY, and found that screening RBUS after a first febrile UTI was not cost-effective when compared to 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 to 72 months.[62]

6.2.2.6 Sanyal 2019

Sanyal 2019 used a decision tree model to compare the cost-effectiveness and budget impact of the management of uncomplicated UTIs in women by initiated community pharmacists versus management initiated by family or emergency physicians. Costs were based on cost data from Canada, and they 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 on the pharmacist-initiated arm and 90% of patients were cured of UTI on the family and emergency physician-initiated arms, though it is not clear which tests were used to assess UTI. They used quality-adjusted-life-months (QALMs) to model health outcomes, but do not explicitly report the QALMs used for different health states and instead report the utilities at the start and end of the 28-day assessment period. Ernst et al.,[63] which was their source, provides more detailed data from which a curve could be fitted to estimate the QALMs.

6.2.2.7 Kassabian 2022

Kassabian 2022 used a decision tree model to perform a cost-effectiveness analysis comparing fosfomycin to nitrofurantoin and trimethoprim-sulfamethoxazole (TMP-SMX) as treatment for uncomplicated UTI from a US perspective, and concluded that fosfomycin may be considered cost-effective, especially if taking account of antibiotic stewardship. 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 TMP-SMX. These estimates were derived using estimates of bacterial susceptibility and the proportions of UTIs caused by different bacteria.

6.2.3 Implications for cost-effectiveness modelling

The available evidence drove our selection of tests and subgroups to model. We furthermore prioritise 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 furthermore 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), and prioritised both over those that only detected UTI.

As summarised under objective 2, Table 6, the only rapid tests with accuracy data were Lodestar DX, Uriscreen, and ULTRoPLEX IFU, none of which 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, and not accuracy of detecting UTI itself. The only culture-based tests with accuracy data which performed AST were Flexicult Human and ID Flexicult, while those that identify 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 so not near patient or in primary care setting. We therefore exclude from modelling. Only one high risk of bias study provided accuracy data on the Uricult tests(Table 10), and this test also only identifies presence of gram-negative bacteria, and was in laboratory setting, so Uricult was not selected for modelling.

We therefore 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 was no

accuracy data so it could not be meaningfully modelled. Final selection of tests is summarised in Table 16.

The populations of interest evaluated by the included studies are summarised in Table 16. [redacted]. Flexicult Human and ID Flexicult were only evaluated in mixed and/or women with uncomplicated UTI. We thus focused evaluations on two populations in which we could model up to 3 tests each.

Mixed population

- Lodestar DX
- Flexicult Human

Women with uncomplicated UTI

- Lodestar DX [redacted])
- Flexicult Human
- ID Flexicult

Table 16 Final selection of tests, and summary of evidence, for modelling based on data availability and potential impact

Test	Rapid or culture	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 high risk of bias	Yes	[redacted]
Flexicult Human	culture-based	AST	3 at unclear, 1 at low	Yes	Women - uncomplicated UTI (2) Mixed (1 study)
ID Flexicult	culture-based	AST	1 at low risk, 1 at unclear	No	Women - uncomplicated UTI (2)
Could be modelled but no comparator in available populations					
Uricult trio	culture-based	identities bacteria	1 at high risk, 1 at unclear risk, 1 at low risk	No	Pregnant (2) Children <16 years (1) Children <24 months (1)

6.3 Evaluating costs, quality of life and cost-effectiveness

Using the conceptual model of Section 6.1 and evidence sources summarised in Section 6.2 we developed a structure and identified the necessary evidence to evaluate costs, quality of life and cost-effectiveness of POCT for UTI. Our model also assesses the reduction in use of

empiric/broad spectrum antibiotics, and therefore antibiotics use overall, as POCT with AST can yield targeted treatment and POCT without AST can indicate when no UTI is present. An NHS and PSS perspective was taken with a life-time 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; analyses 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 Section 6.2.3 has been implemented in the R programming language. Results are not presented from this model as the evidence identified is too limited for results to be meaningful, even for the subset of tests and populations evaluated.

6.3.1 Model structure

Our model structure comprises a decision tree over which the costs and consequences of testing for UTI would play out. Decision trees were the only type of model we identified as being used previously in the UTI literature.[57, 58, 64 59, 60] Key model assumptions are presented in Table 17.

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 doesn't distinguish between first or repeat UTI. We therefore did not adopt a long-term model, such as a cohort Markov model, for the long-term outcomes of Figure 5.

Our decision tree is illustrated in Figure 6. This structure is for rapid POCT that perform AST or identify pathogenic cause (e.g. Astrego PA-100, Lodestar DX), POCT 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.[65] Patients are assumed to either have a true UTI or no underlying UTI. Our conceptualisation is that the POCT with AST or pathogenic cause identification would either identify patients as having UTI and a specific antibiotic to which the patient is susceptible, identify patients as having UTI but not identify a specific antibiotic to which they are susceptible, or identify them as not having UTI. It is assumed that the POCT 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/empiric antibiotics, before targeted treatment is enabled by the results of the test. Under all strategies, if no UTI is detected, the patients 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/empiric antibiotics.

Probabilities of detecting UTI and, when with AST, detecting antibiotic susceptibility, would differ between POCT as per analyses in Section 0.

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

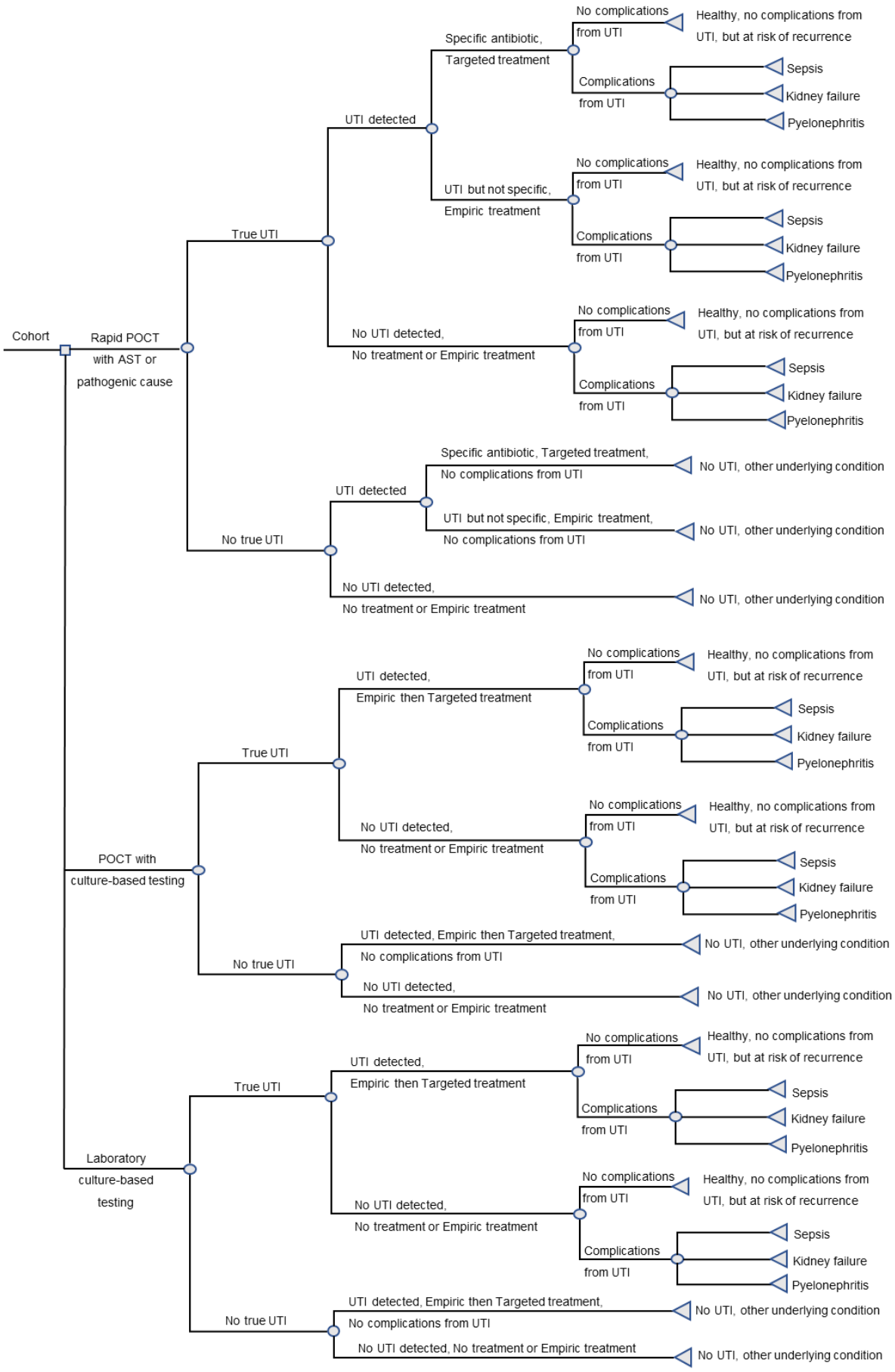
The decision tree then assumes that antibiotics would be assigned accordingly (e.g. targeted if specific susceptibility is known, empiric/broad if unknown, and not treated if known not to be a UTI). Empiric antibiotics are assumed to be potentially followed by targeted treatment if initial antibiotic is unsuccessful and 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 2021, Sadler 2017 and all previous economic models we identified.[57, 64] 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.

Table 17 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 particular exception is that manufacturer submissions note that Astrego can only be used 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 types 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 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 they can only be falsely given broad/empiric antibiotics.
(x)	The UTI is eventually cured by targeted or empiric courses of antibiotics, though patients may suffer complications and remain at risk of recurrent/chronic UTI.

Assumptions of the cost-effectiveness model	
(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 only be treated 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.
(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 antibiotics.
(xv)	Costs and QALYs of complications do not vary by subgroups.

Figure 6 Decision tree structure for short-term modelling*



*"False positives" do not incur further costs or consequences. We assume these branches only incur the cost/disutility of treatment, and that they have no benefit from the POCT. The “UTI but not specific, Empiric/broad treatment” arms include additional costs and QALY losses from where further testing is required to identify and prescribe an effective antibiotic. Recurrence of UTI takes place after decision tree and may include chronic UTI.

6.3.2 Model inputs

Where possible, model inputs were derived from the clinical review, from our additional searches in Section 6.2.2, or from expert opinion. We would recommend further systematic literature reviews and expert elicitation in a full-scale evaluation.

6.3.2.1 Test accuracy parameters

Test accuracy data are summarised in Table 18 but were derived from Section 5.3. Although estimates of sensitivity and specificity for detecting UTI were identified, there was little reliable data identified for the probability of identifying antibiotic susceptibility or pathogenic cause to direct targeted treatment. Sensitivity and specificity for detecting bacteria was identified for Lodestar DX (

Table 12) but these were based on stored rather than fresh urine, which give overestimates of what would be possible in primary care clinical practice.

Table 18 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 a 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, 0.85)	Specificity 0.67 (0.30, 0.90)	No reliable data identified by systematic review	Meta-analysis of women and mixed populations in primary care near patients (Figure 3)
Lodestar DX (Llusern Scientific)	Rapid, identifies pathogenic cause	[redacted]	[redacted]	[redacted] Table 12 [redacted]	[redacted]

Test name	Type of Test (POCT with/without AST)	Probability of correctly detecting a 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
ID Flexicult	Culture-based, with AST	Sensitivity 0.89 (0.84, 0.93	Specificity 0.70 (0.52, 0.84)	No reliable data identified by systematic review	Meta-analysis of women in primary care near patients. (Figure 3)

6.3.2.2 Other model input parameters

Values, distributions and evidence sources for other model input parameters are summarised in Table 19.

Table 19 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* $Beta(\alpha = 2.212762, \beta = 1.475174)$	Wang 2021, Schmiemann 2010[64, 66]	p_uti is different for each patient subgroup Diagnosis of UTI given symptoms was 0.6 (0-1) Wang 2021 and Schmiemann 2010[64, 66]
Probability of correctly detecting a UTI (sensitivity or true positive rate)	p_uti_tp	*See* Table 18	*See* Table 18	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 18	*See* Table 18	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 18	*See* Table 18	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 2015 and NICE NG109, but this did not distinguish between targeted and empiric treatment. [14, 67]

Input	Name used for code/ equations	Value(s) Random distribution	Source of value(s)	Comments
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 NG109, Falagas 2009) Mixed: Assumed same as in women.		p_healthy_emp is different for each test since 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 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
Probability of pyelonephritis on empiric treatment	p_pyelonephritis_emp	Women: 5.6% (NG109, Smaill 2015) Mixed: assume same as in women.		0.01 (0-0.02) in Wang 2021, Ferry 2007, Christiaens 2002.[64, 68, 69] 0.04 in Sadler 2017, for risk of pyelonephritis if clinical cure not achieved, Little 2009.[57, 70] We use Smaill 2015 and NICE NG109 as divided into treated and untreated although it relates to pregnant women and does not distinguish between targeted and empiric treatment.[14, 67]

Input	Name used for code/ equations	Value(s) Random distribution	Source of value(s)	Comments
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 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 since 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 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. (NG109, Smaill 2015)
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 2000 but from Norrby 1990.[58, 71]

Input	Name used for code/ equations	Value(s) Random distribution	Source of value(s)	Comments
Duration side effects from antibiotics		3 days (2-4 days) Normal(mean = 3, SD = 0.5)		Used in Fenwick 2000 but from Carlson 1985.[58, 72]
Overall cost of test	cost_test	[redacted] Flexicult: £48 ID Flexicult: Unavailable use [redacted]	Flexicult: Butler 2018.[8] Lodestar: Manufacturer 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 lab 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 they say nearly 90% (£43.90) are distribution cost. [redacted]. 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.

Input	Name used for code/ equations	Value(s) Random distribution	Source of value(s)	Comments
Cost of follow-up appointment/attention if required to prescribe chosen antibiotic at a later point due to wait length for results	cost_followup_appt	£42	Unit Costs of Health and Social Care 2022 (PSSRU & CHE)	GP appointment cost is £42, including direct care staff costs (nurses)
Overall cost per course of antibiotics	mapped_treatment_costs	Table 20	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_pyelonephritis	£1221.26 (2022 price, inflated from the 2016 price of £986.40)	Sadler 2017.[57]	Sadler 2017 hospitalization cost of pyelonephritis (2016 price): £3992 Sadler 2017 days hospitalization for pyelonephritis: 2 Sadler 2017 outpatient visit cost of pyelonephritis (2016 price): £94 Sadler 2017 Risk of hospitalization if pyelonephritis: 0.20
QALY loss from uncomplicated UTI	qaly_loss_uti		Wang 2021 and Bermingham 2012.[64, 73]	0.68 (0.56-0.72) was QALD for UTI

Input	Name used for code/ equations	Value(s) Random distribution	Source of value(s)	Comments
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
Additional QALY loss from pyelonephritis in the short-term model	qaly_loss_pyelonephritis		Wang 2021 and Bermingham 2012.[64, 73]	0.59 (0.48, 0.64) QALD for pyelonephritis. Duration of treated pyelonephritic attack was 10 days and untreated was 14 days in Whiting 2006 and Barry 1997.[59, 61] 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 2021 and Bermingham 2012. [64, 73]	0.82 (0.58, 0.92) QALD for UTI cured

6.3.3 Health outcomes

In the decision tree, we need to quantify the quality of life with a complicated or uncomplicated UTI, impact on quality of life of testing and of the 3-7 day course of antibiotics, including their adverse events. We did this using utilities, disutilities, and quality adjusted life years (QALYs) over defined time periods. For example, a disutility for antibiotic AE along with a proportion of the cohort expected to suffer these AEs; the QALYs accrued by patients with complicated or uncomplicated UTI over the period of the short-term model. These are summarised in Table 19.

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.

6.3.4 Costs

Costs of testing technologies, staff time to perform the tests, GP appointments, antibiotics courses, managing complicated/uncomplicated UTI, managing each complication, were gathered from evidence sources described in Section 6.2. These were supplemented by routine NHS sources (NHS reference costs, Personal Social Services Research Unit (PSSRU), British National Formulary (BNF)) and discussions with clinical advisors. Costs of antibiotic treatment are summarised in Table 20 while other costs are summarised in Table 19.

Cost of training staff to utilise innovative tests were considered but are a budget impact rather than a cost to include in cost-effectiveness analysis as they relate to cost of setup rather than routine use.

Table 20 Assumptions and sources for costing courses of antibiotic treatment for UTIs

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 and over 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 https://www.nice.org.uk/guidance/ng109/resources/visual-summary-pdf-6544021069	Macrobid 100mg modified-release capsules: £9.50 per 14 capsules	£4.07
Nitrofurantoin	Children aged 3 months and over with a lower UTI (and eGFR ≥45 ml/minute)	Empiric and targeted	3 months to 11 years, 750 micrograms/kg four times a day for 3 days; 12 to 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 https://www.nice.org.uk/guidance/ng109/resources/visual-summary-pdf-6544021069	Macrobid 100mg modified-release capsules: £9.50 per 14 capsules; Nitrofurantoin 50mg tablets: £3.43 per 28 tablets; Nitrofurantoin 50mg capsules: £5.30 per 30 capsules	
Nitrofurantoin	Pregnant women aged 12 years and over 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 https://www.nice.org.uk/guidance/ng109/resources/visual-summary-pdf-6544021069	Macrobid 100mg modified-release capsules: £9.50 per 14 capsules	£9.50
Nitrofurantoin	Men aged 16 years and over 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 https://www.nice.org.uk/guidance/ng109/resources/visual-summary-pdf-6544021069	Macrobid 100mg modified-release capsules: £9.50 per 14 capsules	£9.50

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 and men aged 16 years and over 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 https://www.nice.org.uk/guidance/ng113/resources/visual-summary-pdf-6599495053	Macrobid 100mg modified-release capsules: £9.50 per 14 capsules	£9.50
Cefalexin	Pregnant women aged 12 years and over with a catheter	Empiric	500 mg twice or three times a day for 7 to 10 days	"UTI (catheter): antimicrobial prescribing" NICE guidelines September 2019 https://www.nice.org.uk/guidance/ng113/resources/visual-summary-pdf-6599495053	Cefalexin 500mg tablets: £2.70 per 21 tablets; Cefalexin 500mg capsules: £2.42 per 21 capsules	£1.61 - £3.86
Cefalexin	Non-pregnant women and men aged 16 years and over with acute pyelonephritis	Empiric	500 mg twice or three times a day for 7 to 10 days	"Pyelonephritis (acute): antimicrobial prescribing" NICE guidelines September 2019 https://www.nice.org.uk/guidance/ng111/resources/visual-summary-pdf-6544161037	Cefalexin 500mg tablets: £2.70 per 21 tablets; Cefalexin 500mg capsules: £2.42 per 21 capsules	£1.61 - £3.86
Cefalexin	Pregnant women and men aged 12 years and over with acute pyelonephritis	Empiric	500 mg twice or three times a day for 7 to 10 days	"Pyelonephritis (acute): antimicrobial prescribing" NICE guidelines September 2019 https://www.nice.org.uk/guidance/ng111/resources/visual-summary-pdf-6544161037	Cefalexin 500mg tablets: £2.70 per 21 tablets; Cefalexin 500mg capsules: £2.42 per 21 capsules	£1.61 - £3.86
Fosfomycin	Adults with acute uncomplicated lower UTI	Targeted	3g per 1 dose (granules)	Dosage from BNF	Fosfomycin 3g granules sachets: £4.86 per sachet	

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
Trimethoprim	Women aged 16 years and over with lower UTI	Targeted	200mg twice daily for 3 days	Dosage from BNF	Trimethoprim 200mg tablets: £1.76 per 14 tablets	£0.75
Trimethoprim	Men aged 16 years and over with lower UTI	Targeted	200mg twice daily for 7 days	Dosage from BNF	Trimethoprim 200mg 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 18 years and over with UTI	Targeted	0.5-1g every 6 hours	Dosage from BNF	Ampicillin 500mg capsules: £47.96 per 28 capsules	
Ampicillin	Children with UTI	Targeted	Dosage depends on age	Dosage from BNF		

6.3.5 Analyses

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

6.3.6 Scenario and subgroup analyses

As explained in Section 6.2.3, we only model two populations with three available tests in each:

Mixed population

- Lodestar DX
- Flexicult Human

Women with uncomplicated UTI

- Lodestar DX [redacted]
- 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.

6.4 Summary of evaluation of cost-effectiveness

In Section 6.1 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. Our evaluation of the identified evidence (Section 6.2) and attempt to inform a decision tree implementation of the conceptual model (Section 0) reveal that evidence is too limited for results to be meaningful. This is despite the restriction to a narrow set of tests and subgroups in Section 6.2.3. 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 Sections 4 and 5 was restricted to addressing objectives 1-3 of Section 3, which relate to clinical efficacy, accuracy and technical performance of POCTs. Systematic literature reviews were not conducted on, for example, quality of life in UTI, efficacy of antibiotics for treating UTI, or costs related to complications of UTI. Our pragmatic search of Section 6.2.2 identified 8 previous economic models in UTI but was not a systematic search as no formal PICOS 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 was summarised in Table 18. 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 for the accuracy of detecting specific antibiotic sensitivity.
[redacted]

Table 12 [redacted]

There were more substantial evidence limitations in the other model parameters summarised in Table 19. No evidence on probabilities of sepsis and kidney failure resulting from UTI on targeted antibiotics, empiric antibiotics or no treatment was identified. Probability of pyelonephritis on treatment was identified using NICE guideline NG109 but this did not distinguish between targeted and empiric treatment and related to pregnant women. We would need to assume 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 (Table 20). 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. Total cost per person of the Flexicult test was estimated in Butler 2018, which included administration and interpretation costs, but similar estimates were not available for Lodestar DX or ID Flexicult. [8]
[redacted] 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.

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 the model and use it in a full economic evaluation.

7 DISCUSSION

7.1 Statement of principal findings

There were limited data on the clinical effectiveness of POCT for UTI. The majority of the included studies evaluated culture based tests that take up to 24 hours to give a result: Flexicult Human (4 studies), Uricult trio (3 studies), Dipstreak (2 studies), and ID Flexicult (2 studies). The rapid test Uriscreen was evaluated in 4 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 mins. There were also no data on Chromostreak or Diaslide but these are linked to the Dipstreak test or for Uricult plus which is linked to the Uricult and Uricult trio tests. This limited clinical effectiveness also limited the feasibility of economic evaluation.

Included studies only assessed the following specific populations defined in the scope: women (4 studies, not stratified on age), pregnant women (4 studies), children (4 studies) and those with catheters (1 study). There were no data on any of the other pre-specified populations of interest. This further limited the scope of economic evaluation to these populations. However, those 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 POCT for UTI on clinical outcomes. We identified only two trials, both evaluated Flexicult Human, one compared to standard care and the other to testing with ID Flexicult (which can only tell if UTI is present, it does not give information on antibiotic sensitivity). Both trials were judged at low risk of bias. Neither trial reported a difference in the primary outcome (concordant antibiotic use and appropriate antibiotic prescribing) between intervention groups. Although the study that compared Flexicult Human to 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 different 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 was also limited data on the accuracy of POCT for diagnosing UTI, detecting the pathogenic cause of the infection or for detecting antimicrobial sensitivity. Although 16 studies were included for this objective, individual POCT were each assessed in a maximum of 4 studies. Where there were data from multiple studies for a single test, studies were heterogeneous in terms of setting, population and where the POCT test was performed (near patient or in a laboratory). The limited data suggested that performing the POCT in the laboratory may overestimate accuracy compared to performing the test in a POCT setting, particularly for culture based tests. Using stored urine samples rather than fresh

urine samples was also found to overestimate accuracy in one study that used both types of sample. Some studies were at risk of bias, and so results should be interpreted with caution. Five were judged at high risk of bias due to a large proportion of missing data (3 studies), inclusion of multiple samples from the same patients (1 study) and selected enrolment of patients (1 study). Blinding of the person interpreting the reference standard (usually culture) was often not reported and so 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 summary estimates and what this means for the relative accuracy of the tests.

The Lodestar DX tests showed the greatest potential to have clinical value of the three rapid tests for which data were available.

[redacted] 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 which enables catalase detection followed by hydrogen peroxide to the urine sample, shaking the collection tube and then observing whether a foam ring is 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. Results suggested that both sensitivity and specificity were modest with summary estimates of 74% (95% CI 59, 84) for sensitivity and 70% (95% CI 52, 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 only provides data on whether a UTI is present is less likely to be of value given the poor sensitivity suggested by this study.

There was more data on culture based POCT tests, but these are less likely to be of value in a primary care setting due to the time taken 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 clinicians often start antibiotic treatment while waiting 24 hours for the result (reducing its value in avoiding unnecessary antibiotics). The limited data suggested that Dipstreak (2 studies) and Uricult

(1 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 using samples from secondary care, tested in the laboratory and reported excellent 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, 82) and 70% (52, 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 (4 studies) and ID Flexicult (2 studies) were found to have modest accuracy for the detection of UTI with summary sensitivity 79% (95% CI 72, 85) and 89% (95% CI 84, 93) and summary specificity 67% (95% 30, 90) and 70% (95% CI 52, 84), although these should be interpreted with caution due to 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 87% (95% CI 83, 90), and summary specificity 93% (95% CI 89,95). Two studies of culture based tests provided information on the accuracy of test for 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 reported of Uricult Trio only looked at the detection of the presence of E.Coli infection and reported sensitivity of 60% and specificity of 96%.

There was also very little 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 2 and relates to culture based tests. Three studies evaluated Flexicult Human and two evaluated Uricult Trio. Technical performance data suggested that POCT 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, compared to laboratory tests requiring transportation. The POETIC study included for objective 1, provided additional data on outcomes in the Flexicult Human arm only. This showed the test was quick to perform, 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 sub study 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 whilst awaiting test results, and the expense of maintaining regular stock of tests. Only one study reported data on cost - Flexicult human was reported to cost £48.

[redacted]. There were no other data on costs, and no data on test failure rate or health-related quality of life.

New POCT tests would need to have greater accuracy, be cheaper than standard dipstick tests, or provide additional information to inform treatment than dipstick tests. Although these tests give results within a few minutes, they are only able to suggest whether or not a UTI is present - they do not provide any information on the pathogenic cause or on antimicrobial sensitivity. Six studies provided a direct comparison of POCT tests 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 healthcare costs quality of life impacts arising from UTIs.

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 vs 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 in this narrow comparison, it was decided the results of our economic model would not be meaningful and our findings are limited to the conceptual level.

7.2 Strengths and limitations of the assessment

Our systematic review followed published guidance on the conduct of systematic reviews of diagnostic test accuracy studies[24] and is reported according to PRISMA-2020 guidance[25] and PRISMA-DTA guidance, making our review processes transparent and robust. The protocol was pre-registered on 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 objectives 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 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 publication date; 62 studies were excluded for this reason. All evaluated culture based POCT - the majority of these evaluated Uricult/Uricult trio with a small number evaluating Uriscreen and Diaslide. We conducted a formal assessment of the risk of bias of included studies using the RoB 2 tool for RCTs,[28] the QUADAS-2 tool for diagnostic test accuracy studies,[30] and its extension QUADAS-C[74] 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 details on 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. We calculated summary estimates of sensitivity and specificity across patient subgroups. This assumes that accuracy would not vary by subgroup, however, 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.

We did not include a formal assessment of publication bias due to the small number of included studies, and due to the difficulties in assessing publication bias for diagnostic test accuracy studies where there is no clear threshold for “significance”.[24]

We pre-specified 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.[75, 76] 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 mins, but further data on the accuracy and clinical impact are needed.

[redacted]

A potential limitation of the evidence base is exactly how a UTI should be defined. The gold standard test for UTI is culture, with the concept of significant bacteriuria, usually defined as $>10^5$ CFU/ml, 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.[77] However, there are limitations with 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 result should be interpreted to confirm the presence of absence of UTI,[78, 79] 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 ranging from $\geq 10^3$ CFU to $\geq 10^5$ CFU to define the presence of UTI. In some studies, this was based only on the presence of a single organism, others had different threshold for mixed growth e.g. "Single organism 104 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 – this is likely to have given a more accurate classification of whether or not a UTI was present.

A further problem is potential contamination of urine samples or asymptomatic bacteriuria.[80] Culture will not distinguish between pathogenic and non-pathogenic bacteria, so when bacteria is grown on culture, this 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 reported, most studies included in this review used mid-stream urine samples or urine collection bags in children. Whilst such methods of urine collection do have a greater risk of contamination than other methods such as suprapubic aspiration or catheterisation, this is how urine is likely to be collected in practice and so was appropriate.

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

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 for the accuracy of detecting specific antibiotic sensitivity. Sensitivity and specificity of detecting E.coli estimates for Lodestar DX were identified but not for detecting UTI overall. Evidence on accuracy for Lodestar DX detecting specific pathogens was identified but this was on stored rather than fresh urine, and thus potentially a biased overestimate.

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 8 studies, none of which modelled POCTs and none of which provided all evidence needed to inform our economic evaluation.

We used a broad conceptual model to reflect the influence on the costs, health outcomes and model structures of the choice of populations and subgroups. 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 evidence informing our economic model is too weak for results to be meaningful.

7.3 Uncertainties

Given the limited data available for this appraisal a number of uncertainties remain. These include: accuracy of rapid tests for diagnosing UTI in primary care settings, comparative accuracy of tests, whether accuracy varies according to population, how test interpretation varies between the laboratory and near patient settings, impact of recurrent or chronic UTI on test performance; and economic modelling.

We only identified 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 POCT were designed to be carried out in a near-patient setting, nine studies performed the POCT in a laboratory setting, six of these used samples sent to the laboratory 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 to those done in near patient settings. The

only primary care setting in which studies were conducted were GP practices and antenatal clinics. There were no data in pharmacy settings. Further data is needed on how these tests perform in a near-patient setting.

The limitations of the clinical effectiveness evidence also limited the scope for the economic evaluation. Despite prioritising those tests and subgroups where 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), there was little reliable data identified for the probability of identifying antibiotic susceptibility or pathogenic cause to direct targeted treatment. Sensitivity and specificity of detecting E.coli estimates for Lodestar DX were identified but not for detecting UTI overall.

There were more substantial evidence limitations in the other model parameters summarised in Table 19. No evidence was identified on probabilities of sepsis and kidney failure resulting from UTI on targeted antibiotics, empiric antibiotics or no treatment was identified. Probability of pyelonephritis on treatment was identified using NICE guideline NG109 but this 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 also no evidence on the proportion of patients given antibiotics if their initial test did not detect UTI.

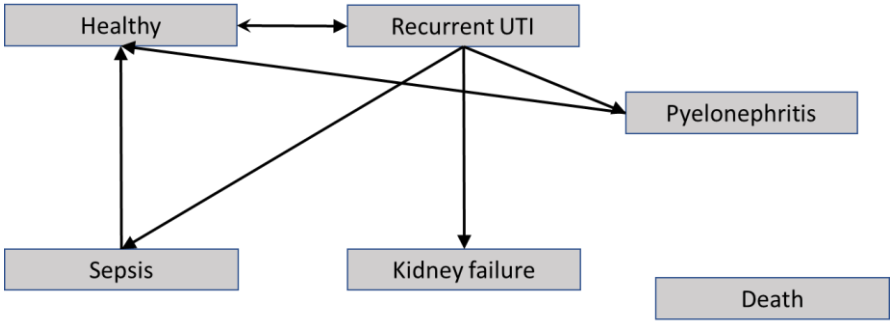
Cost data on POCTs themselves were limited. Total cost per person of the Flexicult test was estimated in Butler 2018, which included administration and interpretation costs, but similar estimates were not available for Lodestar DX or ID Flexicult. [8] The manufacturer of Lodestar DX provided only the price of the test, plus an estimate of 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 17 could be questioned. In particular, the assumption that accuracy does not vary by subgroup could be challenged by Figure 3. For example, pregnant vs catheterised for Uriscreen, specificity of Flexicult Human in mixed vs women, or pregnant vs children for Uricult trio. As further evidence that test accuracy can vary by population, the manufacturer submissions note that Astrego can only be used in women.

Our choice of a decision tree (Figure 6) to represent the conceptual model (Figure 5) 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 UTIs. A Markov model, as illustrated in

Figure 7, could be used to model the long-term consequences of complication branches of our decision tree.

Figure 7 Possible long-term Markov model from decision tree



*Hospitalisation is a factor for each of the complication states. Death is possible from any state.

7.4 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.

7.5 Patient and Public Involvement

We involved two patient representatives with lived experience of UTI in this project. 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 section below about the difference POCT may have for patients with UTI. Involvement of patients had a positive impact on this project, particularly in highlighting the importance of not having to wait for results of tests. Discussions around this topic led to us stratifying our results section into rapid tests and culture-based tests.

7.6 Impact on Patients

The first and most important impact for patients is that a test which can be given, immediately in the doctor's surgery, particularly if it suggests the appropriate antibiotic for

treatment, can relieve symptoms much more quickly and effectively with less impact on AMR.

UTIs can be extremely painful and uncomfortable. They make leaving your house and being away from a toilet very difficult and they therefore impact on the ability of people to manage their everyday lives. For this reason, anything that can make treatment quicker and more effective is immensely valuable to patients. It also means they are less likely to attend Accident and emergency services relieving pressure on those services and reducing the patient's likelihood of coming into contact with other communicable diseases or spending long painful hours waiting for treatment.

The benefit of being able to be diagnosed in your local GP surgery in one visit would have a major impact on people 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 Pandemic more people sought advice and accessed pharmacies and trusted the advice they provided. The fact that pharmacies are in the community and accessible with longer opening hours including weekend opening 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 anti-microbial resistance. This is positive for patients' future treatment of infections. This is also likely to cost less in antibiotic prescribing which would be positive for the NHS.

8 Conclusions

8.1 Implications for practice

There is a clear need for a rapid test that would accurately diagnose a UTI within a short time period in primary care, including GP surgeries of pharmacy settings. Ideally such tests would also provide information on antimicrobial sensitivity, this would allow appropriate targeted antibiotic use, which would mean patients would be treated appropriately more quickly and would limit the total burden of antibiotic prescriptions. 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 the pathogenic cause would also be of value as whilst these would provide direct information on 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, it takes up to 24 hours to produce a result, this is likely to be longer for samples that are taken on Friday as results would then not be available until the following 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 the 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 vs empiric antibiotic efficacy, or costs and quality of life impacts of UTI complications for our model to perform a meaningful comparison. A full evaluation would be needed before any recommendation can be made regarding the cost-effectiveness of POCTs or their ability to impact antibiotic resistance.

Strong evidence that POCT (i) reduce unnecessary antibiotic use; (ii) improve symptoms or (iii) are cost-effective, is needed before such tests are introduced to the NHS.

8.2 Suggested research priorities

Given the paucity of data on POCT test for diagnosing UTI, further studies are needed to determine whether POCT for people with suspected UTI have the potential to be clinically 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) or Lodestar DX(which provides information on pathogenic cause). Studies should also investigate the feasibility of introducing testing within a pharmacy setting, this 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 a primary care (GP surgery and/or Pharmacy) and would randomise GP practices/pharmacies to either “test and treat appropriately” or to standard practice. Outcomes including antibiotic prescribing, symptom duration and costs 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. Studies should either 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 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 vs targeted antibiotic treatment of UTI, efficacy in preventing UTI complications, and both the cost and quality of life impacts of these complications.

9 Acknowledgements

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.

Contributions of authors

Penny Whiting drafted the clinical effectiveness sections of the protocol and led the clinical systematic review. Eve Tomlinson contributed to and checked data extraction and quality assessment for objective 2, extracted data for objective 3, drafted the results sections for objective 2 and 3, and assisted in drafting clinical effectiveness sections of the discussion. Chris Cooper 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 screened and extracted studies and reviewed the company submissions. Hayley Jones provided statistical advice and carried out the meta-analyses of diagnostic accuracy data. Howard Thom drafted the cost-effectiveness section of the report, designed the cost-effectiveness model, and lead the cost-effectiveness assessment. Mary Ward helped design the cost-effectiveness model, gathered input parameters, implemented the model in R, and ran analyses.

Christina Stokes and Samina Begum provided a patient perspective on the project, edited the plain language summary and wrote the section of the report on “Impact on Patients”.

Alastair Hay and Jessica Watson provided clinical advice for the project.

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

9.1 Ethics Statement

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

9.2 Information Governance Statement

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

9.3 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.

10 References

1. Tandogdu Z, Wagenlehner FM. Global epidemiology of urinary tract infections. *Current Opinion in Infectious Diseases* 2016;**29**(1)
2. The 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/).
3. Medina M, Castillo-Pino E. An introduction to the epidemiology and burden of urinary tract infections. *Therapeutic Advances in Urology* 2019;**11**
4. Storme O, Tiran Saucedo J, Garcia-Mora A, Dehesa-Davila M, Naber KG. Risk factors and predisposing conditions for urinary tract infection. *Therapeutic Advances in Urology* 2019;**11**
5. Chronic Urinary Tract Infection Campaign. *What is chronic UTI?* ; 2021. URL: https://cutic.co.uk/what-is-chronic-uti/ (Accessed 29/11/2022).
6. Ades AE, Caldwell DM, Reken S, Welton NJ, Sutton AJ, Dias S. *NICE DSU Technical Support Document 7: Evidence synthesis of treatment efficacy in decision making: a reviewer's checklist*; 2012. (Accessed 22 June 2022) https://www.sheffield.ac.uk/sites/default/files/2022-02/TSD7-reviewer-checklist.final_.08.05.12.pdf:
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-Antimicrobial Resistance* 2020;**2**(2)
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. *British Journal of General Practice* 2018;**68**(669)
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 Family Practice* 2010;**11**
10. The National Institute for Health and Care Excellence (NICE). *Pyelonephritis - acute*. 2021. URL: https://cks.nice.org.uk/topics/pyelonephritis-acute/ (Accessed October 2022).
11. Public Health England. *Diagnosis of urinary tract infections. Quick reference tool for primary care for consultation and local adaptation*. London; 2020. (Accessed October 2022) https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/927195/UTI_diagnostic_flowchart_NICE-October_2020-FINAL.pdf:
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. *Clinical Microbiology and Infection* 2006;**12**(6)
13. Hoffmann T, Peiris R, Mar CD, Cleo G, Glasziou P. Natural history of uncomplicated urinary tract infection without antibiotics: a systematic review. *British Journal of General Practice* 2020;**70**(699)
14. The National Institute for Health and Care Excellence (NICE). *Urinary tract infection (lower): antimicrobial prescribing*; 2018.) https://www.nice.org.uk/guidance/ng109:

15. 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**
16. Chronic UTI Info. *Antibiotics*. 2022. URL: https://www.chronicutiinfo.com/treatment/conventional-medicine/antibiotics/ (Accessed November 2022).
17. Faria R, Hernandez Alava M, Manca A, Wailoo AJ. *NICE DSU Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness for Technology Appraisal: Methods for comparative individual patient data*; 2015. (Accessed June 2022) https://www.sheffield.ac.uk/sites/default/files/2022-02/TSD17-DSU-Observational-data-FINAL.pdf:
18. 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. *European Journal of Clinical Microbiology & Infectious Diseases* 2015;**34**(10)
19. World Health Organization. *Antibiotic resistance*. 2020. URL: https://www.who.int/news-room/fact-sheets/detail/antibiotic-resistance (Accessed November 2022).
20. U. K. Health Security Agency. *English surveillance programme for antimicrobial utilisation and resistance (ESPAUR). Report 2021 to 2022*. London; 2022. (Accessed November 2022) https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1118310/ESPAUR-report-2021-to-2022.pdf:
21. Medicines & Healthcare products Regulatory Agency. *Management and use of IVD point of care test devices*. 2021. URL: https://www.gov.uk/government/publications/in-vitro-diagnostic-point-of-care-test-devices/management-and-use-of-ivd-point-of-care-test-devices (Accessed Novemeber 2022).
22. Centre for Reviews and Dissemination. *CRD's guidance for undertaking reviews in health care. Centre for Reviews and Dissemination*. York: University of York; 2009. (Accessed November 2022) https://wwwyorkacuk/media/crd/Systematic_Reviewspdf.:
23. The National Institute for Health and Care Excellence (NICE). *NICE health technology evaluations: the manual*. 2022. URL: https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation (Accessed November 2022).
24. 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 November 2022).
25. 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. *Journal of Clinical Epidemiology* 2021;**134**
26. 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. *Systematic Reviews* 2017;**6**(1)

27. The National Institute for Health and Care Excellence (NICE). *Diagnostic Assessment Programme manual*. 2011. URL: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-diagnostics-guidance/Diagnostics-assessment-programme-manual.pdf (Accessed November 2022). http://dx.doi.org/https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-diagnostics-guidance/Diagnostics-assessment-programme-manual.pdf
28. 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. *British Medical Journal* 2019;**366**
29. 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. *British Medical Journal* 2016;**355**
30. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB*, et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8)
31. 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**(10)
32. 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**(12)
33. Holm A, Cordoba G, Moller Sorensen T, Rem Jessen L, Frimodt-Moller N, Siersma V*, et al.* 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**(10)
34. 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. *Family Practice* 2017;**34**(4)
35. Holm A, Cordoba G, Sorensen TM, Jessen LR, Frimodt-Moller N, Siersma V*, et al.* Clinical accuracy of point-of-care urine culture in general practice. *Scandinavian Journal of Primary Health Care* 2017;**35**(2)
36. 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**(2)
37. Blom M, Sorensen TL, Espersen F, Frimodt-Moller N. Validation of FLEXICULT SSI-Urinary Kit for use in the primary health care setting. *Scandinavian Journal of Infectious Diseases* 2002;**34**(6)
38. 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? *Scandinavian Journal of Primary Health Care* 2019;**37**(1)
39. Millar L, DeBuque L, Leialoha C, Grandinetti A, Killeen J. Rapid enzymatic urine screening test to detect bacteriuria in pregnancy. *Obstetrics & Gynecology* 2000;**95**(4)
40. Teppa RJ, Roberts JM. The uriscreen test to detect significant asymptomatic bacteriuria during pregnancy. *Journal of the Society for Gynecologic Investigation* 2005;**12**(1)

41. 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. *Family Practice* 2022;**39(4)**
42. 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]. *Revista de Investigacion Clinica* 2002;**54**(6)
43. Anacleto FE, Resontoc LP, Padilla GH. Bedside diagnosis of outpatient childhood urinary tract infection using three-media dipslide culture test. *Pediatric Nephrology* 2009;**24(8)**
44. Greeff A, Jeffery B, Pattinson RC. Uricult trio as a screening test for bacteriuria in pregnancy. *South African Journal of Obstetrics and Gynaecology* 2002;**8(3)**
45. 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 Journal of Pediatric Infectious Diseases* 2010;**17(2)**
46. Mignini L, Carroli G, Abalos E, Widmer M, Amigot S, Nardin JM*, et al.* Accuracy of diagnostic tests to detect asymptomatic bacteriuria during pregnancy. *Obstetrics & Gynecology* 2009;**113**(2 Pt 1)
47. Colodner R, Keness Y. Evaluation of DipStreak containing CNA-MacConkey agar: a new bedside urine culture device. *Israel Medical Association Journal* 2000;**2**(7)
48. Yagupsky P, Rider M, Peled N. Clinical evaluation of a novel chromogenic agar dipslide for diagnosis of urinary tract infections. *European Journal of Clinical Microbiology & Infectious Diseases* 2000;**19**(9)
49. Ayako van der Goes van Naters, Joanna Diggle, Mandy Wootton, Ali Roula, Emma Hayhurst, Jeroen Nieuwland. AIC Detection of bacterial uropathogens in humans using a novel Loop-mediated Isothermal assay and reader device AIC. *Unpublished AIC*
50. Hayward G, Butler C. *TOUCAN*. 2023. URL: https://www.spcr.nihr.ac.uk/research/projects/toucan-platform-for-uti-diagnostic-evaluation (Accessed January 2023).
51. KU Leuven. 2021. ERNIE4: urine and CRP Point-of-care Test in Acutely Ill Children. NCT03835104; URL: https://www.cochranelibrary.com/central/doi/10.1002/central/CN-01965319/full (Accessed November 2022).
52. University of Copenhagen Y. 2015. Diagnostic Accuracy of Point of Care Test of First Voided Urine Compared to Midstream Voided Urine in Primary Care. NCT02585115; URL: https://www.clinicaltrials.gov/ct2/show/NCT02585115?term=Flexicult&draw=2&rank=2 (Accessed November 2022).
53. 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**(2)
54. DRKS00017273. 2019. Management of UTI in German primary care: Feasibility of FLEXICULT™ (MAFL). URL: http://www.drks.de/DRKS00017273).
55. R Core Team. *R: A Language and Environment for Statistical Computing.* . Vienna, Austria: R Foundation; 2019. URL: https://www.R-project.org/ (Accessed December 2022).
56. Carlson KJ. Cost-Effectiveness Analysis of Single-Dose Therapy of Urinary Tract Infection Compared to Conventional Treatment. *European Urology* 1987;**13(suppl 1)**

57. 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**(3)
58. Fenwick EA, Briggs AH, Hawke CI. Management of urinary tract infection in general practice: a cost-effectiveness analysis. *Br J Gen Pract* 2000;**50**(457)
59. 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)
60. 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**(4)
61. 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**(1)
62. 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**(1)
63. 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**
64. Wang R, LaSala C. Role of antibiotic resistance in urinary tract infection management: a cost-effectiveness analysis. *Am J Obstet Gynecol* 2021;**225**(5)
65. Public Health England. *Diagnosis of urinary tract infections: Quick reference tool for primary care for consultation and local adaptation*; 2020. (Accessed Decemeber 2022) https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/927195/UTI_diagnostic_flowchart_NICE-October_2020-FINAL.pdf:
66. 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**(21)
67. Smaill FM, Vazquez JC. Antibiotics for asymptomatic bacteriuria in pregnancy. *Cochrane Database Syst Rev* 2015; 10.1002/14651858.CD000490.pub3(8)
68. 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**(1)
69. 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**(482)
70. 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)
71. Norrby SR. Short-term treatment of uncomplicated lower urinary tract infections in women. *Rev Infect Dis* 1990;**12**(3)
72. Carlson KJ, Mulley AG. Management of acute dysuria. A decision-analysis model of alternative strategies. *Annals of internal medicine* 1985;**102 2**

73. Bermingham SL, Ashe JF. Systematic review of the impact of urinary tract infections on health-related quality of life. *BJU Int* 2012;**110**(11 Pt C)
74. Yang B, Mallett S, Takwoingi Y, Davenport CF, Hyde CJ, Whiting PF*, et al.* QUADAS-C: A Tool for Assessing Risk of Bias in Comparative Diagnostic Accuracy Studies. *Ann Intern Med* 2021;**174**(11)
75. Baltekin Ö, Hammar P, Kovachev P, Myzithra M, Wistrand-Yuen E. Reproducibility of Fully Automated AST for Direct Near Patient Testing. European Congress of Clinical Microbiology & Infectious Diseases,; Lisbon, abstract no. 1059.
76. Baltekin Ö, Boucharin A, Tano E, Andersson DI, Elf J. Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging. *Proceedings of the National Academy of Sciences* 2017;**114**(34)
77. Cohen SN, Kass EH. A Simple Method for Quantitative Urine Culture. *New England Journal of Medicine* 1967;**277**(4)
78. Christopher Cooper, Eve Tomlinson, Catalina Lopez-Manzano, Rachel James, Mary Ward, Jessica Watson*, et al. Point of care tests for urinary tract infections (UTI) to reduce antimicrobial resistance: a systematic review and conceptual economic model to inform Early Value Assessment (EVA) (DAP 69).* . 2022. URL: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022383889).
79. Grabe M, Bishop M, Bjerklund-Johansen T, Botto H, Çek M, Lobel B*, et al.* Guidelines on Urological Infections. 2010
80. Cormican M, Murphy AW, Vellinga A. Interpreting asymptomatic bacteriuria. *BMJ* 2011;**343**
81. Berbik I, Lampe L, Orosz TM. [The uricult test in the diagnosis of urinary tract infections in pregnancy]. *Orvosi Hetilap* 1975;**116**(24)
82. Mayhew MB, Tran E, Choi K, Midic U, Luethy R, Damaraju N*, et al.* Optimization of Genomic Classifiers for Clinical Deployment: Evaluation of Bayesian Optimization to Select Predictive Models of Acute Infection and In-Hospital Mortality. *Pacific Symposium on Biocomputing* 2021;**26**
83. POETIC. 2013. POETIC trial regstration: Point of care testing for urinary tract infection in primary care. ISRCTN65200697; URL: https://www.isrctn.com/ISRCTN65200697 (Accessed November 2022).
84. 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 Family Practice* 2014;**15**
85. 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 Family Practice* 2015;**16**
86. Holm A. 2014. Point of Care Susceptibility Testing in Primary Care. NCT02323087; URL: https://clinicaltrials.gov/ct2/show/NCT02323087).
87. Wikler Matthew A. Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically : Approved standard. abstract no. 1069.
88. Hayhurst E, Nieuwland J, Roula A, van der Goes van Naters A. Tackling the neglected disease in our midst: towards better diagnostics for Urinary Tract Infections at point-of-care. Uppsala antibiotic days conference; 2022; Uppsala, Sweden, abstract no. 1063.

APPENDICES

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 by language.

Resource	Hits
MEDLINE (MEDALL)	526
Embase	416
Cochrane	33
CINHAL	12
Clinical Trials.gov	29
ICTRP	17
Total (prior to deduplication)	1035
- duplicates	-304
N to screen	731

Database: MEDLINE (MEDALL)
Host: Ovid
Data parameters: 1946 to December 02, 2022
Date of search: 5 Dec 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
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.	5070893
18	16 not 17	526

Database: Embase

Host: Ovid

Data parameters: 1974 to 2022 December 02

Date of search: 5 Dec 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/	6258009
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 Dec 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 & allied health (CINAHL)

Host: EBSCOhost

Data parameters: 1981-current

Date of search: 5 Dec 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

Notes: a server-side de-duplication was undertaken at S2 to remove studies included in the MEDLINE database.

Trials registry resources

Clinical Trials.gov
https://www.clinicaltrials.gov/ct2/results/refine?show_xprt=Y
5 Dec 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 Dec 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 Dec 2022

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

Database: Embase

Host: Ovid
Data parameters: 1974 to 2022 December 09
Date of search: 12 Dec 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 Dec 2022

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

Database: Cumulative index to nursing & allied health (CINAHL)

Host: EBSCOhost

Data parameters: 1981-current

Date of search: 12 Dec 2022

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

Trials registry resources

Clinical Trials.gov

12 Dec 2022

(UTRiPLEX* or "Global Access Diagnostics")

ICTRP

12 Dec 2022

(UTRiPLEX* or "Global Access Diagnostics")

Web searching

Searcher: Christopher Cooper

Searcher location: London, UK.

Date of search: 6 Dec 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/	Handsearch 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/	Handsearch 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/products/urinary-tract-infection-testing/	Handsearch 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	Handsearch of the website. Followed by manual review of the TriVerity publications tab.	37/0
Uriscreen	Savyon Diagnostics Ltd	https://www.savyondiagnostics.com/product/uriscreen/	Handsearch of the website	2/0
Diaslide, Dipstreak,	Novamed	https://www.novamed.co.il/culture-device	Handsearch of the website	0/0

Test Name	Manufacturer	Website URL	Search Approach	Results (checked/included)
Chromostreak				
Uricult, Uricult trio and Uricult plus	Aidian; formerly Orion Diagnostica	https://www.aidian.eu/microbiology/uricult/uricult-tests#generally	Handsearch of the website	0/0
UTRiPLEX	Global Access Diagnostics	https://www.globalaccessdx.com/	Handsearch of the website followed by Google overlay search: Flexicult Human site:https://ssidiagnostica.com/	0/0

Appendix 2: List of excluded studies with rationale

Appendix 2.1 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. Journal of Clinical Microbiology. 1992;30(10):2686-91.	Diaslide	Unclear
Edwards B, White RH, Maxted H, Deverill I, White PA. Screening methods for covert bacteriuria in schoolgirls. British Medical Journal. 1975;2(5969):463-7.	Unclear	Unclear
Van Dorsten JP, Bannister ER. Office diagnosis of asymptomatic bacteriuria in pregnant women. American Journal of Obstetrics & Gynecology. 1986;155(4):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. American Journal of Clinical Pathology. 1994;101(1):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]. Japanese Journal of Antibiotics. 1995;48(1):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]. Pediatria Catalana. 1998;58(3):147-50.	Unclear	Unclear
Zoller L, Tobler L. [Comparison of culture count determination with the uricult pour-plate]. Medizinische Laboratorium. 1969;22(9):214-7.	Uricult	Unclear
Breitfellner G. [Experiences with uricult, a new method for the quantitative determination of bacteria in urine]. Wiener Medizinische Wochenschrift. 1970;120(14):235-43.	Uricult	Unclear
Haahr J, Bohn L. [Uricult. A simple method of semiquantitative urine culture]. Ugeskrift for Laeger. 1970;132(29):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 for Laeger. 1970;132(42):1966-70.	Uricult	Unclear
Schmid I, Pletscher E. [Uricult, a simple procedure for the determination of bacterial count in urine]. Medizinische Laboratorium. 1970;23(11):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(1):14-7	Uricult	Unclear
Haahr J, Bohn J. Uricult. A simple method of semi-quantitative culture from urine. Acta Paediatrica Scandinavica. 1971;60(2):245-6.	Uricult	Unclear
Bailey MJ, Neary JT, Notelovitz M. The Uricult dip-slide in significant bacteriuria. South African Medical Journal Suid-Afrikaanse Tydskrif Vir Geneeskunde. 1972;46(37):1323-6.	Uricult	Unclear
Buchanan N. Uricult dip-slide in significant bacteriuria. South African Medical Journal Suid-Afrikaanse Tydskrif Vir Geneeskunde. 1972;46(44):1654.	Uricult	Unclear

Study Details	Test Evaluated	Objective Assessed
Dayer JM, Humair L. [Bacteriuria: importance and value of the semi-quantitative method of Uricult. Comparative study]. Schweizerische Rundschau fur Medizin Praxis. 1972;61(12):384-8.	Uricult	Unclear
Hellwig I. [Demonstrations of urinary tract infections using uricult]. Deutsche Medizinische Wochenschrift. 1972;97(44):1687-9.	Uricult	Unclear
Mongeau JG, Robillard JE, Brousseau Y. Screening for bacteriuria in children: comparison of two dip-tests. Canadian Medical Association Journal. 1972;107(3):227-9.	Uricult	Unclear
Maugeri TL, Cefali M, Galletti G. [Determination of bacteriuria using uricult, a new formula]. Quaderni Sclavo di Diagnostica Clinica e di Laboratorio. 1973;9(4):950-63.	Uricult	Unclear
Bailey MJ, Notelovitz M. Appraisal of the Uricult dip-slide method in the diagnosis of urinary infections. South African Medical Journal Suid-Afrikaanse Tydskrif Vir Geneeskunde. 1973;47(26):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. South African Medical Journal Suid-Afrikaanse Tydskrif Vir Geneeskunde. 1973;47(17):725-7.	Uricult	Unclear
Jackaman FR, Darrell JH, Shackman R. The dip-slide in urology. British Medical Journal. 1973;1(5847):207-8.	Uricult	2
Simplaceanu L, Mosora N, Munteanu E. The Uricult test compared with quantitative bacteriuria in diabetics (Rumanian). [Romanian]. Bacteriologia Virusologia Parazitologia Epidemiologia. 1974;19(5):405-10.	Uricult	2
Steiner PO, Gerber A, Sigrist W. Independent bacteriologic urine examination with the new Enterotube in a regional hospital. [German]. Schweizerische Medizinische Wochenschrift. 1974;104(31):1091-3.	Uricult	2
Narbutowicz B, Kostrzewska K, Krawczynski J. [Detection of bacteriuria by means of the Uricult test]. Pediatria Polska. 1974;49(11):1387-91.	Uricult	Unclear
Mackinnon AE, Strachan CJL, Sleigh JD, Burns MM. Screening for bacteriuria with a dip stick test for urinary glucose. British Journal of Urology. 1974;46(1):101-5.	Uricult	2
Joffe BI, Seftel HC, Distiller LA. Asymptomatic bacteriuria in diabetes mellitus. South African Medical Journal. 1974;48(30):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]. Schweizerische Medizinische Wochenschrift. 1974;104(12):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(1):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(24):1403-6.	Uricult	Unclear
Havlik I. [Screening of asymptomatic bacteriuria in pregnant women by means of Uricult (author's transl)]. Ceskoslovenska Gynekologie. 1975;40(8):581-3.	Uricult	Unclear
Ellner PD, Papachristos T. Detection of bacteriuria by dip-slide. Routine use in a large general hospital. American Journal of Clinical Pathology. 1975;63(4):516-21.	Uricult	2

Study Details	Test Evaluated	Objective Assessed
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]. Polski Tygodnik Lekarski. 1975;30(3):107-8.	Uricult	2
Novakova M, Petracek E. [Personal experience with Uricult]. Zdravotnicka Pracovnice. 1975;25(11):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(24).[81]	Uricult	Unclear
Cvoric A, Zecevic B, Nikolic V, Markovic M. [Determination of bacteriuria by means of Uricult method]. Srpski Arhiv Za Celokupno Lekarstvo. 1976;104(2):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]. Medicinski Pregled. 1976;29(11-12):513-7.	Uricult	Unclear
Duerden BI, Moyes A. Comparison of laboratory methods in the diagnosis of urinary tract infection. Journal of Clinical Pathology. 1976;29(4):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]. Ginekologia Polska. 1977;48(11):961-6	Uricult	Unclear
Golebiowska M, Chlebna-Sokol D, Kostenko D. Uricult test in urinary tract screening of children aged 6 to 36 months. [Polish]. Pediatria Polska. 1977;52(11):1219-22.	Uricult	Unclear
Jojart G, Eder I. [Comparative study of urinary nitrite content and Uricult reactions]. Orvosi Hetilap. 1977;118(33):1975-8.	Uricult	Unclear
Bordt J, Beller FK. Is examination of urinary sediment in prenatal check-up still up-to-date?. [German]. Diagnostik. 1979;12(8):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(4):227-32.	Uricult	2
Emans SJ, Grace E, Masland Jr RP. Asymptomatic bacteriuria in adolescent girls: II. Screening methods. Pediatrics. 1979;64(4):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 for Laeger. 1979;141(22):1477-80.	Uricult	Unclear
Sebbesen O, Nielsen E. Demonstration of bacteriuria with transport agar. Comparison between Uricult and Urotube. [Danish]. Ugeskrift for Laeger. 1979;141(6):375-6.	Uricult	Unclear
Winn WC, Jr., Gillenwater JY. Evaluation of Uricult dip slide in two hospital populations. Urology. 1980;15(1):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. Canadian Medical Association Journal. 1981;124(1):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. Scandinavian Journal of Primary Health Care. 1989;7(2):123-8.	Uricult	Unclear
Lorentzon S, Hovelius B, Miorner H, Tendler M, Aberg A. The diagnosis of bacteriuria during pregnancy. Scandinavian Journal of Primary Health Care. 1990;8(2):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	Uricult	2

Study Details	Test Evaluated	Objective Assessed
urinary infections in infants. [Spanish]. Revista Espanola de Pediatria. 1992;48(283):23-5.		
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]. Anales Espanoles de Pediatria. 1994;41(5):325-8.	Uricult	Unclear
Dalet F, Segovia T. Evaluation of a new agar in Uricult-Trio for rapid detection of Escherichia coli in urine. Journal of Clinical Microbiology. 1995;33(5):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. European Journal of Clinical Microbiology and Infectious Diseases. 1995;14(7):606-9.	Uricult trio	Unclear
Andreu A, Xairo D. [Evaluation of a new method for urine screening based on the study of catalase]. Enfermedades Infecciosas y Microbiologia Clinica. 1991;9(3):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. Journal of Clinical Microbiology. 1992;30(3):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. Diagnostic Microbiology & Infectious Disease. 1993;16(2):93-7.	Uriscreen	Unclear
Nauschuetz WF, Harrison LS, Trevino SB, Becker GR, Benton J. Two rapid urine screens for detection of bacteriuria: an evaluation. Current Microbiology. 1993;26(1):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. Obstetrics & Gynecology. 1996;87(3):410-3.	Uriscreen	Unclear
Palmer LS, Richards I, Kaplan WE. Clinical evaluation of a rapid diagnostic screen (URISCREEN) for bacteriuria in children. Journal of Urology. 1997;157(2):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(4):e41.	Uriscreen	2

Appendix 2.2 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. *Scandinavian Journal of Clinical and Laboratory Investigation* 2000;60(5). Aspevall O, Forsum U, Kjerstadius T, Hallander H. Evaluation of two methods for improving quality of diagnosis of bacteriuria by culture in primary healthcare. *Scandinavian Journal of Clinical & Laboratory Investigation* 2000;60(5).	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 Infectious Diseases. 2017;17(1).	NA	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(3).	NA	Did not evaluate POCT of interest
DRKS00017273. 2019. Management of UTI in German primary care: Feasibility of FLEXICULT™ (MAFL). URL: http://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(1).	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. European Geriatric Medicine 2019;10(5).	NA	Not a primary study (mapping review). References were checked to identify[46]
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. *Scandinavian Journal of Primary Health Care* 2022;40(1).	Flexicult SSI	Analytical validity
KU Leuven. 2015. Urinary Tract Infections in Older Persons Admitted to a Psychogeriatric Ward. NCT02368847; 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 Obstetricia et Gynecologica Scandinavica* 2000;**79**(9).	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. Journal of Clinical Microbiology 2002;40(6) Schaeffer AJ. Evaluation of the DipStreak, a new device with an original streaking mechanism for detection, counting, and	DipStreak	No reference standard for evaluation of accuracy

Study details	Test	Reason for exclusion
presumptive identification of urinary tract pathogens. Journal of Urology 2003;169(4)		
Wigton RS. The Uriscreen test was not better than standard urinalysis and dipstick tests for detecting urinary tract infection in children. *Evidence-Based Medicine* 2000;5(4).	Uriscreen	Not a primary study – secondary report of existing study that was excluded due to publication date of 1999

Appendix 2.3 Studies included in manufacturers 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 less than 30 min using direct single-cell imaging. Proceedings of the National Academy of Sciences. 2017;114(34).	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. (2022). Reproducibility of Fully Automated AST for Direct Near Patient Testing. [Poster presentation] ECCMID 2022, 23-26 April 2022, Lisbon.	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. https://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, Koutelidakis IM, Coyle SM, Cheng HK, Liesenfeld O, Sweeney TE, Giamarellos-Bourboulis EJ. A 29-mRNA host response test from blood accurately distinguishes bacterial and viral infections among emergency department patients. Intensive Care Medicine Experimental. 2021 Jun 18;9(1).[82]	Journal Article	Triverity	Inflammatix Classifier (InSep)	Population - Not UTI
Bauer W, Kappert K, Galtung N, Lehmann D, Wacker J, Cheng HK, Liesenfeld O, Buturovic L, Luethy R, Sweeney TE, Tauber R, Somasundaram R. 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. Critical Care Medicine . 2021 Oct 1;49(10).	Journal Article	Triverity	Inflammatix Classifier (InSep)	Population - Not UTI

Study details	Document type	Manufacturer	Test evaluated	Reason for exclusion
Galtung N, Diehl-Wiesenecker E, Lehmann D, Markmann N, Bergström WH, Wacker J, Liesenfeld O, Mayhew M, Buturovic L, Luethy R, Sweeney TE, Tauber R, Kappert K, Somasundaram R, Bauer W. Prospective validation of a transcriptomic severity classifier among patients with suspected acute infection and sepsis in the emergency department. European Journal of Emergency Medicine. 2022 Oct 1;29(5).	Journal Article	Triverity	Inflammatix Classifier (InSep)	Population - Not UTI
Kostaki A, Wacker JW, Safarika A, Solomonidi N, Katsaros K, Giannikopoulos G, Koutelidakis IM, Hogan CA, Uhle F, Liesenfeld O, Sweeney TE, Giamarellos-Bourboulis EJ. 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 Sep 1;58(3).	Journal Article	Triverity	Inflammatix Classifier (InSep)	Population - Not UTI
Brakenridge SC, Starostik P, Ghita G, Midic U, Darden D, Fenner B, Wacker J, Efron PA, Liesenfeld O, Sweeney TE, Moldawer LL. A Transcriptomic Severity Metric That Predicts Clinical Outcomes in Critically Ill Surgical Sepsis Patients. Critical Care Explorations . 2021 Oct 14;3(10).	Journal Article	Triverity	Inflammatix Classifier (InSep)	Population - Not UTI
Brakenridge SC, Chen U, Loftus T, et al. Evaluation of a Multivalent Transcriptomic Metric for Diagnosing Surgical Sepsis and Estimating Mortality Among Critically Ill Patients. JAMA Network Open Open. 2022;5(7).	Journal Article	Triverity	Inflammatix Classifier (InSep)	Population - Not UTI
Moore AR, Roque J, Shaller BT, Asuni T, Remmel M, Rawling D, Liesenfeld O, Khatri P, Wilson JG, Levitt JE, Sweeney TE, Rogers AJ. Prospective validation of an 11-gene mRNA host response score for mortality risk stratification in the intensive care unit. Scientific Reports. 2021 Jun 22;11(1).	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	Journal Article	Triverity	Inflammatix Classifier (InSep)	General discussion paper on optimization

Study details	Document type	Manufacturer	Test evaluated	Reason for exclusion
the Diagnosis of Acute Infections and Sepsis. Journal of Personalized Medicine 2021;11(8).				
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. Journal of Health Economics & Outcomes Research 2020;7(1).	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. 2019. Test package information. (Accessed January 2023).	Test package information	Uricult	Uricult	Package insert for the test
Uricult. 2019. Test package information (Accessed January 2023).	Test package information	Uricult	Uricult Plus	Package insert for the test
Uricult. 2022. Test package information. (Accessed January 2023).	Test package information	Uricult	Uricult Trio	Package insert for the test
UTRiPLEX. 2022. Rapid Urine Test for Urinary Tract Infection. Instructions for use. Sept 2023. (Accessed January 2023).	Test package information	Utriplex	UTRiPLEX test assay	Package insert for the test

Appendix 3: Data extraction tables

Appendix 3.1: Objective 1

Baseline Details

Study Details	Participants	POCT Test Details	Group 1	Control
Author (Year) Butler (2018)[8, 83, 84] **Study Name** POETIC trial **Country** England, Netherlands, Spain & Wales **Study Design** RCT (individual randomised)«StudyDesign» **Recruitment:** July 2013 to 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.«Exclusion» **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 3 participating regions. **Sample collection:** Urine samples collected using Peezy midstream urine collection kit. «DataExBaselineComments» 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)[33, 85, 86] **Study Name** NA **Country** Denmark **Study Design** RCT (individual randomised)«StudyDesign» **Recruitment:** March 2015 to 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 ≤7days for which the GP suspected uncomplicated UTI, including elderly patients above 65, 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 (since POC culture is read the following day).«Exclusion» **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, 83, 84]	Concordant antibiotic use	Consumption of antibiotic on day 3 (or days 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, 1.20)
	Antibiotic prescribing at initial consultation		267	82.4	282	88.4	OR = 0.56 (0.35, 0.88)
	Antibiotics prescribed to guidelines at initial consultation		156	58.9	166	59.5	OR = 0.99 (0.67, 1.45)
	Patient enablement	Measured using Patient Enablement Instrument at day 14 and 3 months[36]	171	70.1	177	69.7	OR = 0.99 (0.66, 1.48)
	Antibiotic consumed day 3	NR	217	79.2	200	76.6	OR = 1.24 (0.81, 1.89)
	Antibiotic consumed (during 2 weeks)	NR	234	85.1	217	81.6	OR = 1.38 (0.87, 2.19)
	New antibiotic prescription (within 2 weeks)	NR	33	10.3	30	9.7	OR = 1.11 (0.65, 1.89)
	Re-consultation (within 2 weeks)	NR	41	12.9	41	13.2	OR = 0.99 (0.62, 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, 1.81)
	Recurrence of UTI within 3 month period	NR	54	17	69	22.3	OR = 0.72 (0.48, 1.07)
	Duration of symptoms	NR	NA	NA	NA	NA	HR = 1.02 (0.83, 1.25)
	Duration of moderately bad symptoms	NR	NA	NA	NA	NA	HR = 0.98 (0.82, 1.17)
	Overall urinary symptom burden	NR	NA	NA	NA	NA	MD = 0.99 (0.84, 1.19)
	Management changed as result of flexicult	NR	190	63.1	NA	NA	NA
	Change of management	Did not start antibiotic	14	7.4	NA	NA	NA
		Stopped taking antibiotic	10	5.3	NA	NA	NA
		Started taking antibiotic	29	15.3	NA	NA	NA
		Continued with antibiotic	63	33.2	NA	NA	NA
		New antibiotic prescribed	74	38.9	NA	NA	NA
	Time to perform test	Prepare test	NA	NA	NA	NA	9 mins
		Obtain and record result	NA	NA	NA	NA	6 mins
		Discuss result with patient	NA	NA	NA	NA	7 mins
	Cost	Cost per person, including POCT cost in UK	NA	NA	NA	NA	£48

Study	Outcome	Definition	Group 1		Group 2		Effect measure – estimate (95% CI)
			n	%	n	%	
Holm (2017)[33, 85, 86]	Appropriate prescribing	(1) if the patient had UTI in the reference: to prescribe a first-line antibiotic to which the infecting pathogen was susceptible; (2) if the patient had UTI but was allergic to the antibiotic or the pathogen was resistant to all first-line antibiotics: to prescribe a second-line antibiotic or (3) if the patient did not have UTI in the reference: not to prescribe an antibiotic	120	67	121	75	OR = 1.44 (1.03,1.99)
	Symptom free on day 5	NR	NR	NR	NR	NR	OR = 0.91 (0.56, 1.49)
	No significant bacteriuria on day 14	NR	NR	NR	NR	NR	OR = 1.15 (0.62, 2.13)

Risk of Bias

Identify the trial you are examining:	POETIC Butler (2018)[8, 83, 84]

Domain	**Concerns**	**Rationale**
Risk of bias arising from the randomization 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 PP and ITT analysis 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)[33, 85, 86]

Domain	**Concerns**	**Rationale**
Risk of bias arising from the randomization 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.
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 PP and ITT 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 (2), did not fulfil inclusion criteria (7), other (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**

Appendix 3.2: Objective 2

Baseline Details

Study Details	Participants	POCT Test Details	Reference standard
Anacleto(2009)[43] **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 & Population** Secondary care; Uncomplicated UTI Age <16 years **Inclusion criteria** Infants & children age 0 to 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 7 urinary tract; midline defects; failure to thrive; concomitant infections; recurrent UTI; asymptomatic bacteriuria; other co-morbid 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 catheterization was performed. **Target condition** Presence of UTI **Location of test performance** Outpatient department **POCT Test** Uricult trio - Dipslide unscrewed from the tube without being allowed to touch the agar surfaces. Holding the 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 h **Threshold** $\geq 10^4$ CFU	**Reference standard** Culture - standard laboratory culture **Threshold** $\geq 10^4$ CFU

Study Details	Participants	POCT Test Details	Reference standard
Blom (2002)[37] **Country** Denmark **Language** English **Funding** Not reported	**Setting & 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	**Urine sampling method** Not reported **Target condition** Presence of UTI Antimicrobial resistance **Location of test performance** GP surgery – *field trial* **POCT Test** Flexicult™ SSI urinary kit - suspensions of bacteria diluted in 50ml sterile urine to various concentrations. Each suspension was poured into a Flexicult SSI Urinary kit for 1-2s. 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[87] **Threshold** $>10^5$ for UTI diagnosis MIC concentration

Study Details	Participants	POCT Test Details	Reference standard
Bongard(2015)[18] **Country** Wales **Language** English **Funding** Medical Research Council, Cardiff University, European Community's Seventh Framework Programme, R-GNOSIS consortium	**Setting & Population** Laboratory based; Mixed **Inclusion criteria** Fresh urine samples (within ~9 hours) submitted from primary & 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. **Number included (number analysed)** 211(200) **Age** Age <18 years to >65 years – no further details **% Female:** 70	**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 Test** Flexicult™ SSI urinary kit - urine poured to cover all compartments. After ~5s, 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. 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.	**Reference standard** Culture & Microscopy & 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), ≥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.

Study Details	Participants	POCT Test Details	Reference standard
Boon(2022)[41, 51] **Country** Belgium; ERNIE4 study. **Language** English **Funding** Research Foundation Flanders and by a KU Leuven starting grant	**Setting & Population** Primary care; Uncomplicated UTI Age <18 years **Inclusion criteria** Age 3 months-18 years; acute illness of max 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** 5.80-18 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 Test** 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 min after addition of 4 drops of hydrogen peroxide to urine. **POCT Test** Utriplex test (Investigational use, Mologic Ltd, Bedford shire, UK) - measures 3 inflammatory markers - HNE, MMP8 & Cystatin C **Threshold** Visualization of ≥2 test lines after 6 min indicates UTI	**Reference standard** Culture **Threshold** $\geq 10^5$ CFU/mL of a single pathogen

Study Details	Participants	POCT Test Details	Reference standard
Colodner(2000)[47] **Country** Israel **Language** English **Funding** Not reported	**Setting & 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 Test** 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, then incubated overnight for culture evaluation the next day. **Threshold** Evaluated according to manufacturer's chart. Two thresholds evaluated - 10^4 & 10^5 CFU	**Reference standard** Culture - standard culture plates - MacConkey Agar, CAN & 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

Study Details	Participants	POCT Test Details	Reference standard
Greeff(2002)[44] **Country** South Africa **Language** English **Funding** Not reported	**Setting & Population** Antenatal clinics; Screening Pregnant women **Inclusion criteria** Two populations of patients from the Pretoria region were involved: (i) asymptomatic pregnant women attending the antenatal clinic for the first time or presenting in labour; and (ii) pregnant women with symptoms suggestive of UT **Exclusion criteria** NR **Number included (number analysed)** 453(374) **Age –** NR **% Female:** 100	**Urine sampling method** Self-collected midstream urine **Target condition** Presence of UTI **Location of test performance** Antenatal clinic **POCT Test** 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 lab culture **Threshold** $>10^5$ CFU/ml

Study Details	Participants	POCT Test Details	Reference standard
Holm(2017)[35] **Country** Denmark; DTA study nested in Danish RCT[33] **Language** English **Funding** 2016, University of Copenhagen, (b) Læge Sofus Carl Emil Friis og Hustru Olga Doris Friis' Legat and (c) SSI Diagnostika (materials)	**Setting & 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 (POC is read the following day). **Number included (number analysed)** 376 (341) **Age** 48.5 years **% Female**: 100	**Urine sampling method** Midstream urine sample **Target condition** Presence of UTI **Location of test performance** General practice **POCT Test** 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 $\geq 10^3$ CFU/mL for any uropathogen. Inconclusive' labelled as negative. **POCT Test** ID Flexicult; Chromogenic agar allowing identification and quantification of 6 types of bacteria **Threshold** $\geq 10^3$ 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** $\geq 10^3$ CFU/mL for E. coli and S. saprophyticus, $\geq 10^4$ CFU/mL for other typical uropathogens, ≥10^5 for possible uropathogens in accordance with European consensus

Study Details	Participants	POCT Test Details	Reference standard
Hullegie(2017)[34] **Country** Wales, England, Spain and Netherlands. DTA sub-study from POETIC study. [8] **Language** English **Funding** European Community's Seventh Framework Programme and R-GNOSIS consortium	**Setting & Population** Primary care; Uncomplicated UTI; Women **Inclusion criteria** Women randomised to Flexicult arm of POETIC trial; age≥18 years with symptoms of UTI (dysuria, urgency or frequency). **Exclusion criteria** Women who were either terminally ill, were receiving treatment for life-threating cancer, were having severe systemic symptoms or had received antibiotics for UTI within the past four weeks **Number included (number analysed)** 325(312) **Age** 49 **% 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 Test** 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 no. 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/ pre-dominant 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 lab 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

Study Details	Participants	POCT Test Details	Reference standard
Lee(2010)[45] **Country** Korea **Language** Korean - extracted using google translate **Funding** Not reported	**Setting & Population** Secondary care; Uncomplicated UTI Age <24 months **Inclusion criteria** Febrile infants age<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** Midstream urine or urine collection bags **Target condition** Presence of UTI Presence of UTI - caused by E.Coli **Location of test performance** Outpatient setting **POCT Test** 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** $\geq 10^5$ CFU single bacterium; $\geq 10^4$ CFU/ml in patients with symptoms

Study Details	Participants	POCT Test Details	Reference standard
Macias(2002)[42] **Country** Mexico **Language** Spanish **Funding** NR	**Setting & Population** ICU; indwelling catheter **Inclusion criteria** Hospitalised adults; indwelling catheter. **Exclusion criteria** Recognized 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 3-5ml per puncture of the probe. Samples taken every 72hr **Target condition** Presence of UTI **Location of test performance** Not reported but likely in hospital **POCT Test** Uriscreen – 2ml of urine placed in tube with catalyst, to which four drops of H20 added. After mixing gently for five 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 less than 1mm covering the entire surface 3) +++ foam band greater than 1mm	**Reference standard** Culture **Threshold** 10^3 CFU/mL

Study Details	Participants	POCT Test Details	Reference standard
Mignini(2009)[46] **Country** Argentina **Language** English **Funding:** Supported by UNDP/UNFPA/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction.	**Setting & Population** Antenatal clinics; Screening Pregnant women **Inclusion criteria** All women attending antenatal clinics who presented with live foetuses at gestational weeks 12 to 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 3 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 Test** 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** $\geq 10^5$ CFU/mL or higher of a single microorganism or when two different colonies were present but one was 10^5 CFU/mL or higher.	**Reference standard** Culture Classic quantitative culturing in the microbiology lab **Threshold** $\geq 10^5$ CFU/mL or more of a single potential uropathogen or of two organisms not consistent with kin flora were isolated

Study Details	Participants	POCT Test Details	Reference standard
Millar(2000)[39] **Country** 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 & 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 Test** 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** 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** $\geq 10^4$ CFU/ml of single potential uropathogen. Cultures were considered negative if fewer than 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.

Study Details	Participants	POCT Test Details	Reference standard
Pernille(2019)[38, 52] **Country** Denmark **Language** English **Funding** University of Copenhagen, 2016 funds, and The PLU fond (Praktiserende Laegers Undervisningsfond)	**Setting & Population** Primary care; Uncomplicated UTI Women **Inclusion criteria** Women age≥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 MSU analysis **Target condition** Presence of UTI **Location of test performance** Primary care **POCT Test** ID Flexicult **Threshold** >five 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 lab culture **Threshold** $\geq 10^3$ CFU/mL for E. coli and S. saprophyticus, $\geq 10^4$ CFU/mL for other typical uropathogens and $\geq 10^5$ CFU/mL for possible uropathogens. Growth of more than two different colonies (mixed cultures) considered as non-significant growth

Study Details	Participants	POCT Test Details	Reference standard
Teppa(2005)[40] **Country** Venezuela **Language** English **Funding** Not reported	**Setting & 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 Dept of Obstetrics and Gynaecology **POCT Test** Uriscreen – 2mL 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** $\geq 10^5$CFU/mL of single pathogen or any nonuropathogneic bacteria. Contaminated if multiple organisms identified.

Study Details	Participants	POCT Test Details	Reference standard
[redacted]49, 88 [redacted]	[redacted]	[redacted]	[redacted]

Study Details	Participants	POCT Test Details	Reference standard
Yagupsky(2000)[48] **Country** Israel **Language** English **Funding** Not reported	**Setting & 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** Midstream urine samples **Target condition** Presence of UTI Pathogenic cause **Location of test performance** Laboratory **POCT Test** Dipstreak - performed using the Uriselect 3 blood agar configuration, following the manufacturer's instructions. If no growth was observed or the colony count < 10 CFU, plates and DipStreak devices were reincubated for 24 h 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** $\geq 10^5$ 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 nonsignificant growth of other bacteria. Growth of 10^4-10^5 CFU/mL of one or two organisms indicated the need for a repeat culture

Results

Study Details	Population and Setting	POCT Test	Reference standard	Target condition	TP	FP	FN	TN	Sens	Spec	Missing samples/notes
«Author»(«Year») {#«DataExGeneralDataStudyID»}	**Population:** Children (<16 years) **Setting:** «Setting» **Location of test performance:** Near patient setting «Subgroup»	«Test1»	«Refstan»	«Target_condition»	«TP1»	«FP1»	«FN1»	«TN1»	«Sensitivity1»	«Specificity1»	None
Blom(2002)[37]	**Population:** Mixed symptomatic **Setting:** «Setting» **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 5 times (once for each antibiotic)
				Presence of UTI	58	3	17	43	NR	NR	None
Bongard(2015)[18]	**Population:** Mixed **Setting:** «Setting» **Location of test performance:** Laboratory	Flexicult™ SSI urinary kit	Culture & Microscopy	Presence of UTI	39	27	6	128	87	83	None
			Culture & microscopy & spiral plating	Presence of UTI	50	16	4	130	NR	NR	None
			Culture	Antimicrobial resistance	84	2	22	33	NR	NR	2x2 data obtained by summing across all antibiotics
Boon(2022)[41]	**Population:** Children (<18 years) **Setting:** «Setting» **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

Study Details	Population and Setting	POCT Test	Reference standard	Target condition	TP	FP	FN	TN	Sens	Spec	Missing samples/notes
Colodner(2000)[47]	**Population:** Mixed – Fresh urine samples **Setting:** «Setting» **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)[44]	**Population:** Symptomatic pregnant women; Screening pregnant women **Setting:** «Setting»**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	
Holm(2017)[35]	**Population:** Women - uncomplicated UTI **Setting:** «Setting» **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)[34]	**Population:** Women - uncomplicated UTI **Setting:** «Setting» **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 lab definitiion:		74	128	20	67	79	34	

Study Details	Population and Setting	POCT Test	Reference standard	Target condition	TP	FP	FN	TN	Sens	Spec	Missing samples/notes
			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)[45]	**Population:** Children (<16 years) **Setting:** «Setting» **Location of test performance:** Near patient setting	Uricult trio	Culture	Presence of UTI	19	18	13	101	59	85	7 missing samples - 2 patients failed to collect sample, 3 only had urine culture tests performed and 2 patients only performed index test.
				Presence of E.Coli	12	5	8	126	60	96	
Macias(2002([42]	**Population:** Catheterised ICU patients **Setting:** «Setting» **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 greater than 1 mm	35	14	27	32	57	70	
Mignini(2009)[46]	**Population:** Screening - pregnant women **Setting:** «Setting» **Location of test performance:** Laboratory	Uricult	Culture	Presence of UTI	321	8	8	1836	98	100	830 samples excluded due to contamination
Millar(2000)[39]	**Population:** Screening - pregnant women **Setting:** «Setting» **Location of test performance:** Near patient setting	Uriscreen	Culture	Presence of UTI	30	185	13	150	70	45	5/383 samples contaminated & excluded Inter-rater reliability: 28/30 samples interpreted consistently

Study Details	Population and Setting	POCT Test	Reference standard	Target condition	TP	FP	FN	TN	Sens	Spec	Missing samples/notes
Pernille(2019) [38]	**Population:** Women - uncomplicated UTI **Setting:** «Setting» **Location of test performance:** Near patient setting	ID Flexicult	Culture	Presence of UTI - MSU 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 MSU; little impact of delay in analysis
Teppa(2005)[40]	**Population:** Screening - pregnant women **Setting:** «Setting» **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
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Study Details	Population and Setting	POCT Test	Reference standard	Target condition	TP	FP	FN	TN	Sens	Spec	Missing samples/notes
	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
				[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
				[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Yagupsky(2000)[48]	**Population:** Mixed – fresh urine samples **Setting:** «Setting» **Location of test performance:** Laboratory	Dipstreak	Culture	Presence of UTI	270	4	12	509	96	99	275 excluded due to contamination
				Pathogenic cause	211	NA	59	NA	NR	NR	211/270 correctly identified. None incorrectly identified but 59 were not identified

Risk of Bias

Study Details	Anacleto(2009)[43]
Index test:	Uricult Trio

Domain 1: Patient selection	
Consecutive patients; had to have tested positive on LE or nitrite so applicabiltiy 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	
Pre-specified, 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 pre-specified?	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)[37]
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 unlikley that would have been aware of result as likely to have been interpreted first. Pre-specified 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 pre-specified?	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	
1 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)[18]
Index test:	Flexicult Human

Domain 1: Patient selection	
Convenience sample of urines available in the lab	
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 lab 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 pre-specified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**

DOMAIN 3: REFERENCE STANDARD	
Culture & 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 sub-sample assessed for antimicrobial sensitivty - 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)[41, 51]

Comparative review question

Patients:	300 children aged <18
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)	
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?	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 lab 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 pre-specified?	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 un aware 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 lab; 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 v culture comparison but this was due to late introduction of the test.	

Study Details	Colodner(2000)[47]
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 lab 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 pre-specified?	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)[44]
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 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 pre-specified?	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)[35]

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 - explicIty 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 pre-specified?	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)?	NA
Is undergoing one index test unlikely to affect the performance of the other index test(s)?	NA
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 lab 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)[34]
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 likley interpreted blind to lab 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 pre-specified?	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 - 10 cases clinican did not ocmplete CRF, 3 cases test 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)[45]
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	
Pre-specified, 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 pre-specified?	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 & 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
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)[42]
Index test:	Uriscreen

Domain 1: Patient selection	
ICU patients - no details on 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 pre-specified. No information on blinding but test performed before refernece 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 pre-specified?	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)[46]
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 pre-specified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**

DOMAIN 3: REFERENCE STANDARD	
Standard lab 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)[39]
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 pre-specified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Low**

DOMAIN 3: REFERENCE STANDARD	
Standard lab 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)[38, 52]
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 pre-specified?	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	
5 women excluded - 2 unable to deliver sufficient urine; 3 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)[40]
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**

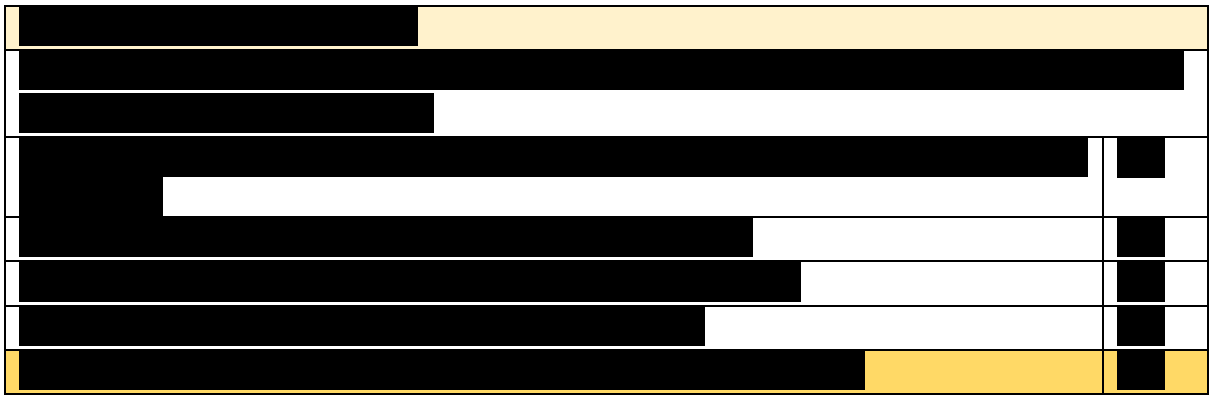
DOMAIN 2: INDEX TEST	
Standard threshold; interpreted before reference standard results available	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
Could the conduct or interpretation of the index 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 2x2 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	

[redacted]	[redacted]
Index test:	Lodestar DX



[redacted]	[redacted]
[redacted]	

Study Details	Yagupsky(2000)[48]
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 perofrmed in same lab 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 pre-specified?	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 perofrmed in same lab 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; Sample 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	

Appendix 3.3: Objective 3

Study details*	Participants & Test	Results
Anacleto(2009)[43] **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 & Population** Secondary care; Uncomplicated UTI Age <16 years **Inclusion criteria** Infants & children age 0 to 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 7 urinary tract; midline defects; failure to thrive; concomitant infections; recurrent UTI; asymptomatic bacteriuria; other co-morbid 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"

Study details*	Participants & Test	Results
Blom (2002)[37] **Country** Denmark **Language** English **Funding** Not reported	**Setting & 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"

Study details*	Participants & Test	Results
Brooks-Howell (2019)[53] **Country** Wales, England, Spain, Netherlands **Language** English **Funding** EU funding as part of the R-GNOSIS program	**Setting & Population** Telephone interviews; Primary care clinicians & 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 lab test (targeted treatment within 24hr instead of 3-4 days) but some concern about possible patient discomfort whilst waiting to obtain results rather than prescribing straight away. • **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.

Study details*	Participants & Test	Results
Butler (2018)[8] **Country** England, Netherlands, Spain & Wales **Language** English **Funding** European Commission Seventh Framework Programme	**Setting & 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.«Exclusion» **Number of eligible patients (randomised)** 654 (653) **Age** 47.6 years (SD=27.6) **Sex** all female **Test** Flexicult™ SSI urinary kit	• **Time to perform test –** 9min to prepare test, 6min to obtain and record result, 7min to discuss result with patient • **Cost –** Cost per person including POCT cost in UK is £48 • **Management change as a result of test –** ○ 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%

Study details*	Participants & Test	Results
Greeff(2002)[44] **Country** South Africa **Language** English **Funding** Not reported	**Setting & Population** Antenatal clinics; Screening Pregnant women **Inclusion criteria** Two populations of patients from the Pretoria region were involved: (i) asymptomatic pregnant women attending the antenatal clinic for the first time or presenting in labour; and (ii) pregnant women with symptoms suggestive of UT **Exclusion criteria** NR **Number included (number analysed)** 453(374) **Age –** NR **% Female:** 100 **Test** Uricult trio	• **Ease of use/ acceptability –** ○ "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" ○ "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"

*All studies were accuracy studies with the exception of Brooks-Howell which employed qualitative thematic analysis of semi-structured interviews

Appendix 4: PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Title page
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Page 4
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Section 1-2
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Section 3
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Section 4.1/ Section 4.2.5
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Section 4.2/ Appendix 1
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Appendix 1
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Section 4.2.3
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Section 4.2.3
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Section 4.1
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Section 4.1/ Appendix 3

Section and Topic	Item #	Checklist item	Location where item is reported
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Section 4.2.4
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Section 4.2.3/ 4.2.5
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Section 4.2.5
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Section 4.2.5
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Section 4.2.5
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Section 4.2.5
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Section 4.2.5
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Not completed
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Not completed
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Not completed
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Section 5.1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Section 5.1
Study characteristics	17	Cite each included study and present its characteristics.	Section 5.1-5.4

Section and Topic	Item #	Checklist item	Location where item is reported
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Section 5.3.1/ Appendix 3.1-3.2
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Section 5.3.2
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Section 5.3.1-5.3.2
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Section 5.3.2
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	NA
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	NA
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	NA
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	NA
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Section 9
	23b	Discuss any limitations of the evidence included in the review.	Section 9.2
	23c	Discuss any limitations of the review processes used.	Section 9.2
	23d	Discuss implications of the results for practice, policy, and future research.	Section 9
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Section 4
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Section 4
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Section 4.3
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Page 2

Section and Topic	Item #	Checklist item	Location where item is reported
Competing interests	26	Declare any competing interests of review authors.	Page 2
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Section 11.3

From: 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. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

Evidence overview: Early value assessment – Point of care tests for urinary tract infections to reduce antimicrobial resistance

This overview summarises the main issues the diagnostics advisory committee needs to consider. It should be read together with the final scope and the early value external assessment report.

1 Aims and scope

Urinary tract infections (UTIs) are a common condition found in primary care and contribute to a large proportion of antibiotic use. Many UTIs, particularly acute uncomplicated UTIs with no risk factors, resolve within a few days. But UTIs often recur over time and can require several courses of antibiotics to treat them. In some cases, UTIs may not respond to treatment and can become chronic. Recurrent infection in adults is defined as at least 3 UTIs per year or 2 UTIs in the last 6 months (Diagnosis of urinary tract infections, published by Public Health England, 2020). There is currently no accepted definition for chronic UTIs, but it is generally described as when a person experiences prolonged periods without remission from symptoms.

Diagnosis of urinary tract infections (published by Public Health England, 2020) sets out several flowcharts to guide diagnosis for people with suspected acute UTIs. Separate pathways are presented for:

- women under 65 years with suspected UTI
- men under 65 years with suspected UTI
- adults who are catheterised or over 65 years with suspected UTI
- infants/children under 16 years with suspected UTI.

The pathways differ in terms of whether initial dipstick testing is done, deciding if urine samples should be sent for further laboratory testing and deciding if or when to prescribe antibiotics.

UTIs are currently diagnosed using a combination of clinical symptoms and dipstick tests (where appropriate). Urine samples may also be sent for laboratory-based tests that identify bacteria and do antimicrobial sensitivity testing (AST; done to determine which antibiotics the infection will respond to).

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. They are rapid and can be done in a GP surgery. But, they may not accurately identify UTIs and are not recommended for all populations (such as men, people aged over 65 years or people who are catheterised). They are also unable to provide information on the bacteria causing the infection or on AST, so laboratory-based tests may also be needed. Laboratory-based tests can take 24 to 72 hours depending on local available facilities, geographical location, and day of sample collection. They may sometimes take longer (up to a week) if there are delays in getting samples to the laboratory or a delay in processing.

Where laboratory-based testing is done, suspected UTIs are often initially treated with empiric (broad spectrum) antibiotics whilst waiting for results. Antibiotics may need to be changed if this initial treatment is unlikely to be effective against the infection, or stopped if no infection is detected.

Point-of-care tests (POCTs) that can improve UTI detection, identify bacteria and do AST quicker than laboratory-based tests could reduce unnecessary or ineffective antibiotic prescribing. This could help reduce the risk of antimicrobial resistance. Improved testing could also lead to a reduced impact of UTIs on patients, reduced laboratory workloads and associated costs, and reduced healthcare resource use associated with mismanaged UTIs. Quicker

use of an appropriate antibiotic could also relieve a person's symptoms much more quickly and effectively.

If implemented, new POCTs may impact the amount of plastic used and waste created in primary or community care settings. Increases in waste production may also increase costs if disposing of materials needs extra equipment or time compared to standard care.

This topic is presented as an early value assessment. The decision questions that would need to be answered in guidance are presented below.

Decision questions

- Do POCTs for people with suspected UTIs have the potential to be clinically and cost-effective to the NHS?
- What evidence is available to support the value proposition outlined in the scope and where are the evidence gaps?

Populations

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)

The populations exclude people with suspected sepsis.

Where data permits, the following subgroups may be considered:

- people with suspected acute UTI
- people with suspected recurrent UTI
- people with suspected chronic UTI
- women under 65
- women over 65
- men under 65
- men under 65

- adults with indwelling urinary catheters
- babies, children and young people under 16
- children under 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 UTIs (for example people with neurogenic bladder, diabetes, polycystic kidney disease or people who are immunocompromised)
- people with suspected pyelonephritis

Interventions

During scoping, several technologies were identified that appeared to meet criteria for inclusion in this assessment but did not have regulatory approval for use in the UK. Such technologies were included in the scope if the company indicated that they expected regulatory approval within 12 months but can only be included in draft or final guidance if they have regulatory approval by the planned draft or final guidance publication date.

The included technologies can all detect bacteria, but some can also identify bacteria or do AST. Only 2 tests can do either of these functions and provide results in under an hour (Astrego PA-100 and Lodestar DX). The EAG divided the tests into rapid tests (that give results in 40 minutes or less) and culture-based tests (that give results in up to 24 hours).

Table 1 – Overview of included point of care tests

Test name	Antibiotics/bacteria targeted	Time to detect bacteria	Time to identify bacteria	Time to result AST	CE-IVD/UKCA marked
Uriscreen	Detects catalase	2	NA	NA	CE-IVD

Test name	Antibiotics/bacteria targeted	Time to detect bacteria	Time to identify bacteria	Time to result AST	CE-IVD/UKCA marked
(Savyon Diagnostics Ltd)	activity as indicator of bacteria	minutes			marked
UTRiPLEX (Global Access Diagnostics)	Detects presence of urinary biomarkers matrix metalloproteinase-8 (MMP8) and human neutrophil elastase (HNE)	6 minutes	NA	NA	Not currently available in the UK UKCA mark expected 2023.
Astrego PA-100 analyser and PA-AST panel U-0501 (Sysmex Astrego)	5 antibiotics (amoxicillin-Clavulanic acid, ciprofloxacin, fosfomycin, nitrofurantoin, trimethoprim)	10 to 15 minutes	NA	30 to 45 minutes for full results	CE-IVD marked Not currently available in the UK but launch planned in 2023
TriVerity (Inflammatix)	Identifies presence type and severity of infection	30 minutes	NA	NA	No CE or UKCA mark Not currently available in the UK
Lodestar DX (Llusern Scientific)	Escherichia coli (E-coli), Klebsiella spp, Proteus mirabilis, Staphylococcus saprophyticus, Enterococcus spp, Pseudomonas aeruginosa	40 minutes	40 minutes	NA	Not currently available in the UK UKCA mark expected 2023.
Flexicult Human	5 antibiotics (mecillinam,	16 to 24 hours	NA	16 to 24 hours	CE-IVD marked

Test name	Antibiotics/bacteria targeted	Time to detect bacteria	Time to identify bacteria	Time to result AST	CE-IVD/UKCA marked
(SSI Diagnostica)	nitrofurantoin, ampicillin, sulfamethizol and trimethoprim).				
Uricult, Uricult trio and Uricult plus (Aidian; formerly Orion Diagnostica)	Uricult: presence of gram-negative bacteria; Uricult plus: also detects enterococci; Uricult trio: also detects gram-negative, β-glucuronidase-producing organisms e.g. E. coli	16 to 24 hours	16 to 24 hours	NA	CE-IVD marked Available in the UK and UKCA mark expected 2023
Diaslide, Dipstreak, Chromostreak (Novamed)	Total bacterial count; presence of gram-negative bacteria; growth of common UTI causing bacteria (E. coli, Proteus, and enterococci) – chromostreak only	18 to 24 hours	18 to 24 hours	NA	Unclear

NA: Not applicable. MMP-8: Matrix metalloproteinase-8. HNE: human neutrophil elastase.

Comparators

The comparators for this assessment are:

- dipstick testing, followed by laboratory-based testing (if necessary; population 1)

  or

- laboratory based testing alone (population 2).

Healthcare settings

The healthcare settings for the intervention are primary or community care.

Further details, including descriptions of the interventions, comparator, care pathway and outcomes, are in the final scope for point of care tests for urinary tract infections to reduce antimicrobial resistance.

2 Summary

Overview of the EAG's systematic review

The external assessment group (EAG) did a systematic review to identify evidence on the clinical effectiveness (objective 1; see section 5.2 in the external assessment report [EAR]), diagnostic accuracy (objective 2; see section 5.3 in the EAR) and technical performance (objective 3; see section 5.4 in the EAR) for each test included in the evaluation. The number of studies identified for each test is in the following table. Find the full systematic review methods and results in section 4 and 5 of the EAR. Data are described further in key issues 1 and 2 of this overview.

Table 2 – Overview of number of studies assessing each rapid test (results in under 40 minutes) for each review objective

Test	Objective 1 (clinical effectiveness)	Objective 2 (accuracy data)	Objective 3 (technical data)
Astrego PA-100 system	0	0	0
Lodestar DX	0	1	0
TriVerity	0	0	0
Uriscreen	0	4	0
UTRiPLEX	0	1	0

Table 3 – Overview of number of studies assessing each culture-based test (results in up to 24 hours) for each review objective

Test	Objective 1 (clinical effectiveness)	Objective 2 (accuracy data)	Objective 3 (technical data)
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

Ongoing studies

The EAG identified ongoing accuracy studies for the Astrego PA-100 AST and Lodestar DX tests. [redacted]. The company stated that the Lodestar DX test will be included in the TOUCAN study, a study evaluating the accuracy of at least 3 POCTs (details of these not yet available) to standard tests in up to 800 women who consult their GP with symptoms of UTI. This study is due to complete in October 2023. The company submission for the Lodestar DX test noted a further ongoing study exploring the potential clinical impact of the Lodestar DX device on treatment decisions with a range of clinicians in Wales. This study is due to be completed in March 2023.

Cost effectiveness

The EAG reviewed published economic evaluations, and also did a pragmatic literature search to identify further published economic evidence for UTIs more broadly. Find more details about the studies included in the EAG's cost-effectiveness analysis in section 6.2.1 and 6.2.2 of the EAR.

The EAG also developed a conceptual model for assessing the cost-effectiveness for POCTs for UTI in comparison to laboratory-based culture testing (with or without dipstick testing; see figure 5 in the EAR). The EAG stated that its model identified pathways for the benefits of POCTs: they could reduce the use of empiric antibiotics and, by reducing the incidence of UTI complications and improving cure rates, reduce healthcare costs quality of life impacts arising from UTIs. Find more detail about the EAG's conceptual economic model in sections 6.1 and 6.2 of the EAR.

The EAG described the structure of a decision-analytic model, based on a decision tree, and identified parameter values that could be used. No results were provided because the EAG considered the evidence identified as too limited for results to be meaningful. Find more detail in sections 6.3.1 the EAR, with a suggested decision tree structure in figure 6.

Cost data was not provided for all tests. Test costs (excluding VAT) were provided by the companies for the Astrego PA-100 system [redacted] and Lodestar DX [redacted]
The EAG identified the costs of Flexicult human per person (£48) from Butler et al. (2018).

Overall, the EAG highlighted a lack of available data, particularly for the more rapid tests. Based on the features of the tests, it considered the Astrego PA-100 system and Lodestar DX as the most promising tests.

3 Issues for consideration

Key issue 1: Impact of tests on antibiotic use

Description of issue

Two studies assessed the impact of POCTs use on antibiotic use (Butler et al. 2018 [POETIC trial] and Holm et al. 2017). Both studies assessed the culture-based test, Flexicult human, which does AST and takes 16 to 24 hours to give results. No data on antibiotic use was identified for all other included tests. A test that provides results in a quicker time may have a different impact on antibiotic use.

Background

During scoping it was highlighted that the extent to which POCTs impact antibiotic prescribing may depend on how quickly results are available and how they can be implemented into clinical decision making in primary and community care settings.

An important potential benefit of POCTs is to reduce the risk of antimicrobial resistance. Technologies that can quickly and more accurately rule out UTIs at the point of care may reduce the number of people receiving antibiotics that do not need them. Technologies that identify bacteria or provide antimicrobial susceptibility results faster than current techniques may help improve appropriate antibiotic prescribing when needed. The included technologies vary in terms of which of the type of information they give (identify if a UTI is present, identify the type of bacteria and do AST; see table 1).

Two randomised controlled trials (RCTs) evaluated the clinical effectiveness of Flexicult human in women aged over 18 years with suspected UTI in primary care. The EAG judged both to be a low risk of bias. The POETIC study (Butler et al., 2018; n=654) was done in England (n=234), Wales (n=219), Netherlands and Spain, and compared Flexicult human with standard care. A Danish trial (Holm et al., 2017; n=376) compared Flexicult

human with ID Flexicult (which identifies the presence of bacteria but does not include an AST component). In the Danish trial, GPs were advised to treat based on the results of the test. But, in the POETIC study, GPs were free to choose how to use the test. This may have included treating people the following day based on the test results, prescribing empirically and reviewing the antibiotic based on the results or providing delayed antibiotics and using the test to guide use of these.

The POETIC study reported a statistically significant difference between groups for antibiotic prescribing at the initial consultation (odds ratio [OR] 0.56, 95% confidence interval [CI] 0.35 to 0.88) but there was no evidence of a difference seen for overall antibiotic prescription or concordant antibiotic use (defined 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 (see table 3). The Danish study reported that 'appropriate prescribing' was higher in the ID Flexicult arm compared to the Flexicult human arm (OR 1.11, 95% CI 1.03, 1.99). No other differences were identified between the groups in either trial, including use of broad spectrum antibiotics. No statistically significant differences in outcomes related to UTI or symptom incidence or duration were reported in either study (see table 8 in the EAR for further detail).

Table 3 – Results of the POETIC trial (evaluating Flexicult human) on clinical outcomes

Outcome	Effect measure – estimate (95% CI)
Concordant antibiotic use	OR = 0.84 (0.58, 1.20)
Antibiotic prescribing at initial consultation	OR = 0.56 (0.35, 0.88)
Antibiotics prescribed to guidelines at initial consultation	OR = 0.99 (0.67, 1.45)
Antibiotic consumed day 3	OR = 1.24 (0.81, 1.89)
Antibiotic consumed (during 2 weeks)	OR = 1.38 (0.87, 2.19)

New antibiotic prescription (within 2 weeks)	OR = 1.11 (0.65, 1.89)
UTI-specific and 1–3 days	Reference
UTI-specific and >3 days	RR = 1.15 (0.71, 1.87)
Broad spectrum and 1–3 days	NA (0 events)
Broad spectrum and >3 days	RR = 1.00 (0.58, 1.75)

CI: confidence interval. NA: not applicable. OR: odds ratio. RR: relative risk. UTI: urinary tract infection.

Find more information about the design, methodology and results of the 2 RCTs in Section 5.2 and Table 8 of the EAR. There is no available data on antibiotic use for other included tests.

In 63% of participants in POETIC patient management was changed as a result of the Flexicult test (Butler et al., 2018). Qualitative feedback from clinicians using Flexicult human in the POETIC trial stated that there was an even split between those that thought it would have no impact on prescribing and those who stated that it had increased their awareness about antibiotic prescribing, and now had more cautious prescribing habits (Brookes-Howell et al., 2019). Most clinicians agreed that the test gave quicker results than laboratory tests (24 hours compared to 3 to 4 days), reassured patients, and had a positive impact on clinician confidence in diagnosing UTI. But, clinicians noted difficulties in test result interpretation, limitations on when it can be used, limited resources to undertake testing, concerns about prolonging patient discomfort whilst waiting for test results and concerns about the potential expense of maintaining 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. Find further detail about the data included for objective 3 in section 5.4 of EAR.

The EAG stated that stronger evidence that point-of care tests reduce inappropriate and unnecessary antibiotic use is needed for all included tests in the assessment, before they can be routinely used in practice.

Questions for committee

- Based on the available evidence, how clear is the potential impact of using the point of care tests on antibiotic prescribing? Would this differ if the test gave results in a quicker time?
- How quickly are results from point of care tests likely to be needed to have the greatest impact on antibiotic use in the NHS?
- What information from tests is likely to be most important to make decisions about antibiotic use (UTI presence, bacterial identification, AST)?
- What measures of antibiotic use are most important to measure to assess potential impact on use and antibiotic resistance (for example, broad versus narrow spectrum, World Health Organisation AWaRe classification or specific type of antibiotic)?

Key issue 2: Accuracy of point of care tests

Description of issue

Few accuracy studies were identified that the EAG did not consider at high or unclear risk of bias. Few studies assessed performance to identify bacteria or assess antimicrobial sensitivity. Five studies directly compared tests to dipsticks (a rapid test initially used to aid UTI diagnosis for certain populations). But, only 3 of these studies evaluated rapid POCTs and the results were mixed. Data was not identified for all groups included in the scope. It is unclear whether estimates of sensitivity and specificity can be generalised to the wider population. Studies were heterogenous in terms of setting, population and where the test was performed.

Background

UTIs are currently diagnosed using a combination of clinical symptoms, dipstick testing (in certain populations) and laboratory-based culture testing

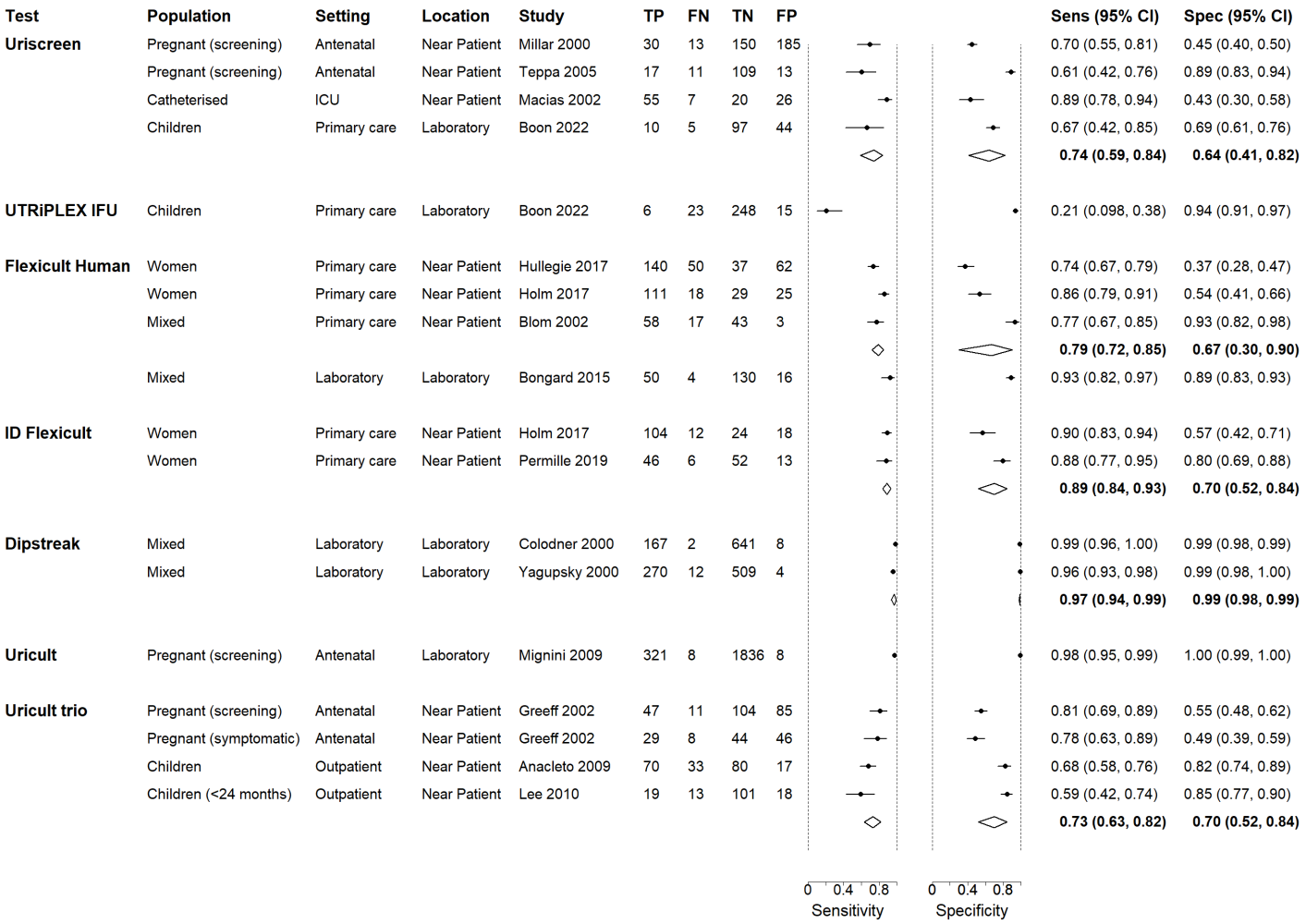
(where appropriate). The Public Health England quick reference tool on the diagnosis of UTIs says that dipstick tests should not be performed in adults who are catheterised or over 65 as these groups often have asymptomatic bacteriuria that is not harmful, can give a positive dipstick result and antibiotics are not beneficial and may cause harm. It also recommends not using dipstick testing in men to rule out infection as they are too unreliable.

The EAG identified 16 studies from 20 articles reporting test accuracy data. Most assessed culture-based tests that take up to 24 hours to provide results. Of the rapid tests, Uriscreen was evaluated in 4 studies, and UTRiPLEX and Lodestar DX were evaluated in 1 study each. Only 3 studies were considered at low risk of bias by the EAG. These were Anacleto et al. (2009), Boon et al. (2022) and Holm et al. (2017). The EAG considered 8 studies to be of unclear risk of bias, mainly due to lack of information of blinding of the interpreter of the reference standard. Find further detail on risk of bias assessment in section 5.3.1 in the EAR. The studies varied in the populations assessed (see table 9 in the EAR and figure 1 in this overview for further detail). Study and summary results for sensitivity and specificity for test accuracy for UTI detection are in the following figure (figure 1).

Three studies evaluating Flexicult human reported data on antimicrobial sensitivity. Estimates of sensitivity ranged from 79% to 90% with a summary estimate of 87% (95% confidence interval [CI] 83 to 90). Estimates of specificity ranged from 72% to 94% with a summary estimate of 93% (95% CI 89 to 95). Two studies (evaluating Lodestar DX and Uricult trio) reported data on the presence of E.Coli. [redacted]

Uricult trio had a reported sensitivity of 60% and specificity of 90%. One study reported that Dipstreak correctly identified the bacteria likely to cause a UTI in 211/270 cases. The EAG stated that there was little reliable data identified on the performance of the tests to identify bacteria to direct targeted treatment.

Figure 1 – Paired forest plots of study and summary estimates of test sensitivity and specificity for the detection of UTI



Six studies directly compared (accuracy of tests was assessed in the same population) POCTs to dipstick testing for LE or nitrite. Rapid tests assessed were UTRiPLEX and Uriscreen. Three studies evaluated Uriscreen (which takes 2 minutes for results) in different populations: people under 18 years, catheterised ICU patients and pregnant women. One study, evaluating Urisceen in people under 18 years (Boon et al. 2022), was considered at low risk of bias by the EAG. It reported higher sensitivity (67% compared to 32%) and lower specificity (69% compared to 86%) than dipstick testing (either nitrite positive or LE positive was considered a positive result). The EAG highlighted that how a positive dipstick test was defined varied across studies (see table 13 in the EAR). One study evaluating UTRiPLEX reported it to be less sensitive but more specific than dipstick testing. Studies reporting on culture-based tests reported that these were more sensitive and more specific than standard dipstick tests, but these tests take much longer to give a result. Find further detail of studies comparing POCTs to dipstick tests in section 5.3.2.10 and Table 13 in the EAR).

Populations specified in the assessed studies included women with uncomplicated UTIs (3 studies), pregnant women (4 studies), people who are catheterised (1 study) and children or infants (1 study under 18 years, 1 study 16 years and 1 study under 24 months). Specific data were not available for all subgroups defined in the scope, for example men and people over 65 years. Five studies included analysed samples from mixed samples with no further detail reported. Three studies specifically stated that people with recurrent UTI were excluded. No further detail was reported on whether people with recurrent or chronic UTI took part in the remaining studies.

The EAG stated that there were insufficient data to investigate whether test performance differed between the populations defined in the scope, or to consider how having recurrent or chronic UTI could impact on test performance.

The EAG and clinical experts stated that culture-based tests may miss infections if bacterial counts in a urine sample are too low, or if the culture medium doesn't favour the specific UTI causing pathogenic bacteria. Clinical experts also noted that urine is filled with multiple bacteria and current laboratory guidelines differ in how culture result should be interpreted to confirm or rule out UTI diagnosis. Nearly all included studies used culture alone as the reference standard with thresholds ranging from 10^3 CFU to 10^5 CFU. Some studies confirmed diagnosis based on the presence of a single type of bacteria, and others had different thresholds for mixed growth.

Questions for committee

- Based on the identified data, how confident are you about how well the technologies are likely to perform if used in the NHS, when compared to tests used in current standard care?

- How generalisable are test accuracy results from identified studies likely to be across the populations included in the scope? For example:

    - women under 65 years, men under 65 years, adults who are catheterised, people over 65 years, infants or children under 16 years with suspected UTI (or other groups, as defined in the scope population; see section 1).
    - people with suspected acute, recurrent or chronic UTI.

Key issue 3: Proposed approach to modelling impact of tests on antibiotic use

Description of issue

The EAG proposed assessing the impact of POCTs on the use of empiric or targeted antibiotics using a decision tree model, which is based on a number of assumptions. There is limited data on the direct impact of tests on antibiotic use.

Background

There was little direct data on the impact of test use on antibiotic prescribing (see key issue 1). The EAG explored modelling the impact of tests on antibiotic use using a decision tree structure. Find the EAG's decision tree in Figure 6 of the EAR. The structure is based on rapid tests that identify bacteria or do AST (such as Lodestar Dx or the Astrego PA-100 system), culture-based POCTs (such as Flexicult human) and laboratory culture-based testing (with or without dipstick testing). Antibiotic use was modelled as being empiric or targeted (that is, use was informed by the results of bacterial identification or AST testing).

Whether and which type of antibiotic was used was modelled based on parameters and features of the tests. Test accuracy parameters for detecting UTI determine how many people get a positive or negative result for UTI. People with positive results (true or false) would get antibiotics and a subset of those who tested negative would also receive empiric antibiotics (no parameter to inform this was identified). Of those testing positive for UTI:

- targeted antibiotics only were assumed to be used for people who were tested by a rapid test which was able to identify a specific antibiotic for targeted treatment (based on test accuracy parameters for bacterial identification or AST)
- empiric antibiotics only were assumed to be used for people tested by a rapid test which was not able to identify a specific antibiotic for targeted treatment
- empiric followed by target antibiotics were assumed to be used by those tested with laboratory-based tests or point of care tests that were not rapid.

The EAG explained that the model would be able to estimate the proportion of patients assigned to empiric antibiotic treatment. Other potential antibiotic-related outcomes were identified at scoping and included measures of antibiotic use such as specific antibiotics based on World Health Organisation

AWaRe classification, changes to prescriptions, rate, dose and duration of antibiotics.

'Rapid' POCTs were defined by the EAG as those giving results in less than 40 minutes. The assumption that tests giving results within this time (and not those over this time limit) would be quick enough to influence initial prescribing decisions may not be the case if implemented in clinical practice.

The amount of antibiotic use would also depend on how many courses were prescribed. The EAG highlighted that no data to inform this parameter were identified. A back-up (delayed) prescription may be used in some case (a prescription with the assumption that it will not be dispensed immediately, but in a few days if symptoms worsen) which may also impact on the extent of antibiotic use. The EAG assumed that the probability of requiring more than 1 course of antibiotics would be higher if empiric antibiotics were used as there would be a higher probability of the first not targeting the correct bacteria.

Test accuracy for UTI detection, bacterial identification and AST can vary between tests, but there were limitations in identified accuracy data and a lack of accuracy data across all populations in the scope (see key issue 2). Find further detail about the EAG's model structure and assumptions in section 6.3.1 of the EAR. A list of key assumptions is described in table 17 of the EAR.

Questions for committee

- Is the approach explored by the EAG suitable to assess the impact of POCTs on antibiotic use? If not, could alternative modelling approaches be used, or potentially would direct data on antibiotic use be needed?

Key issue 4: Modelling impact of test use on health-related outcomes

Description of issue

The EAG explored using a linked evidence approach to assess the impact of tests on longer term health issues. There are a number of evidence gaps where parameters needed for this could not be identified. The approach was based on modelled use of antibiotics (described in key issue 3 above).

Background

Limited evidence was identified on the direct impact of tests on health-related outcomes. Two studies evaluating Flexicult human reported on duration of UTI or symptoms, patient enablement or health resource use. Neither study reported a statistically significant impact of test use.

In the EAG's proposed model, following testing and any antibiotic treatment, people with UTI can either have no complications (but be at ongoing risk of recurrent UTI) or have complications (pyelonephritis, kidney failure or sepsis), modelled as a one-off cost and quality of life decrement. The probability of complications was determined by whether a person received antibiotics or not (no treatment), and if so, which antibiotic type (targeted or empiric). How antibiotic use was modelled is described in the key issue 3. Find further detail on parameters in table 19 of the EAR.

The EAG commented that a lack of data on the benefit of targeted versus empiric antibiotic treatment limited modelling. It highlighted gaps in data for the following parameters (see also table 19 in the EAR):

- probability of sepsis/kidney failure/pyelonephritis on targeted treatment
- probability of sepsis/kidney failure on empiric treatment
- probability of sepsis/kidney failure on no treatment
- costs of treating sepsis and kidney failure

- short term QALY losses for kidney failure and sepsis
- QALY loss from antibiotic AE.

The EAG noted that it did not conduct a systematic literature review for the above parameters, and it is possible that evidence is available.

The EAG stated that testing people with recurrent UTI was modelled by the decision tree as it does not distinguish between first or repeat UTI. People who recover without developing pyelonephritis, kidney failure or sepsis are at risk of recurrent or chronic UTI. The EAG stated that the model currently assumes that patients are eventually cured of their UTI. It noted that the model does not explicitly distinguish between a chronic or acute UTI and people with chronic UTI would return to the decision tree for further testing. Find further detail about the EAG's model structure and assumptions in section 6.3.1 of the EAR.

People having antibiotics could have related adverse events. This would incur a QALY loss, but the EAG did not identify a source for this. The EAG noted that the probability of antibiotic related adverse events may vary by antibiotic. The model does not include long-term impact of unnecessary antibiotic prescription.

The EAG concluded that the limited evidence base limits the use of the model results for decision making. It recommended that further systematic reviews and discussion with clinical experts would be needed for a full economic analysis.

Question for committee

- Is the approach explored by the EAG suitable to assess the impact of POCTs on longer-term health outcomes?

4 Equality considerations

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others.

Key equality issue 1: POCTs may have a different impact on subgroups within the overall population

Description of issue

Data was not identified in all subgroup populations in the scope, and data generated in some groups may not give a good reflection of performance in others. There is a general lack of clinical effectiveness and accuracy data for the populations where dipstick testing is not recommended (men, adults over 65 and adults who are catheterised). There is also a lack of data for people with atypical symptoms or those who many struggle to communicate their symptoms. Ongoing or future studies may not include all groups highlighted in the scope.

Background

Women, pregnant women, older people, and people who are catheterised are more likely to develop a UTI. In adults, people with neurogenic bladder, diabetes, polycystic kidney disease or people who are immunocompromised have a higher risk of complicated UTIs. Dipstick tests that can be used to rule out UTIs are not recommended for men, adults older than 65 and adults who are catheterised. Tests that can more accurately rule out a UTI diagnosis for these groups could have a particular benefit in reducing unnecessary use of antibiotics and side effects resulting from these.

Autistic people, people with neurological disorders (for example dementia) and people who are frail may present with atypical symptoms or struggle to communicate their symptoms with healthcare professionals. People from minority ethnic family backgrounds may experience cultural barriers that may

stop them accessing healthcare for UTIs. Non-English speakers may also have trouble communicating their symptoms which may lead to delays in diagnosis and receiving effective treatment. Tests that can more accurately assess for UTIs may particularly benefit these groups.

Studies assessed the following specific populations defined in the scope: women (4 studies, not stratified on age), pregnant women (4 studies), children (4 studies) and those with catheters (1 study). There were no data on any of the other pre-specified subgroup populations of interest. The EAG stated that there were insufficient data to investigate whether accuracy varied across different populations. It remains unclear whether data from the included subgroups is a good representation of test accuracy across other subgroups included in the assessment. It may be the case that, as with dipstick testing, the test accuracy for UTI diagnosis could be different across different groups.

Question for committee

- Is there anything that would need to be considered in any recommendations for evidence generation or research to address potential equality issues?

5 Overarching issues for committee consideration

Approaches for decision making and recommendations are described in NICE's Early value assessment interim statement. Described below are some preliminary potential benefits and harms of using the technologies, it is anticipated these will be discussed, added to and expanded on in committee.

Potential benefits

- Any improvement in detecting UTIs (compared to current testing that may use no testing or dipstick as initial assessment) and selecting appropriate antibiotics, could reduce a person's symptoms and complications related to the infection. People with recurrent or chronic

UTI may experience symptoms over longer periods of time and a reduction in symptoms may be particularly beneficial for these groups.

- Tests that can more accurately rule out UTIs could reduce side effects from unnecessary antibiotic use. Quicker information on what antibiotics an infection is susceptible to may allow selection of an antibiotic with fewer side effects.

- Any improvements in prescribing could reduce the risk of antimicrobial resistance.

- The implementation of rapid POCTs could lead to reduced laboratory workloads and associated costs and reduced healthcare resource use associated with mismanaged UTIs.

Potential harms

- POCTs may increase primary care costs and resource use. Different POCTs may impact costs and resource use differently as test characteristics, outputs and costs vary.

- POCTs may also cost more than current practice, or require changes to practice, but lead to limited changes in antibiotic use (for example, if results aren’t available quickly enough to change prescribing).

- If delaying initial prescribing of antibiotics (for example, if the result takes longer than a typical GP appointment) this could mean symptoms go untreated for longer and lead to increased patient distress.

- Worse accuracy leading to missed UTIs could decrease antibiotic use but lead to longer suffering with symptoms. Increased overdiagnosis could lead to increased unnecessary use of antibiotics.

- POCTs may impact the amount of plastic used and waste created. Increases in waste production may also increase costs if disposing of materials needs extra equipment or time compared to current practice.

- Some of the included tests have not yet received regulatory approval, are not currently available to the NHS or do not have an available price.

6 Implementation

Potential changes to current care pathway

Changes may be needed to implement technologies into the current care pathway, in which assessment of symptoms with or without a dipstick test can be done quickly. The practicalities of who conducts the test, where the test is conducted, if there is space, and how results are communicated should be considered. If test results take longer than current GP appointments the practicalities of how and when prescriptions are issued and collected should also be considered. Clinical experts noted that electronic prescribing makes this process easier, but this would still be an additional step with risks that people may be lost to follow up.

Clinical experts commented that if the test takes longer to run than a consultation appointment with a prescriber, testing may need to be done in advance of this appointment so results can inform initial prescribing decisions.

Clinical experts also noted that national guidance may need to be updated if new technologies are adopted to ensure clarity on where they fit in the pathway and the impact on empirical prescribing.

Authors

Amy Barr

Topic lead

Emma McCarthy

Associate analyst

Thomas Walker

Technical adviser

Glossary

Antibiotic susceptibility testing

A test performed to determine which antibiotics most effectively treat a urinary tract infection.

Antimicrobial resistance

The loss of effectiveness of any anti-infective medicine, including antiviral, antifungal, antibacterial and antiparasitic medicines.

Back up (delayed) prescribing

A back-up (delayed) prescription is a prescription (which can be post-dated) given to a patient or carer, with the assumption that it will not be dispensed immediately, but in a few days if symptoms worsen.

Biomarker

A naturally occurring molecule, gene or characteristics that can be used to identify an infection or disease.

Complicated UTI

A urinary tract infection with an increased likelihood of complications such as persistent infection, treatment failure and recurrent infection.

Culture-based testing

A method of multiplying bacteria to establish the type and concentration of bacteria in a urine sample. It is typically done in a laboratory.

Dipstick test

A diagnostic test that is dipped into a urine sample and can detect nitrites, leukocytes and red blood cells to inform the likelihood of a urinary tract infection.

Empiric antibiotics

Broad spectrum antibiotics given to treat a suspected urinary tract infection.

Pathogenic bacteria

Bacteria that can cause a urinary tract infection.

Pyelonephritis

A urinary tract infection that affects the kidneys.

Sepsis

A potentially life-threatening condition that occurs when the body's response to an infection damages its own tissues.

Uncomplicated UTI

A urinary tract infection caused by typical pathogens in people with a normal urinary tract and kidney function, and no predisposing co-morbidities.