Diagnostic Assessment Report commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Care Excellence – Final Report

Title: Rapid Tests for Group A Streptococcal infections in people with sore throat

Name of External Assessment Group (EAG) and project lead:

Produced by: Warwick Evidence

Lead author: Hannah Fraser

Co-authors: Daniel Gallacher

Felix Achana

Rachel Court

Sian Taylor-Phillips

Chidozie Nduka

Christopher Stinton

Bex Willans

Paramjit Gill

Hema Mistry

Correspondence to: Dr Hema Mistry

[redacted]

Tel: [redacted]

Email: [redacted]

Date completed: 20 May 2019

Source of funding:
This report was commissioned by the NIHR HTA Programme as project number NIHR127666.

Sian Taylor-Phillips, Christopher Stinton and Hannah Fraser are partly supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care West Midlands at the University Hospitals Birmingham NHS Foundation Trust. Sian Taylor-Phillips is funded by an NIHR Career Development fellowship.

Declared competing interests of the authors:
All authors have no conflicts of interest.

Acknowledgements:
The authors would like to thank Dr Mitul Patel, Professor Michael Moore, Dr Derren Ready, and Mr Mohammed Rafiq, for their expert clinical advice on searching the literature. We are also grateful to Dr Mitul Patel for expert clinical advice during the clinical effectiveness reviewing and data extraction stages and to Dr Derren Ready for comments on the draft report. The authors would like to thank Karoline Munro for helping with the coordinating of the report.

Rider on responsibility for report:
The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:
Rapid Tests for Group A Streptococcal infections in people with sore throat. Diagnostic Assessment Report commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Care Excellence.

Contributions of authors:
Rachel Court (Information specialist) developed the search strategy and undertook searches. Hannah Fraser (Research Associate), Sian Taylor-Phillips (Associate Professor), Chidozie Nduka (Senior Research Fellow), Christopher Stinton (Senior Research Fellow) and Bex Willans (Academic Foundation 2 doctor) conducted the clinical effectiveness systematic review, this included: screening and retrieving papers, assessing against the inclusion criteria, appraising the quality of papers and abstracting data from papers for synthesis. Daniel Gallacher (Research Fellow) conducted the meta-analyses and contributed to both the clinical and cost-effectiveness sections. Felix Achana (Senior Research Fellow) contributed to the cost-effectiveness review and undertook the health economic modelling. Paramjit Gill (Professor of General Practice) provided clinical guidance and helped

develop the model structures. Hema Mistry (Associate Professor) provided project management, conducted the cost-effectiveness review and supervised the economic analysis. All authors were involved in writing draft and final versions of the report.

Contents

Table of Tables

Table of Figures

List of abbreviations

AMR	Antimicrobial Resistance
CEA	Cost-Effectiveness Analysis
CEAC	Cost-Effectiveness Acceptability Curves
CFU/ml	Colony Forming Units per millimeter
CHEERS	Consolidated Health Economic Evaluation Reporting Standards
CI	Confidence Interval
CRD	Centre for Reviews and Disseminations
DAC	Diagnostics Advisory Committee
DARE	Database of Abstracts of Reviews of Effects
EAG	External Assessment Group
ECCMID	European Congress of Clinical Microbiology and Infectious Diseases
EMBASE	Excerpta Medica database
FDA	Food and Drug Administration
FN	False Negative
FP	False Positive
GABHS	Group A beta-haemolytic Streptococcus
GAS	Group A Streptococcus
GP	General Practitioner
HRQoL	Health Related Quality of Life
HTA	Health Technology Assessment
ICER	Incremental Cost-Effectiveness Ratio
INAHTA	International Network of Agencies for Health Technology Assessment
McIsaac	Modified Centor
MFR	Manufacturer
NA	Not Applicable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NHS EED	National Health Service Economic Evaluation Database
NIHR	National Institute for Health Research
NPV	Negative Predictive Value
NR	Not Reported
PCR	Polymerase Chain Reaction
PHE	Public Health for England
PPV	Positive Predictive Value
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSA	Probabilistic Sensitivity Analysis
PSS	Personal Social Services
QALD	Quality-Adjusted Life Days
QALY	Quality-Adjusted Life Years
QoL	Quality of Life
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
RADT	Rapid Antigen Detection Test
RCTs	Randomised Controlled Trials
RePEc	Research Papers in Economics

ROB	Risk of Bias
RTI	Respiratory Tract Infection
ScHARRHUD	School of Health and Related Research Health Utilities Database
STDR	Sore Throat Decision Rules
Strep A	Group A Streptococcus
TN	True Negative
TP	True Positive
UK	United Kingdom

Abstract

Background

Sore throat is a common condition caused by an infection of the airway. Most cases are of a viral nature, however, a substantial number of these infections may be caused by the Group A Streptococcus (Strep A) bacteria. Most sore throats, including viral and bacterial sore throat infections, resolve spontaneously within a few weeks. Point-of-care testing in primary care has been recognised as an emerging technology for aiding targeted antibiotic prescribing in cases of sore throat.

Objective

To systematically review the evidence for 21 point-of-care tests for detecting Strep A bacteria and develop a *de novo* economic model to compare the cost-effectiveness of point-of-care tests in conjunction with clinical scoring tools compared to clinical scoring tools alone in England and Wales.

Review methods

Multiple electronic databases were searched from inception to March 2019. Eligible studies included: people above the age of 5 years presenting with symptoms of a sore throat; comparing point-of-care testing with antibiotic prescribing decisions using clinical scoring tools for Strep A; test accuracy; or cost-effectiveness outcomes. Quality assessment of eligible studies used tailored Quality Assessment of Diagnostic Accuracy Studies – 2 (QUADAS-2) and Consolidated Health Economic Evaluation Reporting Standards (CHEERS). Meta-analysis of sensitivity and specificity were performed for tests with sufficient data.

A *de novo* decision tree model, compared patients managed with point-of-care testing alongside clinical scoring tools with clinical scoring tools alone. Economic models included primary care and hospital management of patients with suspected Group A Strep A bacteria. The model estimated costs (in 2017/2018 prices) and quality-adjusted life years (QALYs) from an National Health Service (NHS) and personal social services (PSS) perspective.

Results

The searches identified 5,919 studies of clinical effectiveness and 6,980 studies of cost-effectiveness, of which included 38 and three studies, respectively. 26 full text articles and abstracts reported on the test accuracy of point-of-care tests compared to biological culture. In the population of interest (patients with Centor/McIsaac scores ≥3 or FeverPAIN ≥4) point estimates were 82.9% to 94.6% for sensitivity and 84.9% to 99.1% for specificity. No information was identified in the elderly population or pharmacy setting, or matching the proposed pathway of care at the recommended National Institute for Health and Care Excellence (NICE) thresholds (Centor/McIsaac > 2). It was not

possible to identify which test is the most accurate due to the paucity of evidence. There was considerable heterogeneity, even for studies performing the same point-of-care test, suggesting that is unlikely any single study will have accurately captured a tests true performance. There is some randomised control trial (RCT) evidence to suggest the use of rapid antigen detection tests may help reduce antibiotic prescribing rates, but there was no evidence on the effect of using molecular technologies. Sensitivity and specificity estimates for each test in each age group and care setting combination were obtained from published literature where available, or from manufacturer documentation if no other sources were available, using meta-analyses where appropriate. Any apparent differences in test accuracy may not be attributable to the tests, and may have been caused by known differences in the studies, latent characteristics, or chance.

Thirteen of the 21 tests for which relevant data were available in final economic modelling; however, there was considerable uncertainty about the cost-effectiveness of the different point-of-care tests for suspected Strep A in both primary and secondary care settings. Uncertainties in the model include parameter inputs and assumptions that increase the cost of testing, and the penalty for antibiotic over-prescriptions. While there is potential for cost-effectiveness in both primary and secondary care settings, key parameter inputs and modelling assumptions need to be confirmed and model findings remain uncertain.

Conclusions

The systematic review and the cost-effectiveness models identified uncertainties around the adoption of point-of-care tests within primary and secondary care settings. Although sensitivity and specificity estimates are promising, we have little information to establish the most accurate point-of-care test.

Future work

Further research is needed to understand the test accuracy of point-of-care tests within the proposed NHS pathway and within comparable settings and patient groups. Future work which considers head-to-head test accuracy studies or randomised controlled trials using multiple point-of-care tests in relevant populations would provide relevant comparator information and help to determine the value of point-of-care testing.

Funding details

This report was commissioned by the NIHR HTA Programme as project number **NIHR127666**.

Word count: 713

Scientific Summary

Background

Sore throat is a common condition caused by an infection of the airway; clinical descriptions of acute sore throat include acute pharyngitis and tonsillitis, both infections of the upper respiratory airway affecting the mucosa. Most cases are viral, however, a substantial number of these infections may be caused by the Group A Streptococcus (GAS or Strep A) bacteria. Most sore throats, resolve spontaneously within a few weeks. An analysis of United Kingdom (UK) primary care usage data identified a reduction in antibiotic prescribing in the UK between 1993 and 2001 for diagnosed episodes of sore throat. Despite this reduction, sore throat and other respiratory tract infections remain a common reason for primary care usage.

Point-of-care testing (rapid antigen detection and molecular tests) in primary care has been recognised as an emerging technology for aiding targeted antibiotic prescribing in cases of sore throat. These tests are intended to be used in addition to clinical scoring systems, such as FeverPAIN and Centor. The purpose of these tests is to increase diagnostic confidence of a suspected Group A Strep infection and guide antimicrobial prescribing decisions in people presenting with an acute sore throat and to contribute to improving antimicrobial stewardship. The tests may be suitable for use in all settings where patients may present with an acute sore throat (Centor scores ≥3, FeverPAIN scores ≥4); this includes both primary and secondary care, and community pharmacies.

The protocol of the review is registered with PROPSPERO as CRD42018118653.

Decision question

The decision problem for this assessment is:

- What is the clinical and cost-effectiveness of rapid antigen detection and molecular tests in patients with high clinical scores (Centor scores ≥3, FeverPAIN scores ≥4), compared to the use of clinical scoring tools alone, for increasing the diagnostic confidence of suspected Group A Streptococcal (GAS or Strep A) infection in people who present with an acute sore throat in primary and secondary care?

Objectives

To systematically review the evidence for the clinical effectiveness of the rapid antigen detection and molecular tests; systematically review existing economic evaluations; and develop a *de novo* economic model to assess the cost-effectiveness of rapid antigen detection and molecular tests in conjunction with clinical scoring tools compared to clinical scoring tools alone in England and Wales.

Methods

Clinical effectiveness and cost-effectiveness systematic reviews

Multiple electronic databases were searched from inception to March 2019 for both the clinical effectiveness and the cost-effectiveness reviews. Supplementary searches were used to identify additional published and unpublished studies. Reference lists of included studies and information provided by the manufacturers of the intervention tests were checked for additional eligible studies.

Two reviewers independently screened and assessed titles and abstracts of all records. Studies were included according to the following criteria:

- Population - People aged 5 years and above presenting with symptoms of an acute sore throat.
- Intervention – Point-of-care tests for Strep A (including rapid antigen detection tests and molecular tests), preferably in those identified as high risk.
- Comparator - Antibiotic prescribing decisions using clinical scoring tools for Strep A such as FeverPAIN or Centor/modified Centor (McIsaac) alone.
- Outcomes – any patient-related outcome, test accuracy or performance, prescribing behaviour, cost-effectiveness estimates.
- Study design - Clinical test accuracy studies that compare the index tests and FeverPAIN/Centor/McIsaac scores to throat swab culture. Studies of head-to-head comparisons of rapid tests were eligible for inclusion if test accuracy statistics were reported for each test. For prescribing behaviour any study design which compared the index test to throat swab culture and/or clinical scores (FeverPAIN/Centor/McIsaac). For cost-effectiveness, any full economic evaluations (or economic models) reporting both cost and outcome estimates.
- Healthcare setting - Primary care (GP clinics, community pharmacies and walk-in centres) and secondary care (urgent care/walk-in centres and emergency departments) settings.

Data were extracted by one reviewer and checked by a second reviewer. Discrepancies were resolved via discussion or by a third reviewer. Evidence were synthesised using narrative review and statistical methods were appropriate. Meta-analyses were undertaken in Stata version 15.

Study quality assessment of eligible studies was undertaken using recognised checklists (tailored Quality Assessment of Diagnostic Accuracy Studies – 2 (QUADAS-2), Cochrane Risk of Bias, JBI Critical Appraisal checklist for analytical cross-sectional studies, and Consolidated Health Economic Evaluation Reporting Standards (CHEERS)).

Cost-effectiveness model

A *de novo* decision tree model was built in Microsoft Excel to assess the cost-effectiveness of rapid antigen detection and molecular tests, in conjunction with clinical scoring tools, compared to the use of clinical scoring tools alone. The base-case economic model included adult patients seen in primary care with suspected GAS infection. The base-case model was adapted to look at the following subgroups with suspected GAS infection: adult patients seen in the hospital, children seen in primary care and children seen in the hospital. The data for the model included prevalence information from the systematic clinical effectiveness review, published literature and expert opinion. The model estimated the mean total costs and mean total QALYs for each rapid antigen detection and molecular tests over a one-year time horizon and adopted a National Health Service (NHS) and personal social services perspective. Costs were in 2017/2018 prices. No discounting of costs and outcomes were performed. Outcomes are reported as incremental cost-effectiveness ratios expressed in terms of cost per quality-adjusted life year gained. A number of sensitivity analyses were undertaken. Probability sensitivity analyses was also undertaken (1,000 model runs).

Results

The searches identified 5,919 studies of clinical effectiveness and 6,980 studies of cost-effectiveness, of which we included 38 and three studies, respectively.

The systematic review of clinical-effectiveness studies identified 38 studies that used the point-of-care tests identified in the National Institute for Health and Care Excellence (NICE) scope and biological culture and/or clinical score as a comparator. These comprised of 26 full text articles, 3 abstracts, 5 manufacturer's submissions (submitted directly to NICE in response to a request for information) and 4 Food and Drug Administration (FDA) documents. There were 26 studies (23 full text articles and 3 abstracts) which reported test accuracy data. The methodological quality of the included studies was poor. In particular, in 65.4% (17/26) studies it was unclear whether the sample was consecutive or convenience. Convenience samples may not provide a true representation of the prevalence of GAS. There was judged to be a high level of bias concerning methods of patient selection. Overall the findings reveal variations in the sensitivity (67.9% to 100%) and specificity (73.3% to 100%) estimates of point-of-care tests. These point estimates were 82.9% to 94.6% for sensitivity and 84.9% to 99.1% for specificity in high-risk populations, including patients with Centor/McIsaac scores > 2, representing the population of interest. These estimates do not account of any of the unpublished manufacturer submissions.

Direct comparison to sore-throat clinical scoring tools revealed that point-of-care tests were generally more specific. However, one methodological limitation concerns the varying way clinical scoring tools have been implemented across the included studies. For instance, different studies apply different clinical score cut-offs when recruiting patients. None of these studies matched the proposed

pathway of care and treatment for patients with acute streptococcal pharyngitis, which would entail evaluating the test accuracy of a combined strategy of sore throat clinical scores at the recommended NICE thresholds (Centor/McIsaac ≥3 or FeverPAIN ≥4) and point-of-care tests. This limitation potentially holds important economic implications as attempts to model this proposed pathway may not be informed by the availability of empirical data. In addition, the overrepresentation of the TestPack Plus Strep A test relative to other point-of-care tests, as well as the overlap of patients across different age groups potentially raises applicability concerns in the economic model.

Data for test accuracy were sparse for each combination of test, population and setting. There were very few head-to-head (direct) comparison studies between index tests. It was not possible to identify which test is the most accurate due to the lack of available evidence. There was a large degree of heterogeneity among results for studies performing the same rapid test. Where a test is reported in several studies its accuracy may appear lower compared to tests reported in only a single study, particularly those at high risk of bias or unpublished methods, so we report on volume and quality of data available as well as accuracy estimates. The heterogeneity introduced by the differing characteristics of the studies, further confounded attempts to produce meaningful estimates of test performance, such as care setting, age group, throat score restriction and disease prevalence. Due to the potential heterogeneity, estimates for the sensitivity and specificity of each test were stratified by age group, throat score and care setting, though a lack of evidence meant generalisations had to be made for the majority of estimates.

Of the scoped secondary outcomes, our search only identified studies discussing antibiotic prescribing rates and appropriateness (n=12). There is some RCT evidence to suggest the use of rapid antigen detection tests may help reduce antibiotic prescribing rates, but there was no evidence on the effect of using molecular technologies. If a test was proven to be extremely accurate, then it is plausible that clinical staff would trust the outcomes. No information was found on number of appointments required per episode, morbidity, mortality, onward transmission of infection, health-related quality of life, patient satisfaction with the test or healthcare professional satisfaction with the test.

Cost-effectiveness

The systematic review of cost-effectiveness studies identified three studies that used the rapid antigen detection tests as identified in the NICE scope and were classed as economic evaluations. Two studies had some notable limitations and could not be fully data extracted. The one study that allowed a full data extraction, was classed as a high quality economic evaluation when checked against the CHEERS reporting tool.

Thirteen of the twenty-one tests listed in the NICE scope had relevant data on test accuracy and costs, to be included in the final economic modelling. In the base-case analysis, which included adult patients seen in primary care with suspected GAS infection, the economic model found considerable uncertainty about the cost-effectiveness of the different point-of-care tests for suspected GAS infection. This finding was also seen in the other economic models which were adapted (adult patients seen in the hospital, children seen in primary care and children seen in the hospital). Important uncertainties in the model include parameter inputs and assumptions that increase the (i) cost of testing (acquisition cost of test, additional clinician time for administering and processing test results, and cost of confirmatory throat culture for those testing negative) and ii) penalty for antibiotic over-prescription/unnecessary antibiotic use (acquisition cost of antibiotic and probabilities for penicillin-induced anaphylaxis and rash).

Discussion and Conclusions

Main findings

The systematic review and cost-effectiveness model identify uncertainties around the adoption of point-of-care tests within the NHS. The available evidence is heterogeneous in populations studied, design, methods and analysis. Although sensitivity and specificity estimates are promising, we have little information on the best point-of-care test to use. While there is potential for the point-of-care tests to be cost-effectiveness in both primary and secondary care settings, key parameter inputs and modelling assumptions need to be confirmed and model findings remain uncertain.

Strengths and limitations

Strengths of the work include a robust and comprehensive systematic review (literature search, data extraction and analysis) strategy and the building of a *de novo* decision tree model to assess cost-effectiveness.

No studies on point-of-care use in a pharmacy setting or in the elderly population were retrieved. Additionally, no study matched the proposed pathway of care and treatment for patients with acute streptococcal pharyngitis, which would entail evaluating the test accuracy of a combined strategy of sore throat clinical scores at the recommended NICE thresholds (Centor/McIsaac ≥3 or FeverPAIN ≥4) and point-of-care tests in the age groups defined in the scope.

Although the economic model represented the clinical care pathway in the NHS, practice and management will vary from site to site (within and across both primary care and secondary care settings). The modelling may have underestimated the costs as we did not take into account the

different strains of GAS which may have influenced test performance and alter the profile of complications, seasonality of GAS infection and the onward transmission of the infection.

Implications for healthcare

Our findings indicate that point-of-care testing were not cost-effective within the current thresholds and should be viewed cautiously by clinicians and policy makers, in view of the poor quality of the evidence available to us. Healthcare professionals should be mindful of the potential variation in performance of the different testing methods and strategies in their day-to-day practice.

Research priorities

Further research is needed to understand the test accuracy of point-of-care tests within the proposed NHS pathway and within comparable settings and patient groups. Future work which considers head-to-head test accuracy studies or randomised controlled trials using multiple point-of-care tests in relevant populations would provide relevant comparator information and determine the value of point-of-care testing.

Word count: 2,128

Plain English Summary

Sore throat is a common condition caused by an infection of the airway. Most cases are of a viral nature, however, a substantial number of these infections may be caused by the Group A Streptococcus (Strep A) bacteria. Most sore throats, including viral and bacterial sore throat infections, resolve spontaneously within a few weeks. Currently, National Institute for Health and Care Excellence (NICE) guidance recommends the use of clinical scoring tools to identify patients for whom antibiotic treatment is appropriate. In spite of this recommendation, there is a huge variation in antibiotic prescription for sore throat in general practice.

Ideally, a throat swab culture should be undertaken to identify the organism causing the infection in cases where diagnosis is uncertain. However, this takes time, causing potential delays in administering the correct treatment.

Our review considered evidence for the test accuracy and cost-effectiveness of 21 point-of-care tests for detecting Strep A bacteria. We built an economic model, predicting costs and benefits for adults and children in a primary care or hospital setting, to help inform how best to manage patients. The findings will support NICE to make recommendations about the use of point-of-care tests for detecting Strep A bacteria in the National Health Service (NHS) in England and Wales.

The clinical effectiveness review found 38 relevant studies, of these 26 reported on the accuracy of the point-of-care tests. These studies found wide variation in the accuracy of the test. The quality of the evidence was weak and there was little information on all 21 tests. As the studies were all so different, it is not possible to identify which test is the most accurate.

The economic model found considerable uncertainty about how costs and benefits would change if point-of-care tests were introduced in both primary care or in hospital settings. Further research is needed to see whether point-of-care testing provides value for money.

Word count: 312

1 Introduction

1.1 Description of the health problem

Sore throat is a common condition;[1, 2] clinical descriptions of acute sore throat include acute pharyngitis and tonsillitis, both infections of the upper respiratory airway affecting the mucosa.[3, 4] In a Scottish survey, 31% of respondents reported having experienced a severe sore throat in the past 12 months.[1] Symptoms of sore throat include pain in the throat and may also include fever or headache, however not all patients will require or seek medical advice and/or treatment for these symptoms. An analysis of United Kingdom (UK) primary care usage data identified a reduction in the UK between 1993 and 2001 for diagnosed episodes of sore throat.[2] This finding may suggest changes in patient behaviour regarding self-care, changes in general practitioner (GP) diagnosis and recording of sore throat or actual change in prevalence of sore throat, although there is no evidence to support this theory. Despite this reduction, sore throat and other respiratory tract infections remain a common reason for primary care usage; a quarter of the population will visit their GP because of a respiratory tract infection (RTI) each year.[5]

In the UK, diagnosis of sore throat is currently based mainly on clinical assessment and it is recommended by National Institute for Health and Care Excellence (NICE) that the FeverPAIN or Centor criteria are also used. The FeverPAIN and Centor tools were designed to predict Strep A (Centor, FeverPAIN), C (FeverPAIN) and G (FeverPAIN),[6, 7] and have been proposed as methods by which clinicians can identify which patients are most likely to benefit from antibiotic use for sore throat.[8] This is because sore throat is often a self-limiting illness, most cases have a viral aetiology and therefore antibiotics would not be an effective treatment in these instances. Additionally, as antibiotics only reduce duration of symptoms by a very short period this must be traded off against the side effects. Around 5-17% of sore throats are due to a bacterial infection, typically Group A beta-haemolytic streptococcus (GABHS), also known as 'Streptococcus pyogenes' or 'Group A streptococcus' or 'GAS' or 'Strep A'.[5, 8] Expert advice suggests that bacterial sore throat can also be caused by Group C and Group G streptococci, however GAS is thought to account for around 80% of bacterial throat infections, and Group C and G Strep for around 20%. Most cases of Strep A infection

resolve without complications and in fact, many people carry the bacteria without experiencing illness. Despite these factors, most patients presenting with sore throat in the UK will be given antibiotics in primary care.[9, 10] Although rates of antibiotic prescribing for sore throat declined between 1993 and 2001, more recent prescribing data for 2011 remains close to the 2001 figure, with median practice prescribed antibiotics for sore throat at 60%.[2, 9] RTIs, which include sore throats, account for a large proportion of antibiotic use in general practice in the UK (approximately 60%).[8]

There are clinical and epidemiological reasons why clinicians may prescribe antibiotics for sore throat in the absence of microbiological confirmation; the first is practical. The current reference standard, culture of the bacteria grown from a throat swab, take over 18 hours for a result.[11] Where clinicians suspect GAS infection based on clinical judgement and use of the FeverPAIN or Centor criteria, there is an opportunity to reduce the risk or harm caused by complications such as tonsillitis, pharyngitis, scarlet fever, impetigo, erysipelas (an infection in the upper layer of the skin), glomerular nephritis, rheumatic fever, cellulitis and pneumonia. Some vulnerable patient groups such as those who are immunocompromised are at higher risk of developing invasive GAS infection. To prevent onward transmission, current Public Health for England (PHE) guidance on invasive Strep A infection management indicates use of antibiotics in close contacts of people who have invasive GAS infection if they have symptoms of Strep A infection themselves, such as sore throat or are in a particular risk group or setting.[11] Although these factors must be considered in understanding the reasons for use of antibiotics to treat sore throat in the absence of more accurate diagnosis, another factor that impacts use is patient demand. Although patient attendances for minor ailments at GP surgeries has reduced, when patients do visit their GP there is an expectation of intervention and this is increasingly the case.[11] Furthermore, RTIs account for a high proportion of working days lost in the UK, in 2016 almost a quarter (24.8%, 34 million days), so ensuring patients receive appropriate and timely treatment also has an economic impact on the economy and on patients.[12] This rationale and demands need to be balanced however with the aforementioned statistics regarding the low prevalence of bacterial infection in sore throat and the risk of antimicrobial resistance (AMR).

Overuse or inappropriate use of antibiotics can lead to bacteria developing resistance, leading to an emergence of multi-drug resistant pathogens, which are increasingly difficult to treat. AMR could contribute to an estimated 10 million deaths every year globally by 2050 and a global productivity cost of £66 trillion.[11] In response to this threat, 'antimicrobial stewardship' has been a central strategy adopted by the Chief Medical Officer and NICE.[11, 13] Point-of-care testing in primary care has been recognised as an emerging technology for aiding targeted antibiotic prescribing in cases of sore throat, by supporting clinicians with diagnosis and to communicate appropriate use of antibiotics to patients.[14] Several technologies have been developed for point-of-care testing in primary care for appropriate administration of antibiotics to those that would benefit and to prevent delay and associated complications.

The NICE Diagnostic Advisory Committee (DAC) is tasked with providing guidance to the National Health Service (NHS) about the use of point-of-care tests for the detection of GAS in sore throat infections. To inform the DAC, the external assessment group (EAG) has provided this assessment of the clinical accuracy and cost-effectiveness of point-of-care tests for the detection of GAS as a replacement or adjunct for standard assessment procedures. The potential value of the point-of-care tests is in rapidly determining the presence and nature of a bacterial infection.

1.1.1 Aetiology and pathology

Most sore throats are caused by an infection, mainly viral, and so are typically spread from person to person via respiratory droplets; non-infectious causes are uncommon. [3] In the case of infectious causes, viruses, bacteria or fungi invade the upper respiratory mucosa, causing a local inflammatory response.[4] Complications associated with sore throat caused by infection are rare, however GAS infection has a small risk of the following complications:[3]

- Otitis media
- Acute sinusitis
- Peri-tonsillar abscess
- Rheumatic fever and post-Streptococcal glomerulonephritis are also complications associated with Strep A throat infection, however these are extremely rare in developed countries.

- invasive GAS, if the bacteria move from the throat into a sterile body site (which can lead to severe infections, sepsis and Streptococcal toxic shock syndrome).

Children are most likely to carry or be infected by GAS, however people aged over 65 or those whose immune system is compromised (e.g. people living with HIV, diabetes, heart disease, cancer, using high dose steroids or intravenous drugs) are at higher risk of developing invasive Strep A infection.[15]

Fusobacterium necrophorum infection affecting the pharynx or tonsils can (very rarely) lead to Lemierre disease (sepsis and jugular vein thrombosis).[3]

1.1.2 Diagnosis and care pathway

Figure 1 depicts the care pathway for assessing and treating a sore throat as outlined in the NICE's antimicrobial prescribing guidance on sore throat (NG84).[5] Most uncomplicated sore throats are managed without seeking medical advice and will tend to resolve within one week.[10] Suggested conservative measures include simple analgesia, maintaining hydration, salt gargling and throat lozenges. In selected cases where a GP, or a pharmacist, or a healthcare practitioner in the secondary care, such as in Accident and Emergency, feels that the patient may benefit from antibiotics, the prescriber should apply either the FeverPAIN or Centor scores to guide their decision-making. The NICE antimicrobial prescribing guideline on acute sore throat does not make any recommendations about using point-of-care tests or throat cultures to confirm GAS infection.[5]

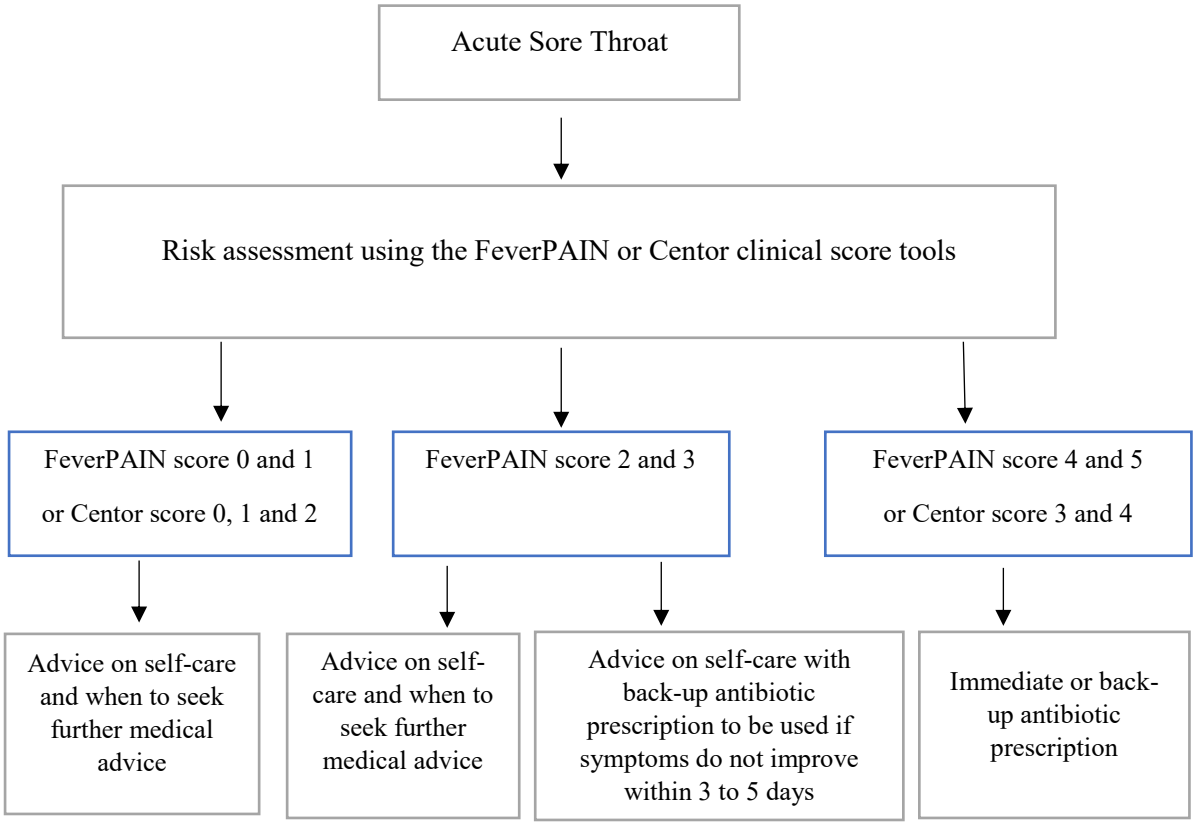


Figure 1: Diagnostic and care pathway for managing acute sore throat in patients who are not at high-risk of complications

1.1.3 Significance to the NHS and current service cost

The significance of sore throat and inappropriate use of antibiotics to the NHS broadly falls into two categories; the first is associated with healthcare use directly due to sore throat and the second is the impact of inappropriate use of antibiotics contributing to AMR.

RTIs, including sore throat, account for a large proportion of primary care use and antibiotic prescribing.[10] However, there is already evidence that the majority of patients prefer to self-medicate minor ailments such as sore throat where they feel able to do so.[1, 2, 13] For example, a visit to the GP practice for a diagnosis and treatment for sore throat, incurs the cost of the visit to a GP practice and any treatment prescribed. In addition, in the current system where GPs can use the FeverPAIN or Centor criteria to inform antibiotic prescribing, there is the potential cost of additional healthcare use for patients whose condition does not improve or who develop complications due to ineffective or no treatment being prescribed. The risk of

complications however is low and current prescribing activity suggests overuse rather than underuse of antibiotics for sore throat. Another cost associated with the current system is laboratory costs where the reference standard for diagnosis is used, mainly throat swabs sent for culture.

Although these costs and the impact of minor ailment use on the NHS is a key consideration, the primary aim of the intervention being considered is to reduce inappropriate antibiotic prescribing. Doing so could support a reduction in promotors of AMR. The main antibiotic prescribed by general practice is penicillin and these are the first line treatment currently recommended by NICE for suspected GAS throat infection.[8, 14] Across Europe, an estimated 25,000 people die each year as a result of hospital infections caused by the five most common resistant bacteria and a parliamentary report estimated the annual cost to the NHS to be £180million per year. [16] While it is often possible at present to use alternative treatments to treat resistant infection, costs of treatment and risk of mortality are likely to approximately double for a resistant infection.[16] One study investigating the cost of a 10-month outbreak of a type of antibiotic-resistant bacteria (carbapenemase-producing enterobacteria) found the total cost to be close to £1,000,000. The main cost was missed revenue from cancellation of planned surgical procedures due to ward closures and lack of bed space. Other costs were associated with additional staff time, increased length of patient stay in hospital, screening, bed and ward closure, contact precautions, anti-ineffective costs, HPV decontamination, ward-based monitors.[16] In addition to healthcare costs and risk of litigation associated with AMR related harm, there is a wider societal cost to lost productivity and reduced quality of life for patients suffering the effects of AMR infections.

1.2 Clear definition of interventions

There are rapid tests for the Group A Streptococcus bacteria, which are intended to be used in addition to clinical scoring systems, such as FeverPAIN and Centor. The purpose is to increase diagnostic confidence of a suspected Strep A infection and guide antimicrobial prescribing decisions in people presenting with an acute sore throat and to contribute to improving antimicrobial stewardship. The tests may be suitable for use in all settings where patients may present with an acute sore throat; this includes both primary and secondary care, and community pharmacies.[11]

Twenty-one rapid tests for GAS detection are available. The tests use either immunoassay detection methods (rapid antigen detection tests) or molecular methods (polymerase chain reaction [PCR] or isothermal nucleic acid amplification). The tests listed below were identified from the NICE scope on point-of-care testing in primary care for Strep A infection in sore throat.

1.3 Comparative technical overview of the point-of-care tests for Strep A

Seventeen rapid antigen detection tests were identified, and their product properties are summarised below (see Table 1). For each test, the limit of detection has been defined as the lowest concentration of Strep A in a sample that can be distinguished from negative samples. Of these, 16 tests use lateral flow techniques (also known as immunochromatographic or immunofluorescent assays), and one test is a turbidimetric immunoassay.

The lateral flow (immunochromatographic and immunofluorescence) tests require a throat swab which is typically placed into a specimen extraction tube and mixed with reagents to extract the sample from the swab. The swab is discarded and then either a test strip is immersed in the extracted solution or drops of the extracted solution are added to the sample well of a test cassette. The sample then migrates along the test strip or cassette, with any GAS antigens present in the sample binding to immobilised Strep A antibodies in the test strip or cassette. When Strep A is present at levels above the detection limit of the test, a line appears in the test line region of the strip or cassette. A control line shows technical success of the test. Results should be discarded when the control line indicates that the test has failed (that is, no line appears in the control line region). Depending on the technology, the results are read by either visual inspection or by using an automated test reader device.

The turbidimetric immunoassay has similar sample collection and extraction steps to the lateral flow tests, but the extracted solution is placed into a cuvette which is prefilled with reagents. This contains rabbit anti-Strep A antibodies which bind to GAS antigens present in

the sample. The Quikread go instrument measures the absorbance of each cuvette and converts the absorbance value into a positive or negative result.

Several of the companies recommend that negative rapid antigen detection test results are confirmed by microbiological culture of a throat swab.

Table 1: Rapid antigen detection tests – product descriptions and properties from manufacturers data

Product	Test format and supply	Method	Limit of Detection	Description of results	Time to result (minutes)[a]
Clearview Exact Strep A cassette* (Abbott)	25 individually pouched test cassettes	Lateral flow (immunochromatographic)	$5x10^4$ organisms/test	Positive results are indicated by 2 lines: one in the control region (C) and the other in the test region (T). Read by visual inspection	5
Clearview Exact Strep A dipstick – test* strip (Abbott)	25 test kits Dipstick	Lateral flow (immunochromatographic)	$5x10^4$ organisms/test	Positive results are indicated by 2 lines: one in the control region (C) and the other in the test region (T). Read by visual inspection	5
BD Veritor Plus system group A Strep Assay - cassette (Beckton Dickinson)	30 test kits Test cassette	Lateral flow (immunochromatographic)	Strain 12384: $1x10^5$ CFU/ml Strain 19615: $5x10^4$ CFU/ml Strain 25663: $2x10^5$ CFU/ml	Analysed by a BD Veritor system analyser module. Results are displayed visually	5
Strep A rapid test - cassette (Biopanda Reagents)	20 test cassettes	Lateral flow (immunochromatographic)	$1x10^5$ organisms/swab	Positive results are indicated by 2 lines: one in the control region (C) and the other in the test region (T). Read by visual inspection	5
Strep A rapid test – test strip (Biopanda Reagents)	No information provided	Lateral flow (immunochromatographic)	$1x10^5$ organisms/swab	Positive results are indicated by 2 lines: one in the control region (C) and the other in the test region (T). Read by visual inspection	5
NADAL Strep A - test strip (nal von minden GmbH)	40 test strips including controls, 50 test strips (tube) including controls, as well as positive and negative control vials.	Lateral flow (immunochromatographic)	$1.5x10^5$ organisms/swab	Positive results are indicated by 2 lines: one in the control region (C) and the other in the test region (T). Read by visual inspection	5
NADAL Strep A - cassette (nal von minden GmbH)	20 test cassettes including controls as well as positive and negative control vials	Lateral flow (immunochromatographic)	$1.5x10^5$ organisms/swab	Positive results are indicated by 2 lines: one in the control region (C) and the other in the test region (T). Read by visual inspection	5

Product	Test format and supply	Method	Limit of Detection	Description of results	Time to result (minutes)[a]
NADAL Strep A plus - cassette (nal von minden GmbH)	20 pack cassettes including controls and 5 pack cassettes including controls	Lateral flow (immunochromatographic)	$1.5x10^5$ organisms/swab	Positive results are indicated by 2 lines: one in the control region (C) and the other in the test region (T). Read by visual inspection	5
NADAL Strep A plus - test strip (nal von minden GmbH)	40 test strips	Lateral flow (immunochromatographic)	$1.5x10^5$ organisms/swab	Positive results are indicated by 2 lines: one in the control region (C) and the other in the test region (T). Read by visual inspection	5
NADAL Strep A scan test - cassette (nal von minden GmbH)	20 pack cassettes including controls	Lateral flow (immunochromatographic)	$1.5x10^5$ organisms/swab	Extracted solution is placed into the test cassette, with the Colibri placed on top. Analysed using a Colibri reader and Colibri USB and software	5
OSOM Strep A test – test strip (Sekisui Diagnostics)	50-test pack	Lateral flow (immunochromatographic)	Not known	Positive results are indicated by 2 lines: one in the control region (C) and the other in the test region (T). Read by visual inspection	5
QuikRead Go Strep A test kit (Orion Diagnostica)	50 tests including controls	Turbidimetric immunoassay	$7x10^4$ CFU/swab	Analysed using the QuikRead Go instrument	<7
Alere TestPack Plus Strep A - cassette (Abbott)	20 or 40 tests	Lateral flow (immunochromatographic)	Not known	Positive results are indicated by 2 lines: one in the control region (C) and the other in the test region (T). Read by visual inspection	5
Bionexia Strep A plus - cassette (Biomerieux)	25 test cassettes	Lateral flow (immunochromatographic)	$1x10^4$ organisms/swab	Positive results are indicated by 2 lines: one in the control region (C) and the other in the test region (T). Read by visual inspection	5
Bionexia Strep A dipstick – test strip (Biomerieux)	25 test strips	Lateral flow (immunochromatographic)	Not known	Positive results are indicated by 2 lines: one in the control region (C) and the other in the test	5

Product	Test format and supply	Method	Limit of Detection	Description of results	Time to result (minutes)[a]
				region (T). Read by visual inspection	
Biosynex Strep A - cassette (Biosynex)	Not reported	Lateral flow (immunochromatographic)	$1x10^5$ bacteria/swab	Positive results are indicated by 2 lines: one in the control region (C) and the other in the test region (T). Read by visual inspection	5
Sofia Strep A FIA (Quidel)	25 cassettes, including positive and negative control vials	Lateral flow (immunofluorescence)	Strain Bruno [CIP 104226]: $1.86x10^4$ CFU/test Strain CDC-SS-1402: $9.24x10^3$ CFU/test Strain CDC-SS-1460: $2.34x10^4$ CFU/test	Analysed using the Sofia analyser which interprets the immunofluorescent signal using on-board method-specific algorithms. Results are displayed on screen as positive, negative or invalid.	5-6

CFU/ml Colorny forming units per millimeter

Clearview Exact Strep A cassette and Clearview Exact Strep A dipstick – test strip (both from Abbott) have been updated and replaced with the Clearview Exact 2

Four molecular tests were identified which use nucleic acid amplification techniques, either polymerase chain reaction (PCR) or isothermal nucleic acid amplification, to amplify and detect a specific fragment of the GAS genome (see Table 2). In each test, any GAS DNA present in the sample is labelled during the reaction, producing fluorescent light, which is monitored by a reader. If fluorescence reaches a specific threshold, the test is considered positive. If the threshold is not reached during the set time (usually up to 15 minutes), the test is negative.

Table 2: Molecular tests – product descriptions and properties

Product	Test supply and format	Method	Analyser	Limit of Detection	Description of results	Time to result (minutes)[a]
Alere i Strep A (Abbott)	24 test kits	Isothermal nucleic acid amplification	Alere i instrument	Strain: ATCC12344 4.2 CFU/ml ATCC19615 41.8 CFU/ml	Alere I instrument heats, mixes and detects, then presents results automatically on the digital display	<8
Alere i Strep A 2 [ID NOW Strep A 2]* (Abbott)	Information not available	Isothermal nucleic acid amplification	Alere i instrument	Not provided by manufacturer	Alere I instrument heats, mixes and detects, then presents results automatically on the digital display	<6
Cobas Strep A Assay (Roche Diagnostics)	Strep A Assay box of 20	Polymerase chain reaction	Cobas Liat analyser	Strain: ATCC BAA-946 5 CFU/mlATCC BAA-1066 10 CFU/mlATCC 12370 10 CFU/ml ATCC 700294 20 CFU/ml	Results displayed digitally	<15

Product	Test supply and format	Method	Analyser	Limit of Detection	Description of results	Time to result (minutes)[a]
Xpert Xpress Strep A (Cepheid)	Each kit contains sufficient reagents to process 10 specimens or quality control samples	Polymerase chain reaction	GeneXpert system	Strain: ATCC BAA-946 ATCC 19615 9–18 CFU/mL in a transport medium or 3–6 CFU/test.	Results displayed digitally	≥18

*The Alere i and Alere i Strep A 2 have now been replaced with the ID NOW Strep A 2

1.4 Anticipated position of point-of-care tests in the treatment pathway

The population of interest is people aged 5 and over presenting to healthcare providers in a primary (GP surgeries, community pharmacies and walk-in centres) or secondary care (urgent care/walk-in centres and emergency departments) setting with symptoms of an acute sore throat. These patients are identified as being more likely (FeverPAIN score of 2 or 3) or most likely (FeverPAIN score of 4 or 5, or a Centor score of 3 or 4) to benefit from an antibiotic by a clinical scoring tool. Relevant subgroups to be evaluated may include children (aged 5 to 14), adults (aged 15 to 75), and the elderly (adults over the age of 75 years). In elderly patients, the infection is more likely to be invasive and have a higher associated mortality rate.

1.5 Comparator

The comparator will be antibiotic prescribing based on clinical judgment and clinical scoring tools alone for GAS. However, the literature search for the comparator arm may also result in evidence referring to clinical scoring for Group C and Group G streptococci. The clinical scoring tools which may be used in NHS practice are FeverPAIN and Centor/modified Centor (McIsaac). These criteria are based on research evidence that assessed the individual and combination of sore throat symptoms most likely to be present in patients with clinically confirmed streptococcal infection (whether GAS or non-GAS).

FeverPAIN

The FeverPAIN clinical scoring tool includes the following variables:

- Clinical history
    - Sore throat (none; mild; moderate; severe)
    - Cough or cold symptoms (none; mild; moderate; severe)
    - Muscle aches (none; mild; moderate; severe)
    - Fever in last 24 hours (yes; no)
    - Onset of illness (0-3 days; 4-7 days; 7+ days)

- Clinical examination
    - Cervical glands (none; 1-2cm; >2cm)
    - Inflamed tonsils (none; mild; moderate; severe)
    - Pus on tonsils (yes; no)

The result of FeverPAIN is presented as a score ranging from 0 to 5, with one point assigned for each symptom present.

Centor

The Centor clinical scoring tool includes the following variables:

- Cough (yes; no)
- Exudate or swelling on tonsils (no; yes)
- Tender/swollen anterior cervical lymph nodes (no; yes)
- Temperature >38°C (no; yes)

Expert advice suggests that the McIsaac (modified Centor) clinical scoring tool may also be used. The McIsaac score adjusts the Centor score to account for the higher incidence of GAS in children and reduced incidence in older adults. This adds age criteria (3-14 years; 15-44 years; ≥45 years) and adds one point for those aged under 15 and subtracts one point for those aged 45 and over. The result of Centor/modified Centor is presented as a score ranging from 0 to 4, with one point assigned for each symptom present.[17]

1.6 Reference standard

The reference standard for assessing the test accuracy of point-of-care tests for Strep A infections is microbiological culture of throat swabs using standard blood agar or Streptococcal Selective agar as the culture medium. In the latter, antibiotics can be added to the standard blood agar to suppress the normal pharyngeal microflora, thus improving the yield of the Strep A infections. However, there is no consensus on the preferred medium.[18]

Throat swab culture remains the best reference standard for diagnosing streptococcal pharyngitis. Assuming strict adherence to standard operating guidelines for obtaining throat

swab samples and processing cultures should result in very high accuracy. However, where the validity of a negative culture is in doubt, a repeat culture is performed and trumps empirical antibiotic treatment.[18] The accuracy of culture cannot be guaranteed to be 100% with several studies observing discordance with PCR or other measures.[19-21]

In recent studies, PCR techniques were used as arbitrators of discordant results between throat culture and point-of-care tests.[19, 22, 23] A threshold quantity of viable organisms must be exceeded in order for culture to be positive, whereas PCR-based tests are able to detect the genome of organisms irrespective of their viability. However, PCR cannot distinguish between acute Strep A pharyngitis and asymptomatic pharyngeal carriage, and therefore may detect carriage in the absence of a streptococcal infection. Therefore, our reference standard does not include PCR. Further, some of the index tests are PCR based, and so a PCR based reference standard would be biased in favour of these index tests. Where such arbitration using PCR is reported we have included in this report, but the main analysis uses culture as the reference standard.

2 Definition of the decision problem

2.1 Decision question

This report undertaken for the NICE Diagnostics Assessment Programme examines the clinical and cost-effectiveness of point-of-care tests for diagnosing Group A Streptococcal infections in people who present with an acute sore throat in primary and secondary care settings. The report will help NICE to make recommendations about how well the tests work and whether the benefits are worth the cost of the tests, when used in the NHS in England. The assessment also considers other outcomes including antibiotic prescription behaviour, clinical improvement in patients' symptoms and costs associated with treatment based on evidence identified through systematic literature searches.

The decision question for this project is:

What is the clinical and cost-effectiveness of rapid antigen detection and molecular tests in patients with high clinical scores (Centor scores ≥3, FeverPAIN scores of ≥4), compared to the use of clinical scoring tools alone, for increasing the diagnostic confidence of suspected Group A Streptococcal infection in people who present with an acute sore throat in primary and secondary care?

2.1.1 Overall aim of the assessment

The overall aim of this report was to present evidence on the clinical- and cost-effectiveness of rapid antigen detection and molecular tests in those with high clinical scores (Centor scores ≥3, FeverPAIN scores of ≥4), compared to the use of clinical scoring tools alone, for increasing the diagnostic confidence of suspected Group A Streptococcal infection in people who present with an acute sore throat in primary and secondary care.

2.1.2 Objectives

- To systematically review the evidence for the clinical effectiveness of selected rapid tests for Group A Streptococcal infections in people with a sore throat.
- To systematically review existing economic evaluations and develop a *de novo* economic model to assess the cost-effectiveness of rapid tests in conjunction with

clinical scoring tools for Group A Streptococcal infections compared to clinical scoring tools alone.

3 Clinical effectiveness review

3.1 Methods

3.1.1 Search strategies for clinical effectiveness

The search strategy for the clinical effectiveness review is detailed in Appendix 1. An iterative procedure was used to develop the database search strategies, building on the scoping searches undertaken by NICE for this assessment and the searches underpinning the related Medtech innovation briefing published by NICE in 2018.[15] Database searches were run in November and December 2018 and were updated in March 2019. No date or language limits were applied. Grey literature searches were undertaken in February and March 2019.

Briefly, the search strategy included:

- Databases: MEDLINE (Ovid), MEDLINE In-Process & Other Non-Indexed Citations (Ovid), MEDLINE Epub Ahead of Print (Ovid), MEDLINE Daily update (Ovid); Embase (Ovid); Cochrane Database of Systematic Reviews (Wiley); CENTRAL (Wiley); Database of Abstracts of Reviews of Effects (DARE) (Centre for Reviews and Disseminations (CRD)); Health Technology Assessment database (CRD); Science Citation Index and Conference Proceedings (Web of Science); PROSPERO International Prospective Register of Systematic Reviews (CRD)
- Trial database: ClinicalTrials.gov
- Reference lists of relevant reviews and included studies
- Online resources of health services research organisations and regulatory bodies: International Network of Agencies for Health Technology Assessment (INAHTA) http://www.inahta.org/; Food and Drug Administration (FDA) medical devices; FDA CLIA - Clinical Laboratory Improvement Amendments database; European Commission medical devices
- Online resources of selected professional societies and conferences: British Society for Antimicrobial Chemotherapy; British Infection Association; Public Health England; British Society for Antimicrobial Chemotherapy; Royal College of Pathologists; Streptococcal biology conference; Lancefield International Symposium on Streptococci and Streptococcal Diseases; Federation of Infection Societies conference; The European Congress of Clinical Microbiology and Infectious

(ECCMID); Microbiology Society Conference; American Society of Microbiology; Association of Clinical Biochemistry and Laboratory medicine

- Online resources of manufacturers of the included rapid tests.

3.1.2 Inclusion and exclusion of relevant studies

3.1.2.1 Inclusion criteria

Studies that satisfied the following criteria were included:

Population	People aged 5 years and above presenting with symptoms of an acute sore throat. Where possible relevant subgroups evaluated included children (aged 5 to 14), adults (aged 15 to 75), and the elderly (adults over the age of 75 years), however mixed populations were acceptable. Studies of children under the age of 5 could be included providing 90% or more were above this age.
Intervention	Point-of-care tests for GAS (including rapid antigen detection tests and molecular tests as described in Table 1 and Table 2).
Comparator	Clinical scoring tools (such as FeverPAIN, Centor or McIsaac) The reference standard for assessing the test accuracy of rapid tests is microbiological culture of throat swabs.
Outcome	Outcomes of test performance • Test accuracy: sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV). Where possible evaluated by relevant clinical scores (Centor/McIsaac ≥3, FeverPAIN ≥4) • Discordant results with throat culture • Test failure rates • Time to antimicrobial prescribing decision • Changes to antimicrobial prescribing decision • Number of appointments required per episode • Number of delayed or immediate antibiotic prescriptions issued Clinical outcomes:

	• Morbidity, including post-GAS infection complications such as rheumatic fever and side-effects from antibiotic therapy • Mortality • Contribution to antimicrobial stewardship and onward transmission of infection Patient reported outcomes: • Health-related quality of life • Patient satisfaction with test and antimicrobial prescribing decision • Healthcare professional satisfaction with test and antimicrobial prescribing decision Costs
Study design	For test accuracy data: Clinical test accuracy studies that compare the index tests (point-of-care tests for GAS) to throat swab culture. Studies of head-to-head comparisons of rapid tests were eligible for inclusion if test accuracy statistics were reported for each test. For data on other clinical outcomes: Any study design comparing the index tests (point-of-care tests for Strep A) to throat swab culture and/or clinical scoring tools (Centor, McIsaac or FeverPAIN)
Healthcare setting	Primary care (GP clinics, community pharmacies and walk-in centres) and secondary care (urgent care/walk-in centres and emergency departments) settings.

3.1.2.2 **Exclusion criteria**

Population	Patients without acute sore throat Patients with existing comorbidities Patients with known invasive Group A Strep infection.

Intervention	Other point-of-care tests which are not listed in the NICE scope
Comparator	For test accuracy data: No comparison of index test versus throat culture reported For other outcomes: No comparison of index test versus throat culture or clinical scoring tools (Centor, McIsaac or FeverPAIN)
Study design	Reviews, biological studies, case reports, editorials and opinions, poster presentations without supporting abstracts, non-English language reports, meeting abstracts without sufficient information to produce 2x2 contingency tables for test performance
Date	Studies published before 1998 (keeping in line with the 1998 directive of the European parliament requiring all in-vitro diagnostic devices to have a CE marking)
Setting	Hospital inpatient

3.1.3 Study selection strategy

All publications identified in searches from all sources were collated in Endnote and de-duplicated. Two reviewers independently screened the titles and abstracts of all records identified by the searches (Cohen's Kappa = 0.997) and discrepancies were resolved through discussion. Full copies of all studies deemed potentially relevant were obtained and two reviewers independently assessed these for inclusion; any disagreements were resolved by consensus or discussion with a third reviewer. Records excluded at full text stage and reasons for exclusion were documented.

3.1.4 Data extraction strategy

All data were extracted by one reviewer, using a piloted data extraction form. A second reviewer checked the extracted data on test accuracy (2x2 table, sensitivity, specificity, positive predictive value, and negative predictive value), whereas a third reviewer checked other extracted data. Any disagreements were resolved by consensus. A sample data extraction form used in this review is available in Appendix 2. Test accuracy statistics for rapid/index tests were derived from data extracted onto 2x2 contingency tables in the format shown in Table 3. As shown, A represents the number of patients positive for GAS by rapid

test and throat culture (true positives); B represents the number of patients positive for GAS by rapid test but not throat culture (false positives); C represents the number of patients negative for GAS by rapid test but positive by throat culture (false negatives); and D represents the number of patients negative for GAS by rapid test and throat culture (true negatives). Sensitivity was calculated as A/(A+C); specificity as D/(B+D); positive predictive value (PPV) as A/(A+B); and negative predictive value (NPV) as D/(C+D). Similarly, using data extracted in the formats shown in Table 4 and Table 5, we calculated accuracy statistics for the current pathway (Centor / McIssac / FeverPAIN scores) based on NICE thresholds. Where PCR techniques were employed to arbitrate discordant results between microbiological culture and rapid tests we report the PCR results for the discordant cases. We also extracted test accuracy data for each index test with culture as reference standard in studies of head-to-head (direct) comparisons of index tests. Data on other outcomes of test performance, morbidity, antibiotic prescribing behaviour, population characteristics, and settings were also extracted using the extraction form.

Table 3: 2x2 contingency table for rapid test versus throat culture

	Culture +	Culture -	Total
Index test +	A	B	A+B
Index test -	C	D	C+D
Total	A+C	B+D	A+B+C+D

Table 4: 2x2 contingency table for Centor/Modified Centor versus throat culture

	Culture +	Culture -	Total
Centor/McIsaac score ≥3	A	B	A+B
Centor/McIsaac score <3	C	D	C+D
Total	A+C	B+D	A+B+C+D

Table 5: 2x2 contingency table for FeverPAIN versus throat culture

	Culture +	Culture -	Total
FeverPAIN ≥ 4	A	B	A+B
FeverPAIN <4	C	D	C+D
Total	A+C	B+D	A+B+C+D

3.1.5 Quality assessment strategy for test accuracy studies

Quality assessment of eligible test accuracy studies was undertaken with a tailored Quality Assessment of Diagnostic Accuracy Studies – 2 (QUADAS-2) tool. Methodological quality was assessed by a single reviewer and findings were checked by a second reviewer. Disagreements were resolved by consensus or use of a third reviewer.

Quality assessment aimed to assess the risk of bias and applicability concerns of included studies where one (or more) of our 21 scoped tests were the index test(s), and biological throat culture was the reference standard. Additional tests outside of scope were not quality appraised.

Modifications to tailor the QUADAS-2 form to the research question in terms of the risk of bias assessment were as follows (see Appendix 4 for the tailored QUADAS-2 form and guidance notes).

3.1.5.1 **Patient selection domain:**

Two further signalling questions were added to this domain. The first was: *Were selection criteria clearly described?* It is important that the correct patient groups were included within the studies. Patients under the age of 5 follow a different NICE clinical pathway[24] owing to them being more likely to present with a sore throat, less likely to be able to articulate their symptoms and it is less likely a throat swab can be obtained. Likewise, a clinical score (such as Centor or FeverPAIN) should be reported, with patients only included if they have a score of above 3 on Centor or above 4 on FeverPAIN. Those with lower scores may be systematically different and therefore test accuracy may also differ introducing bias.

Including patients under the age of 5 and with a low clinical score also raises applicability concerns.

The second signalling question that was added was: *Were patients seen in an ambulatory care setting?* Patients seen as inpatients may vary in severity and have comorbidities affecting their diagnosis.

3.1.5.2 **Index test domain:**

Two questions were added within this domain. The first was: *Was a separate swab undertaken for the index test?* This question was added as manufacturer's specifications require separate swabs be taken for index and reference standard tests. Using one swab for multiple purposes may reduce the quantity of sample for testing and thus affect the accuracy of the test. The second was: Is the test reading objective? Some of the tests require a subjective reading of whether a line, indicating a positive result, has appeared. Due to this, there is always a high level of bias in any rapid test which requires a determination of the result by the human reader. Tests with automated readings have been shown to have improved specificity and reduce operator errors, especially in unclear results.[25]

3.1.5.3 **Comparator domain:**

One additional signalling question was added in this domain: *Was a separate swab taken for throat culture testing?* Using one swab for multiple purposes may reduce the quantity of sample for testing and thus affect the accuracy of the test. Under this domain the directions for taking a throat culture specimen were clarified based upon the PHE guidelines on UK Standards for Microbiology Investigations.[26]

3.1.5.4 **Flow and timing:**

Two further signalling questions were added to the flow and timing domain. The first was: *Were both index test(s) and reference standard (and comparator where included) all carried out at the same appointment?* The swabs for a rapid test and culture should be done at the

same appointment. The levels of GAS are likely to vary by day, so taking a later sample could introduce systematic bias.

The additional signalling question was: *Were both index test(s) and reference standard (and comparator where included) all carried out prior to commencement of antibiotics?* Patients should not have been treated with antibiotics prior to testing as antibiotics are likely to have reduced the amount of GAS present.

3.1.5.5 Quality appraisal strategy for studies of prescribing behaviour and clinical outcomes

Quality appraisal for studies of prescribing behaviour and clinical outcomes used two different tools. The Cochrane Risk of Bias (ROB) tool for Randomised Controlled Trials (RCTs)[27] and the Joanna Briggs Institute Critical Appraisal Checklist for Analytical Cross-sectional Studies.[28] Methodological quality was assessed by a single reviewer and findings were checked by a second reviewer. Disagreements were resolved by consensus or use of a third reviewer.

3.1.6 Assessment of test accuracy

In order to assess the accuracy of the point-of-care tests, we planned to conduct a series of meta-analyses on the available data. Data from studies which either presented 2x2 tables for one of the index tests was compared to culture or provided information that allowed calculation of the 2x2 table were included in the meta-analyses.

The median age of participants was used to categorise each study into one of the three age groups of interest, with two reviewers discussing where the categorisation was not straightforward. Setting was also considered to inform the age categorisation where necessary (e.g. if study was conducted within a paediatric department). The setting of each study was treated as a categorical variable, indicating primary care (healthcare centre, GP clinic or primary care clinic), secondary care (emergency department, private pediatric clinic, outpatient clinic, urgent care clinic or walk in centre), and pharmacy setting or mixed.

For the purpose of the meta-analysis the throat-score of the population was dichotomised to 0 if the study population included patients who have scores below the threshold set in the scope, and 1 if the study population matches the scope (Centor/McIsaac ≥ 3 or FeverPAIN ≥ 4). Alternative throat score classification of study populations was also considered, using the categories of a population matching the scope (as above), a population restricted by throat score but still including patients not in the scope (e.g. Centor = 2), and a population without any restriction by throat score.

3.1.7 Methods of analysis/synthesis

We planned to use bivariate models to perform each meta-analysis, as they allow simultaneous estimation of both sensitivity and specificity, accounting for correlation between the two measures. Where at least 5 studies existed for a test, we used a random effects model to allow for deviation in test performance across each study. If the random effects model failed to provide reliable estimates, or if there were between 2 and 5 studies a fixed effects model was fitted. If bivariate models failed to converge or did not provide meaningful estimates, then univariate analyses were performed instead, considering sensitivity and specificity individually. [29, 30] Where bivariate models were used, a comparison to the equivalent univariate models was made, and any difference noted. It was not anticipated that any meaningful difference between the two model types would be observed given the small data available.

For index tests that had just one study, a meta-analysis was not conducted. The impact of age, setting and prevalence on test performance were all assessed through the meta-analysis of relevant subgroups. NICE advised the EAG against meta-analysis across rapid tests from different manufacturers.

3.2 Clinical Effectiveness Results

3.2.1 Search results

Figure 2 is a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram that illustrates the study selection process for the clinical

effectiveness review. The search identified 5,919 records through database and other searches. Following duplicate removal, we screened 3,309 records of which 3,072 were excluded by their titles and abstracts, leaving 237 assessed for their eligibility to be included in the review. 199 studies were subsequently excluded with reason leaving 38 studies (26 full texts [6, 19, 22, 23, 31-52], 3 abstracts [53-55], 5 manufacturers' studies (submitted directly to NICE in response to a request for information) and 4 FDA documents[56-59]). The most common reason for exclusion at this stage was not reporting any of the rapid tests listed in the scope. The full list of excluded studies with reasons for exclusion can be found in Appendix 3.

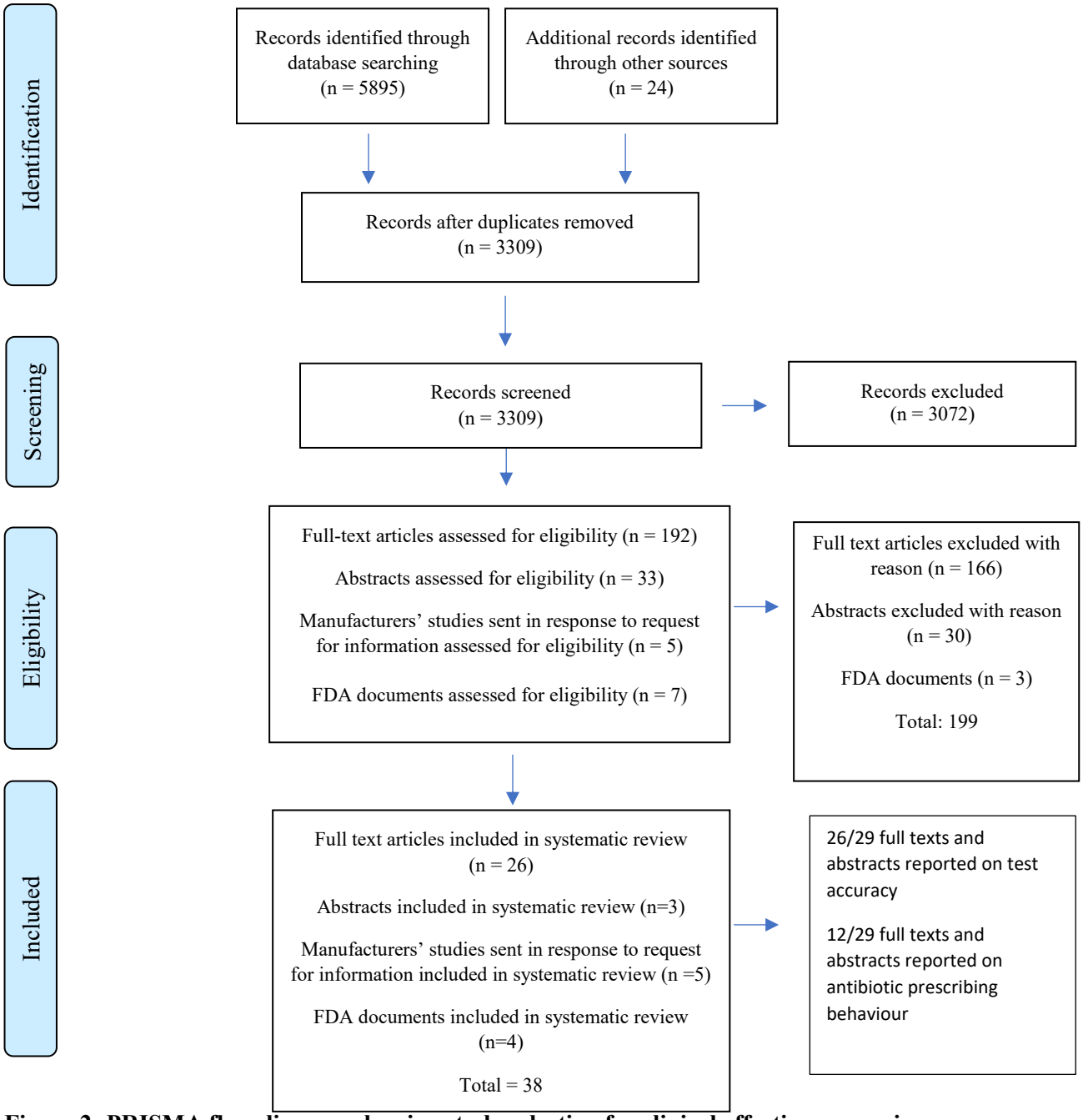


Figure 2: PRISMA flow diagram showing study selection for clinical effectiveness review

3.2.2 Study characteristics

Characteristics of the 38 studies included in the clinical effectiveness review are described in Figure 3, Figure 4 and Table 6. Of the 29 studies (full texts and abstracts)[6, 19, 22, 23, 31-55] identified by the search, 26 of the studies reported test accuracy data.[19, 22, 23, 31-49, 51, 54, 55, 60] Three of the identified studies only reported other outcomes (such as antibiotic prescribing rates) and did not report test accuracy.[6, 50, 52] In addition there were 5 studies sent by manufacturers in response to a request for information by NICE and 4 FDA documents retrieved.[56-59]

The tests, their settings, the populations they cover, and the head-to-head studies are illustrated Figure 3 and Figure 4.

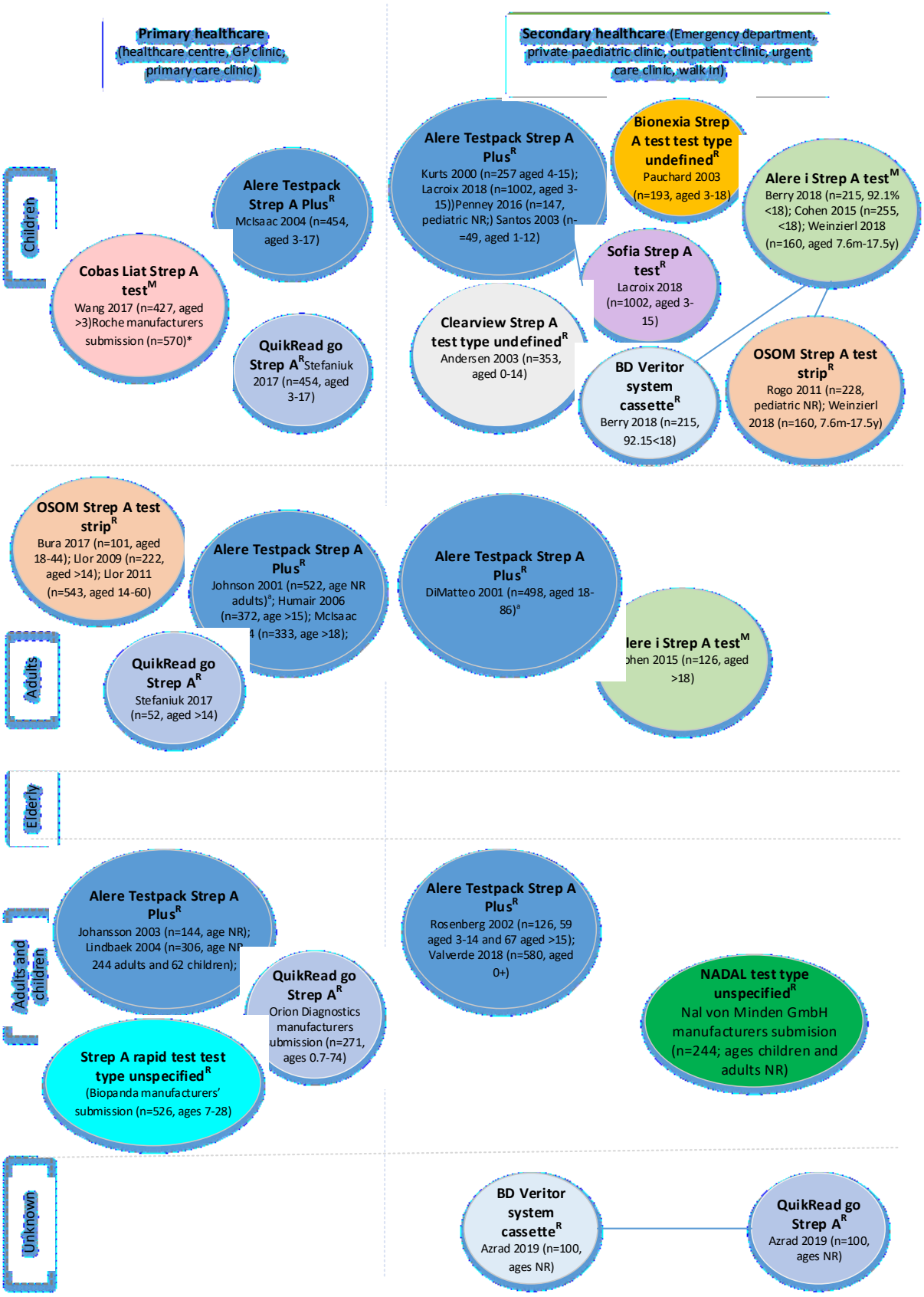


Figure 3: Diagram of studies by test type, setting and population for included test accuracy studies with extractable 2x2 data

Note: Lines between tests indicate head-to-head (direct) comparisons.

Figure 4: Diagram of studies by test type, setting and population for included test accuracy studies with extractable 2x2 data

Table 6: Characteristics of the included studies

Study Reference	Data Source	Setting	Study Population				Index test	Comparison with Centor / McIsaac / FeverPAIN scores	Throat swab Culture Medium	Outcomes	Test accuracy in high-risk sub-populations with Centor / McIsaac scores ≥3 or FeverPAIN score ≥4
			N	Age group as reported	Sore throat clinical score criteria	GAS prevalence (%)					
Published articles and abstracts											
Anderson 2003 [53]	Abstract	Secondary	353	Children (0 to 14 years)	No criteria reported. Used clinical symptoms	15	Clearview Strep A	No	NR	Test accuracy	NR
Azrad 2019 [31]	Published article	Secondary	100	NR	No criteria reported. Used clinical symptoms	2525	BD veritor system QuikRead Go Strep A test kit (Orion Diagnostica)	No	Strep Selective Agar	Test accuracy	NR
Berry 2018 [19]	Published article	Secondary	215	Children (age range not reported)	NR	19.5	Alere i Strep A test BD veritor system	No	Blood agar	Test accuracy Antibiotic prescribing behaviour	NR NR
Bird 2018 [32]	Published article	Secondary	395	Children	McIsaac ≥3	NR or calculable	Bionexia Strep A	Yes/Centor	NA	Test accuracy Antibiotic prescribing behaviour	NR
Bura 2017 [33]	Published article	Primary	101	Adults (18 to 44 years)	Centor ≥ 2	22.7	OSOM Strep A test (Sekisui diagnostics)	Yes/Centor	Blood agar	Test accuracy	No

Study Reference	Data Source	Setting	Study Population				Index test	Comparison with Centor / McIsaac / FeverPAIN scores	Throat swab Culture Medium	Outcomes	Test accuracy in high-risk sub-populations with Centor / McIsaac scores ≥3 or FeverPAIN score ≥4
			N	Age group as reported	Sore throat clinical score criteria	GAS prevalence (%)					
										Antibiotic prescribing behaviour	
Cohen 2015 [34]	Published article	Secondary	481	Children (median age = 11 yrs)	McIsaac all scores	30.3	Alere i Strep A test	Yes/McIsaac	Blood agar	Test accuracy	No
DiMatteo 2011 [35]	Published article	Secondary	383	Adults (18 to 86 years)	Centor ≥ 1	NR or calculable	Alere TestPack Plus Strep A (Abbott)	Yes/McIsaac	Strep Selective Agar	Test accuracy	NR
Humair 2006 [36]	Published article	Primary	224	Adults (15 to 65 years)	Centor = 2 Centor > 2	46.9	Alere TestPack Plus Strep A (Abbott)	Yes/Centor	Blood agar	Test accuracy, Antibiotic prescribing behaviour	Yes
Johannson 2003 [37]	Published article	Primary	144	Mixed (children and adults)	NR	31.4	Alere TestPack Plus Strep A (Abbott)	No	NR	Test accuracy, Antibiotic prescribing behaviour	NR
Johnson 2001 [38]	Published article	Primary	522	Adults (median age = 26 years)	No criteria reported. Used clinical symptoms	NR or calculable	Alere TestPack Plus Strep A (Abbott)	No	Blood agar	Test accuracy	NR
Kurtz 2000 [39]	Published article	Secondary	257	Children (4 to 15 years)	No criteria reported.	31.1	Alere TestPack Plus	No	Blood agar		NR

Study Reference	Data Source	Setting	Study Population				Index test	Comparison with Centor / McIsaac / FeverPAIN scores	Throat swab Culture Medium	Outcomes	Test accuracy in high-risk sub-populations with Centor / McIsaac scores ≥3 or FeverPAIN score ≥4
			N	Age group as reported	Sore throat clinical score criteria	GAS prevalence (%)					
					Used clinical symptoms		Strep A (Abbott)			Test accuracy	
Lacroix 2018 [22]	Published article	Secondary	1002	Children	McIsaac ≥ 2	38	Sofia Strep A FIA (Quidel) Alere Testpack Plus Strep A test (Abbott)	No	Blood agar	Test accuracy	No
Lindbæk 2004 [40]	Published article	Primary	306	Adults (median age = 23.9 years)	NR	35.9	Alere TestPack Plus Strep A (Abbott)	No	Strep Selective Agar	Test accuracy	NR
Little 2013 [6]	Published article	Primary	1760	Mixed (age ≥ 3 years)	FeverPAIN ≥ 1	NR or calculable	Alere TestPack Plus Strep A (Abbott)	Yes	None	Antibiotic prescribing behaviour	No
Llor 2009 [41]	Published article	Primary	222	Adults (median age = 30.6 years)	Centor ≥ 2	21.2	OSOM Strep A (Genzyme)	Yes/Centor	Blood agar	Test accuracy	No
Llor 2011 [42]	Published article	Primary	116	Adults (median age = 31.7 years)	Centor =1 Centor =2	16.7	OSOM Strep A test	Yes/Centor	Blood agar	Test accuracy, Antibiotic prescribing behaviour	Yes
McIsaac 2004 [43]	Published article	Primary	787	Children (3 to 17 years)	McIsaac all scores	29	TestPack Plus Strep A Test (Abbott)	Yes/McIsaac	Blood agar	Test accuracy, Antibiotic	

Study Reference	Data Source	Setting	Study Population				Index test	Comparison with Centor / McIsaac / FeverPAIN scores	Throat swab Culture Medium	Outcomes	Test accuracy in high-risk sub-populations with Centor / McIsaac scores ≥3 or FeverPAIN score ≥4
			N	Age group as reported	Sore throat clinical score criteria	GAS prevalence (%)					
				Adults (≥ 18 years) results reported separately by group						prescribing behaviour	No
Nerbrand 2002 [44]	Published article	Primary	615	Mixed (children and adults	No criteria reported. Used clinical symptoms	21.1	TestPack Plus Strep A Test (Abbott)	No	Blood agar	Test accuracy	NR
Pauchard 2013 [54]	Abstract	Secondary	193	Children (3 to 18 years)	McIsaac > 2	37	Strep A rapid test (Biopanda)	Yes/McIsaac	NR	Test accuracy	NR
Penney 2016 [45]	Published article	Secondary	147	Children (mean age = 8.8 years)	No criteria reported. Used clinical symptoms	40.1	Alere TestPack Plus Strep A Test (Abbott)	No	Strep selective agar	Test accuracy	NR
Rogo 2011 [46]	Published article	Secondary	228	Children	No criteria reported. Used clinical symptoms	28.9	OSOM Strep A test (Genzyme)	No	Blood agar	Test accuracy	NR
Rosenberg 2002 [47]	Published article	Secondary	126	Mixed (children and adults)	Centor all scores	25.4	TestPack Plus Strep A Test (Abbott)	Yes/Centor	Blood agar	Test accuracy Antibiotic prescribing behaviour	NR

Study Reference	Data Source	Setting	Study Population				Index test	Comparison with Centor / McIsaac / FeverPAIN scores	Throat swab Culture Medium	Outcomes	Test accuracy in high-risk sub-populations with Centor / McIsaac scores ≥3 or FeverPAIN score ≥4
			N	Age group as reported	Sore throat clinical score criteria	GAS prevalence (%)					
Santos 2003 [48]	Published article	Secondary	49	Children (1 to 12 years)	No criteria reported. Used clinical symptoms	30	Alere TestPack Plus Strep A (Abbott)	No	Blood agar	Test accuracy	NR
Stefaniuk 2017 [49]	Published article	Primary	44	Children	McIsaac/ Centor all scores	26.3	QuikRead Go Strep A test kit (Orion Diagnostica)	Yes/Centor	Blood agar	Test accuracy Antibiotic prescribing behaviour	No
			96	Adult + Children	McIsaac/ Centor all scores	22.4					
Thornley 2016 [50]	Published article	Pharmacy	149	NR	Centor > 2	24.2	OSOM Strep A test (Sekisui)	Yes/Centor	None	Antibiotic prescribing behaviour	NA
Valverde 2018 [55]	Abstract	Secondary	580	Mixed (age ≥ 0 yr)	NR	NR or calculable	TestPack Plus Strep A Test	No	Blood agar	Test accuracy	NR
Wang 2017 [23]	Published article	Primary	427	Children	Centor ≥ 1	30.2	Cobas Liat Strep A Assay (Roche)	No	Culture medium NR	Test accuracy	No
Weinzierl 2018 [51]	Published article	Secondary	160	Children (median age = 6.5 years)	NR	38	OSOM Strep A test	No	Blood agar	Test accuracy	NR
							Alere I Strep A test				

Study Reference	Data Source	Setting	Study Population				Index test	Comparison with Centor / McIsaac / FeverPAIN scores	Throat swab Culture Medium	Outcomes	Test accuracy in high-risk sub-populations with Centor / McIsaac scores ≥3 or FeverPAIN score ≥4
			N	Age group as reported	Sore throat clinical score criteria	GAS prevalence (%)					
Worrall 2007 [52]	Published article	Primary	533	NR	Centor all scores	NR or calculable	Clearview Exact Strep A (Wampole)	Yes/Centor	NA	Antibiotic prescribing behaviour	NA
Manufacturer's studies provided in responses to request by NICE											
Biopanda	Manufacturer's information	Secondary	160	6.5	NA	23.2	Alere I Strep A test	No	Blood agar	Test accuracy	NR
Cepheid	Manufacturer's information	Primary	577	NR	NA	25.6	Xpert Xpress	Yes/Centor	NA	Test accuracy	NR
nal von minden GmbH	Manufacturer's information	Unknown	244	Mixed (Adults and children)	NA	34.4	NADAL Strep A test Unspecified	No	Blood agar	Test accuracy	NR
Orion Diagnostica	Manufacturer's information	Primary	271	NR	NA	32.8	QuikRead Go Strep A test kit (Orion Diagnostica)	No	Strep Selective Agar	Test accuracy	NR
Roche Diagnostics	Manufacturer's information	Mixed	570	Mixed (age ≥ 3 years)	NA	30.4	Cobas Liat Strep A Assay (Roche)	No	Blood agar	Test accuracy	NR
FDA Documents											
Abbott [58]	FDA document	Mixed	981	NR	NA	20.2	Alere I Strep A2 test	No	Blood agar	Test accuracy	NR

Study Reference	Data Source	Setting	Study Population				Index test	Comparison with Centor / McIsaac / FeverPAIN scores	Throat swab Culture Medium	Outcomes	Test accuracy in high-risk sub-populations with Centor / McIsaac scores ≥3 or FeverPAIN score ≥4
			N	Age group as reported	Sore throat clinical score criteria	GAS prevalence (%)					
Beckton Dickinson [56]	FDA document	Mixed	796	Mixed (age ≥ 0 yr)	NA	18.7	BD veritor system	No	Blood agar	Test accuracy	NR
Cepheid [59]	FDA document	Mixed	618	NR	NA	25.6	Xpert Xpress Strep A (Cepheid)	No	Culture - medium NR	Test accuracy	NR
Quidel [57]	FDA document	Mixed	736	NR	NA	17.4	Sofia Strep A FIA (Quidel)	No	Blood agar	Test accuracy	NR

NA = Not applicable; NR = Not reported

3.2.2.1 **Population**

The 38 included studies comprised ~14,000 symptomatic participants. Prevalence of Strep A ranged from 15% to 49%, with no clear demographic or clinical patterns accounting for this variation.[36, 53] Similarly, prevalence estimates of GAS were no more or less likely to be higher in secondary or primary care settings. The study population comprised adults and children, however the exact proportions are unknown as they were not reported in about half of the included studies. In most of the included studies, participants aged under 18 years were identified as children. In fact, only two studies met the age criterion for children (ages 5 to 14 years) as defined in the protocol and scope.[47, 48] Hence, studies which included children under five years as well as older than age 5 were included in the present review. More so, only two studies met the age criterion for adults (age ≥ 15 years) as defined in the protocol and scope, and therefore the findings of the review may only be applicable to a mixed population.[36, 49]

All 38 studies included patients with a sore throat, however other clinical characteristics were insufficiently reported across most of the included studies. For instance, sore throat clinical scores (e.g. Centor/McIsaac/FeverPAIN scores) were reported in 16 studies.[6, 22, 32-36, 41-43, 47, 49, 50, 52, 54] Of these 16 studies, 2 exclusively included patients with high clinical scores (Centor ≥3, FeverPAIN ≥4). [6, 50]Both of these studies were on prescribing behaviours. However, there were 2 test accuracy studies which included patients with lower clinical scores (Centor scores <3) but reported test accuracy results separately by Centor score.[36, 42]

Recent antibiotic use prior to enrolment was considered in eight included studies, and patients without any recent use of antibiotics prior to recruitment were eligible for inclusion in these studies.[23, 32, 33, 39, 42, 44, 45, 47]

3.2.2.2 **Index tests**

There were more studies evaluating rapid antigen detection tests (76%, 29/38) than molecular tests (18%, 7/38) or studies comparing both rapid test and molecular tests (5%, 2/38). For instance, the TestPack Plus Strep A (Abbott) was the most common antigen detection test evaluated in 13 studies (excluding unpublished studies conducted by the manufacturers).[6, 22,

[35-40, 43-45, 48, 55] Conversely, the only molecular test evaluated in a peer-reviewed journal article was the PCR-based Cobas Liat Strep A Assay (Roche Diagnostics).[23]

As shown in Table 6, there were four studies providing head-to-head comparisons of index tests: BD Veritor System (Beckton Dickinson) and QuikRead Go (Orion Diagnostica);[31] Alere i Strep A (Abbott) and BD Veritor System (Beckton Dickinson);[19] Alere i Strep A (Abbott) and Sofia Strep A FIA (Quidel);[22] and Alere i Strep A (Abbott) and OSOM Strep A.[51] Essentially, each index test was compared to throat culture as the reference standard in order to obtain test accuracy.

The search strategy revealed test accuracy studies of OSOM Ultra Strep A (Genzyme & Sekisui).[61, 62] However, these studies were subsequently excluded because the EAG could not confirm whether it is the same as the OSOM Strep A test (Genzyme & Sekisui) which is listed among the scoped rapid tests. Similarly, it was unclear if Sofia Strep A+ Plus FIA (Quidel)[63] and OSROM Strep A (Sekisui)[64] were identical to Sofia Strep A FIA (Quidel) and OSOM Strep A (Sekisui) respectively, hence studies of the former were excluded.

3.2.2.3 **Comparator and Reference Standard**

Index tests were compared to Centor, McIsaac or FeverPAIN scoring tools in twelve studies.[6, 32, 33, 35, 42, 43, 47, 49, 50, 52, 54, 65] However, only six of these studies directly compared test accuracy between clinical scoring tools and point-of-care tests.[36, 41-43, 49] Only 2 reported test accuracy in patients with high clinical scores (Centor ≥3, FeverPAIN ≥4).[36, 42]

The culture medium used for the reference standard (blood agar or strep selective agar) was reported in all but five studies.[23, 37, 53, 54] Neither the manufacturer's submissions (submitted directly to NICE in response to a request for information) or FDA studies provided information on index tests compared to clinical scoring tools.

3.2.2.4 **Outcomes**

There were 38 studies included across all outcomes. 26 published articles (full texts and abstracts) reported test accuracy data (of which 7 also reported on antibiotic prescribing rates and 5 report on test failure rate), plus an additional five submissions from manufacturers and four FDA documents.

Five studies had insufficient data to construct 2x2 contingency tables to ascertain the accuracy of index tests with microbiological throat culture as the reference standard.[6, 32, 44, 50, 52] These studies were further excluded from the assessment of test accuracy in section 3.2.6. An attempt to verify at least some of the discrepant results between rapid tests and microbiological culture was undertaken in only five studies.[19, 22, 23, 34, 40] Antibiotic prescribing behaviour was reported in twelve studies. [6, 19, 32, 33, 36, 37, 42, 43, 47, 49, 50, 52] None of the other outcomes in the scope or protocol were reported in any of the included studies.

3.2.2.5 **Setting**

Participants in the included studies were recruited from GP/primary care clinics/family practices,[6, 37, 38, 40-44] community pharmacies,[50] paediatrics clinics,[39, 46, 51, 53] paediatric emergency departments,[22, 45, 54] hospital outpatients departments,[19, 48] and emergency departments.[32, 47] There were two multi-centre studies with mixed populations from primary and secondary care settings: Cohen et al. (2015)[34] sampled patients from the emergency department (secondary care) and urgent care clinics (primary care); Wang et al. (2017)[23] sampled patients from paediatric clinics (secondary care) and family practices (primary care).

Only one unpublished study supplied by the manufacturers confirmed study setting (Orion Diagnostica, primary care). The remaining unpublished studies conducted by manufacturers may have included mixed populations from primary and secondary care settings, however this is purely speculative as study settings were not reported in these studies. However, these studies provide no evidence to suggest any recruitment of in-patients.

3.2.2.6 **Study design**

The 26 published studies on test accuracy were comprised of 1 RCT [42] and 25 cohort studies.[19, 22, 23, 31-41, 43, 45-49, 52-55, 66] It was unclear what study design had been undertaken in any of the unpublished studies provided by the manufacturers or the FDA.

The 12 studies which provided data on antibiotic prescribing rates were comprised of 3 RCTs,[6, 42, 52] 1 before and after cohort study[19] and 8 one-armed cohort studies.[32, 33, 36, 37, 43, 47, 49, 50]

3.2.3 Quality considerations of included studies

The assessment of risk of bias and applicability for the 26 included test accuracy studies[19, 22, 23, 31-43, 45-49, 52-55, 66] using the QUADAS-2 tool are summarised in Table 7 and Figure 5. Four of the included studies compared two index tests relevant to this review so there are 30 quality assessment ratings for the index test domains. Likewise, one study included two different culture mediums as their reference standard so there are 27 quality assessment ratings across the reference standard domains.

3.2.3.1 Risk of bias for test accuracy studies

In general, the methodological and reporting quality of the included studies was poor, with risk of bias considered high in two or more domains for 13 studies (50%).[19, 31-33, 35, 37-40, 43, 46-48] No study was at low risk of bias in all four domains.

In 65.4% of studies (17/26)[19, 22, 23, 32, 34, 39, 40, 43, 44, 46, 48, 49, 51, 53-55, 67] it was not clear whether patients were consecutively included or a convenience sample had been chosen and only 15.4% (4/26 studies) [36, 41, 42, 45] were rated as having a low risk of bias in the patient selection domain (domain 1: patient selection). The selection process in the remaining 19.2% (5/26 studies) [31, 33, 38, 47] was rated as high risk of bias with studies clearly reporting convenience samples, case control designs, or having made inappropriate exclusions from the eligible screening population.

The key risks of bias were surrounding how the index test was undertaken (22/30 domain high risk, 74.2%, 22/26 studies).[19, 22, 32, 33, 35-43, 45-48, 52-55, 66] Although all the included studies were on pre-developed tests which had in built thresholds, in many cases, use of the index test required a subjective reading by a clinician (domain 2: index tests). There were further concerns that studies often used the same swab intended for the index test to first streak the agar for biological culture, rather than taking an additional swab sample. Using one swab for

multiple purposes may reduce the amount of the sample and underestimate the accuracy of the test.

Unclear or incomplete reporting was common in the reference standard domain (domain 3: reference standard). In all studies time taken to process the biological culture exceeded that of the rapid test, with biological cultures generally reported 48 hours following sample collection. However, many studies did not state that laboratory staff were blinded to the results of the index test or reference standard (domain 3: reference standard 13/27, 48.1%).[23, 31, 32, 36, 37, 41, 42, 44, 45, 48, 49, 53, 54] There was a high risk of bias in 22.2% of the studies (6/27)[19, 35, 38-40, 46] because the methods of biological culture testing did not match current UK guidelines.[26]

The flow of patients through the studies was rated at high risk of bias in 31% of studies (8/26, domain 4: flow and timing).[31-33, 35, 37, 38, 43, 48] The majority of these (62.5%, 5/8 studies)[31-33, 38, 43, 48] had incomplete testing and made exclusions from the analysis. However, in two of these studies only some patients received the reference standard (partial verification bias). In one study[35] only patients with negative rapid test results received the reference standard and in the other[37] only those with positive rapid test results were given the reference standard. The use of antibiotics was a further concern, with one study directly reporting 61 patients taking antibiotics at the time of testing[31] and 82% (9/10) of unclear ratings were linked to prior/current antibiotic use not being reported.[19, 34, 39, 42, 46, 49, 51, 53-55]

3.2.3.2 **Applicability of study findings for test accuracy studies**

The applicability of study findings was assessed in regards to three domains: patient selection, index test (rapid or molecular test), and reference standard (biological culture). There were significant concerns regarding the applicability of the studies to UK practice for patient selection in 22 of the 26 studies (85%; domain 1: patient selection).[19, 22, 23, 31-34, 37-41, 45-49, 52-55, 66] In the UK the test would only be given following an assessment using a clinical scoring tool such as Centor or FeverPAIN. The rapid test would only be given in people with Centor scores of three or more, and FeverPAIN scores of four or more. In all 22 studies either a clinical scoring tool was not used, or if used, patients were included with scores lower than UK cut-offs and test accuracy data was not reported separately by score. In addition 17 of the 22 studies (77%) [19,

[22, 23, 32, 34, 37, 39, 45-49, 52-55, 66] included children under the age of five. Children under the age of five follow a different clinical pathway due to difference in the presentation of symptoms and difficulties around communication and sample collection.[24] Concerns regarding the applicability of the index test was low for the majority of the studies (21/30 domains, 70%, 18 studies) [19, 22, 23, 31, 33, 34, 36, 38-42, 45-49, 66], with studies reporting that the tests were carried out according to manufacturer's guidelines. The eight remaining studies [32, 35, 37, 43, 51, 53-55] were rated unclear as this was not specified (domain 2: index test). Only four studies (4/27, 14.8%) [35, 37, 39, 40] were rated as having high concern for the applicability with respect to the reference standard (due to deviations from UK guidelines on the undertaking of appropriate culture methods with respect to agar type, incubation period, or atmosphere; domain 3: reference standard).

Table 7: Judgement of risk of bias and applicability of included studies

Study	RISK OF BIAS						APPLICABILITY CONCERNS				
	Patient Selection	Index Test	Additional Index test	Reference Standard	Additional Reference Standard	Flow & Timing	Patient Selection	Index Test	Additional Index test	Reference Standard	Additional Reference Standard
Andersen 2003 [53]	Unclear	High	NA	Unclear	NA	Unclear	High	Unclear	NA	Unclear	NA
Azrad 2019 [31]	High	Low	Low	Unclear	NA	High	High	Low	Low	Unclear	NA
Berry 2018 [19]	Unclear	High	Low	High	NA	Unclear	High	Low	Low	Unclear	NA
Bird 2018 [32]	Unclear	High	NA	Unclear	NA	High	High	Unclear	NA	Unclear	NA
Bura 2017 [33]	High	High	NA	Low	NA	High	High	Low	NA	Low	NA
Cohen 2015 [34]	Unclear	Low	NA	Low	NA	Unclear	High	Low	NA	Low	NA
Dimatteo 2011 [35]	High	High	NA	High	NA	High	Low	Unclear	NA	High	NA
Humair 2006 [36]	Low	High	NA	Unclear	NA	Low	Low	Low	NA	Low	NA
Johansson 2003 [37]	Unclear	High	NA	Unclear	NA	High	High	Unclear	NA	Unclear	NA
Johnson 2001 [38]	High	High	NA	High	NA	High	High	Low	NA	High	NA
Kurtz 2000 [39]	Unclear	High	NA	High	Low	Unclear	High	Low	NA	High	Low
Lacroix 2018 [22]	Unclear	Low	High	Low	NA	Low	High	Low	Low	Low	NA
Lindbaek 2004 [40]	Unclear	High	NA	High	NA	Low	High	Low	NA	High	NA
Llor 2009 [41]	Low	High	NA	Unclear	NA	Low	High	Low	NA	Unclear	NA
Llor 2011 [42]	Low	High	NA	Unclear	NA	Unclear	Low	Low	NA	Low	NA
McIsaac 2004 [43]	Unclear	High	NA	Low	NA	High	Unclear	Unclear	NA	Low	NA
Nerbrand 2002 [44]	Unclear	High	NA	Unclear	NA	Low	High	Low	NA	Low	NA
Pauchard 2013 [54]	Unclear	High	NA	Unclear	NA	Unclear	High	Unclear	NA	Unclear	NA
Penney 2016 [45]	Low	High	NA	Unclear	NA	Low	High	Low	NA	Unclear	NA
Rogo 2011 [46]	Unclear	High	NA	High	NA	Unclear	High	Low	NA	Unclear	NA
Rosenberg 2002 [47]	High	High	NA	Low	NA	Low	High	Low	NA	Low	NA
Santos 2003 [48]	Unclear	High	NA	Unclear	NA	High	High	Low	NA	Low	NA

Study	RISK OF BIAS						APPLICABILITY CONCERNS				
	Patient Selection	Index Test	Additional Index test	Reference Standard	Additional Reference Standard	Flow & Timing	Patient Selection	Index Test	Additional Index test	Reference Standard	Additional Reference Standard
Stefaniuk 2017 [49]	Unclear	Low	NA	Unclear	NA	Unclear	High	Low	NA	Low	NA
Valverde 2018 [55]	Unclear	High	NA	Low	NA	Unclear	High	Unclear	NA	Low	NA
Wang 2017 [23]	Unclear	Low	NA	Unclear	NA	Low	High	Low	NA	Unclear	NA
Weinzierl 2018 [51]	Unclear	Low	High	Low	NA	Unclear	High	Unclear	Unclear	Low	NA

NA – not applicable (the study did not have this additional test)

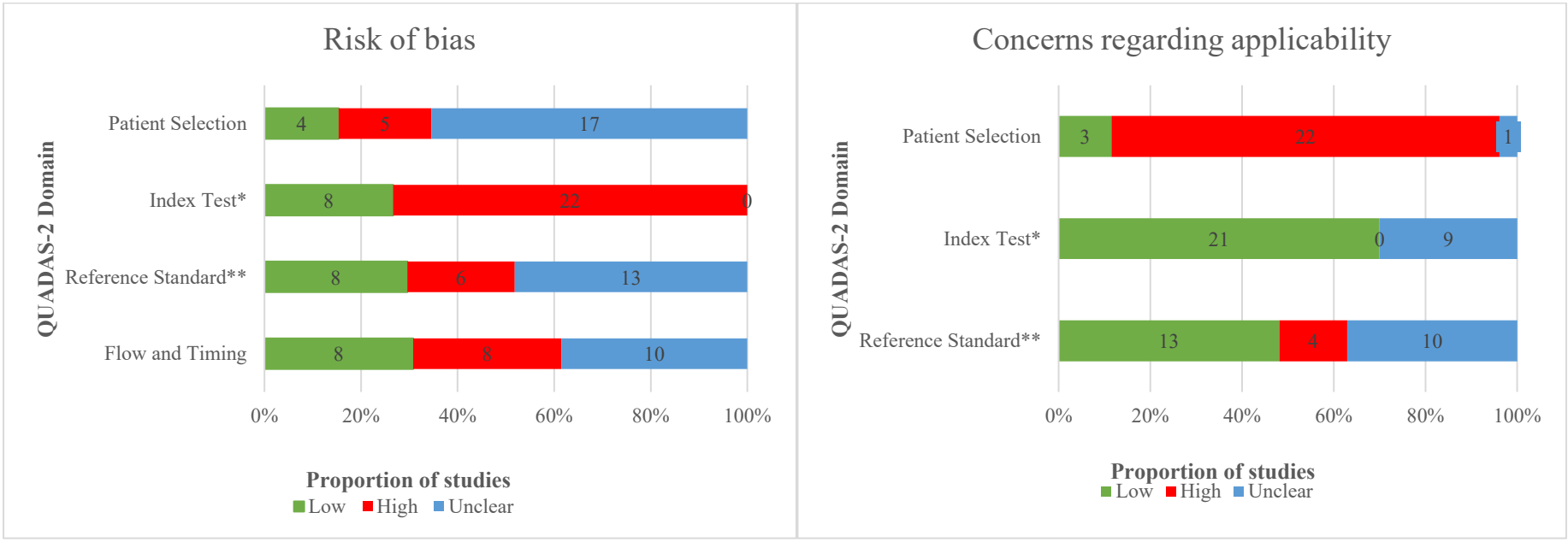


Figure 5: Concerns regarding bias and applicability in included studies

* Four studies included two index tests relevant to this review. ** One study included two reference standards (culture methods) relevant to this review

3.2.4 Assessment of studies of prescribing behaviour and clinical outcomes

There were 12 studies which reported on antibiotic prescribing behaviour. [6, 19, 32, 33, 36, 37, 42, 43, 47, 49, 50, 52] Of these, three studies were RCTs[6, 42, 52] and were quality appraised using the Cochrane ROB tool for RCTs.[27] There were six studies (including one before and after study) which were single arm cohorts and have been appraised using the JBI Critical Appraisal checklist [28] ENREF 28 for analytical cross-sectional studies.[19, 33, 37, 47, 49, 68] The remaining 3 studies were either one-armed cohort studies using pre-determined guidelines to hypothetically estimate prescribing behaviour and offer no information on what happened in the real world or on what Clinicians would do.[36, 43, 50] These studies were not quality appraised and have been briefly summarised later in the results (section 3.2.8.3).

3.2.4.1 **RCTs**

Risk of Bias of the included trials is shown in Figure 6 and Table 8. The domains regarding blinding were removed as we were interested in test-treat trials measuring prescribing decisions with and without rapid tests. Therefore clinicians could not be blinded to test results, and we considered blinding to which exact test was used to be unnecessary in this context. In general, the methodological quality of the RCTs was fair, with all studies having at least one domain rated as unclear. There was unclear risk of bias in four domains across the three studies (random sequence generation, allocation concealment, incomplete outcome data and selective outcome reporting). This was due to insufficient information presented on which to make an assessment. The remaining applicable domains were judge to be at low risk of bias.

Table 8: Judgement of risk of bias of included randomised controlled trials

Study	Random sequence generation	Allocation concealment	Incomplete outcome data	Selective outcome reporting	Other bias
Little 2013[6]	Low	Low	Unclear	Unclear	Low
Llor 2011[42]	Low	Low	Low	Unclear	Low
Worrall 2007[52]	Unclear	Unclear	Low	Unclear	Low

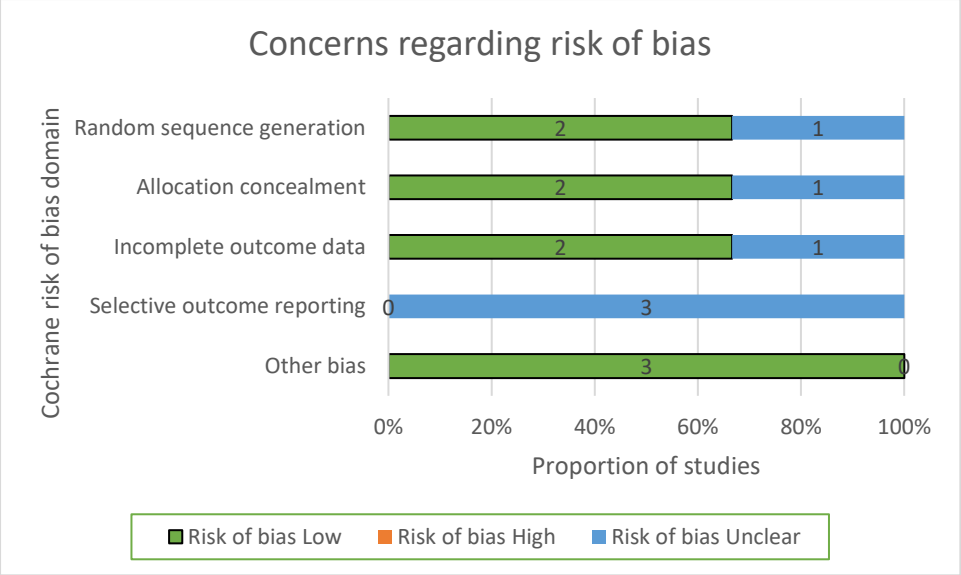


Figure 6 Concerns regarding the risk of bias of included randomised controlled trials

3.2.4.2 Cohort studies

Risk of bias in the included cohort studies is shown in Figure 7 and Table 9. No study had high methodological quality across all areas. There was low methodological quality regarding criteria for inclusion in 83% of studies (5/6) and details regarding the study subjects in 33% of studies (2/6).[19, 32, 37, 47, 49] These studies reported the details of the patients, but provided no information on the details of those who are making the prescribing decisions. The outcomes of interest in these studies were prescribing behaviour. The measurement of prescribing behaviour considered to be valid and reliable was recording in medical records, only 33% of the studies clearly reported this.[19, 33] A confounder in the studies was current antibiotic use. 33% (2/6) studies did not clearly specify current or recent antibiotic use as an excluding factor.

Table 9: Judgement of risk of bias of included non-RCT studies

Study	Were the criteria for inclusion in the sample clearly defined?	Were the study subjects and the setting described in detail?	Was the exposure measured in a valid and reliable way?	Were objective, standard criteria used for measurement of the condition?	Were confounding factors identified?	Were strategies to deal with confounding factors stated?	Were the outcomes measured in a valid and reliable way?	Was statistical analysis appropriate?
Berry 2018 [19]	No	Yes	Yes	No	Unclear	Unclear	Yes	Yes
Bird 2018 [32]	No	Yes	Yes	Yes	Yes	Yes	No	Yes
Bura 2017 [33]	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Johansson 2003 [37]	No	No	Yes	Unclear	Yes	Yes	No	Yes
Rosenberg 2002 [47]	No	Yes	Yes	Yes	Yes	Yes	Unclear	Yes
Stefaniuk 2017 [49]	No	No	Yes	Yes	Unclear	Unclear	Unclear	Yes

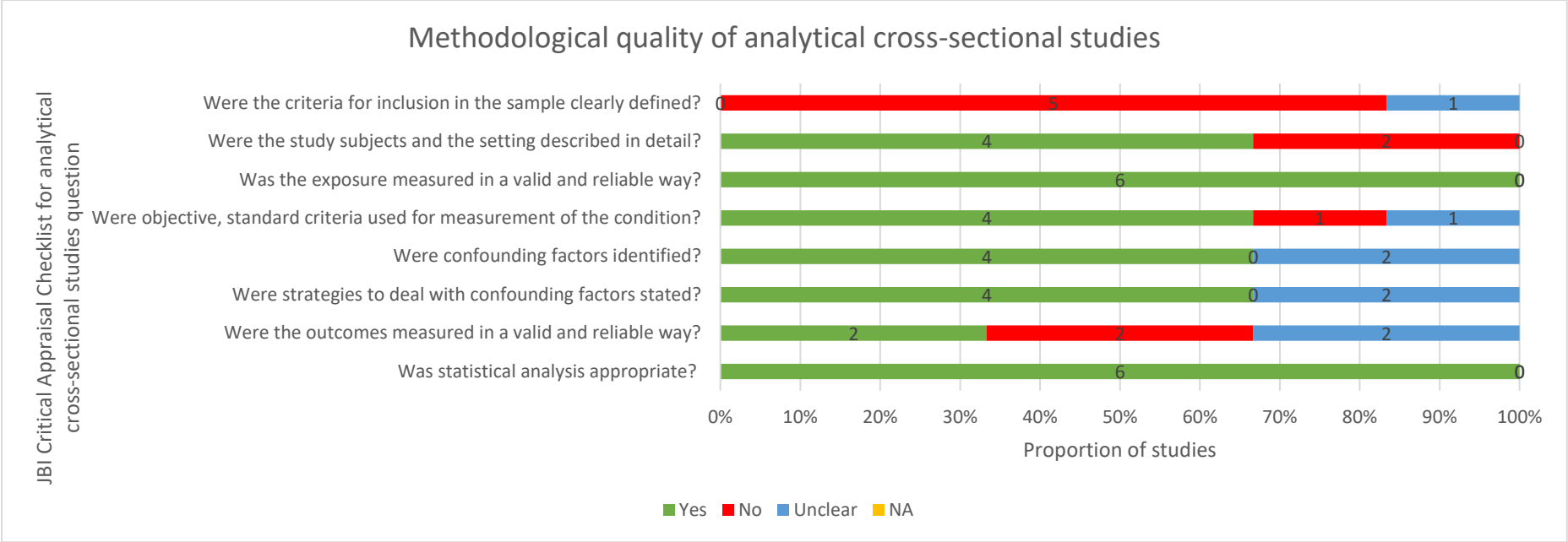


Figure 7: Methodological quality of included analytical cross-sectional studies

3.2.5 Current pathway (clinical scoring tools only)

3.2.5.1 Accuracy of clinical scoring tools with culture as reference standard

Accuracy statistics for Centor[36, 41, 42, 49] and McIsaac scores[43, 49, 54] with microbiological culture as the reference standard are presented in Table 10. The results show wide variations in the test accuracy of sore throat clinical scoring tools. Specificity point estimates were reported between 17.2% and 64.8%, while sensitivity point estimates were reported between 73.5% and 97.2%. This suggests that these tools might be better at identifying people who do have Streptococcus than they are at identifying people who do not.

Rosenberg et al. (2002)[47] and Johansson et al. (2003)[37] also reported accuracy statistics for sore throat symptoms with culture as reference standard. However, these studies provided insufficient data to construct 2x2 contingency tables using the recommended clinical scoring threshold (section 3.1.4). The use of different clinical scoring tools, age selection criteria, clinical score inclusion criteria, and settings across the seven contributing studies precluded any pooling.

3.2.5.2 Accuracy of clinical scoring tools split by age group

Two [43, 49] of the six studies included a mixed population of adults and children. In the study by McIsaac et al. (2004)[43] threshold of >2 (Modified Centor/McIsaac score) produced a sensitivity estimate of 88.4% (95% Confidence Interval (CI) 82.0 to 92.8) and a specificity estimate of 23.4% (95% CI 18.8 to 28.7) in children aged 3 to 17 years, and sensitivity estimate of 76.7% (95% CI 65.1 to 85.8) and a specificity estimate of 43.9% (95% CI 37.8 to 50.1) in adults aged 18 years or older.

In the study by Stefaniuk et al. (2017)[49] a threshold of >2 (Modified Centor/McIsaac score) produced a sensitivity estimate of 100% (95% CI 80.0 to 100.0), and a specificity of 8.3% (95% CI 1.5 to 28.5) in children aged 1 to 14 years, and a sensitivity estimate of 73.9% (95% CI 51.3 to 88.9), and a specificity estimate of 41.4% (95% CI 24.1 to 60.9) in participants aged 15 years and older. As previously discussed (section 3.2.2.1), this overlap across age groups potentially limits subgroup analysis. However, none of the other six studies included patients under 14 years of age.

3.2.5.3 Accuracy of clinical scoring tools split by primary/secondary care setting

Patients were recruited from primary care settings in five [36, 41-43, 49] of the six studies. Details of the primary care setting studies are outlined in Table 10. In brief, these studies provided point estimates of sensitivity of 0.74 – 0.86 and specificity of 0.25 – 0.65. The single study

from a secondary care setting [54] reported a higher point estimates for sensitivity (97.2% [95% CI 89.3 to 99.5]) and a lower point estimate for specificity (17.2% [95% CI 11.2% to 25.3%]), albeit with overlapping confidence intervals with some of the primary care setting studies, compared to the other five studies. This may have been due to setting or other sources of heterogeneity between studies.

Table 10: Accuracy of clinical scores with culture as the reference standard

Study Reference	GAS Prevalence (%)	Setting	Clinical score	Test Accuracy Statistics				
				Culture +	culture -	Total	Sensitivity	Specificity
Humair 2006 [36]	46.9	Primary care/GP clinic	Centor score ≥3	105	119	224	0.750 (0.678 to 0.822)	0.487 (0.423 to 0.551)
			Centor score <3	35	113	148		
			Total	140	232	372		
Llor 2009 [41]	21.2	Primary care/GP clinic	Centor score ≥3	47	104	151	0.855 (0.761 to 0.948)	0.377 (0.304 to 0.451)
			Centor score <3	8	63	71		
			Total	55	167	222		
Llor 2011 [42]	16.7	Primary care/GP clinic	Centor score ≥3	36	80	116	0.735 (0.587 to 0.846)	0.648 (0.581 to 0.709)
			Centor score <3	13	147	160		
			Total	49	227	276		
McIsaac 2004 [43]	29	Primary care/GP clinic	McIsaac score ≥3	193	375	568	0.846 (0.800 to 0.893)	0.329 (0.290 to 0.368)
			McIsaac score <3	35	184	219		
			Total	228	559	787		
Pauchard 2013 [54]	37	Hospital	McIsaac score ≥3	69	101	170	0.972 (0.893 to 0.995)	0.172 (0.112 to 0.253)
			McIsaac score <2	2	21	23		
			Total	71	122	193		
Stefaniuk 2017 [49]	22.4	Primary care/GP clinic	Centor/McIsaac score ≥3	37	39	76	0.861 (0.714 to 0.942)	0.250 (0.145 to 0.392)
			Centor/McIsaac score <2	6	13	19		
			Total	43	52	95		

3.2.5.4 Accuracy of clinical scoring tools using PCR to resolve discordant cases

No analysis of discordant results between sore throat clinical scores and microbiological culture was undertaken in any of the included studies.

3.2.6 Point-of-care/index tests

3.2.6.1 Accuracy of point-of-care tests with culture as reference standard

The systematic review identified 35 pieces of literature that provided evidence comparing the performance of 18 of the named index tests to culture. These included 23 peer-reviewed papers, 3 abstracts, 5 manufacturer responses (submitted directly to NICE in response to a request for information) and 4 FDA reports. Two studies reported results that were inconsistent which prevented the construction of a reliable 2x2 table and were excluded during the data extraction.[32, 44] A summary of the final 33 pieces of literature can be found in Table 11. The sources provided by the manufacturers were not peer-reviewed, and neither were three abstracts. The sources identified from FDA reports received some scrutiny from the FDA. The remaining 21 studies were published in peer-reviewed journals. All sensitivity and specificity estimates are presented alongside their 95% confidence interval. Meta-analyses were performed where appropriate, a summary of which can be found in Figure 9.

Clearview Exact Strep A Cassette and Clearview Exact Strep A Dipstick (Abbott)

The only evidence related to the Clearview Exact Strep A Cassette and Dipstick was provided by Andersen et al. [53], which did not report which version of the test they used. Andersen et al. reported a sensitivity of 0.68 (95% CI 0.55, 0.81) and a specificity of 0.95 (95% CI 0.93, 0.98) when examining children presenting in a secondary care setting.

BD Veritor Plus System (Beckton Dickinson)

Azrad et al. [31] and Berry et al. [19] both presented results for the BD Veritor Plus System compared to culture in a secondary care setting. Azrad et al. did not report the age group, whilst Berry et al. looked at children. The sensitivities of the test were 0.80 (0.59, 0.92) and

0.76 (0.60, 0.87), and the specificities were 0.79 (0.67, 0.87) and 0.94 (0.89, 0.97), by Azrad et al. and Berry et al. respectively. Beckton Dickinson provided data to the FDA which estimated a sensitivity of 0.97 (0.92, 0.99), and a specificity of 0.96 (0.94, 0.97).[56]

Univariate models were fitted to the two studies for the BD Veritor Plus System. The models estimated a sensitivity of 0.78 (0.67, 0.87), and a specificity of 0.90 (0.86, 0.93).

Strep A Rapid Test Cassette (Biopanda)

The only evidence related to the Strep A Rapid Test Cassette was provided by Biopanda in response to a request for information by NICE. They reported a sensitivity of 0.95 (0.89, 0.98) and a specificity of 0.98 (0.96, 0.99), in a population of children and adults in a primary care setting.

NADAL Strep A Strip, NADAL Strep A Cassette, NADAL Strep A Plus Cassette, NADAL Strep A Plus Strip and NADAL Strep A Scan (nal von minden GmbH)

The only evidence related to the NADAL Strep A Cassettes, Strips and Scan tests was provided by nal von minden GmbH in response to a request for information by NICE, and did not distinguish between any of the NADAL varieties. They reported a sensitivity of 0.98 (0.91, 1.00) and a specificity of 0.98 (0.93, 0.99) from a study performed in a secondary care setting including both children and adults.

OSOM Strep A Strip (Sekisui)

Five studies compared the OSOM Strep A Strip to culture. [33, 41, 42, 46, 51] Bura et al., [33] Llor et al. [41] and Llor et al.[42] all examined adult populations presenting at primary care centres and reported sensitivities of 0.96 (0.76, 1.00), 0.95 (0.85, 0.99) and 0.90 (0.78, 0.97), and specificities of 0.97 (0.90, 1.00), 0.92 (0.86, 0.95) and 0.94 (0.90, 0.97), respectively.

Meanwhile Rogo et al.[46] and Weinzierl et al.[51] examined children in secondary care, with respective sensitivities of 0.98 (0.91, 1.00) and 0.89 (0.77, 0.95), and specificities of 0.99 (0.96, 1.00) and 0.91 (0.83, 0.96).

Despite having five sources of data, a bivariate model failed to converge for the OSOM test. However, univariate models did converge. These models estimated a sensitivity of 0.94 (0.89, 0.98) and a specificity of 0.95 (0.91, 0.98).

QuikRead Go Strep A Kit (Orion)

Azrad et al.[31] and Stefaniuk et al.[49] both compared the accuracy of the QuikRead Go Strep A Kit to culture, and reported respective sensitivities of 0.80 (0.59, 0.92) and 0.91 (0.78, 0.97), and specificities of 0.73 (0.62, 0.83) and 0.85 (0.72, 0.93). Azrad et al. investigated both children and adult patients in a primary care setting, whereas the data from Stefaniuk et al. reflected a secondary care setting but did not report the ages of the patients. Orion also provided data from their own study in response to a request for information by NICE, which estimated a sensitivity 0.83 (0.73, 0.90) and a specificity of 0.97 (0.93, 0.99) for children and adults in primary care.

Univariate models were fitted to the two studies that investigated the QuikRead Go test. The resulting sensitivity was 0.87 (0.78, 0.95) and the specificity was 0.78 (0.71, 0.85).

Alere TestPack Plus Cassette (Abbott)

There were twelve published studies which compared the accuracy of the Alere TestPack Plus Cassette to culture. Two studies did not report sufficient data to estimate a complete 2x2 table.[35, 38] Four of the remaining studies were conducted in a primary care setting. One of these was in an adult population: Humair et al.[36] estimated a sensitivity of 0.91 (0.86, 0.95), and specificity of 0.95 (0.92, 0.98). The other primary care-based studies combined children and adult populations: Lindbaek et al.[40] (sensitivity: of 0.94 [0.90, 0.99], specificity: 0.86 [0.80, 0.91]), Johannson et al.[37] (sensitivity: 0.87 [0.74, 0.94], specificity: 0.96 [0.89, 0.99]) and (sensitivity: 0.83 [0.77, 0.87], specificity: 0.99 [0.98, 1.00]).

Six other studies were conducted in secondary care settings, three of which assessed children without any restriction from a clinical tool score.[39, 45, 48] Kurtz et al., Penney et al. and Santos et al.[39, 45, 48] reported sensitivities of 0.80 (0.71, 0.89), 0.76 (0.65, 0.87) and 0.73 (0.45, 0.91)

respectively. Their specificities were 0.93 (0.89, 0.97), 1.00 (0.95, 1.00) and 0.94 (0.79, 0.99). Lacroix et al.[22] also examined children in secondary care, but restricted the study population to having a McIsaac score ≥ 2. [22] The sensitivity was 0.76 (0.71, 0.80) and specificity was 0.97 (0.95, 0.98). Two studies examined both children and adults in secondary care. Rosenberg et al. estimated a sensitivity of 0.75 (0.56, 0.88) and a specificity of 0.99 (0.93, 1.00), whilst Valverde et al. estimated 0.92 (0.87, 0.95) and 0.93 (0.90, 0.95) respectively. [47, 55]

For the Alere TestPack Plus test a bivariate model was fitted to meta-analyse all studies. The model suggested the test had a sensitivity of 0.85 (0.79, 0.90) and a specificity of 0.96 (0.94, 0.98). Univariate models were also investigated and were identical to two decimal places.

Bionexia Strep A Dipstick (Biomerieux)

Only one abstract presented data for Biomerieux's Bionexia Strep A Dipstick. Pauchard et al. performed a study in children in a secondary care setting, and estimated a sensitivity of 0.85 (0.74, 0.92) and a specificity of 0.91 (0.84, 0.95).[54]

Sofia Strep A FIA (Quidel)

One peer-reviewed study presented data comparing the Sofia Strep A FIA to culture. Lacroix et al.[22] used the test on children in secondary care, and estimated a sensitivity of 0.85 (0.81, 0.89) and a specificity of 0.95 (0.93, 0.97). Quidel also provided data from their own study to the FDA, which estimated a sensitivity of 0.91 (0.84, 0.95) and a specificity of 0.96 (0.94, 0.97).[57]

Alere i Strep A (Abbott)

Three studies compared the Alere i Strep A test to culture, all in a secondary care setting. Berry et al.[19] and Weinzierl et al.[51] looked only at children, and estimated respective sensitivities of 1.00 (0.90, 1.00) and 0.98 (0.90, 1.00), and specificities of 0.91 (0.86, 0.95) and 1.00 (0.95, 1.00). Cohen 2015[34] examined in both children and adults, and produced respective estimates of sensitivity and specificity of 0.96 (0.91, 0.98) and 0.95 (0.91, 0.97).

When meta-analysed, the three studies using the Alere i Strep A test yielded a sensitivity of 0.98 (0.95, 1.00), and a specificity of 0.96 (0.90, 1.00).

Alere i Strep A 2 (Abbott)

Only manufacturer information submitted to the FDA was available for the Alere i Strep A 2 test, which reported a sensitivity of 0.98 (0.95, 0.99) and a specificity of 0.93 (0.91, 0.95), but did not report the age of patients or the care setting.[58]

Cobas Strep A Assay on Liat (Roche)

There were two sources of data comparing the Cobas Strep A Assay on Liat test to culture. Wang et al.[23] performed the test in children in a primary care setting, and estimated a sensitivity of 0.98 (0.93, 0.99) and a specificity of 0.94 (0.91, 0.96). The manufacturer (Roche) provided the other source in response to a request for information by NICE, which produced estimates of sensitivity and specificity of 0.98 (0.95, 1.00) and 0.94 (0.91, 0.96), respectively. Roche stated that the data they provided overlapped with the Wang et al. study. The data supplied by Roche were identical to the data available from the FDA for this test.

Xpert Xpress Strep A (Cepheid)

Only manufacturer information was available for the Xpert Xpress Strep A test by Cepheid, which was provided in response to a request for information by NICE. The data provided by the manufacturer reported a sensitivity of 1.00 (0.97, 1.00), and a specificity of 0.94 (0.91, 0.96). This differed slightly from the information available from the FDA which had a sensitivity of 0.99 (0.96, 1.00) and a specificity of 0.94 (0.91, 0.96).[59] Due to the differences in sample size and the resolution of discordant samples, we have treated these sources as two independent studies, but it is not clear if there is overlap in patients.

Biosynex Strep A Cassette (Biosynex), Strep A Rapid Test Strip (Biopanda) and Bionexia Strep A Plus Cassette (Biomerieux)

No data were identified for any of the following tests:

- Biosynex Strep A Cassette test (Biosynex)
- Bionexia Strep A Plus Cassette test. (Biomerieux)

- Strep A Rapid Test Strip (Biopanda)

Figure 7 and Figure 8 present the sensitivity and specificity for all studies that had complete 2x2 data. Data were only available for eighteen tests, and just seven tests were used in more than one independent study. Ignoring manufacturer and FDA sources of data, this reduces to ten tests with published data and five tests having more than one independent study.

Note that where studies provided performance data by subgroups, this was incorporated into the relevant analyses when producing estimates to feed into the cost-effectiveness modelling. It is clear that there is a large degree of heterogeneity between the studies, and it is difficult to attribute any observed differences in test performance to the tests themselves. The confidence intervals in the figures may differ slightly to the table, due to differences in their method of calculation.

It is apparent that the data sourced from the manufacturer responses (submitted directly to NICE in response to a request for information) and FDA submissions consistently provided higher estimates of sensitivity and specificity than the peer-reviewed studies. This support the view that the manufacturer data may be at high risk of bias, and any cost-effectiveness analyses incorporating them may be unreliable.

Table 11: Summary of available evidence by test

Study	Care Setting	Age group	Clinical Tool Score Restriction	Strep A infections Prevalence	Ref Type	N	TP	FN	FP	TN	Accuracy Data
Clearview Exact Strep A Cassette - Abbott											
Andersen 2003 (abstract) [53]	Secondary	Children	None	15.0%	NR	353	36	17	15	285	Sens = 0.68 (0.55, 0.81) Spec = 0.95 (0.93, 0.98) PPV = 0.71 (0.58, 0.83) NPV = 0.94 (0.92, 0.97)
Clearview Exact Strep A Dipstick - Abbott											
Andersen 2003 [53]	Secondary	Children	None	15.0%	NR	353	36	17	15	285	Sens = 0.68 (0.55, 0.81) Spec = 0.95 (0.93, 0.98) PPV = 0.71 (0.58, 0.83) NPV = 0.94 (0.92, 0.97)
BD Veritor Plus System - Beckton Dickinson											
Azrad 2019 [31]	Secondary	NR	None	25.0%	Strep selective agar	100	20	5	16	59	Sens = 0.80 (0.59, 0.92) Spec = 0.79 (0.67, 0.87) PPV = 0.56 (0.38, 0.72) NPV = 0.92 (0.82, 0.97)
Beckton Dickinson [FDA] [56] *	NR	Children and Adults	None	18.7%	Blood agar	796	144	5	29	618	Sens = 0.97 (0.92, 0.99) Spec = 0.96 (0.94, 0.97) PPV = 0.83 (0.77, 0.88) NPV = 0.99 (0.98, 1.00)
Berry 2018 [19]	Secondary	Children	None	19.5%	Blood agar	215	32	10	11	162	Sens = 0.76 (0.60, 0.87) Spec = 0.94 (0.89, 0.97) PPV = 0.74 (0.56, 0.86) NPV = 0.94 (0.89, 0.97)
Strep A Rapid Test Cassette - Biopanda											
Biopanda [MFR] *	Primary	Children and Adults	None	23.2%	Blood agar	526	116	6	9	395	Sens = 0.95 (0.89, 0.98) Spec = 0.98 (0.96, 0.99) PPV = 0.85 (0.79,0.90) NPV = 0.99 (0.97, 0.99)
Strep A Rapid Test Strip - Biopanda											
(no data)											

Study	Care Setting	Age group	Clinical Tool Score Restriction	Strep A infections Prevalence	Ref Type	N	TP	FN	FP	TN	Accuracy Data
NADAL Strep A Strip - nal von minden GmbH											
nal von minden GmbH [MFR] *	Secondary	Children and Adults	None	34.4%	Blood agar	244	82	2	4	156	Sens = 0.98 (0.91, 1.00) Spec = 0.98 (0.93, 0.99) PPV = 0.95 (0.88, 0.99) NPV = 0.99 (0.95, 1.00)
NADAL Strep A Cassette - nal von minden GmbH											
nal von minden GmbH [MFR] *	Secondary	Children and Adults	None	34.4%	Blood agar	244	82	2	4	156	Sens = 0.98 (0.91, 1.00) Spec = 0.98 (0.93, 0.99) PPV = 0.95 (0.88, 0.99) NPV = 0.99 (0.95, 1.00)
NADAL Strep A Plus Cassette - nal von minden GmbH											
nal von minden GmbH [MFR] *	Secondary	Children and Adults	None	34.4%	Blood agar	244	82	2	4	156	Sens = 0.98 (0.91, 1.00) Spec = 0.98 (0.93, 0.99) PPV = 0.95 (0.88, 0.99) NPV = 0.99 (0.95, 1.00)
NADAL Strep A Plus Strip - nal von minden GmbH											
nal von minden GmbH [MFR] *	Secondary	Children and Adults	None	34.4%	Blood agar	244	82	2	4	156	Sens = 0.98 (0.91, 1.00) Spec = 0.98 (0.93, 0.99) PPV = 0.95 (0.88, 0.99) NPV = 0.99 (0.95, 1.00)
NADAL Strep A Scan - nal von minden GmbH											
nal von minden GmbH [MFR] *	Secondary	Children and Adults	None	34.4%	Blood agar	244	82	2	4	156	Sens = 0.98 (0.91, 1.00) Spec = 0.98 (0.93, 0.99) PPV = 0.95 (0.88, 0.99) NPV = 0.99 (0.95, 1.00)
OSOM Strep A Strip - Sekisui											
Bura 2017 [33]	Primary	Adults	Centor ≥ 2	22.7%	Blood agar	101	22	1	2	76	Sens = 0.96 (0.76, 1.00) Spec = 0.97 (0.90, 1.00) PPV = 0.92 (0.72, 0.99) NPV = 0.99 (0.92, 1.00)
Llor 2009 [41]	Primary	Adults	Centor ≥ 2	24.8%	Blood agar	222	52	3	14	153	Sens = 0.95 (0.85, 0.99) Spec = 0.92 (0.86, 0.95)

Study	Care Setting	Age group	Clinical Tool Score Restriction	Strep A infections Prevalence	Ref Type	N	TP	FN	FP	TN	Accuracy Data
											PPV = 0.79 (0.69, 0.86) NPV = 0.98 (0.94, 0.99)
Llor 2011 [42]	Primary	Adults	Centor ≥ 2 #	17.8%	Blood agar	276	44	5	14	213	Sens = 0.90 (0.78, 0.97) Spec = 0.94 (0.90, 0.97) PPV = 0.76 (0.65, 0.84) NPV = 0.98 (0.95, 0.99)
Rogo 2011 [46]	Secondary	Children	None	28.9%	Blood agar	228	65	1	1	161	Sens = 0.98 (0.91, 1.00) Spec = 0.99 (0.96, 1.00) PPV = 0.98 (0.91, 1.00) NPV = 0.99 (0.96, 1.00)
Weinzierl 2018 [51]	Secondary	Children	None	38.1%	Blood agar	160	54	7	9	90	Sens = 0.89 (0.77, 0.95) Spec = 0.91 (0.83, 0.96) PPV = 0.86 (0.74, 0.93) NPV = 0.93 (0.85, 0.97)
QuikRead Go Strep A Kit - Orion											
Azrad 2019 [31]	Secondary	NR	None	25.0%	Strep selective agar	100	20	5	20	55	Sens = 0.80 (0.59, 0.92) Spec = 0.73 (0.62, 0.83) PPV = 0.50 (0.34, 0.66) NPV = 0.92 (0.81, 0.97)
Orion [MFR] *	Primary	Children and Adults	None	32.8%	Strep selective agar	271	74	15	5	177	Sens = 0.83 (0.73, 0.90) Spec = 0.97 (0.93, 0.99) PPV = 0.95 (0.88, 0.99) NPV = 0.92 (0.87, 0.95)
Stefaniuk 2017 [49]	Primary	Children and Adults #	None	45.3%	Blood agar	95	39	4	8	44	Sens = 0.91 (0.78, 0.97) Spec = 0.85 (0.72, 0.93) PPV = 0.83 (0.72, 0.90) NPV = 0.92 (0.81, 0.97)
Alere TestPack Plus Cassette - Abbott											
DiMatteo 2001 [35]	Secondary	Adults	Centor ≥ 1	NR	Strep selective agar	NR	NR	22	NR	361	NPV = 0.94 (0.91, 0.96)

Study	Care Setting	Age group	Clinical Tool Score Restriction	Strep A infections Prevalence	Ref Type	N	TP	FN	FP	TN	Accuracy Data
Humair 2006 [36]	Primary	Adults	Centor ≥ 2 #	37.6%	Blood agar	372	128	12	11	221	Sens = 0.91 (0.86, 0.95) Spec = 0.95 (0.92, 0.98) PPV = 0.92 (0.87, 0.95) NPV = 0.95 (0.91, 0.97)
Johannson 2003 [37]	Primary	Children and Adults	None	31.4%	NR	144	46	7	4	87	Sens = 0.87 (0.74, 0.94) Spec = 0.96 (0.89, 0.99) PPV = 0.92 (0.80, 0.97) NPV = 0.93 (0.85, 0.97)
Johnson 2001 [38]	Primary	Adults	None	NR	Blood agar	NR	445	NR	77	NR	PPV = 0.85 (0.82, 0.88)
Kurtz 2000 [39]	Secondary	Children	None	31.1%	Blood agar	257	64	16	13	164	Sens = 0.80 (0.71, 0.89) Spec = 0.93 (0.89, 0.97) PPV = 0.83 (0.75, 0.92) NPV = 0.91 (0.87, 0.95)
Lacroix 2018 [22]	Secondary	Children	McIsaac ≥ 2	35.7%	Blood agar	1002	271	87	21	623	Sens = 0.76 (0.71, 0.80) Spec = 0.97 (0.95, 0.98) PPV = 0.93 (0.89, 0.95) NPV = 0.88 (0.85, 0.90)
Lindbaek 2004 [40]	Primary	Children and Adults	None	35.9%	Strep selective agar	306	106	4	27	169	Sens = 0.94 (0.90, 0.99) Spec = 0.86 (0.80, 0.91) PPV = 0.80 (0.72, 0.86) NPV = 0.98 (0.94, 0.99)
McIsaac 2004 [43]	Primary	Children and Adults #	McIsaac ≥ 2	29.0%	Blood agar	787	189	39	5	554	Sens = 0.83 (0.77, 0.88) Spec = 0.99 (0.98, 1.00) PPV = 0.97 (0.94, 0.99) NPV = 0.93 (0.91, 0.95)
Penney 2016 [45]	Secondary	Children	None	40.1%	Strep selective agar	147	45	14	0	88	Sens = 0.76 (0.65, 0.87) Spec = 1.00 (0.95, 1.00) PPV = 1.00 (0.90, 1.00) NPV = 0.86 (0.78, 0.92)
Rosenberg 2002 [47]	Secondary	Children and Adults	None	25.4%	Blood agar	126	24	8	1	93	Sens = 0.75 (0.56, 0.88) Spec = 0.99 (0.93, 1.00) PPV = 0.96 (0.78, 1.00)

Study	Care Setting	Age group	Clinical Tool Score Restriction	Strep A infections Prevalence	Ref Type	N	TP	FN	FP	TN	Accuracy Data
											NPV = 0.92 (0.85, 0.96)
Santos 2003 [48]	Secondary	Children	None	30.6%	Blood agar	49	11	4	2	32	Sens = 0.73 (0.45, 0.91) Spec = 0.94 (0.79, 0.99) PPV = 0.85 (0.54, 0.97) NPV = 0.89 (0.73, 0.96)
Valverde 2018 (abstract) [55]	Secondary	Children and Adults	None	40.0%	Blood agar	580	181	16	27	356	Sens = 0.92 (0.87, 0.95) Spec = 0.93 (0.90, 0.95) PPV = 0.87 (0.82, 0.91) NPV = 0.96 (0.93, 0.97)
Bionexia Strep A Plus Cassette - Biomerieux											
(no data)											
Bionexia Strep A Dipstick - Biomerieux											
Pauchard 2013 (abstract) [54]	Secondary	Children	None	36.8%	NR	193	60	11	11	111	Sens = 0.85 (0.74, 0.92) Spec = 0.91 (0.84, 0.95) PPV = 0.85 (0.76, 0.93) NPV = 0.91 (0.86, 0.96)
Biosynex Strep A Cassette											
(no data)											
Sofia Strep A FIA - Quidel											
Lacroix 2018 [22]	Secondary	Children	McIsaac ≥ 2	35.7%	Blood agar	1002	305	53	31	613	Sens = 0.85 (0.81, 0.89) Spec = 0.95 (0.93, 0.97) PPV = 0.91 (0.87, 0.94) NPV = 0.92 (0.90, 0.94)
Quidel [FDA] * [57]	NR	NR	None	17.4%	Blood agar	736	116	12	24	584	Sens = 0.91 (0.84, 0.95) Spec = 0.96 (0.94, 0.97) PPV = 0.83 (0.75, 0.89) NPV = 0.98 (0.96, 0.99)
Alere i Strep A - Abbott											
Berry 2018 [19]	Secondary	Children	None	19.5%	Blood agar	215	42	0	15	158	Sens = 1.00 (0.90, 1.00) Spec = 0.91 (0.86, 0.95) PPV = 0.74 (0.60, 0.84) NPV = 1.00 (0.97, 1.00)

Study	Care Setting	Age group	Clinical Tool Score Restriction	Strep A infections Prevalence	Ref Type	N	TP	FN	FP	TN	Accuracy Data
Cohen 2015 [34]	Secondary	Children and Adults #	None	30.3%	Blood agar	481	141	6	18	316	Sens = 0.96 (0.91, 0.98) Spec = 0.95 (0.91, 0.97) PPV = 0.89 (0.82, 0.93) NPV = 0.98 (0.96, 0.99)
Weinzierl 2018 [51]	Secondary	Children	None	38.1%	Blood agar	160	60	1	0	99	Sens = 0.98 (0.90, 1.00) Spec = 1.00 (0.95, 1.00) PPV = 1.00 (0.93, 1.00) NPV = 0.99 (0.94, 1.00)
Alere i Strep A 2 - Abbott											
Abbott [FDA] * [58]	NR	NR	None	20.2%	Blood agar	981	195	3	52	731	Sens = 0.98 (0.95, 0.99) Spec = 0.93 (0.91, 0.95) PPV = 0.79 (0.73, 0.84) NPV = 1.00 (0.99, 1.00)
Cobas Strep A Assay on Liat - Roche											
Roche [MFR] *	NR	Children and Adults	None	30.4%	Blood agar	570	170	3	23	374	Sens = 0.98 (0.95, 1.00) Spec = 0.94 (0.91, 0.96) PPV = 0.88 (0.82, 0.92) NPV = 0.99 (0.97, 1.00)
Wang 2017 [23]	Primary	Children	Centor ≥ 1	30.2%	NR	427	126	3	20	278	Sens = 0.98 (0.93, 0.99) Spec = 0.93 (0.90, 0.96) PPV = 0.86 (0.79, 0.91) NPV = 0.99 (0.97, 1.00)
Xpert Xpress Strep A - Cepheid											
Cepheid [MFR] *	NR	NR	None	23.9%	NR	577	138	0	26	413	Sens = 1.00 (0.97, 1.00) Spec = 0.94 (0.91, 0.96) PPV = 0.84 (0.79, 0.89) NPV = 1.00 (0.99, 1.00)
Cepheid [FDA] * [59]	Primary and Secondary	Children and Adults	None	25.6%	NR	618	157	1	27	433	Sens = 0.99 (0.96, 1.00) Spec = 0.94 (0.91, 0.96) PPV = 0.85 (0.79, 0.90) NPV = 1.00 (0.99, 1.00)

Key:

* indicates submission was provided by company, and data were not included in primary meta-analysis.

# indicates that data were presented for subgroups of interest.

FDA: Food and Drug Administration (USA), FN: false negative, FP: false positive, MFR: manufacturer, N: number of samples analysed, NPV: negative predictive value, PPV: positive predictive value, Sens: sensitivity, Spec: specificity, TN: true negative, TP: true positive.

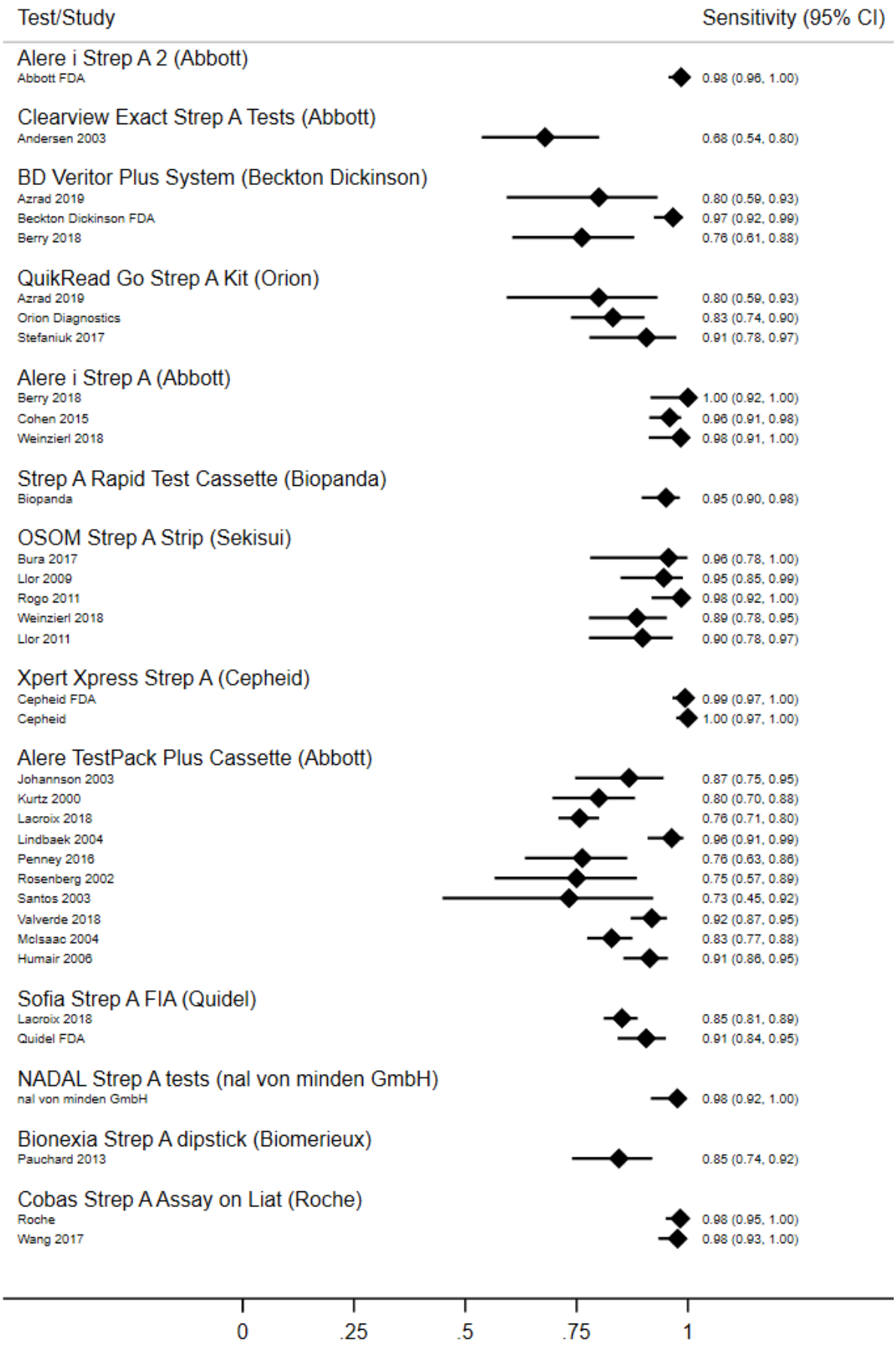


Figure 7: Study level data for the studies included in the meta-analysis of test accuracy

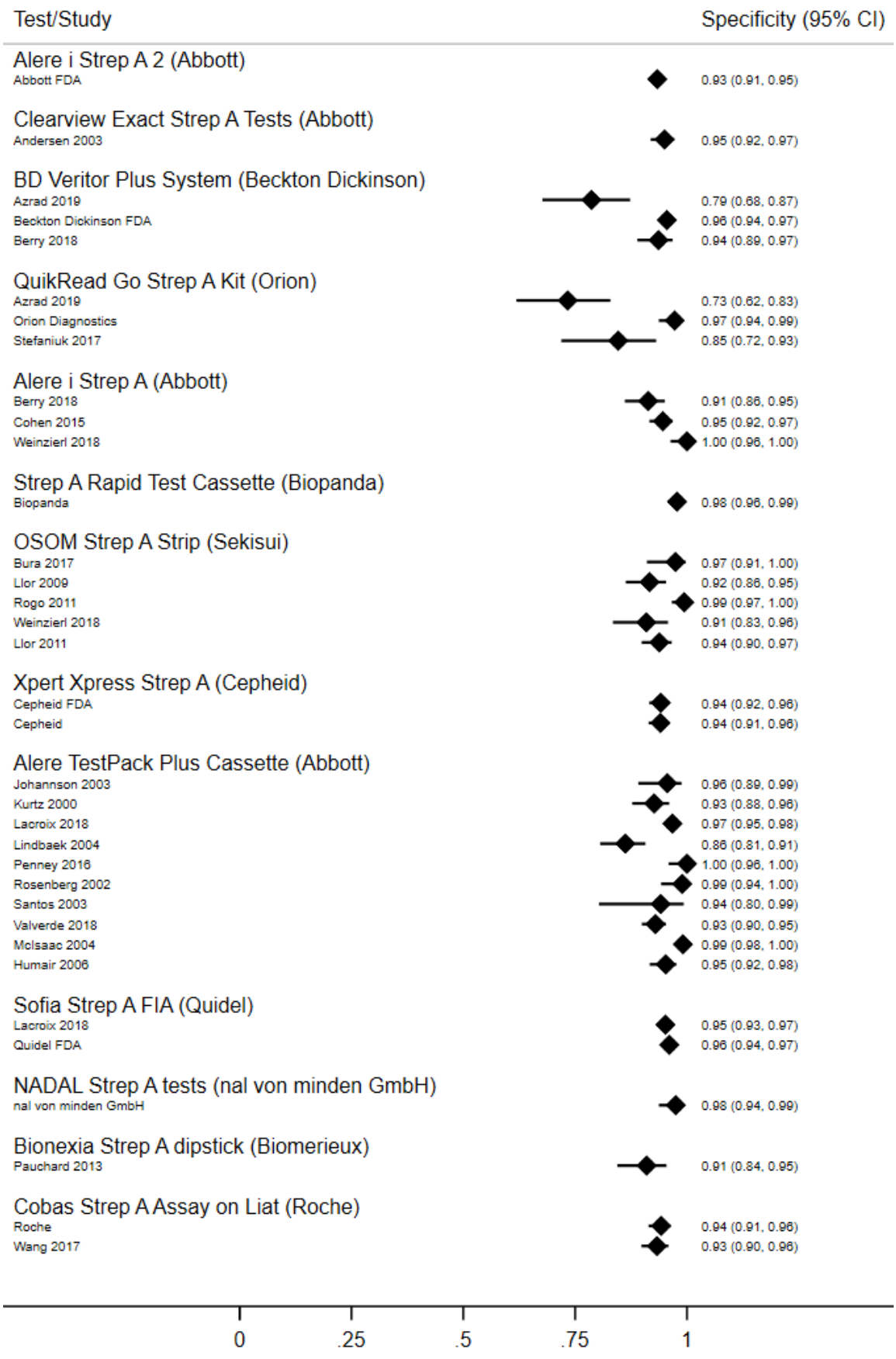


Figure 8: Study level data for the studies included in the meta-analysis of test accuracy

3.2.6.2 **Head-to-head (direct) comparison between tests**

Initially, we sought to identify whether there was evidence to support the hypothesis that the tests might have different test accuracy. Due to the large degree of inter-study variability, the most informative studies were those which performed multiple tests on the same patient population, of which there were four.

Azrad et al. compared both the BD Veritor System (Beckton Dickinson) and QuikRead Go Strep A Kit (Orion) tests to culture for 100 patients.[31] The BD Veritor System had a sensitivity of 0.80 (0.59, 0.92) and specificity of 0.79 (0.67, 0.87). The QuikRead Go test had an identical sensitivity of 0.80 (0.59, 0.92) and a slightly lower point estimate for specificity of 0.73 with overlapping confidence intervals (0.62, 0.83).

Berry et al. compared both the BD Veritor System and the Alere i Strep A tests to culture.[19] The tests performed differently, with the BD Veritor System having a sensitivity of 0.76 (0.60, 0.87) and a specificity of 0.94 (0.89, 0.97) and Alere i Strep A having a sensitivity of 1.00 (0.90, 1.00) and a specificity of 0.91 (0.86, 0.95).

Lacroix et al. investigated both the Alere TestPack Plus and the Sofia Strep A FIA tests.[22] Again the tests performed differently, with the Alere TestPack Plus having lower detection rates with a sensitivity of 0.76 (0.71, 0.80) and a specificity of 0.97 (0.95, 0.98). Meanwhile the Sofia Strep A FIA had a sensitivity of 0.85 (0.81, 0.89) and a specificity of 0.95 (0.93, 0.97).

Finally, Weinzierl et al. assessed the Alere i Strep A and the OSOM Strep A Strip tests.[51] The OSOM Strep A Strip had a sensitivity of 0.89 (0.77, 0.95) and a specificity of 0.91 (0.83, 0.96), whereas the Alere i Strep A test had a sensitivity of 0.98 (0.90, 1.00) and a specificity of 1.00 (0.95, 1.00).

Conclusion: There is insufficient evidence to perform a meaningful comparison of the rapid tests, or to establish any reliable hierarchy of test performance. Whilst some tests may perform similarly, the existing evidence does not allow identification of any clear groups of tests, and it is likely that there is some variation in accuracy of the 21 tests. There is considerable heterogeneity, potentially caused by the differences in study design and population.

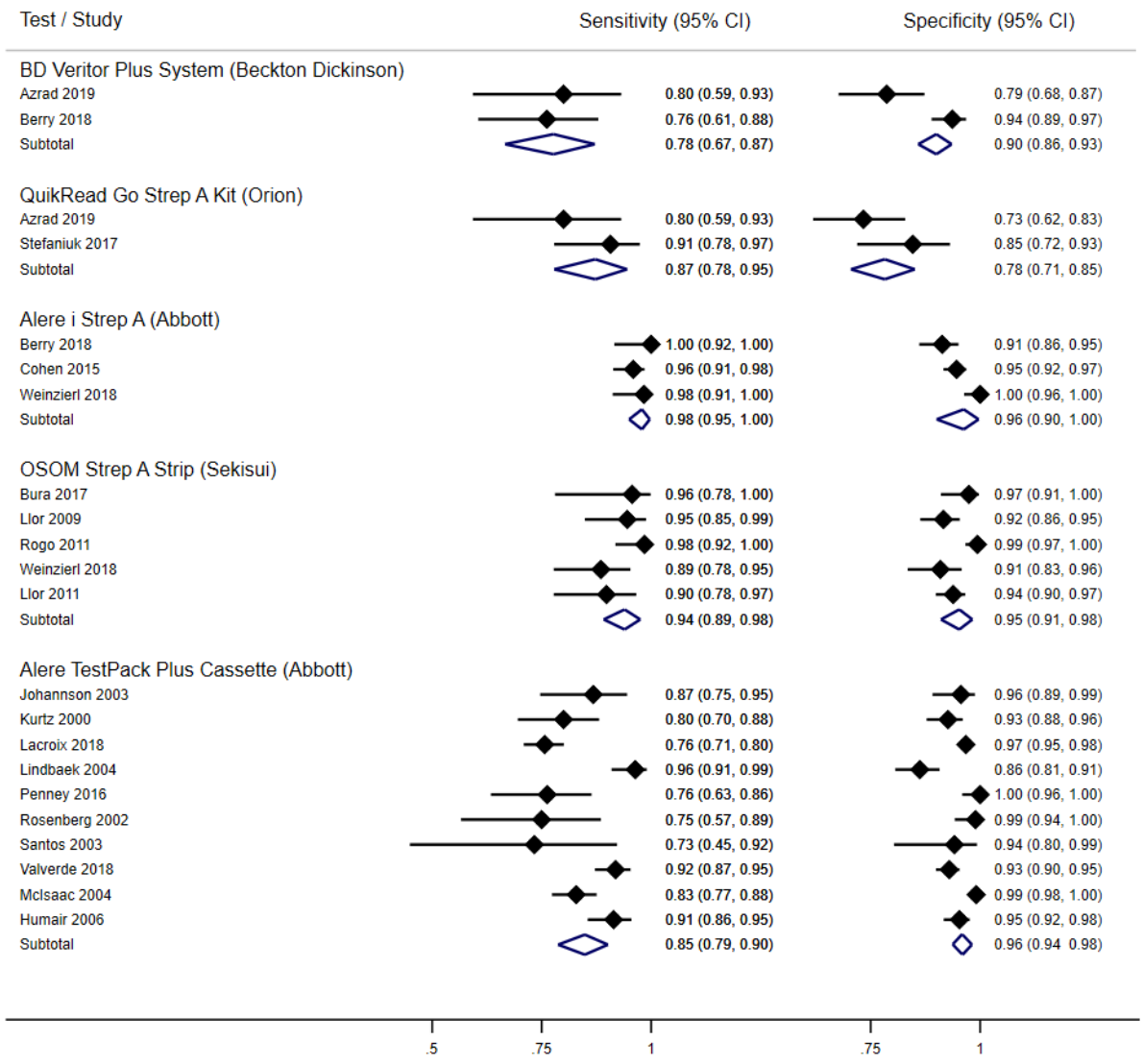


Figure 9: Summary of meta analyses performed on tests with multiple studies excluding manufacturer responses and FDA reports.

3.2.6.3 Accuracy of point-of-care tests in population at high-risk of Strep A infection as defined by sore throat clinical scores

The primary population of interest in this review is patients with high clinical scores (Centor ≥3, FeverPAIN ≥4). We report test accuracy data in that population, and by the patient's score according to a clinical measuring tool such as Centor or McIsaac. The majority of studies either did not place or did not report placing a restriction of the clinical scoring tool on their patient populations.

Eight studies present results based on some restriction of Centor or McIsaac (either ≥ 1 or 2 or 3), which informed for four tests. A summary of evidence can be found in Table 12.

Only two studies presented data for populations which matched the NICE scope, that is either having a Centor or McIsaac score of three or greater or a FeverPAIN score of four or greater.[36, 42] We dichotomised the data from these studies into patients meeting the scope based on throat score, and patients not meeting the scope.

Humair et al. investigated the Alere TestPack Plus test in adults presenting in a primary care setting, with a Centor score ≥ 2.[36] In the Centor=2 and Centor >2 subgroups, the sensitivities were 0.80 (0.63, 0.92) and 0.95 (0.89, 0.98), and the specificities were 0.96 (0.91, 0.99) and 0.94 (0.88, 0.98) respectively. The subgroups had 148 and 224 patients respectively.

Llor et al investigated adult patients in a primary care setting with a Centor score ≥ 1 when assessing the performance of the OSOM Strep A Strip.[42] In the population with a Centor = 1 or 2, consisting of 116 patients, the OSOM Strep A strip had a sensitivity of 0.85 (0.55, 0.98) and a specificity of 0.93 (0.87, 0.96). in the population with a Centor score > 2, with 160 patients, the test had a sensitivity of 0.92 (0.76, 0.98) and a specificity of 0.96 (0.89, 0.99).

The remaining data for studies which restricted their population by throat score are presented below.

OSOM Strep A Strip

Three studies compared the OSOM Strep A Strip in a restricted population. Bura et al.[33] and Llor et al.[41] both focused on patients with a Centor score ≥ 2, and reported sensitivities of 0.96 (0.76, 1.00) and 0.95 (0.85, 0.99), and specificities of 0.97 (0.90, 1.00) and 0.92 (0.86, 0.95).

Meanwhile Llor et al.[42] considered patients with a Centor score ≥ 1 and reported sensitivities of 0.90 (0.78, 0.97), and specificities of 0.94 (0.90, 0.97), respectively.

Alere TestPack Plus Cassette

Four studies investigating the Alere TestPack Plus restricted their population by throat score. DiMatteo et al. looked only at patients with a Centor score ≥ 1 but did not present complete 2x2 information and so sensitivity and specificity could not be calculated.[35] Lacroix et al. and McIsaac et al. both examined test performance in patients with MsIsaac scores ≥ 2.[22, 43] The former estimated a sensitivity of 0.76 (0.71, 0.80) and a specificity of 0.97 (0.95, 0.98), whilst the latter estimated a sensitivity of 0.83 (0.77, 0.88) and a specificity of 0.99 (0.98, 1.00).

Humair et al. also considered only patients with a Centor score ≥ 2, but also presented results by score subgroup mentioned earlier.[36] In the full population, a sensitivity of 0.91 (0.86, 0.95) and specificity 0.95 (0.92, 0.98) were reported.

Sofia Strep A FIA

One study compared Sofia Strep A FIA to culture and restricted their population by throat score. Lacroix et al. used Sofia Strep A FIA in patients with a McIsaac score ≥ 2.[22] In this population the test had a sensitivity of 0.85 (0.81, 0.89) and a specificity of 0.95 (0.93, 0.97).

Cobas Strep A Assay on Liat

One study compared Cobas Strep A Assay on Liat to culture in patients restricted by throat score. Wang et al. used the test in patients with a Centor score ≥ 1.[23] In this population the test had a sensitivity of 0.98 (0.93, 0.99) and a specificity of 0.93 (0.90, 0.96).

Conclusion: The limited evidence suggests that some tests may have a higher sensitivity in patient populations that have a higher score according to a clinical tool, such as Centor.

Table 12: Summary of data informing on test performance in studies that restricted their population by clinical throat score.

Study	Care Setting	Age group	Clinical Tool Score Restriction	Strep A infections Prevalence	Ref Type	N	TP	FN	FP	TN	Accuracy Data
OSOM Strep A Strip - Sekisui											
Bura 2017 [33]	Primary	Adults	Centor ≥ 2	22.7%	Blood agar	101	22	1	2	76	Sens = 0.96 (0.76, 1.00) Spec = 0.97 (0.90, 1.00) PPV = 0.92 (0.72, 0.99) NPV = 0.99 (0.92, 1.00)
Llor 2009 [41]	Primary	Adults	Centor ≥ 2	24.8%	Blood agar	222	52	3	14	153	Sens = 0.95 (0.85, 0.99) Spec = 0.92 (0.86, 0.95) PPV = 0.79 (0.69, 0.86) NPV = 0.98 (0.94, 0.99)
Llor 2011 [42]	Primary	Adults	Centor ≥ 1	17.8%	Blood agar	276	44	5	14	213	Sens = 0.90 (0.78, 0.97) Spec = 0.94 (0.90, 0.97) PPV = 0.76 (0.65, 0.84) NPV = 0.98 (0.95, 0.99)
Llor 2011 [42]	Primary	Adults	Centor > 2	31.0%	Blood agar	116	33	3	3	77	Sens = 0.92 (0.76, 0.98) Spec = 0.96 (0.89, 0.99) PPV = 0.92 (0.78, 0.97) NPV = 0.96 (0.90, 0.99)
Llor 2011 [42]	Primary	Adults	Centor = 1 or 2	8.1%	Blood agar	160	11	2	11	136	Sens = 0.85 (0.55, 0.98) Spec = 0.93 (0.87, 0.96) PPV = 0.50 (0.35, 0.65) NPV = 0.99 (0.95, 1.00)
Alere TestPack Plus Cassette - Abbott											
DiMatteo 2001 [35]	Secondary	Adults	Centor ≥ 1	NR	Strep selective agar			22		361	NPV = 0.94 (0.91, 0.96)
Humair 2006 [36]	Primary	Adults	Centor ≥ 2	37.6%	Blood agar	372	128	12	11	221	Sens = 0.91 (0.86, 0.95) Spec = 0.95 (0.92, 0.98) PPV = 0.92 (0.87, 0.95) NPV = 0.95 (0.91, 0.97)
Humair 2006 [36]	Primary	Adults	Centor = 2	23.6%	Blood agar	148	28	7	4	109	Sens = 0.80 (0.63, 0.92) Spec = 0.96 (0.91, 0.99)

											PPV = 0.88 (0.73, 0.95) NPV = 0.94 (0.89, 0.97)
Humair 2006 [36]	Primary	Adults	Centor > 2	46.9%	Blood agar	224	100	5	7	112	Sens = 0.95 (0.89, 0.98) Spec = 0.94 (0.88, 0.98) PPV = 0.93 (0.87, 0.97) NPV = 0.96 (0.90, 0.98)
Lacroix 2018 [22]	Secondary	Children	McIsaac ≥ 2	35.7%	Blood agar	1002	271	87	21	623	Sens = 0.76 (0.71, 0.80) Spec = 0.97 (0.95, 0.98) PPV = 0.93 (0.89, 0.95) NPV = 0.88 (0.85, 0.90)
McIsaac 2004 [43]	Primary	Children and Adults #	McIsaac ≥ 2	29.0%	Blood agar	787	189	39	5	554	Sens = 0.83 (0.77, 0.88) Spec = 0.99 (0.98, 1.00) PPV = 0.97 (0.94, 0.99) NPV = 0.93 (0.91, 0.95)
Sofia Strep A FIA - Quidel											
Lacroix 2018 [22]	Secondary	Children	McIsaac ≥ 2	35.7%	Blood agar	1002	305	53	31	613	Sens = 0.85 (0.81, 0.89) Spec = 0.95 (0.93, 0.97) PPV = 0.91 (0.87, 0.94) NPV = 0.92 (0.90, 0.94)
Cobas Strep A Assay on Liat - Roche											
Wang 2017 [23]	Primary	Children	Centor ≥ 1	30.2%	NR	427	126	3	20	278	Sens = 0.98 (0.93, 0.99) Spec = 0.93 (0.90, 0.96) PPV = 0.86 (0.79, 0.91) NPV = 0.99 (0.97, 1.00)

Key: FN: false negative, FP: false positive, NPV: negative predictive value, PPV: positive predictive value, Sens: sensitivity, Spec: specificity, TN: true negative, TP: true positive.

3.2.6.4 **Accuracy of point-of-care tests split by age group**

We sought to identify whether there was evidence to support the hypothesis that the tests might have different performance characteristics based on the age group on which the test is being used. No studies categorized their age groups identical to the NICE scope, and so we classified them into children and adult populations where possible, or a combination of children and adults. No studies presented results specific to a 60+ year old population, though patients in this category may have been included within an "adult" population. Seven studies concentrated on exclusively adult populations, providing accuracy data for two tests.[33, 35, 36, 38, 40-42] Ten studies looked exclusively at children providing data for nine tests.[19, 22, 23, 39, 45, 46, 48, 51, 53, 54] Three studies considered both adults and children, and presented accuracy data for them separately, allowing a within-trial comparison to be made.[34, 43, 49] Each of these three studies investigated a different test.

Cohen et al. examined both adults and children when investigating the accuracy of the Alere i Strep A test.[34] In children the test had a sensitivity of 0.96 (0.91, 0.99) and a specificity of 0.93 (0.89, 0.96). In adults, the sensitivity was 0.95 (0.74, 1.00) and the specificity was 0.97 (0.92, 0.99).

McIsaac et al. examined the Alere TestPack Plus test in children and adult populations, presenting the results by age category.[43] In children, the sensitivity was 0.86 (0.79, 0.91) and the specificity was 0.99 (0.97, 1.00). In adults, the sensitivity was 0.77 (0.65, 0.86) with a specificity of 0.99 (0.97, 1.00).

Stefaniuk et al. used the QuikRead Go Strep A Kit test in both adults and children[49] In children the test had a sensitivity of 0.80 (0.56, 0.94) and a specificity of 0.91, (0.72, 0.99) were estimated. In adults, the test sensitivity was 1.00 (0.86, 0.95) and specificity was 0.79 (0.60, 0.92).

Further age specific results are presented below and in Table 13.

Clearview Exact Strep A Cassette and Clearview Exact Strep A Dipstick (Abbott)

Only data for a child population were available for the Clearview Exact Strep A Cassette and Dipstick tests, and was provided by Andersen et al.[53], which did not distinguish between the cassette and dipstick varieties. Andersen et al. reported a sensitivity of 0.68 (0.55, 0.81) and a specificity of 0.95 (0.93, 0.98).

BD Veritor Plus System (Beckton Dickinson)

The only age-specific test accuracy data for the BD Veritor Plus System was in children, and published by Berry et al.[19] The sensitivity of the test was 0.76 (0.60, 0.87), and the specificity was 0.94 (0.89, 0.97).

OSOM Strep A Strip (Sekisui)

Three studies presented data for the OSOM Strep A Strip in adult patients.[33, 41, 42] Bura et al., Llor et al. and Llor et al. reported respective sensitivities of 0.96 (0.76, 1.00), 0.95 (0.85, 0.99) and 0.90 (0.78, 0.97), and specificities of 0.97 (0.90, 1.00), 0.92 (0.86, 0.95) and 0.94 (0.90, 0.97). Rogo et al. and Weinzierl et al. both studied only children, and estimated sensitivities of 0.98 (0.91, 1.00) and 0.89 (0.77, 0.95), and specificities of 0.99 (0.96, 1.00) and 0.91 (0.83, 0.96), respectively.[46, 51]

QuikRead Go Strep A Kit (Orion)

Stefaniuk et al. examined both adults and children.[49] In children a sensitivity of 0.80 (0.56, 0.94) and specificity of 0.91, (0.72, 0.99) were estimated. In adults, the test sensitivity was 1.00 (0.86, 0.95) and specificity was 0.79 (0.60. 0.92).

Alere TestPack Plus (Abbott)

Three studies used the Alere TestPack Plus test in adult populations.[35, 36, 38] DiMatteo et al. and Johnson et al. did not provide complete results and sensitivity and specificity could not be calculated. Humair et al. did provide sufficient information and the test's sensitivity was 0.91 (0.86, 0.95). The specificity was 0.95 (0.92, 0.98).

Four studies used the test in child populations only.[22, 39, 45, 48] The lowest sensitivity was reported by Santos et al, (0.73 [0.45, 0.91]) and the highest by Kurtz et al. (0.80 [0.71, 0.89]). The sensitivities ranged from 0.93 (0.89, 0.97) as reported by Kurtz et al. to 1.00 (0.95, 1.00) by Penney et al. McIsaac et al. performed the test in both groups.[43] In children, the sensitivity was 0.86 (0.79, 0.91) and the specificity was 0.99 (0.97, 1.00). In adults, the sensitivity was 0.77 (0.65, 0.86) with a specificity of 0.99 (0.97, 1.00).

Bionexia Strep A Dipstick (Biomerieux)

Only data for children were available for Biomerieux's Bionexia Strep A Dipstick. Pauchard et al.[54] estimated a sensitivity of 0.85 (0.74, 0.92) and a specificity of 0.91 (0.84, 0.95).

Sofia Strep A FIA (Quidel)

One study compared Sofia Strep A FIA to culture in children, with no adult data available.[22] Lacroix et al. reported a sensitivity of 0.85 (0.81, 0.89) and a specificity of 0.95 (0.93, 0.97).

Alere i Strep A (Abbot)

Two studies presented data for the Alere i Strep A test for child populations.[19, 51] Berry et al. and Weinzierl et al. reported respective sensitivities of 1.00 (0.90, 1.00) and 0.98 (0.90, 1.00), and specificities of 0.91 (0.86, 0.95) and 1.00 (0.95, 1.00).

Cohen et al. examined both adults and children, and presented results by age group.[34] In children the test had a sensitivity of 0.96 (0.91, 0.99) and a specificity of 0.93 (0.89, 0.96). In adults, the sensitivity was 0.95 (0.74, 1.00) and the specificity was 0.97 (0.92, 0.99).

Cobas Strep A Assay on Liat (Roche)

Only data for a child population were available for the Cobas Strep A Assay on Liat test.[23] Wang et al. reported the test had a sensitivity of 0.98 (0.93, 0.99) and a specificity of 0.93 (0.90, 0.96).

Meta-analyses were performed to compare the accuracy estimates of the child and adult populations for both the OSOM Strep A Strip and Alere TestPack Plus tests, as these were the only tests with sufficient data.

For the TestPack Plus, in children the sensitivity was estimated as 0.80 (0.74, 0.84) and the specificity as 0.98 (0.95, 0.99). In adults, the sensitivity was estimated as 0.87 (0.82, 0.91) and the specificity as 0.98 (0.96, 0.99).

For OSOM, the dichotomisation of studies into the two age categories was identical to the dichotomisation for primary and secondary care settings. Univariate models fitted to the children/secondary care data estimated a sensitivity of 0.95 (0.90, 0.98) and a specificity of 0.97 (0.95, 0.99). Models fitted to the adult/primary care data estimated a sensitivity of 0.93 (0.88, 0.97) and a specificity of 0.94 (0.91, 0.97).

Conclusion: It is unclear whether test accuracy varies based on the age of the population in which it is being used. Further evidence is required.

Table 13: Summary of test data for age groups of interest

Study	Care Setting	Age group	Clinical Tool Score Restriction	Strep A infections Prevalence	Ref Type	N	TP	FN	FP	TN	Accuracy Data
Clearview Exact Strep A Cassette - Abbott											
Andersen 2003 [53]	Secondary	Children	None	15.0%	NR	353	36	17	15	285	Sens = 0.68 (0.55, 0.81) Spec = 0.95 (0.93, 0.98) PPV = 0.71 (0.58, 0.83) NPV = 0.94 (0.92, 0.97)
Clearview Exact Strep A Dipstick - Abbott											
Andersen 2003 [53]	Secondary	Children	None	15.0%	NR	353	36	17	15	285	Sens = 0.68 (0.55, 0.81) Spec = 0.95 (0.93, 0.98) PPV = 0.71 (0.58, 0.83) NPV = 0.94 (0.92, 0.97)
BD Veritor Plus System – Beckton Dickinson											
Berry 2018 [19]	Secondary	Children	None	19.5%	Blood agar	215	32	10	11	162	Sens = 0.76 (0.60, 0.87) Spec = 0.94 (0.89, 0.97) PPV = 0.74 (0.56, 0.86) NPV = 0.94 (0.89, 0.97)
OSOM Strep A Strip - Sekisui											
Bura 2017 [33]	Primary	Adults	Centor ≥ 2	22.7%	Blood agar	101	22	1	2	76	Sens = 0.96 (0.76, 1.00) Spec = 0.97 (0.90, 1.00) PPV = 0.92 (0.72, 0.99) NPV = 0.99 (0.92, 1.00)
Llor 2009 [41]	Primary	Adults	Centor ≥ 2	24.8%	Blood agar	222	52	3	14	153	Sens = 0.95 (0.85, 0.99) Spec = 0.92 (0.86, 0.95) PPV = 0.79 (0.69, 0.86) NPV = 0.98 (0.94, 0.99)
Llor 2011 [42]	Primary	Adults	Centor ≥ 2 #	17.8%	Blood agar	276	44	5	14	213	Sens = 0.90 (0.78, 0.97) Spec = 0.94 (0.90, 0.97) PPV = 0.76 (0.65, 0.84) NPV = 0.98 (0.95, 0.99)
Rogo 2011 [46]	Secondary	Children	None	28.9%	Blood agar	228	65	1	1	161	Sens = 0.98 (0.91, 1.00) Spec = 0.99 (0.96, 1.00)

Study	Care Setting	Age group	Clinical Tool Score Restriction	Strep A infections Prevalence	Ref Type	N	TP	FN	FP	TN	Accuracy Data
											PPV = 0.98 (0.91, 1.00) NPV = 0.99 (0.96, 1.00)
Weinzierl 2018 [51]	Secondary	Children	None	38.1%	Blood agar	160	54	7	9	90	Sens = 0.89 (0.77, 0.95) Spec = 0.91 (0.83, 0.96) PPV = 0.86 (0.74, 0.93) NPV = 0.93 (0.85, 0.97)
QuikRead Go Strep A Kit - Orion											
Stefaniuk 2017 [49]	Primary	Children	None	46.5%	Blood agar	43	16	4	2	21	Sens = 0.80 (0.56, 0.94) Spec = 0.91 (0.72, 0.99) PPV = 0.89 (0.68, 0.97) NPV = 0.84 (0.68, 0.93)
Stefaniuk 2017 [49]	Primary	Adults	None	44.2%	Blood agar	52	23	0	6	23	Sens = 1.00 (0.85, 1.00) Spec = 0.79 (0.60, 0.92) PPV = 0.79 (0.65, 0.89) NPV = 1.00 (0.85, 1.00)
Alere TestPack Plus Cassette - Abbott											
DiMatteo 2001 [35]	Secondary	Adults	Centor ≥ 1	NR	Strep selective agar			22		361	NPV = 0.94 (0.91, 0.96)
Humair 2006 [36]	Primary	Adults	Centor ≥ 2 #	37.6%	Blood agar	372	128	12	11	221	Sens = 0.91 (0.86, 0.95) Spec = 0.95 (0.92, 0.98) PPV = 0.92 (0.87, 0.95) NPV = 0.95 (0.91, 0.97)
Johnson 2001 [38]	Primary	Adults	None	NR	Blood agar		445		77		PPV = 0.85 (0.82, 0.88)
Kurtz 2000 [39]	Secondary	Children	None	31.1%	Blood agar	257	64	16	13	164	Sens = 0.80 (0.71, 0.89) Spec = 0.93 (0.89, 0.97) PPV = 0.83 (0.75, 0.92) NPV = 0.91 (0.87, 0.95)
Lacroix 2018 [22]	Secondary	Children	McIsaac ≥ 2	35.7%	Blood agar	1002	271	87	21	623	Sens = 0.76 (0.71, 0.80) Spec = 0.97 (0.95, 0.98) PPV = 0.93 (0.89, 0.95)

Study	Care Setting	Age group	Clinical Tool Score Restriction	Strep A infections Prevalence	Ref Type	N	TP	FN	FP	TN	Accuracy Data
											NPV = 0.88 (0.85, 0.90)
McIsaac 2004 [43]	Primary	Children	McIsaac ≥ 2	34.1%	Blood agar	454	133	22	3	296	Sens = 0.86 (0.79, 0.91) Spec = 0.99 (0.97, 1.00) PPV = 0.98 (0.93, 0.99) NPV = 0.93 (0.90, 0.95)
McIsaac 2004 [43]	Primary	Adults	McIsaac ≥ 2	21.9%	Blood agar	333	56	17	2	258	Sens = 0.77 (0.65, 0.86) Spec = 0.99 (0.97, 1.00) PPV = 0.97 (0.88, 0.99) NPV = 0.94 (0.91, 0.96)
Penney 2016 [45]	Secondary	Children	None	40.1%	Strep selective agar	147	45	14	0	88	Sens = 0.76 (0.65, 0.87) Spec = 1.00 (0.95, 1.00) PPV = 1.00 (0.90, 1.00) NPV = 0.86 (0.78, 0.92)
Santos 2003 [48]	Secondary	Children	None	30.6%	Blood agar	49	11	4	2	32	Sens = 0.73 (0.45, 0.91) Spec = 0.94 (0.79, 0.99) PPV = 0.85 (0.54, 0.97) NPV = 0.89 (0.73, 0.96)
Bionexia Strep A Dipstick - Biomerieux											
Pauchard 2013 [54]	Secondary	Children	None	36.8%	NR	193	60	11	11	111	Sens = 0.85 (0.74, 0.92) Spec = 0.91 (0.84, 0.95) PPV = 0.85 (0.76, 0.93) NPV = 0.91 (0.86, 0.96)
Sofia Strep A FIA - Quidel											
Lacroix 2018 [22]	Secondary	Children	McIsaac ≥ 2	35.7%	Blood agar	1002	305	53	31	613	Sens = 0.85 (0.81, 0.89) Spec = 0.95 (0.93, 0.97) PPV = 0.91 (0.87, 0.94) NPV = 0.92 (0.90, 0.94)
Alere i Strep A - Abbott											
Berry 2018 [19]	Secondary	Children	None	19.5%	Blood agar	215	42	0	15	158	Sens = 1.00 (0.90, 1.00) Spec = 0.91 (0.86, 0.95) PPV = 0.74 (0.60, 0.84) NPV = 1.00 (0.97, 1.00)

Study	Care Setting	Age group	Clinical Tool Score Restriction	Strep A infections Prevalence	Ref Type	N	TP	FN	FP	TN	Accuracy Data
Cohen 2015 [34]	Secondary	Children	None		Blood agar	355	123	5	15	212	Sens = 0.96 (0.91, 0.99) Spec = 0.93 (0.89, 0.96) PPV = 0.89 (0.83, 0.93) NPV = 0.98 (0.95, 0.99)
Cohen 2015 [34]	Secondary	Adults	None		Blood agar	126	18	1	3	104	Sens = 0.95 (0.74, 1.00) Spec = 0.97 (0.92, 0.99) PPV = 0.86 (0.66, 0.95) NPV = 0.99 (0.94, 1.00)
Weinzierl 2018 [51]	Secondary	Children	None	38.1%	Blood agar	160	60	1	0	99	Sens = 0.98 (0.90, 1.00) Spec = 1.00 (0.95, 1.00) PPV = 1.00 (0.93, 1.00) NPV = 0.99 (0.94, 1.00)
Cobas Strep A Assay on Liat - Roche											
Wang 2017 [23]	Primary	Children	Centor ≥ 1	30.2%	NR	427	126	3	20	278	Sens = 0.98 (0.93, 0.99) Spec = 0.93 (0.90, 0.96) PPV = 0.86 (0.79, 0.91) NPV = 0.99 (0.97, 1.00)

Key: FN: false negative, FP: false positive, NPV: negative predictive value, PPV: positive predictive value, Sens: sensitivity, Spec: specificity, TN: true negative, TP: true positive.

3.2.6.5 Accuracy of point-of-care tests split by primary/secondary care setting

We sought to identify whether there was evidence to support the hypothesis that the tests might have different performance characteristics based on the setting in which the test is being used. No studies provided a breakdown of results comparing test accuracy between primary and secondary setting. Fourteen studies considered patients in a secondary care setting, providing data for nine tests. Ten studies looked at patients at primary care settings, covering four tests. A summary of care-setting related data can be found in Table 14.

Clearview Exact Strep A Cassette and Clearview Exact Strep A Dipstick (Abbott)

Only data in a hospital setting were available for the Clearview Exact Strep A Cassette and Dipstick tests, and was provided by Andersen et al.[53], which did not distinguish between the cassette and dipstick varieties. Andersen et al. reported a sensitivity of 0.68 (0.55, 0.81) and a specificity of 0.95 (0.93, 0.98).

BD Veritor Plus System (Beckton Dickinson)

Azrad et al.[31] and Berry et al.[19] both presented results for the BD Veritor Plus System in a hospital setting. The sensitivities of the test were 0.80 (0.59, 0.92) and 0.76 (0.60, 0.87), and the specificities were 0.79 (0.67, 0.87) and 0.94 (0.89, 0.97), by Azrad et al. and Berry et al. respectively.

NADAL Strep A Strip, NADAL Strep A Cassette, NADAL Strep A Plus Cassette, NADAL Strep A Plus Strip and NADAL Strep A Scan (nal von minden GmbH)

Only evidence from a secondary care setting was available for the NADAL tests, which did not distinguish between any of the varieties. The manufacturer reported a sensitivity of 0.98 (0.91, 1.00) and a specificity of 0.98 (0.93, 0.99).

Strep A Rapid Test Cassette (Biopanda)

The data provided by Biopanda for the Strep A Rapid Test was reportedly from a primary care setting. The sensitivity was 0.95 (0.0.89, 0.98) and the specificity was 0.98 (0.96, 0.99).

OSOM Strep A Strip (Sekisui)

Three studies presented data for the OSOM Strep A Strip in a primary care setting.[33, 41, 42] Bura et al., Llor et al. and Llor et al. reported respective sensitivities of 0.96 (0.76, 1.00), 0.95 (0.85, 0.99) and 0.90 (0.78, 0.97), and specificities of 0.97 (0.90, 1.00), 0.92 (0.86, 0.95) and 0.94 (0.90, 0.97). Rogo et al. and Weinzierl et al. both used the test in a hospital setting, and estimated sensitivities of 0.98 (0.91, 1.00) and 0.89 (0.77, 0.95), and specificities of 0.99 (0.96, 1.00) and 0.91 (0.83, 0.96), respectively.[46, 51]

QuikRead Go Strep A Kit (Orion)

Azrad et al. compared the performance of the QuikRead Go Strep A Kit to culture in a hospital setting, and reported a sensitivity of 0.80 (0.59, 0.92), and specificity of 0.73 (0.62, 0.83).[31] Stefaniuk et al. looked in a primary care setting, and reported a sensitivity 0.91 (0.78, 0.97) and specificity 0.85 (0.72, 0.93).[49] The data provided by Orion was also reported as being from a primary care setting, and estimated a sensitivity of 0.83 (0.73, 0.90) and a specificity of 0.97 (0.93, 0.99).

Alere TestPack Plus Cassette (Abbott)

There were seven published studies which compared the performance of the Alere TestPack Plus Cassette to culture in a secondary care setting. One study did not report sufficient data to complete a 2x2 table.[35] Rosenberg et al. and Valverde et al. both examined a combination of children and adults, estimating sensitivities of 0.75 (0.56, 0.88) and 0.92 (0.87, 0.95), and specificities of 0.99 (0.93, 1.00) and 0.93 (0.90, 0.95) respectively.[47, 55] The four remaining studies only included children.[22, 39, 45, 48] The sensitivities ranged from 0.73 (0.45, 0.91)[48] to 0.80 (0.71, 0.90)[39] and the specificities ranged from 0.93 (0.89, 0.97)[39] to 1.00 (0.95, 1.00).[45]

Five studies reported the accuracy of the Alere TestPack Plus Cassette in primary care. One did not present complete 2x2 data.[38] One reported for adult populations: Humair et al. estimated a sensitivity of 0.91 (0.86, 0.95), and a specificity of 0.95 (0.92, 0.98).[36] Lindbaek et al., Johannson et al. and McIsaac et al. combined adults and children, and reported

respective sensitivities of 0.94 (0.90, 0.99), 0.87 (0.74, 0.94) and 0.83 (0.77, 0.87) alongside specificities of 0.86 (0.80, 0.91), 0.96 (0.89, 0.99) and 0.99 (0.98, 1.00).[37, 40, 43]

Bionexia Strep A Dipstick (Biomerieux)

Only data for a hospital setting were available for Biomerieux's Bionexia Strep A Dipstick. Pauchard et al. estimated a sensitivity of 0.85 (0.74, 0.92) and a specificity of 0.91 (0.84, 0.95).[54]

Sofia Strep A FIA (Quidel)

One study compared Sofia Strep A FIA to culture in hospital setting, with no GP data available.[22] Lacroix et al. reported a sensitivity of 0.85 (0.81, 0.89) and a specificity of 0.95 (0.93, 0.97).

Alere i Strep A (Abbott)

Three studies compared the Alere i Strep A test to culture in a hospital setting. Berry et al.[19] and Weinzierl et al.[51] looked only at children, and estimated respective sensitivities of 1.00 (0.90, 1.00) and 0.98 (0.90, 1.00), and specificities of 0.91 (0.86, 0.95) and 1.00 (0.95, 1.00). Cohen 2015[34] examined in both children and adults, and produced respective estimates of sensitivity and specificity of 0.96 (0.91, 0.98) and 0.95 (0.91, 0.97).

Cobas Strep A Assay on Liat (Roche)

Only data from a GP setting were available for the Cobas Strep A Assay on Liat test.[23] Wang et al. reported the test had a sensitivity of 0.98 (0.93, 0.99) and a specificity of 0.93 (0.90, 0.96).

Meta-analyses were performed to indirectly compare the accuracy estimates of the child and adult populations for both the OSOM Strep A Strip and Alere TestPack Plus tests, as these were the only tests with sufficient data.

Firstly, fitted to data for the TestPack Plus test, univariate models estimated a sensitivity of 0.90 (0.83, 0.96) and specificity of 0.95 (0.88, 0.99) in a primary setting, compared to a sensitivity of 0.80 (0.71, 0.88) and specificity of 0.97 (0.94, 0.99) in secondary care.

The OSOM test also had sufficient studies to perform univariate meta-analyses. However, the dichotomisation of studies into primary and secondary care settings was identical to the age dichotomisation. Univariate models fitted to the children/secondary care data estimated a sensitivity of 0.95 (0.90, 0.98) and a specificity of 0.97 (0.95, 0.99). Models fitted to the adult/primary care data estimated a sensitivity of 0.93 (0.88, 0.97) and a specificity of 0.94 (0.91, 0.97).

Conclusion: Test performance may vary depending on the care setting in which the test is being used. Further evidence is required.

Table 14: Summary of test performance data by care setting

Study	Care Setting	Age group	Clinical Tool Score Restriction	Strep A infections Prevalence	Ref Type	N	TP	FN	FP	TN	Accuracy Data
Clearview Exact Strep A Cassette - Abbott											
Andersen 2003 [53]	Secondary	Children	None	15.0%	NR	353	36	17	15	285	Sens = 0.68 (0.55, 0.81) Spec = 0.95 (0.93, 0.98) PPV = 0.71 (0.58, 0.83) NPV = 0.94 (0.92, 0.97)
Clearview Exact Strep A Dipstick - Abbott											
Andersen 2003 [53]	Secondary	Children	None	15.0%	NR	353	36	17	15	285	Sens = 0.68 (0.55, 0.81) Spec = 0.95 (0.93, 0.98) PPV = 0.71 (0.58, 0.83) NPV = 0.94 (0.92, 0.97)
BD Veritor Plus System - Beckton Dickinson											
Azrad 2019 [31]	Secondary	NR	None	25.0%	Strep selective agar	100	20	5	16	59	Sens = 0.80 (0.59, 0.92) Spec = 0.79 (0.67, 0.87) PPV = 0.56 (0.38, 0.72) NPV = 0.92 (0.82, 0.97)
Berry 2018 [19]	Secondary	Children	None	19.5%	Blood agar	215	32	10	11	162	Sens = 0.76 (0.60, 0.87) Spec = 0.94 (0.89, 0.97) PPV = 0.74 (0.56, 0.86) NPV = 0.94 (0.89, 0.97)
NADAL Strep A Strip - nal von minden GmbH											
nal von minden GmbH [MFR] *	Secondary	Children and Adults	None	34.4%	Blood agar	244	82	2	4	156	Sens = 0.98 (0.91, 1.00) Spec = 0.98 (0.93, 0.99) PPV = 0.95 (0.88, 0.99) NPV = 0.99 (0.95, 1.00)
NADAL Strep A Cassette - nal von minden GmbH											
nal von minden GmbH [MFR] *	Secondary	Children and Adults	None	34.4%	Blood agar	244	82	2	4	156	Sens = 0.98 (0.91, 1.00) Spec = 0.98 (0.93, 0.99) PPV = 0.95 (0.88, 0.99) NPV = 0.99 (0.95, 1.00)
NADAL Strep A Plus Cassette - nal von minden GmbH											

Study	Care Setting	Age group	Clinical Tool Score Restriction	Strep A infections Prevalence	Ref Type	N	TP	FN	FP	TN	Accuracy Data
nal von minden GmbH [MFR] *	Secondary	Children and Adults	None	34.4%	Blood agar	244	82	2	4	156	Sens = 0.98 (0.91, 1.00) Spec = 0.98 (0.93, 0.99) PPV = 0.95 (0.88, 0.99) NPV = 0.99 (0.95, 1.00)
NADAL Strep A Plus Strip - nal von minden GmbH											
nal von minden GmbH [MFR] *	Secondary	Children and Adults	None	34.4%	Blood agar	244	82	2	4	156	Sens = 0.98 (0.91, 1.00) Spec = 0.98 (0.93, 0.99) PPV = 0.95 (0.88, 0.99) NPV = 0.99 (0.95, 1.00)
NADAL Strep A Scan - nal von minden GmbH											
nal von minden GmbH [MFR] *	Secondary	Children and Adults	None	34.4%	Blood agar	244	82	2	4	156	Sens = 0.98 (0.91, 1.00) Spec = 0.98 (0.93, 0.99) PPV = 0.95 (0.88, 0.99) NPV = 0.99 (0.95, 1.00)
OSOM Strep A Strip - Sekisui											
Bura 2017 [33]	Primary	Adults	Centor ≥ 2	22.7%	Blood agar	101	22	1	2	76	Sens = 0.96 (0.76, 1.00) Spec = 0.97 (0.90, 1.00) PPV = 0.92 (0.72, 0.99) NPV = 0.99 (0.92, 1.00)
Llor 2009 [41]	Primary	Adults	Centor ≥ 2	24.8%	Blood agar	222	52	3	14	153	Sens = 0.95 (0.85, 0.99) Spec = 0.92 (0.86, 0.95) PPV = 0.79 (0.69, 0.86) NPV = 0.98 (0.94, 0.99)
Llor 2011 [42]	Primary	Adults	Centor ≥ 2 #	17.8%	Blood agar	276	44	5	14	213	Sens = 0.90 (0.78, 0.97) Spec = 0.94 (0.90, 0.97) PPV = 0.76 (0.65, 0.84) NPV = 0.98 (0.95, 0.99)
Rogo 2011 [46]	Secondary	Children	None	28.9%	Blood agar	228	65	1	1	161	Sens = 0.98 (0.91, 1.00) Spec = 0.99 (0.96, 1.00) PPV = 0.98 (0.91, 1.00) NPV = 0.99 (0.96, 1.00)

Study	Care Setting	Age group	Clinical Tool Score Restriction	Strep A infections Prevalence	Ref Type	N	TP	FN	FP	TN	Accuracy Data
Weinzierl 2018 [51]	Secondary	Children	None	38.1%	Blood agar	160	54	7	9	90	Sens = 0.89 (0.77, 0.95) Spec = 0.91 (0.83, 0.96) PPV = 0.86 (0.74, 0.93) NPV = 0.93 (0.85, 0.97)
QuikRead Go Strep A Kit - Orion											
Azrad 2019 [31]	Secondary	NR	None	25.0%	Strep selective agar	100	20	5	20	55	Sens = 0.80 (0.59, 0.92) Spec = 0.73 (0.62, 0.83) PPV = 0.50 (0.34, 0.66) NPV = 0.92 (0.81, 0.97)
Stefaniuk 2017 [49]	Primary	Children and Adults #	None	45.3%	Blood agar	95	39	4	8	44	Sens = 0.91 (0.78, 0.97) Spec = 0.85 (0.72, 0.93) PPV = 0.83 (0.72, 0.90) NPV = 0.92 (0.81, 0.97)
Alere TestPack Plus Cassette - Abbott											
DiMatteo 2001 [35]	Secondary	Adults	Centor ≥ 1	NR	Strep selective agar	NR	NR	22	NR	361	NPV = 0.94 (0.91, 0.96)
Humair 2006 [36]	Primary	Adults	Centor ≥ 2 #	37.6%	Blood agar	372	128	12	11	221	Sens = 0.91 (0.86, 0.95) Spec = 0.95 (0.92, 0.98) PPV = 0.92 (0.87, 0.95) NPV = 0.95 (0.91, 0.97)
Johannson 2003 [37]	Primary	Children and Adults	None	31.4%	NR	144	46	7	4	87	Sens = 0.87 (0.74, 0.94) Spec = 0.96 (0.89, 0.99) PPV = 0.92 (0.80, 0.97) NPV = 0.93 (0.85, 0.97)
Johnson 2001 [38]	Primary	Adults	None	NR	Blood agar	NR	445	NR	77	NR	PPV = 0.85 (0.82, 0.88)
Kurtz 2000 [39]	Secondary	Children	None	31.1%	Blood agar	257	64	16	13	164	Sens = 0.80 (0.71, 0.89) Spec = 0.93 (0.89, 0.97) PPV = 0.83 (0.75, 0.92) NPV = 0.91 (0.87, 0.95)

Study	Care Setting	Age group	Clinical Tool Score Restriction	Strep A infections Prevalence	Ref Type	N	TP	FN	FP	TN	Accuracy Data
Lacroix 2018 [22]	Secondary	Children	McIsaac ≥ 2	35.7%	Blood agar	1002	271	87	21	623	Sens = 0.76 (0.71, 0.80) Spec = 0.97 (0.95, 0.98) PPV = 0.93 (0.89, 0.95) NPV = 0.88 (0.85, 0.90)
Lindbaek 2004 [40]	Primary	Children and Adults	None	35.9%	Strep selective agar	306	106	4	27	169	Sens = 0.94 (0.90, 0.99) Spec = 0.86 (0.80, 0.91) PPV = 0.80 (0.72, 0.86) NPV = 0.98 (0.94, 0.99)
McIsaac 2004 [43]	Primary	Children and Adults #	McIsaac ≥ 2	29.0%	Blood agar	787	189	39	5	554	Sens = 0.83 (0.77, 0.88) Spec = 0.99 (0.98, 1.00) PPV = 0.97 (0.94, 0.99) NPV = 0.93 (0.91, 0.95)
Penney 2016 [45]	Secondary	Children	None	40.1%	Strep selective agar	147	45	14	0	88	Sens = 0.76 (0.65, 0.87) Spec = 1.00 (0.95, 1.00) PPV = 1.00 (0.90, 1.00) NPV = 0.86 (0.78, 0.92)
Rosenberg 2002 [47]	Secondary	Children and Adults	None	25.4%	Blood agar	126	24	8	1	93	Sens = 0.75 (0.56, 0.88) Spec = 0.99 (0.93, 1.00) PPV = 0.96 (0.78, 1.00) NPV = 0.92 (0.85, 0.96)
Santos 2003 [48]	Secondary	Children	None	30.6%	Blood agar	49	11	4	2	32	Sens = 0.73 (0.45, 0.91) Spec = 0.94 (0.79, 0.99) PPV = 0.85 (0.54, 0.97) NPV = 0.89 (0.73, 0.96)
Valverde 2018 [55]	Secondary	Children and Adults	None	40.0%	Blood agar	580	181	16	27	356	Sens = 0.92 (0.87, 0.95) Spec = 0.93 (0.90, 0.95) PPV = 0.87 (0.82, 0.91) NPV = 0.96 (0.93, 0.97)
Bionexia Strep A Dipstick - Biomerieux											
Pauchard 2013 [54]	Secondary	Children	None	36.8%	NR	193	60	11	11	111	Sens = 0.85 (0.74, 0.92) Spec = 0.91 (0.84, 0.95) PPV = 0.85 (0.76, 0.93)

Study	Care Setting	Age group	Clinical Tool Score Restriction	Strep A infections Prevalence	Ref Type	N	TP	FN	FP	TN	Accuracy Data
											NPV = 0.91 (0.86, 0.96)
Sofia Strep A FIA - Quidel											
Lacroix 2018 [22]	Secondary	Children	McIsaac ≥ 2	35.7%	Blood agar	1002	305	53	31	613	Sens = 0.85 (0.81, 0.89) Spec = 0.95 (0.93, 0.97) PPV = 0.91 (0.87, 0.94) NPV = 0.92 (0.90, 0.94)
Alere i Strep A - Abbott											
Berry 2018 [19]	Secondary	Children	None	19.5%	Blood agar	215	42	0	15	158	Sens = 1.00 (0.90, 1.00) Spec = 0.91 (0.86, 0.95) PPV = 0.74 (0.60, 0.84) NPV = 1.00 (0.97, 1.00)
Cohen 2015 [34]	Secondary	Children and Adults #	None	30.3%	Blood agar	481	141	6	18	316	Sens = 0.96 (0.91, 0.98) Spec = 0.95 (0.91, 0.97) PPV = 0.89 (0.82, 0.93) NPV = 0.98 (0.96, 0.99)
Weinzierl 2018 [51]	Secondary	Children	None	38.1%	Blood agar	160	60	1	0	99	Sens = 0.98 (0.90, 1.00) Spec = 1.00 (0.95, 1.00) PPV = 1.00 (0.93, 1.00) NPV = 0.99 (0.94, 1.00)
Cobas Strep A Assay on Liat - Roche											
Wang 2017 [23]	Primary	Children	Centor ≥ 1	30.2%	NR	427	126	3	20	278	Sens = 0.98 (0.93, 0.99) Spec = 0.93 (0.90, 0.96) PPV = 0.86 (0.79, 0.91) NPV = 0.99 (0.97, 1.00)

Key: FN: false negative, FP: false positive, NPV: negative predictive value, PPV: positive predictive value, Sens: sensitivity, Spec: specificity, TN: true negative, TP: true positive.

3.2.6.6 **Estimates of test accuracy for cost-effectiveness modelling**

Having established that a number of factors may influence test accuracy, we sought to provide estimates for each test to be used in the cost-effectiveness modelling. This is consistent with the findings of Leeflang et al.[69] Ideally, estimates would have come from a meta-analysis of several studies specific to scope population, by age group and setting. However, the evidence base was not sufficient to do this. In total there were 21 tests x 3 age groups x 3 settings = 189 pairs of sensitivity and specificity estimates required. However, no data were available specific to the elderly or the pharmacy setting, nor for three of the tests, meaning just 18 x 2 x 2 = 72 potential pairs of estimates. Each estimate came from a combination of five studies or fewer. Factoring in the observed variation in test accuracy between studies alongside the scant evidence base, there is a significant likelihood that the final estimates may not be representative of the tests' true accuracy. There is a significant risk that a test with a larger evidence base published in peer-reviewed journal articles may be disadvantaged in comparison to a test where there is only unpublished manufacturer information at high risk of bias.

We prioritised information from published studies (i.e. not those in manufacturer (submitted directly to NICE in response to a request for information) and FDA documents) where data were available for patients restricted by throat score as per the scope. This provided accuracy data for one pair of estimates, and relaxing the age group restriction provided another pair of estimates. It was necessary to relax the throat score restriction to obtain further estimates. An additional 13 pairs of estimates were obtained from studies which matched the age and care setting of the test. One further pair of estimates were obtained by using estimates from a mixed age population for an adult population. Relaxing the care setting and age restrictions allowed estimation of 24 pairs of test accuracy estimates for children and adult populations. Where multiple options for considering relaxing either age group or setting differences between studies and target population, factors such as sample size and number of studies were also considered. Studies in manufacturer responses to NICE and in FDA documents were only included if no other evidence were available for a specific test. Where these data are used we consider the analysis to be at extremely high risk of bias and we do not recommend these are sufficient to underpin any clinical decisions. The data from neither of these two sources matched a subgroup of interest or were restricted by throat score, but relaxing the ages and care settings provided estimates for a further 32 pairs.

A summary of the studies providing evidence to each estimate can be found in Table 15.

Table 15: Summary of studies providing estimates of test performance for economic modelling, colour coded for reliability.

Test	Primary Children	Primary Adult	Secondary Children	Secondary Adult
Clearview Exact Strep A cassette	1 abstract (Andersen 2003, n = 353) – wrong setting, right age, wrong score restriction	1 abstract (Andersen 2003, n = 353) – wrong setting, wrong age, wrong score restriction	1 abstract (Andersen 2003, n = 353) – right setting, right age, wrong score restriction	1 abstract (Andersen 2003, n = 353) – right setting, wrong age, wrong score restriction
Clearview Exact Strep A dipstick – test strip	1 abstract (Andersen 2003, n = 353) – wrong setting, right age, wrong score restriction	1 abstract (Andersen 2003, n = 353) – wrong setting, wrong age, wrong score restriction	1 abstract (Andersen 2003) – right setting, right age, wrong score restriction	1 abstract (Andersen 2003, n = 353) – right setting, wrong age, wrong score restriction
BD Veritor Plus system group A Strep Assay - cassette	1 study (Berry 2018, n = 215) – wrong setting, right age, wrong score restriction	2 studies (Berry 2018, n = 215; Azrad 2019, n = 100) – wrong setting, wrong age, wrong score restriction	1 study (Berry 2018, n = 215) – right setting, right age, wrong score restriction	2 studies (Berry 2018, n = 215; Azrad 2019, n=100) – wrong setting, wrong age, wrong score restriction
Strep A rapid test - cassette	1 mfr response (Biopanda, n = 526) – right setting, right age, wrong score restriction	1 mfr response (Biopanda, n = 526) – right setting, wrong age, wrong score restriction	1 mfr response (Biopanda, n = 526) – wrong setting, right age, wrong score restriction	1 response study (Biopanda, n = 526) – wrong setting, wrong age, wrong score restriction
Strep A rapid test – test strip	(no data)	(no data)	(no data)	(no data)
NADAL Strep A - test strip	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction	1 mfr response (Nal von minden GmbH, n = 244) – right setting, wrong age, wrong score restriction	1 mfr response (Nal von minden GmbH, n = 244) – right setting, wrong age, wrong score restriction
NADAL Strep A - cassette	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction

NADAL Strep A plus - cassette	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction
NADAL Strep A plus - test strip	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction
NADAL Strep A scan test - cassette	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction	1 mfr response (Nal von minden GmbH, n = 244) – wrong setting, wrong age, wrong score restriction
OSOM Strep A test – test strip	1 study (Llor 2011, n = 116) – right setting, wrong age, right score restriction	3 studies (Bura 2017, n = 101; Llor 2009, n = 222; Llor 2011, n = 276) – right setting, right age, wrong score restriction	2 studies (Rogo 2011, n = 228; Weinzierl 2018, n = 160) – right setting, right age, wrong score restriction	5 studies (Bura 2017, n = 101; Llor 2009, n = 222; Llor 2011, n = 276, Rogo 2011, n = 228; Weinzierl 2018, n = 160) – wrong setting, wrong age, wrong score restriction
QuikRead Go Strep A test kit	1 study (Stefaniuk, n = 43) – right setting, right age, wrong score restriction	1 study (Stefaniuk, n = 52) – right setting, right age, wrong score restriction	2 studies (Azrad 2019, n = 100; Stefaniuk 2017, n = 95) – wrong setting, wrong age, wrong score restriction	2 studies (Azrad 2019, n = 100; Stefaniuk 2017, n = 95) – wrong setting, wrong age, wrong score restriction
Alere TestPack Plus Strep A - cassette	1 study (McIsaac 2004, n = 494) – right setting, right age, wrong score restriction	1 study (Humair 2006, n = 224) – right setting, right age, right score restriction	4 studies (Kurtz 2000, n = 257; Lacroix 2018, n = 1002; Penney 2016, n = 147; Santos 2003, n = 49) – right setting, right age, wrong score restriction	1 study and 1 abstract (Rosenberg 2002, n = 126; Valverde 2018, n = 580) – right setting, wrong age, wrong score restriction
Bionexia Strep A plus - cassette	(no data)	(no data)	(no data)	(no data)
Bionexia Strep A dipstick – test strip	1 abstract (Pauchard 2013, n = 193) – wrong setting, right age,	1 abstract (Pauchard 2013, n = 193) – wrong setting, wrong	1 abstract (Pauchard 2013, n = 193) – right setting, right age,	1 abstract (Pauchard 2013, n = 193) – wrong setting, wrong

	wrong score restriction	age, wrong score restriction	wrong score restriction	age, wrong score restriction
Biosynex Strep A - cassette	(no data)	(no data)	(no data)	(no data)
Sofia Strep A FIA	1 study (Lacroix 2018, n = 1002) – wrong setting, right age, wrong score restriction	1 study (Lacroix 2018, n = 1002) – wrong setting, wrong age, wrong score restriction	1 study (Lacroix 2018, n = 1002) – right setting, right age, wrong score restriction	1 study (Lacroix 2018, n = 1002) – wrong setting, wrong age, wrong score restriction
Alere i Strep A	3 studies (Berry 2018, n=215 ; Cohen 2015, n = 355; Weinzierl 2018, n = 160) – wrong setting, right age, wrong score restriction	1 study (Cohen 2015, n = 126) – wrong setting, right age, wrong score restriction	3 studies (Berry 2018, n=215 ; Cohen 2015, n = 355; Weinzierl 2018, n = 160) – right setting, right age, wrong score restriction	1 study (Cohen 2015, n = 126) – right setting, right age, wrong score restriction
Alere i Strep A 2	1 FDA study (Alere, n = 981) – wrong setting, wrong age, wrong score restriction	1 FDA study (Alere, n = 981) – wrong setting, wrong age, wrong score restriction	1 FDA study (Alere, n = 981) – wrong setting, wrong age, wrong score restriction	1 FDA study (Alere, n = 981) – wrong setting, wrong age, wrong score restriction
Cobas Strep A Assay on Liat system	1 study (Wang 2017, n = 427) – right setting, right age, wrong score restriction	1 study (Wang 2017, n = 427) – right setting, wrong age, wrong score restriction	1 study (Wang 2017, n = 427) – wrong setting, right age, wrong score restriction	1 study (Wang 2017, n = 427) – wrong setting, wrong age, wrong score restriction
Xpert Xpress Strep A	1 FDA report and 1 mfr response (Cepheid, n = 618 and 577) – wrong setting, wrong age, wrong score restriction	1 FDA report and 1 mfr response (Cepheid, n = 618 and 577) – wrong setting, wrong age, wrong score restriction	1 FDA report and 1 mfr response (Cepheid, n = 618 and 577) – wrong setting, wrong age, wrong score restriction	1 FDA report and 1 mfr response (Cepheid, n = 618 and 577) – wrong setting, wrong age, wrong score restriction

Estimates were colour coded by number of studies, sample size of studies, source of information and population relevance. Red indicates highly unreliable data, green indicates multiple reliable and relevant sources. Abbreviations: FDA – Food and Drug Administration (USA), mfr - manufacturer, n – number of patients. Mfr response refers to unpublished accuracy data from the manufacturer provided within the NICE review process.

3.2.6.7 **Accuracy of point-of-care tests using PCR to resolve discordant cases**

Discordant results between point-of-care tests and culture were resolved using PCR in four studies.[19, 23, 34, 70] All discrepant results between a point-of-care test and culture (POC positive culture negative, and vice versa) were analysed in two of these studies.[19, 23]

All twenty samples that were Cobas Liat Strep A-positive but culture-negative were confirmed positive by PCR and bidirectional sequencing, and all three samples that were Cobas Liat Strep A negative and reference culture positive were confirmed positive by PCR and bidirectional sequencing . [23] Wang et al. also examined the discrepancies between the TestPack Plus Strep A test and culture. All discordant results, that is the 20 cases positive by the test and negative by culture and the three cases that test negative and culture positive, were positive according to PCR.

In evaluating the accuracy of the BD Veritor system, Berry et al. also identified 21 discordant results with throat culture, including 11 positive on the index test but not culture, and 10 positive on culture but not the index test. [19] PCR detected GAS in 6 of the 11 results that were positive by BD Veritor System but negative by culture. PCR detected GAS in all 11 samples that were negative by the BD Veritor system but positive by culture. In the same population, Berry et al. found that 14 of the 15 results which were positive for Alere i Strep A test and negative for culture had a positive PCR. There were no reported occasions where Alere i Strep A test gave a negative result when culture gave a positive result.

Similarly, Cohen et al. (2015)[34] and Lacroix et al. (2018)[22] only analysed some of the discrepancies between a point-of-care test and culture. Cohen et al. (2015)[34] identified a total of 24 discordant results between the Alere i Strep A test and culture. There were 18 positive samples on the Alere i Strep A test and not on culture, 13 of which were confirmed positive by PCR, whereas the other 5 results were PCR-negative. Four of the six cases which were positive on culture but not on Alere i Strep A were confirmed negative by PCR.

Lacroix et al. [22] found 84 discordant results between Sofia Strep A FIA and culture. (31 false positives and 53 false negatives). Eleven of the 31 false-positive samples were missing, hence PCR assays could not be performed for these samples. Eleven of those with samples present were confirmed positive by PCR and nine were negative by PCR. Lacroix et al. also found 21 results positive by TestPack Plus Strep A but negative by culture, nine of which were confirmed positive by PCR. Eight were confirmed PCR-negative, leaving four missing samples, which precluded additional PCR assays. Lacroix et al. did not provide test-specific results for the cases that were negative by rapid test and positive by culture.

Table 16 summarizes the key findings from these analyses.

Interestingly, Lindbaek et al. (2004)[40] used a second culture medium (a liquid medium/broth) to resolve discrepant results between the TestPack Plus Strep A test (Abbott) and microbiological culture (strep selective agar). In this study, the second culture medium (Colistin and oxolinic acid [COBA] + TSA sheep + SXT + Lim broth + 1st culture medium) detected Strep A in 17 of 27 (63%) patients who previously tested positive by the Alere TestPack Plus Strep A test but negative by the first culture medium (Columbia agar + horse blood + COBA).

Table 16: Accuracy of point-of-care tests using PCR to arbitrate discordant results with culture

Study Reference	Index test	2x2 contingency tables		
		PCR+	PCR-	Total
Berry 2018 [19]	Alere i Strep A test + Culture -	14	1	15
	Alere i Strep A test - Culture +	0	0	0
	Total	14	1	15
Berry 2018 [19]	BD Veritor system + Culture -	6 (Berry et al also report 5)	5	11
	BD Veritor system - Culture +	10	0	10
	Total	16	5	21
Wang 2017 [23]	Cobas Liat Strep A Assay + Culture -	20	0	20
	Cobas Liat Strep A Assay - Culture +	3	0	3
	Total	23	0	23
Cohen 2015 [34]	Alere i Strep A+ Culture -	13	5	18
	Alere i Strep A- Culture +	2	4	6
	Total	15	9	24
Lacroix 2018 [22]	Sofia Strep A FIA+ Culture -	11	9	31 (11 missing samples)
	Sofia Strep A FIA- Culture +	NR	NR	53
	Total	NR	NR	84
	TestPack Plus Strep A + Culture -	9	8	21

Study Reference	Index test	2x2 contingency tables		
		PCR+	PCR-	Total
Lacroix 2018 [22]				(4 missing samples)
	TestPack Plus Strep A - Culture +	NR	NR	87
	Total	NR	NR	108

3.2.6.8 Direct comparison of point-of-care test accuracy with clinical scores

Six studies directly compared test accuracies between point-of-care tests and clinical scores.[36, 41-43, 49, 54] The results are summarized in Table 17**.** Sensitivity point estimates for clinical scores were higher compared to rapid tests in two studies.[43, 54] However, point estimates for sensitivity and particularly specificity of rapid tests (including TestPack Plus Strep A, OSOM Strep A, QuikRead Go Strep A tests) were generally higher. Sensitivity (82.9% to 94.6%) and specificity (84.9% to 99.1%) point estimates of point-of-care tests were consistently high when compared to point estimates for clinical scores (Sensitivity 73.5% to 97.2%; Specificity 17.2% to 64.8%).

3.2.6.9 Test Failure rate

Five studies reported on test failure rate.[22, 34, 35, 40, 51] These five studies reported on three different point-of-care tests (Alere i, Testpack Strep A Plus and Sofia FIA Strep A). For the Alere i test, the test failure rate ranged from 0%-2.8%.[34, 51] The TestPack Strep A plus test failure rate ranged from 0.3-1.3%.[35, 40] Whereas the Sofia FIA strep test failure rate was reported as 4.7%.[22] Differences could be due to environmental factors such as staff training as opposed to issues with the tests.

Table 17: Direct comparison of point-of-care test accuracy with clinical scores

Study Reference	Clinical score	Test Accuracy Statistics for clinical scores					Test Accuracy Statistics for index tests					
		culture +	culture -	Total	Sensitivity	Specificity	Index test	culture +	culture -	Total	Sensitivity	Specificity
Humair 2006 [36]	Centor score > 2	105	119	224	0.750 (0.678 to 0.822)	0.487 (0.423 to 0.551)	TestPack Plus Strep A+	128	11	139	0.914 (0.852 to 0.953)	0.953 (0.914 to 0.947)
	Centor score ≤ 2	35	113	148			TestPack Plus Strep A-	12	221	233		
	Total	140	232	372			Total	140	232	372		
Llor 2009 [41]	Centor score > 2	47	104	151	0.855 (0.761 to 0.948)	0.377 (0.304 to 0.451)	OSOM Strep A+	52	14	66	0.946 (0.839 to 0.986)	0.916 (0.861 to 0.952)
	Centor score ≤ 2	8	63	71			OSOM Strep A-	3	153	156		
	Total	55	167	222			Total	55	167	222		
Llor 2011 [42]	Centor score > 2	36	80	116	0.735 (0.587 to 0.846)	0.648 (0.581 to 0.709)	OSOM Strep A+	44	14	58	0.898 (0.770 to 0.962)	0.938 (0.897 to 0.933)
	Centor score ≤ 2	13	147	160			OSOM Strep A-	5	213	218		
	Total	49	227	276			Total	49	227	276		
McIsaac 2004 [43]	McIsaac score > 2	193	375	568	0.847 (0.792 to 0.889)	0.329 (0.291 to 0.370)	TestPack Plus Strep A+	189	5	194	0.829 (0.772 to 0.874)	0.991 (0.978 to 0.997)
	McIsaac score ≤ 2	35	184	219			TestPack Plus Strep A-	39	554	593		
	Total	228	559	787			Total	228	559	787		
Pauchard 2013 [54]	McIsaac score > 2	69	101	170	0.972 (0.893 to 0.995)	0.172 (0.112 to 0.253)	Strep A Rapid Test +	60	11	71	0.845 (0.735 to 0.914)	0.910 (0.841 to 0.952)
	McIsaac score ≤ 2	2	21	23			Strep A rapid Test -	11	111	122		
	Total	71	122	193			Total	71	122	193		
Stefaniuk 2017 [49]	Centor score > 2	37	39	76	0.861 (0.714 to 0.942)	0.250 (0.145 to 0.392)	QuikRead Go Strep A+	39	8	47	0.907 (0.770 to 0.970)	0.849 (0.719 to 0.928)

Study Reference	Clinical score	Test Accuracy Statistics for clinical scores					Test Accuracy Statistics for index tests					
		culture +	culture -	Total	Sensitivity	Specificity	Index test	culture +	culture -	Total	Sensitivity	Specificity
	Centor score ≤ 2	6	13	19			QikRead Go Strep A-	4	44	48		
	Total	43	52	95			Total	43	52	95		

3.2.7 Proposed pathway (combined strategy of clinical score and point-of-care tests)

3.2.7.1 Test accuracy of combined clinical score and point-of-care test with culture as reference standard

None of the included studies evaluated the accuracy of a combined strategy of a sore throat clinical score (at the recommended NICE cut-offs of Centor / McIsaac score ≥3 or FeverPAIN ≥4) with point-of-care test. This would require the combination of the two methods into a single procedure, where positive results are produced by individuals with both a high clinical score and a positive point-of-care test, and negative results are given either by patients with a low clinical score or patients with a high score but a negative point-of-care test. As shown in Table 18, Rosenberg et al. (2002)[47] provides the only available evidence that attempts to match the proposed pathway, but not at the recommended Centor cut-off.

Table 18: Accuracy of combined Centor score of 2 / 3 and rapid testing with culture as reference standard

Study Reference	Combined strategy	Test Accuracy statistics for clinical scores				
		culture +	culture -	Total	Sensitivity for patients with Centor score 2 or 3	Specificity for patients with Centor score 2 or 3
Rosenberg 2002 [47]	Centor score of 2 or 3 AND TestPack Plus Strep A+	12	0	12	For patients with Centor score 2 or 3: 0.80 (0.52, 0.96) Overall: 0.88 (0.71, 0.96)	For patients with Centor score 2 or 3: 1.00 (0.92, 1.00) Overall: 0.78 (0.68, 0.86)
	Centor score of 2 or 3 AND TestPack Plus Strep A–	3	44	47		
	Centor score <2, AND no rapid test	1	29	30		
	Centor score 4, or 5 AND no rapid test	16	21	37		
	Total	32	94	126		

3.2.8 Other outcomes

No information was found on number of appointments required per episode, morbidity, mortality, onward transmission of infection, health-related quality of life, patient satisfaction with the test or healthcare professional satisfaction with the test.

Twelve studies reported on antibiotic prescribing behaviours. RCTs and before and after studies have been described in Table 19, Table 20 and Figure 10. The remaining 8 studies which included one armed cohorts or hypothetical antibiotic management are briefly summarised in the text.

Table 19: RCTs on antibiotic prescribing behaviour

Study reference	Country	Index test	Study details	Antibiotic prescribing behaviour
Little 2013[6]	UK	Alere TestPack Plus (IMI TestPack)	3 armed trial with a delayed antibiotics arm (clinical assessment without a tool), a clinical tool arm and a rapid test following clinical tool arm. Clinicians given guidance to follow on prescribing. Arm 1: Delayed antibiotics control arm - depending severity of presentation patients were either given antibiotics, given no antibiotics or given a delayed prescription to collect after three to five days if symptoms didn't improve or worsened. Arm 2: Clinical score arm - patients assessed using FeverPAIN. Scores of 0 -1 were not offered antibiotics. Immediate antibiotics were offered for scores above 4 and for scores of 2 or 3 delayed antibiotics were offered. Arm 3: RADT arm - Those with score of clinical score 0 or 1 were not offered antibiotics or an RADT, those with a score of 2 were offered delayed antibiotics and those with scores above 3 were	Antibiotics offered immediately or a delayed prescription to 89% (185/207) in delayed prescription control arm, to 59% (124/211) in the clinical score arm and 40% (86/213) in the clinical score plus RADT arm. Use of antibiotics ascertained from the patients with incomplete responses as follows: 46% (75/164) used antibiotics in the delayed prescription arm compared to 37% (60/161) in the clinical score arm and 35% (58/164) in the clinical score plus RADT arm

			given an RADT. All those with negative RADT were not offered antibiotics	
Llor 2011[42]	Spain	OSOM Strep A test	2 arm cluster randomised trial. Healthcare centres randomised to intervention (RADT) arm or control arm (management with clinical criteria only).	Control arm GPs prescribed antibiotics in 64% (168/262) patients compared to 44% (123/281) RADT arm. Of the 60 test positive cases 59 were given antibiotics (98%). In those who the test was negative 69/225 were given antibiotics (31%) Across both trial arms, antibiotic treatment was 'inappropriate' (as culture was negative) in 40% (210/526) of patients, and in 3% (16/526) of patients antibiotics were not prescribed when culture was positive. 153 of these cases occurred in the control arm, and 73 in the RADT arm, category of inappropriate decision (over or underprescribing) is not reported by trial arm.
Worrall 2007[52]	Canada	Clearview Exact Strep A	4 armed trial: control arm using clinician's independent decisions as usual practice, arm using sore throat decision rules (STDR, ≤1 no need for antibiotics, 2 decisions made by the clinician, 3 or 4 antibiotics needed), arm using a rapid test (RADT), and arm using both STDR and RADT (≤1 no need for antibiotics, 2 RADT, 3 or 4 antibiotics needed.. Clinician's were recommended to follow the guidance but it was not enforced	46.7% (247/533) of patients received antibiotics. 58% (82/141) usual practice, 55% (94/170) with Centor score alone compared to 27% (32/120) with rapid antigen testing alone and 38% (39/102) with combined rapid antigen testing and Centor score

3.2.8.1 **Antibiotic prescribing behaviours: RCT evidence**

There were three randomised controlled trials (RCTs) reporting on antibiotic use. All three trials found higher antibiotic prescription rates or use in control arms with no point-of-care test compared to those given a point-of-care test.

In the UK RCT in primary care by Little et al[6] patients (mean ages 29 and 31 years across arms, no age range provided) in a primary healthcare setting were randomly assigned to a delayed antibiotics control arm, a clinical score arm or a rapid antigen detection tests (RADT) arm (IMI TestPack, later known as Alere). In the delayed antibiotics control arm, depending on the severity of their presentation patients were either given antibiotics, given no antibiotics or given a delayed prescription to collect after three to five days if symptoms didn't improve or worsened. This control group was there to represent current UK practice at the time. In the clinical score arm patients were assessed using the FeverPAIN clinical scoring tool. Patients with scores of 0 or 1 were not offered antibiotics. Immediate antibiotics were offered for scores above 4 and for scores of 2 or 3 delayed antibiotics were offered. In the RADT group all patients also received the clinical scoring tool. Those with score of 0 or 1 were not offered antibiotics of RADT, those with a score of 2 were offered delayed antibiotics and those with scores above 3 were given an RADT. All those with negative RADT were not offered antibiotics. There were 207 patients in the delayed prescribing arm, of which 79% (164/207) received a delayed prescription, 10% received no antibiotics (21/207) and 10% (21/207) received immediate antibiotics. In the clinical score arm 41% (87/211) received a delayed prescription, 41% (87/211) received no antibiotics and 16% (33/211) received immediate antibiotics. In the RADT there were fewer delayed prescription decisions, with only 23% (48/213) patients receiving a delayed prescription, 59% (126/213) patients were offered no antibiotics and 18% (38/213) were given immediate antibiotics. Patients reported antibiotic use of 46% (75/164) in the delayed prescription arm, 37% (60/161) in the clinical score arm, and 35% (58/164) in the clinical score plus RADT arm. The total numbers in each arm were considerably lower for antibiotic use, indicating significant loss to follow up, so these numbers should be interpreted with caution. Likewise, symptom severity was worse in the control arm, so effect sizes may be overestimates. This was a UK based trial based in a primary healthcare setting. For this reason it is likely to generalise to the UK population.

The second trial was by Llor et al.[42] They included patients over the age of 14 (mean age 31.7 years) visiting primary healthcare centres across Spain. This was a cluster randomised controlled trial with the centre as the unit of randomisation. This form of randomisation can be prone to imbalancing baseline characteristics of patients, however, authors reported no significant differences in baseline characteristics (such as gender, mean age and by clinical symptoms) between the participants across intervention and control arms. Patients were randomised to either a control arm in which patients were only assessed using clinical criteria (Centor) or an intervention arm where patients were assessed with both a Centor score and a RADT (OSOM Strep A test). In total 54% (291/543) patients were prescribed antibiotics. Antibiotics were more likely to be prescribed in the clinical score only arm, with GPs prescribing antibiotics in 64% (168/262) patients compared to 44% (123/281) in the RADT arm. There was a correlation between Centor score and antibiotic prescription rates across both groups, with more antibiotics prescribed to those with higher scores (score of 4 - 80% antibiotics [37/46] in intervention arm and 96% [43/35] in the control arm compared to 16% [4/70] in intervention arm and 33.% [20/61] in control arm). In the subgroup of interest to the UK population (those with Centor scores of three or more) 74% (90/122) were given antibiotics in the intervention arm compared to 85% (100/119) in the control arm. Antibiotic appropriateness is also discussed in the trial. 98% (59/60) patients with a positive RADT were given antibiotics and 31% (69/225) with a negative test results and received antibiotics. The authors determined that treatment was inappropriate (based on culture results) in 43% of patients (226/526) with 210 unnecessary prescriptions and 16 untreated cases. 153 of these cases occurred in the control arm, and 73 in the RADT arm however the category of inappropriate decision (over or underprescribing) is not reported by trial arm.

The third trial was a four-armed cluster randomised trial in Canada by Worrall et al.[52] The trial included 40 physicians who were asked to consecutively recruit adult patients (aged 19 years or older, no further details reported). There was a control arm using usual clinical practice, an intervention arm using sore throat decision rules (STDR, modified Centor), an intervention arm using a rapid test (RADT) and an intervention arm using both STDR and RADT. In the STDR group clinical scores of ≤1 no antibiotics were recommended, scores of 3 or 4 antibiotics were recommended and scores of 2 the prescribing decision lay with clinicians. In the combined STDR and RADT group, RADT was only used for scores of 2. It is implied although not explicitly stated that all those in the RADT arm received an RADT.

They found that 47% (247/533) of patients received antibiotics. By arm, 58% (82/141) of patients received antibiotics from the usual practice arm, compared to 55% (94/170) with Centor score alone, 27% (32/120) with rapid antigen testing alone and 38% (39/102) with combined rapid antigen testing and Centor score. As this was a cluster-randomised trial and each arm included only 8-10 doctors, differences could be due to differences between doctors rather than between strategies. Additionally, it may be due to differences in patients across arms. The study only reports on the characteristics of the physicians, we have no baseline patient data. Finally, as the Canadian medical system differs to the UK system so the results may not be generalizable.

There was no RCT evidence on molecular technologies and antibiotic prescribing rates.

Table 20: Before and after study on antibiotic prescribing behaviours

Study reference	Country	Index test	Study details	Antibiotic prescribing behaviour
Bird 2018[32]	UK	Bionexia Strep A	Prospective cohort before and after study. Baseline antibiotic prescribing data were collected retrospectively from October-November 2014 (method of diagnosis in this phase is not reported) and compared (following introduction of a new algorithm, RADT for those with a McIsaac score of more than 3) to rates in August-November 2015 and September-November 2016. Only positive RADT given antibiotics but clinicians could prescribe if they still had high clinical suspicion of GAS pharyngitis	Following implementation of an algorithm combining McIsaac scores and Bionexia Strep A rapid testing, antibiotic prescribing rates fell steeply from 79% (166/210) at baseline to 24% (51/214) in year one and 28.2% (51/181) for the second year.

3.2.8.2 **Antibiotic prescribing behaviours: Before and after studies**

There was one study which was a before and after study assessing antibiotic prescribing rates. The study by Bird et al[32] analysed children (6 months to 16 years) presenting to a UK pediatric emergency department with a sore throat. The study compared baseline data from

October and November 2014 to prescribing rates the following two years (August to November 2015, and September to November 2016) following the implementation of using both a McIsaac score and an RADT. Baseline data were collected retrospectively from a departmental audit, when it is implied that the method of diagnosis was just clinician examination, with the aim to assess the impact of a clinical scoring system and rapid test on prescribing rates. A rapid test could only be requested if there was a McIsaac score of 3 or more. Following implementation, antibiotic prescribing rates fell steeply from 79% (166/210) at baseline to 24% (51/214) in year one and 28% (51/181) for the second year. However, seasonality may be a confounding factor, with higher prescribing rates over the later autumn months (October and November) than late summer (August and September). Likewise, there may be some regression to the mean, as the high initial prescribing rates may have prompted the study but may be subject to fluctuations.

There were no two armed cohort studies analysing molecular technologies and antibiotic prescribing rates.

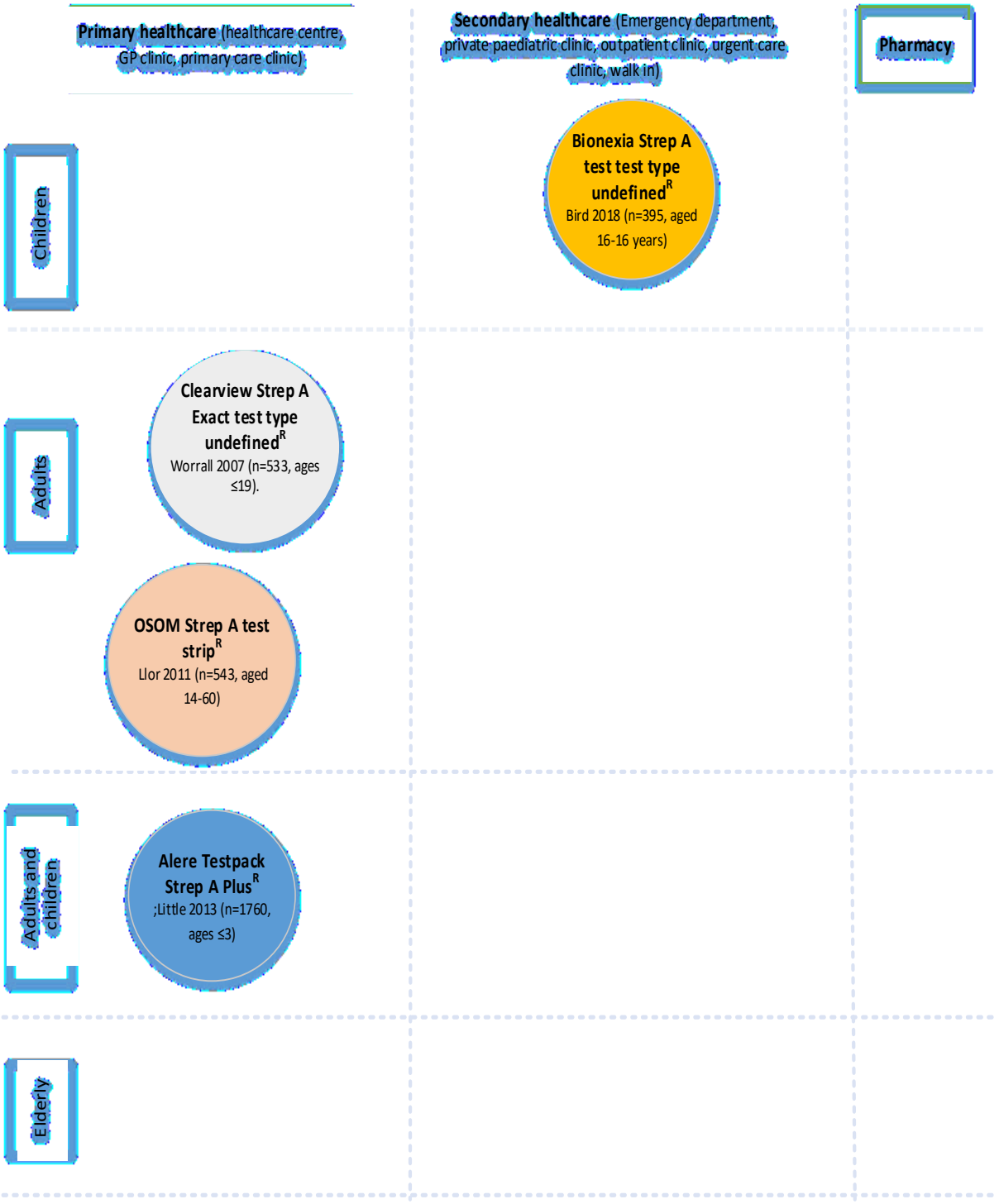


Figure 10: Diagram of studies which show comparative data (RCTs and before and after studies) on antibiotic prescribing rates by test type, setting and population

Note: Lines between tests indicate head-to-head (direct) comparisons (none were found).

3.2.8.3 **Antibiotic prescribing behaviours: Other study designs**

There were an additional eight studies which reported antibiotic prescribing behaviours in single arm cohorts.[19, 33, 36, 37, 43, 47, 49, 50] No comparative data were possible within these study designs, only hypothetical comparisons so all results within this section should be interpreted with caution, and considered less informative than the RCT results. In these cohort studies patients received the same intervention however authors also determined hypothetical management scenarios and compared how this would have affected antibiotic prescribing rates. Of the 8 studies, 3 of the studies provided hypothetical rules, so does not help inform us on real world behaviour. These 3 studies have been included and briefly summarised but were not quality appraised. 5 studies report on either what happened in the real world, or what clinician's reported they would do. These studies suggested that using a rapid test would decrease antibiotic use by as little as 9% up to 74%.[19, 33, 37, 47, 49]

Two of the five single arm cohorts reported on real world behaviour. The first study, by Stefaniuk et al[49], examined children and adults in a primary care setting in Poland. 46% (44/96) of the study group were children aged 3-14 years and 25% (24/96) were aged 31-35 years (overall mean age was not provided). 98% (46/47) of patients with a positive QuikRead Go Strep A test received antibiotics and 24% (12/49) patients with both a negative rapid test and culture were treated with antibiotics.

The second study reporting on real world behaviour was by Berry et al[19] compared BD Veritor testing to Alere i testing and a chart review to determine hypothetical impact of results on antibiotic use . The study took place in pediatric outpatient clinics (mean age not reported) in the USA. Prescribing decisions were made with knowledge of the BD Veritor test results, but not the Alere i test or culture. The authors found 34% (73/215) were prescribed antibiotics, of these 25 were prescribed at clinic visit and were later deemed to be inappropriate treatment (on the basis of culture results). Of these 20/25 (80%) were negative on BD Veritor, Alere i and culture, 5 were positive only with BD Veritor. Of the 215 who did not receive antibiotics, 13 BD veritor negative cases were identified by the authors as potential missed cases on the basis of PCR and Alere i positive results, of which 6 received antibiotics within 6 days of the original appointment. These analyses provide descriptive behaviour data using BD Veritor test, but cannot be used to compare Alere and BD Veritor

for appropriateness of prescribing behaviour as decisions were made using BD Veritor and not Alere i. This study using Alere i was the only study to use a molecular technology, and no prescribing behaviours were based upon it, hence there is no evidence on molecular technologies and antibiotic prescribing rates.

Three of the single arm cohort studies reported on hypothetical scenarios based around clinician's decisions.[33, 37, 47] Bura et al[33] examined a cohort of adults (median 26 years, range 18-44) in primary care in Poland with Centor scores of above 2 (this was a case-control design for test accuracy outcomes, but cohort for prescribing behaviour). All patients and controls were given a rapid test and culture. GPs could then choose whether to give antibiotic therapy. It is stated that this choice was not influenced by the research team however we cannot be certain of this as they were aware of the rapid test result. Clinicians were aware of the Centor score at the time of antibiotic prescribing. They found 58% (59/101) patients received an antibiotic. All RADT positive cases received treatment including 2 who were culture negative. In addition 46% (35/77) of test negative cases received antibiotics. They determined that 40% (23/59) of cases received an unnecessary antibiotic prescription. Unnecessary has been defined here as being culture negative. The authors also gave hypothetical management scenarios based upon different Centor scores and scenarios. Antibiotics would be prescribed to 29% (11/38) with a Centor score of 2, 62% (23/37) with a Centor score of 3 and 96% (25/26) with a score of 4. They surmised that 23% (23/101) would have been treated using positive culture results alone and 24% (24/101) would have been treated using a rapid test, meaning one person was mistakenly given antibiotics. However, 54% of those given antibiotics were treated for non-GAS. From the control group, one person would have been treated with antibiotics, additionally other forms of streptococcus were identified in 13 people from this group.

The study by Rosenberg et al.[47] was a one armed prospective observational cohort in which all patients were given a clinical score (Centor), rapid test and culture. The study was on patients older than 3 years (47% [59/126] aged 3-14, 50% [63/126] aged 15-44 and 3% [4/126] aged 45 and above), presenting to an emergency department in Canada. They also report physician's clinical impressions and their hypothetical management. Authors report on score alone, physician examination alone, rapid test alone or rapid test for clinical scores

above 3. They found that physicians prescribed antibiotics to 37% (46/126) of patients, after obtaining the results of the rapid test, of these 18 had negative culture results. They hypothesised that 20% (25/126) would have received antibiotics in the rapid test group, compared to 29% (37/126) in the clinical score group.

The last study by Johannson et al.[37] was a prospective observational single armed cohort in which all patients received both a rapid test and culture and these results were compared to hypothetical management suggestions made by physicians. They included adult patients (aged 25-44 years, mean age not reported) reporting to primary health care centres in Sweden. Physicians also clinically assessed patients, and gave hypothetical management suggestions based upon their level of certainty for Strep A (absolutely positive, positive, possibly positive, possibly negative, negative, absolutely negative). No results are clearly provided, however 26% (24/94) patients with a negative rapid test received treatment, it is unclear how many of these were culture positive.

There were three additional studies which reported on hypothetical prescribing decisions based on assumptions about doctors behaviour, however no real world decisions were reported and doctors were not asked about behaviour.[36, 43, 50]

3.3 Summary of the clinical effectiveness findings and implications for the health economic model

Overall the findings reveal wide variations in the sensitivity (67.9% to 100%) and specificity (73.3% to 100%) estimates of point-of-care tests. These estimates were 82.9% to 94.6% for sensitivity and 84.9% to 99.1% for specificity in high-risk populations, including patients with Centor/McIsaac scores > 2, which represents the population of interest. These estimates do not account of any of the unpublished manufacturer submissions.

Clinical scoring tools (FeverPAIN and Centor) have been proposed as a method by which clinicians can identify which patients are most likely to benefit from antibiotic use for sore throat.[8] These tools were developed to predict Strep A (Centor, FeverPAIN), C (FeverPAIN) and G (FeverPAIN). Most studies making direct comparisons between sore-throat clinical scoring tools and point-of-care tests indicated that sensitivity estimates were higher for the point-of-care tests, and that specificity was generally comparable between the two approaches.

A methodological limitation of the clinical scoring tools concerns the varying way they have been implemented across the included studies. For instance, different studies apply different clinical score cut-offs when recruiting patients. None of these studies matched the proposed pathway of care and treatment for patients with acute streptococcal pharyngitis, which would entail evaluating the test accuracy of a combined strategy of sore throat clinical scores at the recommended NICE thresholds (Centor / McIsaac ≥3 or FeverPAIN ≥4) and point-of-care tests. This limitation potentially holds important economic implications as attempts to model this proposed pathway may not be informed by the availability of empirical data. In addition, the overrepresentation of the TestPack Plus Strep A test relative to other point-of-care tests, as well as the overlap of patients across different age groups potentially raises applicability concerns in the economic model.

Investigation of discordant results between index tests and the reference standard of culture was available for several studies using PCR or culture. This analysis indicated that using culture as the reference standard may have resulted in underestimating sensitivity (specificity

estimates derived used PCR were too variable to draw conclusions about potential over/underestimation by culture). However, PCR can detect indolent GAS so the extent of this is unclear.

Data for test accuracy were sparse for each combination of test, population and setting. There were very few head-to-head (direct) comparison studies between index tests. There was heterogeneity between studies, the cause of which is unclear due to a lack of direct comparison data of different age groups, settings or tests within the same study.

Test accuracy point estimates in manufacturers submissions may be systematically higher than in the peer reviewed literature and the study characteristics are often unclear. Therefore, there is a risk of making inappropriate comparisons between tests in the economic model where one test has a range of peer-reviewed publications and another only has manufacturer data.

With the exception of a single study using the Sofia Strep A FIA test (failure rate = 4.7%),[22] failure rates for point-of-care tests were generally low (0 to 2.8%) and unlikely to hold any major implications for the economic model, especially the data for this outcome have not been reported in most of the included studies.

There was no evidence found on time to antimicrobial prescribing decision, number of appointments required per episode, and onward transmission of infection.

The findings also suggest that RADT may help reduce antibiotic prescription rates in patients who receive these tests compared to patients assessed using only a clinical scoring tool. The three randomised controlled trials addressing this question all found up to 30% fewer antibiotics were prescribed following the administration of a RADT. There were no studies identified assessing the use of molecular technologies and antibiotic prescription rates.

4 Cost-effectiveness

4.1 Systematic review of existing cost-effectiveness studies

4.1.1 Introduction

This chapter will explore and review all published cost-effectiveness studies including any existing economic models of the use of different rapid antigen detection or molecular tests (as listed in the final scope and protocol for detection of GAS in detail). Studies providing resource use, costs, utilities and probabilities, useful to inform economic modelling were also identified.

4.1.2 Methods

4.1.2.1 Search strategy

A comprehensive search of the literature for published economic evaluations (including any existing models), cost studies and quality of life (utility) studies was performed. The systematic search included searching the following electronic databases during January 2019 (on 22nd, 29th and 30th January 2019) and an updated search was conducted on all databases during March 2019 (on 7th and 13th March 2019):

- MEDLINE and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions (Ovid);
- Excerpta Medica database (Embase) (Ovid);
- National Health Service Economic Evaluation Database (NHS EED) and Health Technology Assessment (HTA) database (CRD);
- Science Citation Index and Conference Proceedings Science (Web of Science);
- Cost-Effectiveness Analysis (CEA) registry;
- EconPapers (Research Papers in Economics (RePEc)); and
- School of Health and Related Research Health Utilities Database (ScHARRHUD).

The search terms included economic and quality of life (QoL) terms combined with either sore throat or GAS. No date limits were applied and databases were searched from inception. The search strategy was developed by an experienced information specialist, based on the clinical effectiveness review and with input from a health economist. Details of the full

search strategies are provided in Appendix 5. In addition to these searches, any relevant cost-effectiveness studies identified during the clinical effectiveness review were brought to the attention of the reviewers and assessed for eligibility alongside the results of this review.

4.1.2.2 **Assessment of eligibility**

Citations and abstracts from the electronic online databases were exported into a citation software package (X7, Thomson Reuters, CA, USA) and duplicate records identified and removed. Two reviewers independently reviewed titles and abstracts to identify potentially relevant papers for inclusion. Any discrepancies were resolved by discussion.

4.1.2.3 **Inclusion criteria**

Only studies meeting the following inclusion criteria were included in the review.

- Study type: fully published economic evaluations (including any economic models).
- Population: people aged 5 and over presenting to healthcare providers in a primary care (GP surgeries, community pharmacies and walk-in centres) or secondary care (urgent care/walk-in centres and emergency departments) setting with symptoms of an acute sore throat.
- Intervention: 17 rapid antigen detection tests or 4 molecular tests (as described in section 1.3).
- Comparator: antibiotic prescribing decisions using clinical judgement and a clinical scoring tool such as FeverPAIN or Centor.
- Outcomes: cost-benefit or cost-consequences or cost-effectiveness or cost-utility studies reporting outcomes as cost-consequence measures or clinical effectiveness measures or utility measures (utility, EQ-5D or SF-6D score or QALYs).

4.1.2.4 **Exclusion criteria**

Studies meeting the following exclusion criteria were excluded from the review:

- Non-English-language publications
- Studies not in humans
- Studies not in GAS or sore throat

- Studies with the wrong test or no specified test
- Studies which were not full economic evaluations (incremental costs and incremental benefits)

Studies which provided useful information for the economic model such as resource use, costs, utilities and probabilities were retained but were not included in this review.

4.1.2.5 **Data extraction**

Data extraction was carried out by one reviewer using standardised data extraction sheets and was then checked by a second reviewer. Data extracted included the following information:

- study details: study title, author names, source of publication, language and publication type;
- baseline characteristics: population (and subgroups), intervention, comparators, outcomes, study design, setting and location and type of economic evaluation;
- methods: study perspective, time horizon, discount rate, measurement of effectiveness, measurement and valuation preference-based outcomes, resource use and costs, currency, price date and conversion, model type, assumptions and analytical methods;
- results: study parameters, incremental costs and outcomes and reporting of uncertainty;
- discussion: study findings, limitations, generalisability and conclusions; and
- other: sources of funding, conflicts of interest and any comments.

4.1.2.6 **Data synthesis**

Information extracted from the included studies were summarised and tabulated. Findings from individual studies were compared narratively.

4.1.2.7 **Quality assessment**

The quality of full economic evaluation studies that were identified were assessed using the consolidated health economic evaluation reporting standards (CHEERS) checklist by one reviewer and cross-checked by a second reviewer. The CHEERS checklist comprises six dimensions (including title and abstract, introduction, methods, results, discussion and other)

and under these dimensions, a series of questions check whether the criteria have been clearly reported.[71] If the studies included any model-based economic evaluations, they were further critically appraised using the framework on quality assessment for economic modelling developed by Phillips et al (2004).[72] The framework assesses models under the dimensions of structure, data and consistency and whether the criteria has been clearly reported.

4.1.3 Results

4.1.3.1 Search results

The literature search identified 6,980 records through electronic database searches and other sources. After removing duplicates, 2,756 records were screened for inclusion. One article was found via our clinical effectiveness search. Based on title and abstract sift only, 2,737 records were excluded. The remaining nineteen records were included for full-text screening. A further sixteen articles were excluded at the full-text stage, as these studies did not contain a full economic evaluation or specify the right test, see Appendix 6 for further comments. Only three articles were data extracted and quality assessed.

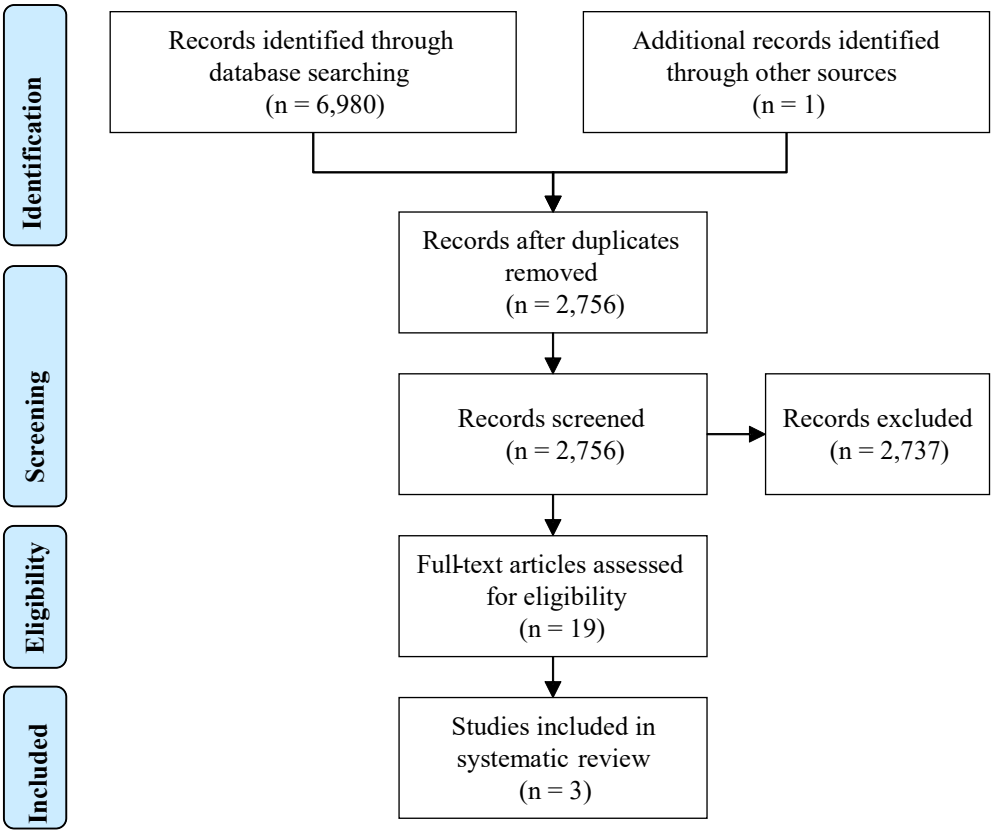


Figure 11: Prisma flow diagram for economic evaluation studies

The literature search identified three studies which had evidence pertaining to incremental costs and outcomes: Bura et al (2017);[33] Humair et al (2006);[36] and Little et al (2014).[73]

The economic information from the first two studies have been summarised below as there was not enough information for a full data extraction. These two studies did not explicitly state the following: study perspective, time horizon, type of economic evaluation, measurement of effectiveness or analytical methods. Bura et al [33] was a prospective case-control study consisting of 101 adults (aged 18-44 years) who went to GP clinics in Poland because of sore throat lasting no longer than 7 days. Controls (n = 101) were volunteers from the same area, who were matched to cases according to their age and sex. The study was conducted over one year. The OSOM Strep A test (Sekisui diagnostics) in conjunction with throat culture was compared with Centor and throat culture to confirm presence of GAS. The costs of diagnosing and treating GAS included: symptomatic treatment, test cost (€1.39), single culture to identify GAS, antibiotic therapy and anti-microbial medications. Economic analysis of five strategies were compared for treating patients with GAS in terms of cost per patient with appropriate Group A Strep treatment ranged from €2.89 (for treat only rapid antigen detection test positive cases) to €6.93 (for treat only Group A Strep + (culture-positive) cases). The authors concluded that the use of the rapid test significantly increases the number of people with GAS related pharyngitis to be treated with antibiotics.

Humair et al (2006)[36] was a prospective cohort study consisting of 372 adults (aged 15-65 years) who were treated at a GP clinic in Switzerland. The Alere TestPack Plus Strep A (Abbott) was compared with throat culture. A decision tree model was used to compare antibiotic prescription for five strategies. Information used in the decision model included antibiotic rate for appropriate use, overuse in patients without GAS, underuse in patients with GAS, appropriate treatment for patients with Group A Strep and without treatment in patients without Group A Strep. The model did not consider QoL, complications or adverse drug effects. Costs were in US$ in 2002 prices. Costs included 10-day course of penicillin, test cost $5.00 and $18.00 for throat culture. The authors found that systematic throat culture had the highest rates of appropriate treatment; whereas empirical treatment in patients with clinical scores of 3 or 4 resulted in the most antibiotic overuse. The cost per case

appropriately treated ranged from $15.30 (systematic rapid test) to $32.40 (systematic throat culture). Sensitivity analyses were performed to check the robustness of results. The authors concluded that the rapid test is a valid test for diagnosis of GAS.

Little et al (2014)[73] conducted an economic analysis alongside a RCT in the UK which included both adults and children with acute sore throat who were seen in primary care clinics (see Table 21). They compared randomised patients to targeted antibiotic use according to (1) delayed antibiotics (control group), (2) clinical score using FeverPAIN or (3) RADT - Alere TestPack Plus Strep A (Abbott) was used according to clinical score. The analysis was from an NHS perspective and the time horizon was short (14 and 28 days), hence long-term effects were not captured. Health related quality of life (HRQoL) was evaluated using EQ-5D. Quality-adjusted life years (QALYs) were adjusted for baseline differences and were calculated using mean EQ-5D scores obtained from the 14-day diary records. It was assumed that the HRQoL changes linearly over time. The analysis included a cost-effectiveness analysis (cost per change in symptom severity) and a cost-utility analysis (cost per QALY). Cost-effectiveness acceptability curves (CEACs) was generated using bootstrapping with 5,000 samples.

The mean symptom scores were adjusted for baseline differences and for the cost-effectiveness analysis, the clinical score group dominated both the delayed antibiotic group and the RADT group, as it was more clinically effective (lower symptom score) and less costly. However, the point estimate of symptom score and the corresponding 95% confidence intervals for clinical score and RADT groups were quite close. The CEAC showed that if the value of a point change in the symptom score was varied between £0 and £500, and it was found that over the entire range the clinical score group was most likely to be cost-effective. In the cost-utility analysis, the delayed group was dominated by the clinical score group for both the timeframes. The incremental cost-effectiveness ratio (ICER) for RADT group compared to clinical score group was £74,286 for the 14-day time frame and £24,528 for the 28-day time frame.

Table 21: Data extraction for cost-effectiveness studies

Study details	
Study title	PRImary care Streptococcal Management (PRISM) study: in vitro study, diagnostic cohorts and a pragmatic adaptive randomised controlled trial with nested qualitative study and cost-effectiveness study
First author	Paul Little
Co-authors	Richard Hobbs, Michael Moore, David Mant, Ian Williamson, Cliodna McNulty, Gemma Lasseter, MY Edith Cheng, Geraldine Leydon, Lisa McDermott, David Turner, Rafael Pinedo-Villanueva, James Raftery, Paul Glasziou and Mark Mullee on behalf of the PRISM investigators
Source of publication	Health Technology Assessment. 2014 Vol 18(6)
Language	English Language
Publication type	Original article
Inclusion criteria/study eligibility/PICOS	
Population (and subgroups)	Patients aged ≥ 3 years, who had acute sore throat
Intervention(s)	Rapid antigen detection tests (RADTs) used with clinical score (FeverPAIN). All patients received the clinical scoring tool. Those with score of 0 or 1 were not offered antibiotics or a RADT, those with a score of 2 were offered delayed antibiotics and those with scores above 3 were given an RADT. All those with negative RADT were not offered antibiotics.
Comparator(s)	Delayed antibiotics (control group) or clinical score only. In the control group, depending on the severity of their presentation patients were either given antibiotics, given no antibiotics or given a delayed prescription to collect after 3-5 days if symptoms didn't improve or worsened. In the clinical score group patients were assessed using the FeverPAIN clinical scoring tool. Patients with scores of 0 or 1 were not offered antibiotics. Immediate antibiotics were offered for scores above 4 and for scores of 2 or 3 delayed antibiotics were offered.
Outcome(s)	Point change in symptom severity score (primary outcome measure in trial) and quality-adjusted life years (QALYs) based on EQ-5D, The symptom severity score is a two-item score (sore throat, difficulty swallowing), each symptom was scored 0 = no problem to 6 = as bad as it can be. A higher score indicates worse symptoms.
Study design	Economic analysis alongside a clinical trial
Setting and location	GP clinics in south and central England
Type of economic evaluation	Cost-effectiveness and cost-utility analysis
Methods	
Study perspective	NHS perspective
Time horizon	14 days and 28 days (1 month) after randomisation
Discount rate	Not applicable
Measurement of effectiveness	EQ-5D measured completed at baseline and 14 days after recruitment and recorded in a patient-completed diary.

Measurement and valuation of preference-based outcomes	EQ-5D values were scored using the standard UK tariff.
Resource use and costs	Resource use data were obtained from GP case notes and from study clinicians. Data included GP and nurse practitioner visits; antibiotics; practice visits for complications of infections and antibiotic complications; and hospital admissions related to infections. Costs included test costs, staff time, medications, complications and hospital admissions. Unit costs were obtained from the Unit Costs of Health and Social Care, NHS reference costs and NHS drug tariff. The costs associated with the clinical score plus the test comprised the additional time required to provide the intervention as well as the cost of the RADT (£3.25 per test; £65 for 20 tests).
Currency, price date and conversion	Costs are in 2010/11 prices in UK pound sterling
Model type	None, as it was based on trial data
Assumptions	EQ-5D for the end of the 28-day follow-up period were not available; therefore, the values obtained at the end of the 14-day period were assumed to persist to the end of the study period, that is the last value obtained was carried forward for 14 days.
Analytical methods	Incremental costs and outcomes presented
Results	
Study parameters	Mean and 95% CIs were generated for use cost variables. Mean values (with 95% CI) for outcome variables (both symptom score and QALYs), were estimated using regression equations controlling for baseline characteristics (fever and baseline symptoms).
Incremental costs and outcomes	<u>Cost-effectiveness analysis</u> • Mean cost in each group was £44 (clinical score; n=167); £49 (RADT; n=163) and £51 (delayed prescribing; n=168). • Mean point estimate in severity score was 2.83 (clinical score); 2.84 (RADT) and 3.15 (delayed prescribing). • Overall results showed that the clinical score group dominated the other two groups, being more clinically effective (having a lower symptom score) and less costly. <u>Cost-utility analysis (complete case analysis, n=257)</u> • Mean cost in each group was £46 (clinical score); £49 (RADT) and £50 (delayed prescribing). • Mean QALY at 28 days was 0.0174 (clinical score); 0.0175 (RADT) and 0.0171 (delayed prescribing). • As the QALY gain is marginally higher in the test group than in the clinical score group, the RADT generates additional QALYs at £24,528 per QALY. Delayed prescribing was dominated.
Characterising uncertainty	Bootstrapping using 5000 samples was used to generate cost-effectiveness acceptability curves. Bootstrapping was also used to generate scatterplots on the cost-effectiveness plane. At a value of £30,000 per QALY, the probabilities that the three groups were cost-effective were 28%, 38% and 35%, for the delayed prescribing, clinical score and RADT groups, respectively, for the 28-day QALY gain.

Discussion	
Study findings	The clinical scoring tool (FeverPAIN) was effective in helping to reduce symptoms and the costs in all three groups were similar. The cost–utility analysis was less clear, as QALY differences were very small generating wide CIs. The CEACs for the cost–utility study indicate that clinical score is most likely to be cost-effective over all values, however, they also indicate considerable uncertainty.
Limitations	• 14-day diary had EQ-5D data for only two time points (0 and 14 days) • The smaller QALY data set may not be representative of the larger group of individuals in the cost-effectiveness study • Timeframe short, hence longer-term impacts were not known • No indirect costs were estimated
Generalisability	The generalisability of the analysis may be limited to the unit costs used in the analysis.
Other	
Source of funding	The study was funded by the NIHR Health Technology Assessment programme
Conflicts of interest	None declared
Comments	• None
Authors conclusion	
Using a clinical score appears to be an efficient use of health-care resources compared with either delayed antibiotic prescribing or the use of a RADT combined with a clinical score.	
Reviewer's conclusion	
The authors used appropriate economic methods for the study. **Name of first reviewer:** Hema Mistry; **Name of second reviewer:** Felix Achana	

4.1.3.2 **Quality assessment**

The quality of the reporting of the economic analysis provided by the Little et al (2014)[73] study was assessed using the 25-point CHEERS checklist [71] and is provided in Table 22. The article was comprehensively reported with 22 of the 25 statements (88.0%) were a yes, one statement (4.0%) was not completed and two statements (8.0%) did not apply.

Table 22. CHEERS quality assessment checklist for economic evaluation studies

Assessment	**Little et al (2014)**
Title	Y
Abstract	Y
Introduction	
Background and objectives	Y
Methods	
Target population and subgroups	Y
Setting and location	Y
Study perspective	Y
Comparators	Y
Time horizon	Y
Discount rate	N/A

Assessment	Little et al (2014)
Choice of health outcomes	Y
Measurement of effectiveness	Y
Measurement and valuation of preference-based outcomes	Y
Estimating resources and costs	Y
Currency, price date, and conversion	Y
Choice of model	N/A
Assumptions	Y
Analytical methods	Y
Results	
Study parameters	Y
Incremental costs and outcomes	Y
Characterising uncertainty	Y
Discussion	
Study findings	Y
Limitations	Y
Generalizability	Y
Other	
Source of funding	Y
Conflicts of interest	N
N- No; N/A- Not Applicable; Y- Yes	

4.1.4 Summary

The cost-effectiveness search highlighted three studies that used the rapid antigen detection tests as identified in the NICE scope and were classed as economic evaluations. Of these, three studies only one allowed a full data extraction and was classed as a high quality economic evaluation when checked against the CHEERS reporting tool. In the next chapter, we build a *de novo* economic model comparing the different tests identified in the NICE scope for the various settings for patients with Group A Strep.

4.2 Cost-effectiveness methods and results

4.2.1 Modelled population

The population of interest is people aged 5 and over presenting to healthcare providers in a primary (GP surgeries, community pharmacies and walk-in centres) or secondary care (urgent care/walk-in centres and emergency departments) setting with symptoms of an acute sore throat. These patients are identified as being more likely (FeverPAIN score of 2 or 3) or most likely (FeverPAIN score of 4 or 5, or a Centor score of 3 or 4) to benefit from an antibiotic by a clinical scoring tool. Potential subgroups identified in the NICE scope include children (aged 5 to 14), adults (aged 15 to 75), and the elderly (over the age of 75 years). However,

the analyses have been restricted to adults and children due to lack of evidence on test accuracy among the elderly patient population.

4.2.2 Model structure

A decision tree model from the perspective of the UK NHS and Personal Social Service (PSS) was developed to estimate the costs and quality-adjusted life-years (QALYs) associated with point-of-care testing in conjunction with clinical scoring tools such as the Centor and FeverPAIN score for GAS compared with clinical assessment incorporating clinical scoring tools alone (usual care).[74]

The model structure as depicted in Figure 12, Figure 13, and Figure 14 makes use of a decision tree to model potential care pathways associated with a suspected Strep A infection/sore throat presentation under the intervention (point-of-care testing and clinical scoring tools) and usual care (clinical scoring tools alone) conditions.

Previous economic evaluations of management strategies for streptococcal pharyngitis have estimated up to 76.5 in quality-adjusted life-days which could be lost as a result of rare but serious complications of the infection such acute rheumatic fever.[75-77] Thus, for this economic model we have assumed a one-year time horizon where we model only one-episode of GAS per patient and we have assumed that this time horizon is sufficient to capture the impact of rare but serious complications of the infection on economic costs and outcomes. This differs with the stated time horizon of 14-days originally conceived in the EAG protocol for this self-limiting illness for which majority cases would be expected to resolve satisfactorily.

The model takes account of the prevalence of disease in the modelled population, the test accuracy of clinical scoring algorithms and point-of-care tests, the proportion of patients treated with immediate and delayed antibiotics given a positive or negative clinical score and/or test result (prescribing behaviour of treating clinicians) and the probability of developing important but rare complications of the infection (i.e. suppurative complications such as peritonsillar abscess, quinsy[78]) and non-suppurative complications such as acute

rheumatic fever. [78] Penicillin-induced rash and anaphylactic complications of penicillin are incorporated as adverse effects of treatment.[78, 79]

The model estimated costs in 2017/2018 prices. Economic costs accrued over the modelled time horizon from resource use associated with simulated care pathways. They include the cost of the point-of-care tests (including additional cost of confirmatory throat culture for a negative test result), GP consultations, antimicrobial therapy and treatment for GAS-related complications and the unwanted effects of penicillin. QALYs are calculated as a weighted sum of the difference between utility decrements associated with GAS infection and related complications and the general UK population utility norms, weighted by the modelled time horizon in years. No discounting was applied to costs and benefits due to the one-year time horizon.

The base-case analysis assumes that patients presenting with suspected GAS in the usual care arm receive immediate or delayed antimicrobial treatment based on clinical assessment and outcome of clinical scoring algorithm indicating possible GAS infection. We assumed a score ≥3 on the Centor (equivalent to FeverPAIN score ≥4) as the threshold for commencing immediate antibiotics (or testing for those in the intervention arm) as shown in Figure 1 and in line with recent NICE guidance on antimicrobial prescribing for acute sore throat infections.[8]

We explored the impact of alternative thresholds (Centor score ≥2 and ≥1) for commencing antibiotic treatment and on testing. These alternative thresholds have differing performance (sensitivity and specificity) to the Centor score ≥3, hence could be considered as assessing an alternative performance of the Centor tool. For the intervention arm, we assumed that patients presenting with suspected GAS will be screened first using clinical scoring tool for signs and symptoms of the infection. Those screening positive (i.e. Centor score ≥3 and FeverPAIN ≥4) are offered a point-of-care test followed by immediate antibiotics if testing indicates positive GAS infection. Those screening negative according to clinical scoring algorithm or test are offered delayed antibiotic prescription with probability of 0.49 and 0.29 in the usual care and test arms respectively based on the PRISM trial data.[6]

Over the one-year time horizon, patients with suspected GAS infection receiving either immediate or delayed antibiotics can make a complete recovery or go on to develop complications requiring a period of hospital stay. The risk of developing serious complications related to GAS are modelled as a function of antimicrobial treatment so that those correctly diagnosed and appropriately treated present a lower risk of serious GAS complications compared with those incorrectly diagnosed who receive no antimicrobial treatment. Separate models (each with same underlying structure depicted in Figure 12 to Figure 14) are specified for adults and children in primary and secondary care settings respectively.

Details of methodology used to derive parameter inputs and the data sources used to inform estimates are discussed in the sections below.

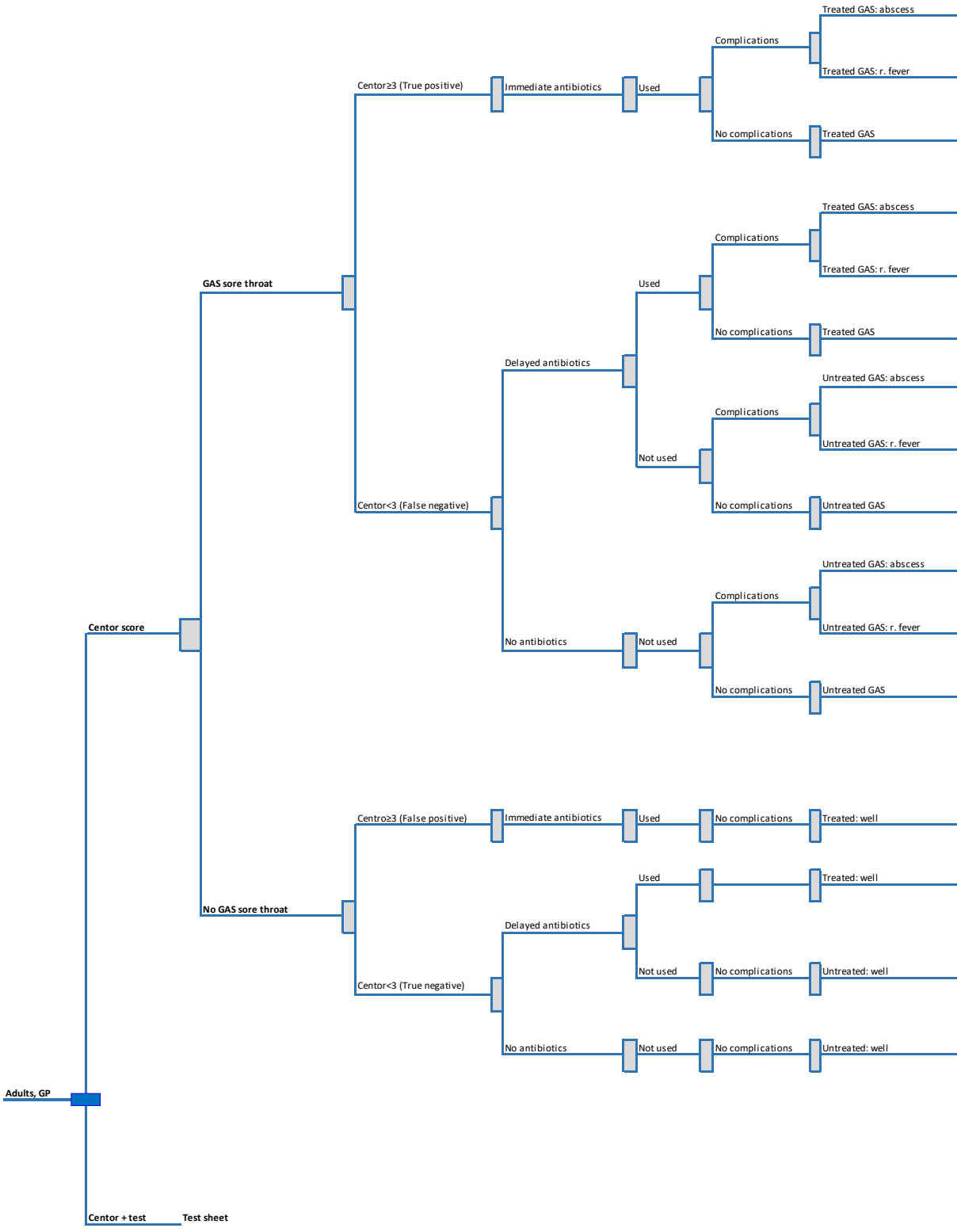


Figure 12: Strep A DAR Model part 1

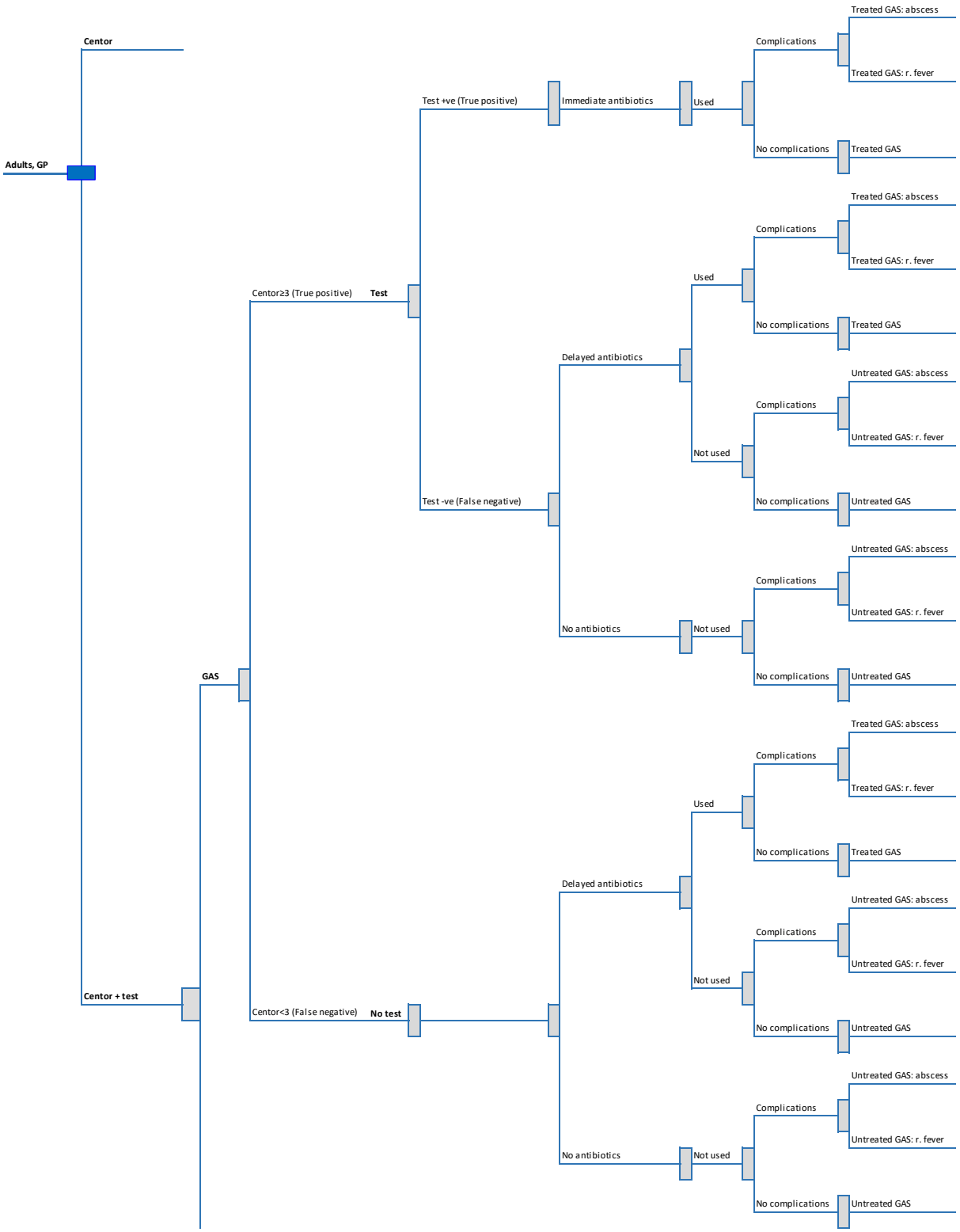


Figure 13: Strep A DAR Model part 2

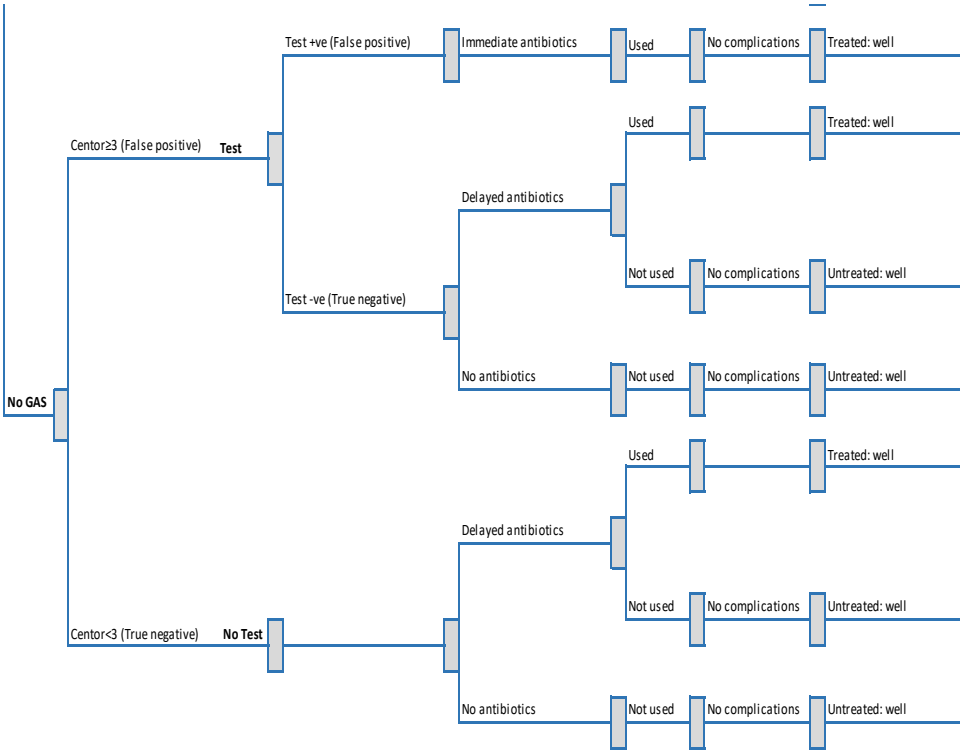


Figure 14: Strep A DAR Model part 3

4.2.3 Effectiveness evidence used in the economic model

4.2.3.1 Accuracy of clinical scoring algorithms (all models)

Accuracy in the usual care arm was based on estimates of sensitivity and specificity of the Centor score taken from published meta-analysis of 12 studies by Aalbers et al (2011).[80] Table 23 summarises the reported estimates of sensitivity and specificity of the Centor score at cut-offs of ≥1, ≥2, ≥3 and 4 for positive GAS infection. The base-case model used the estimates at cut-off ≥3 for a positive result and <3 for a negative result. At this threshold, the Centor score has sensitivity of 49% (95% CI 38% to 60%) and specificity of 82% (95% CI 72% to 88%). Alternative thresholds on the Centor score were explored in sensitivity analyses.[80] However, we were unable to evaluate the FeverPAIN clinical score due to lack of accuracy estimates in a format suitable for the economic model (i.e. sensitivity and specificity of the FeverPAIN at cut-off of ≥4).

Table 23: Diagnostic accuracy of the Centor score based on meta-analysis of 12 studies reported by Aalbers et al (2011)[80]

Centor threshold for positive GAS infection	Sensitivity (95% CI)	Specificity (95% CI)	N primary studies included in meta-analysis	Distributional form in model
≥ 1	0.95 (0.91 to 0.97)	0.18 (0.12 to 0.26)	11	Normal (logit scale)
≥ 2	0.79 (0.71 to 0.86)	0.55 (0.45 to 0.65)	12	Normal (logit scale)
≥ 3	0.49 (0.38 to 0.60)	0.82 (0.72 to 0.88)	11	Normal (logit scale)
4	0.18 (0.12 to 0.27)	0.95 (0.92 to 0.97)	11	Normal (logit scale)

4.2.3.2 Accuracy of point-of-care tests

Estimates of test accuracy for the point-of-care tests obtained from our systematic review with and without meta-analyses are summarised in Table 24 by test, clinical setting (primary care) and patient population (adults and children). When no studies reporting accuracy data was identified in our systematic review, we obtained the estimates from either the manufacturer website or from the manufacturer submissions (submitted directly to NICE in response to a request for information). Test accuracy data were available for 6 (28.6%) of the 21 tests from published sources identified in the clinical effectiveness review, a further 4 tests (19%) had accuracy data from both published sources and manufacturer's submission, 6 tests (27.6%) had only manufacturer's data and 2 tests (9.5%) had FDA data. Test accuracy data were not available for the 3 (14.3%) remaining tests (Biopanda's Strep A rapid test strip, Bionexia Strep A cassette and Bionexia Strep A plus cassette). Two of the three tests (Bionexia Strep A cassette and Bionexia Strep A plus cassette) were excluded from the economic modelling of individual tests due to lack of test accuracy data. Biopanda's Strep A rapid test strip accuracy was assumed to be equal to cassette version of this test for which accuracy estimates was available. In general, estimates of sensitivity and specificity obtained from the published sources tended to be variable and lower than those provided by the manufacturer. For example, whilst sensitivity of point-of-care testing in adults based on the published sources ranged from 68% for Abbott's Clearview Exact Strep A cassette to 100% for QuikRead Go Strep A test kit; estimates provided in the manufacturer's submission ranged from 95% for Biopanda Reagents Strep A rapid test cassette to 98% for nal von minden's NADAL Strep A test's. Similar trend in specificity is observed with the manufacturers' estimates being generally much higher than estimates based on published

data. Thus, the source of test accuracy data is likely to be an important driver of cost-effectiveness. The economic models presented here which are based solely on manufacturers test accuracy data, with no peer-reviewed published data are likely to overestimate test accuracy, and therefore the results of these models cannot be reliably interpreted.

Table 24: Test accuracy of point-of-care tests used in economic model in primary care

Test ID	Test Name	Manufacturer	Sensitivity (95%CI)	Specificity (95%CI)	Distribution	Data source
Adults						
1	Clearview Exact Strep A cassette	Abbott	0.68 (0.54, 0.8)	0.95 (0.92, 0.97)	Normal (logit)	1 abstract (Andersen 2003)
2	Clearview Exact Strep A dipstick - test strip	Abbott	0.68 (0.54, 0.8)	0.95 (0.92, 0.97)	Normal (logit)	1 abstract (Andersen 2003)
3	BD Veritor Plus system group A Strep Assay - cassette	Beckton Dickinson	0.78 (0.67, 0.87)	0.9 (0.86, 0.93)	Normal (logit)	2 studies (Berry 2018; Azrad 2019)
4	Strep A rapid test - cassette	Biopanda Reagents	0.95 (0.9, 0.98)	0.98 (0.96, 0.99)	Normal (logit)	1 manufacturer response to NICE
5	Strep A rapid test - test strip	Biopanda Reagents				No data
6	NADAL Strep A - test strip	nal von minden GmbH	0.98 (0.92, 1)	0.98 (0.94, 0.99)	Normal (logit)	1 manufacturer response to NICE
7	NADAL Strep A - cassette	nal von minden GmbH	0.98 (0.92, 1)	0.98 (0.94, 0.99)	Normal (logit)	1 manufacturer response to NICE
8	NADAL Strep A plus - cassette	nal von minden GmbH	0.98 (0.92, 1)	0.98 (0.94, 0.99)	Normal (logit)	1 manufacturer response to NICE
9	NADAL Strep A plus - test strip	nal von minden GmbH	0.98 (0.92, 1)	0.98 (0.94, 0.99)	Normal (logit)	1 manufacturer response to NICE
10	NADAL Strep A scan test - cassette	nal von minden GmbH	0.98 (0.92, 1)	0.98 (0.94, 0.99)	Normal (logit)	1 manufacturer response to NICE
11	OSOM Strep A test - test strip	Sekisui Diagnostics	0.92 (0.76, 0.98)	0.96 (0.89, 0.99)	Normal (logit)	3 studies (Llor 2011; Llor 2009; Bura 2017)
12	QuikRead Go Strep A test kit	Orion Diagnostica	1 (0.85, 1)	0.79 (0.6, 0.92)	Normal (logit)	1 study (Stefaniuk 2017)
13	Alere TestPack Plus Strep A - cassette	Abbott	0.95 (0.89, 0.98)	0.94 (0.88, 0.98)	Normal (logit)	1 study (Humair 2006)
14	Bionexia Strep A plus - cassette	Biomerieux				No data
15	Bionexia Strep A dipstick - test strip	Biomerieux	0.85 (0.74, 0.92)	0.91 (0.84, 0.95)	Normal (logit)	1 abstract (Pauchard 2003)
16	Biosynex Strep A - cassette	Biosynex				No data

Test ID	Test Name	Manufacturer	Sensitivity (95%CI)	Specificity (95%CI)	Distribution	Data source
17	Sofia Strep A FIA	Quidel	0.85 (0.81, 0.89)	0.95 (0.93, 0.97)	Normal (logit)	1 study (Lacroix 2018)
18	Alere i Strep A	Abbott	0.95 (0.74, 1)	0.97 (0.92, 0.99)	Normal (logit)	1 study (Cohen 2015)
19	Alere i Strep A 2	Abbott	0.98 (0.96, 1)	0.93 (0.91, 0.95)	Normal (logit)	1 FDA Report
20	Cobas Strep A Assay on Liat system	Roche Diagnostics	0.98 (0.93, 1)	0.93 (0.9, 0.96)	Normal (logit)	1 study (Wang 2017)
21	Xpert Xpress Strep A	Cepheid	1 (0.99, 1)	0.94 (0.92, 0.96)	Normal (logit)	1 manufacturer response to NICE and 1 FDA report
Children						
1	Clearview Exact Strep A cassette	Abbott	0.68 (0.54, 0.8)	0.95 (0.92, 0.97)	Normal (logit)	1 study (Andersen 2003)
2	Clearview Exact Strep A dipstick - test strip	Abbott	0.68 (0.54, 0.8)	0.95 (0.92, 0.97)	Normal (logit)	1 study (Andersen 2003)
3	BD Veritor Plus system group A Strep Assay - cassette	Beckton Dickinson	0.76 (0.61, 0.88)	0.94 (0.89, 0.97)	Normal (logit)	1 study (Berry 2018)
4	Strep A rapid test - cassette	Biopanda Reagents	0.95 (0.9, 0.98)	0.98 (0.96, 0.99)	Normal (logit)	1 manufacturer response to NICE
5	Strep A rapid test - test strip	Biopanda Reagents				No data
6	NADAL Strep A - test strip	nal von minden GmbH	0.98 (0.92, 1)	0.98 (0.94, 0.99)	Normal (logit)	1 manufacturer response to NICE
7	NADAL Strep A - cassette	nal von minden GmbH	0.98 (0.92, 1)	0.98 (0.94, 0.99)	Normal (logit)	1 manufacturer response to NICE
8	NADAL Strep A plus - cassette	nal von minden GmbH	0.98 (0.92, 1)	0.98 (0.94, 0.99)	Normal (logit)	1 manufacturer response to NICE
9	NADAL Strep A plus - test strip	nal von minden GmbH	0.98 (0.92, 1)	0.98 (0.94, 0.99)	Normal (logit)	1 manufacturer response to NICE
10	NADAL Strep A scan test - cassette	nal von minden GmbH	0.98 (0.92, 1)	0.98 (0.94, 0.99)	Normal (logit)	1 manufacturer response to NICE
11	OSOM Strep A test – test strip	Sekisui Diagnostics	0.94 (0.89, 0.98)	0.95 (0.91, 0.98)	Normal (logit)	1 study (Llor 2011)
12	QuikRead Go Strep A test kit	Orion Diagnostica	0.80 (0.56, 0.94)	0.91 (0.72, 0.99)	Normal (logit)	1 study (Stefaniuk 2017)
13	Alere TestPack Plus Strep A - cassette	Abbott	0.86 (0.79, 0.91)	0.99 (0.97, 1)	Normal (logit)	1 study (McIsaac 2004)

Test ID	Test Name	Manufacturer	Sensitivity (95%CI)	Specificity (95%CI)	Distribution	Data source
14	Bionexia Strep A plus - cassette	Biomerieux				No data
15	Bionexia Strep A dipstick – test strip	Biomerieux	0.85 (0.74, 0.92)	0.91 (0.84, 0.95)	Normal (logit)	1 abstract (Pauchard 2013)
16	Biosynex Strep A - cassette	Biosynex				No data
17	Sofia Strep A FIA	Quidel	0.85 (0.81, 0.89)	0.95 (0.93, 0.97)	Normal (logit)	1 study (Lacroix 2018)
18	Alere i Strep A	Abbott	0.98 (0.95, 1)	0.96 (0.89, 1)	Normal (logit)	3 studies (Berry 2018; Cohen 2015; Weinzierl 2018)
19	Alere i Strep A 2	Abbott	0.98 (0.96, 1)	0.93 (0.91, 0.95)	Normal (logit)	1 FDA Report
20	Cobas Strep A Assay on Liat system	Roche Diagnostics	0.98 (0.93, 1)	0.93 (0.9, 0.96)	Normal (logit)	1 study (Wang 2017)
21	Xpert Xpress Strep A	Cepheid	1 (0.99, 1)	0.94 (0.92, 0.96)	Normal (logit)	1 manufacturer response to NICE and 1 FDA report

4.2.4 Prevalence of GAS infection in the modelled population

Data on the adult prevalence of GAS in the UK was available in one study out of the 38 published studies, abstracts, and reports submitted by manufacturers to NICE included in our test accuracy effectiveness review. The study by Little et al. from the review[6] (with additional data from the full HTA report[73]) found a prevalence rate of 34% (95% CI 31% to 38%) for pathogenic streptococci infection among 204/597 patients aged ≥5 years presenting in UK primary care settings. Of these, 136 (66.7%) were GAS. This gives a GAS prevalence rate of 22.7% (136/597). This study did not consecutively recruit patients, meaning there may be bias in the sample which could have affected the true prevalence rate. As there were no UK adult studies in secondary care, this estimate was used across both primary and secondary settings (see Table 25).

There were no clear UK estimates for prevalence in children from the systematic review, a median value from 3 non-UK children in primary care only studies was calculated.[23, 43, 49] The median value was 30.2%.

Table 25: Prevalence of GAS by settings, country and population

Systematic review data			**Estimate used in model**			
Patient population and clinical settings	**No. of studies**	**Median prevalence (range) %**	**Central estimate**	**SE**	**Distribution**	**Source**
Adults						
Primary and secondary care	1[6]	22.6	0.226	0.051	Beta	Systematic review
Children						
Primary and secondary care	3[23, 43, 49]	30.2 (26.3,34.1)	30.2	0.015	Beta	Systematic review

4.2.5 Treatment related probabilities and complication rates

Treatment related probabilities and complication rates following GAS used in the economic model are presented in Table 26. The proportion of patients attending repeat consultations for sore throat infections (used to inform calculation of treatment costs) was obtained from Little et al (2013b).[81] In this large cohort study of UK patients presenting in primary care with sore throat, a total of 889 (14.2%) repeat consultations for new or resolved symptoms were reported among 13,288 adults and adolescents.

In base-case models, the probability of commencing antibiotic treatment given a positive clinical score (defined as Centor score ≥3) in the usual care arm or a positive clinical score and test result in the intervention arm was set to 1 based on the prescribing behaviour of general practitioners reported in the PRISM trial.[6] The probability of a delayed prescription given negative clinical score (defined as Centor score <3 in the base-case) was set to 0.51 based on data suggesting that 91/178 patients in the clinical score arm of PRISM trial with FeverPAIN score <4 were offered a delayed prescription[6] with the assumption that a Centor score <3 is equivalent to a FeverPAIN score <4. The probability of a delayed prescription given negative test was set to 0.273 based on the PRISM data (48/174 patients in the clinical score + test arm were given delayed prescription). The probability of antibiotic use among those receiving a delayed prescription was set to 0.46 based on PRISM data showing reported antibiotic used among the 75/164 patients in the control arm who were offered delayed prescription.

Complications for treated (i.e. antibiotics given) and untreated (no antibiotics) GAS infections were also estimated based on another Little et al. study:[81] 78 and 75 complications (quinsy, sinusitis, otitis media and cellulitis) were reported among 5,932 treated and 4,974 untreated individuals generating a complication rate of 1.3% and 1.5% respectively. As this study did not report rates for rare but important non-suppurative sequalae of GAS sore throat such as acute rheumatic fever,[82] we assumed that majority of complications were suppurative in nature with only a tiny proportion of patients (no more than 0.01%) going on to develop non-suppurative sequalae. The impact of this assumption on the cost-effectiveness estimates was assessed by halving and doubling it in sensitivity analyses. We assumed that 2% of patients prescribed antibiotics (100% of those prescribed immediate antibiotics and 46% of those prescribed delayed antibiotics) will go on to develop penicillin-induced rash and 0.1% will develop penicillin-induced anaphylaxis/sepsis based on estimates reported in previous economic evaluation of diagnostic and treatment strategies for adults with streptococcal pharyngitis.[75] Sensitivity analysis explored the impact of halving and doubling complications associated with penicillin use on the base-case cost-effectiveness.

Table 26: Probabilities used in the economic model

Description of parameter	Mean	SE[1]	Distribution	Source
GP practice				
Proportion attending repeat GP consultation following GAS infection	0.142	0.007	Beta	Little 2013b[81]
Antibiotic prescribing probabilities				
Probability antibiotics given Centor score ≥3 or positive test (immediate prescription)	1			Little *et al.* 2013a[6]
Probability antibiotics given Centor score <3 (delayed prescription, usual care arm)	0.51	0.026	Beta	Little *et al.* 2013a[6]
Probability antibiotics given negative test (delayed prescription, intervention arm)	0.267	0.014	Beta	Little *et al.* 2013a[6]
Probability antibiotics use given delayed prescription	0.46	0.023	Beta	Little *et al.* 2013a[6]
Complication rates following GAS infection				
Probability of complication given antibiotics (treated infection)	0.013	0.0005	Beta	Little 2013b[81]
Probability of complications given no antibiotics (untreated infection)	0.015	0.0007	Beta	Little 2013b[81]
Proportion of complications that are non-suppurative (i.e. rheumatic fever)	0.0001			Analyst assumption
Adverse effects of penicillin				
Penicillin-induced rash	0.02		Beta	Neuner et al 2003 [75]
Penicillin-induced anaphylaxis/sepsis	0.0001		Beta	Neuner et al 2003 [75]
[1]Standed error (SE) derived assuming upper and lower bound equals to 10% of mean/central estimate				

4.2.6 Health utility and estimation of QALY gains

Table 27 presents estimates of health utilities used to inform the economic model. A mean baseline utility of 0.863 equal to the mean utility norm for the general UK adult population[83] was assumed for the modelled adult population treated in primary and secondary care. For the children population models, we assume a mean utility of 0.94 equivalent to mean UK utility norm for the under 25 year population,[83] the closest age group to children. Utility decrements associated with GAS and related complications such as development of peritoncillar abscess, rheumatic fever and anaphylactic complications of penicillin were obtained from previously published economic evaluations of diagnostic and management strategies for adults with pharyngitis.[75, 77] The two studies reported losses of 0.15 and 0.25 in quality-adjusted life days for treated and untreated sore throat infections whilst related complications such as acute rheumatic fever, penicillin-induced anaphylaxis (sepsis), peritoncillar abscess and penicillin-induced rash were associated with the greatest health impact with estimates of 76.5, 9, 5 and 0.65 in quality-adjusted life-days lost respectively. These estimates translate into utility decrements of 0.000411 (0.15/365) and 0.000685 (0.25/365) for treated and untreated

GAS infection, 0.0017 (0.65/365) for penicillin-induced rash, 0.0037 (5/365) for peritonsillar abscess, 0.025 (9/365) for penicillin-induced sepsis and 0.209 (76.5/364) for rheumatic fever respectively. Quality adjusted-life-years (QALYs) were calculated at the end of each pathway in the model by subtracting from the baseline utility of 0.863 (or 0.94 for the children model), the utility decrements associated with all outcomes that occur in the modelled pathway (assuming utility decrements are additive) weighted by modelled time horizon in years (i.e. 365/365 for the one-year base-case time horizon). Disutility associated with unwanted effect of penicillin (rash and anaphylaxis) were added to care pathways associated with treated infection (immediate or delayed antibiotic use) weighted by the respective event probability (0.00003671). For example, the total QALY accrued from uncomplicated GAS infection with complete resolution following immediate antibiotic treatment would be equal to (0.863-0.000411-0.00003671)*1 = 0.8595311 QALYs over the one-year time horizon considered in the base-case analysis for adults. Similarly, if this infection had resulted in a subsequent complication (e.g. an abscess), then the total QALY estimate would be slightly lower at (0.863-0.000411-0.0137-0.00003671)*1= 0.8458311.

Table 27: Utilities

Utility/disutility	Mean	SE	Distribution	Source
Baseline (UK population norm, adults)	0.863	0.044	Beta	Kind et al 1998[83]
Baseline (UK population norm, children)	0.94	0.048	Beta	Kind et al 1998[83]
Utility decrement associated with untreated infection	0.000685	0.00005	Beta	Neuner et al 2003[75]
Utility decrement associated with treated infection	0.000411	0.00003	Beta	Neuner et al 2003[75]
Utility decrement associated with penicillin-induced rash	0.0017	0.0001	Beta	Neuner et al 2003[75]
Utility decrement associated with abscess	0.0137	0.0007	Beta	Neuner et al 2003[75]
Utility decrement associated with penicillin-induced anaphylaxis (sepsis)	0.025	0.0013	Beta	Neuner et al 2003[75]
Utility decrement associated with rheumatic fever	0.209	0.011	Beta	Neuner et al 2003[75]

4.2.7 Health and social care costs

4.2.7.1 Cost of tests

Table 28 presents the unit cost for each point-of-care test and estimates of resource use in terms of the additional GP time required to administer and process test results. Cost data were

available for 13 (62 %) of the 21 tests considered in the NICE scope. The majority of the costs were provided by the manufacturers (submitted directly to NICE in response to a request for information) and ranged from £0.77 per test for the Biopanda's Strep A rapid test strip to £75.03 inclusive of VAT (2017/2018 prices) for Cobas Strep A Assay on Liat system supplied by Roche Diagnostics. Unit costs for Abbott's Clearview Exact Strep A tests were obtained from the NHS supply chain catalogue at £1.92 per test for the Clearview Strep A dipstick - test strip and £2.72 for the cassette version.[84] The duration of additional GP time for processing test results were estimated based on information provided in the manufacturer's submission and ranged from 5-12 minutes. Costs associated with additional GP time for processing test results are included in the base-case analysis. The cost of confirmatory swab culture following negative test result are calculated as part of the costs associated with modelled pathways in the intervention arm details of which are given in the next section.

Table 28: Test costs

Test ID	Test Name	Cost	Test process time	Source
1	Clearview Exact Strep A cassette (Abbott)	£2.72	0	NHS Supply chain catalogue (NPC =HHH2552)[84]
2	Clearview Exact Strep A dipstick – test strip (Abbott)	£1.92	0	Medisave UK Ltd.[85]
3	BD Veritor Plus system group A Strep Assay - cassette (Beckton Dickinson)			Test cost not available
4	Strep A rapid test - cassette (Biopanda Reagents)	£0.98	5	Manufacturer's submission*
5	Strep A rapid test – test strip (Biopanda Reagents)	£0.77	0	Manufacturer's submission*
6	NADAL Strep A - test strip (nal von minden GmbH)	£1.44	5	Manufacturer's submission*
7	NADAL Strep A - cassette (nal von minden GmbH)	£1.68	5	Manufacturer's submission*
8	NADAL Strep A plus - cassette (nal von minden GmbH)	£1.8	5	Manufacturer's submission*
9	NADAL Strep A plus - test strip (nal von minden GmbH)	£1.56	5	Manufacturer's submission*
10	NADAL Strep A scan test - cassette (nal von minden GmbH)	£2.28	5	Manufacturer's submission*
11	OSOM Strep A test – test strip (Sekisui Diagnostics)			Test cost not available
12	QuikRead Go Strep A test kit (Orion Diagnostica)	£5.02	5	Manufacturer's submission*
13	Alere TestPack Plus Strep A - cassette (Abbott)	£3.24	5	Manufacturer's submission*
14	Bionexia Strep A plus - cassette (Biomerieux)			Test cost not available
15	Bionexia Strep A dipstick – test strip (Biomerieux)			Test cost not available
16	Biosynex Strep A - cassette (Biosynex)			Test cost not available
17	Sofia Strep A FIA (Quidel)			Test cost not available
18	ALERE i Strep A (Abbott)			Test cost not available
19	ALERE i Strep A 2 (Abbott)			Test cost not available
20	Cobas Strep A Assay on Liat system (Roche Diagnostics)	£75.03	6	Manufacturer's submission*
21	Xpert Xpress Strep A (Cepheid)	£4.92	12	Manufacturer's submission*

*submitted directly to NICE in response to a request for information

4.2.8 Treatment costs

Unit costs of healthcare service use associated with modelled care pathways are summarised in Table 29. As described in 4.2.2 above, the base-case models incorporate three treatment options for patients presenting with suspected GAS infection in primary care and secondary care settings: immediate antibiotics (option one), delayed antibiotics with reported use (option two) and delayed antibiotics that have not been used or no antibiotics offered (option three). All three options account for repeat GP consultations at 14.2%[6] over the modelled time horizon, with a typical GP consultation lasting 9.22 minutes at an average cost of £4.02 per minute[86] and pain relief (500 mg paracetamol) costing £0.74 per 32-tablet pack,[87] but the options differed in the way antibiotics are prescribed.

Under options one and two, patients incur cost of antibiotics at £0.91 per treatment course (Phenoxymethylpenicillin 250mg, 28-tablets pack)[87] and costs associated with managing adverse effects of penicillin: penicillin-induced rash (assumed seen by GP at additional expense, (£4.02 per minute) and switched to erythromycin 500mg at £10 per treatment course [87] weighted by 0.02, the probability of a rash) and penicillin-induced anaphylaxis estimated at £1,744 based on data reported in a 2017 cost of sepsis study[88] (see below) weighted by 0.0001, the probability of sepsis).[75] No costs associated with antibiotic use are included under option three (delayed antibiotics prescription given but not used), however, we assume that 14.2% of patients attended a repeat consultation[81] and will use the delayed antibiotics prescription under this option.

Confirmatory swab culture costing £8[89] were added to options two and three for patients with a negative test result (intervention only) but not to option one, as patients with a positive test result receive immediate antibiotics. On average the estimated treatment costs based on these assumptions and repeat consultation rate of 14.2% were £44.89 (option one – intervention and usual care arms and option two – usual care arm), £52.89 (option two – intervention arm including confirmatory culture costs), £43.78 (option three – usual care arm), and £51.77 (option three intervention arm including confirmatory culture costs) (see Table 29).

The cost of sepsis was estimated to be £1,744 based on data reported in a study,[88] which estimated that 93,973 adults would need treatment for sepsis in UK hospitals at annual total

cost £163,949,055 (Table 29). The cost of treating GAS related abscess was estimated at £1,571 based on the NHS reference cost for a tonsillectomy in adults 19 years and over with a HealthCare Resource Groups code CA60A.[89] The cost of treating acute rheumatic fever was estimated at £1,772.44 based on the NHS reference cost for Other Acquired Cardiac Conditions with CC Score 6-8 and HealthCare Resource Groups code EB14C.[89]

Table 29: Treatment costs (2017/18 price year)

Treatment Costs	Mean	SE	Distribution	Source
Antibiotic (Phenoxymethylpenicillin 250mg, 28-tablets pack)	£0.91	£0.046	Gamma	BNF 72 (2017)[87]
Pain relief (paracetamol 500mg, 32-tablets pack)	£0.74	£0.037	Gamma	BNF 72 (2017)[87]
GP consultation (9.22 minutes)	£37.4	£1.91	Gamma	PSSRU Unit costs 2017[86]
Throat culture (swab)	£8.00	£0.41	Gamma	2017 reference costs[89]
Penicillin induced rash (switch to Erythromycin 500mg)	£10.00	£0.51	Gamma	BNF 72 (2017)[87]
Treatment costs, sepsis	£1,744.64	£89.01	Gamma	Derived from data reported in Hex et al. 2017[88]
Treatment modality costs (assumptions)				
Treatment option 1 (usual care and intervention arms) and option 2 (usual care arm): assume immediate/delayed antibiotics (£0.91) at initial consultation (£37.43); 14.2% reconsultations during which patients get paracetamol (£5.42) + weighted treatment costs penicillin side-effects (£1.12 per patient).	£44.89			Derived from other treatment costs
Treatment option 2 (intervention arm): assume antibiotics (£0.91) given at initial consultation (£37.43); 14.2% reconsultations during which patients paracetamol (£5.42), weighted treatment costs penicillin side-effects (£1.12 per patient) and confirmatory culture (£8).	£52.89			Derived from other treatment costs
Treatment option 3 (usual care arm): assume paracetamol (£0.74) at initial consultation (£37.43) and delayed antibiotic use among the 14.2% attending repeat consultation (£5.60).	£43.78			Derived from other treatment costs
Treatment option 3 (intervention arm): assume paracetamol (£0.74) at initial consultation (£37.43), delayed antibiotic use among the 14.2% attending repeat consultation (£5.60) and confirmatory throat culture (£8)	£51.77			Derived from other treatment costs
Complication of GAS, costs				
Treatment costs, abscess	£1,571.28	£80	Gamma	2017 Reference costs (Tonsillectomy, 19 years and

Treatment Costs	Mean	SE	Distribution	Source
				over (HRG CA60A)[89]
Treatment costs, acute rheumatic fever	£1,772.44	£90.43	Gamma	2017 Reference costs (Other Acquired Cardiac Conditions with CC Score 6-8 with HRG code EB14C)[89]

4.2.9 Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) were conducted to explore the impact of parameter uncertainty on base-case cost-effectiveness of point-of-care testing for GAS infection. The PSA was implemented via Monte Carlo simulations involving 1,000 draws for all model inputs except for the acquisition cost of the tests which were entered as deterministic values. This enabled us to simulate 1,000 replicates of the base-case ICER (displayed on cost-effectiveness planes) and calculate the probability of cost-effectiveness at threshold values ranging from £0 to £100,000 per QALY gained (CEACs). The sensitivity and specificity of the clinical scoring algorithm and individual test point-of-care tests were assumed to be drawn from separate normal distributions on the logit-scale, as the relatively small number of studies reporting test-specific accuracy data precluded joint synthesis sensitivity and specificity and estimation of the between-study correlation (Table 23 and Table 24). Prevalence, probabilities and utility values (Table 25 to Table 27) were assumed to be drawn from a Beta distribution reflecting scale of measurement for quantities constrained to lie in the interval 0-1. Costs were assigned a Gamma distribution. Generally, the uncertainty surrounding input parameters (standard errors and confidence intervals) were generally not available, therefore we assumed a 10% of the mean as equivalent to lower and upper 95% confidence limits and calculated the standard errors assuming approximate normal distribution.

4.2.10 Base-case analyses

The main base-case model was based on the adult population in a primary care setting. This model was then adapted for adults in a secondary care setting, for children in a primary care setting, and for children in a secondary care setting.

4.2.11 Adult primary care model: base-case analysis results

The base-case cost-effectiveness results for adults treated in primary care are presented in Table 30 for 13 of the 21 tests for which test accuracy and cost data were available. The rate at which incremental QALYs accrued over the one-year modelled time horizon was small, thus estimates of simulated costs and QALYs were multiplied by 1,000 to aid clarity in presentation of incremental estimates in the result tables and texts. The mean simulated costs under base-case assumptions was £49,147 per 1,000 individuals treated in primary care under usual care practice and ranged from £50,353 per 1000 individuals in the test group using the Biopanda Reagents's Strep A rapid test strip to £74,932 per 1,000 individuals using the Cobas Strep A Assay on Liat system, Roche Diagnostics. The corresponding estimated mean QALYs were 859.825 per 1,000 individuals under usual care practice and ranged between 859.821 QALYs per 1,000 individuals in the intervention group using Abbott's Clearview Exact Strep A cassette or test strip to 859.829 QALYs per 1,000 individuals using Cepheid's Xpert Xpress Strep A tests. In terms of incremental cost-effectiveness, the base-case estimates suggest usual care was cheaper and generated marginally more QALYs than (and therefore dominated) both cassette and strip versions of Abbott's Clearview Exact Strep A test. Incremental cost-effectiveness ratios (ICERs) for the remaining eleven tests suggest testing was more costly and more effective than usual care with ICERs range from £388,314 per QALY gained for Biopanda's Strep A rapid test strip to £7,059,731 per QALY gained for Roche Diagnostics's Cobas Strep A Assay on Liat system compared with usual care.

Table 30: Adult primary care model: Base-case cost-effectiveness results

TestID	Test Name	Mean costs / 1000 individuals	Mean QALYs / 1000 individuals	Inc. Costs / 1000 individuals	Inc. QALYS / 1000 individuals	ICER versus usual care
	Usual care (Clinical scoring based on Centor ≥3 plus clinical assessment)	£49,147	859.82458955	£0	0.0000000	-
1	Clearview Exact Strep A cassette (Abbott)	£51,103	859.82063008	£1,957	-0.0039595	Dominated
2	Clearview Exact Strep A dipstick – test strip (Abbott)	£50,903	859.82063008	£1,757	-0.0039595	Dominated
3	BD Veritor Plus system group A Strep Assay - cassette (Beckton Dickinson)					
4	Strep A rapid test - cassette (Biopanda Reagents)	£55,482	859.82769587	£6,335	0.0031063	£2,039,376
5	Strep A rapid test – test strip (Biopanda Reagents)*	£50,353	859.82769587	£1,206	0.0031063	£388,314
6	NADAL Strep A - test strip (nal von minden GmbH)	£55,564	859.82846603	£6,418	0.0038765	£1,655,497
7	NADAL Strep A - cassette (nal von minden GmbH)	£55,624	859.82846603	£6,478	0.0038765	£1,670,977
8	NADAL Strep A plus - cassette (nal von minden GmbH)	£55,654	859.82846603	£6,508	0.0038765	£1,678,719
9	NADAL Strep A plus - test strip (nal von minden GmbH)	£55,594	859.82846603	£6,448	0.0038765	£1,663,238
10	NADAL Strep A scan test - cassette (nal von minden GmbH)	£55,774	859.82846603	£6,628	0.0038765	£1,709,682
11	OSOM Strep A test – test strip (Sekisui Diagnostics)					
12	QuikRead Go Strep A test kit (Orion Diagnostica)	£56,253	859.82810269	£7,106	0.0035131	£2,022,721
13	Alere TestPack Plus Strep A - cassette (Abbott)	£50,931	859.82751669	£1,785	0.0029271	£609,714
14	Bionexia Strep A plus - cassette (Biomerieux)					
15	Bionexia Strep A dipstick – test strip (Biomerieux)					
16	Biosynex Strep A - cassette (Biosynex)					
17	Sofia Strep A FIA (Quidel)					
18	ALERE i Strep A (Abbott)					
19	ALERE i Strep A 2 (Abbott)					
20	Cobas Strep A Assay on Liat system (Roche Diagnostics)	£74,932	859.82824206	£25,786	0.0036525	£7,059,731
21	Xpert Xpress Strep A (Cepheid)	£63,491	859.82854357	£14,344	0.0039540	£3,627,808

[1]Test accuracy data for Biopanda Reagents Strep A rapid test – test strip not available, assumed the strip has same test accuracy as the cassette version of the test

Note: Missing cost-effectiveness estimates are for the tests where either test accuracy or cost data were not available

4.2.12 Adult primary care model: probabilistic sensitivity analyses

Table 31 below presents probabilistic estimates for adults presenting in primary care. The probabilistic estimates were very similar to the deterministic base-case results with ICERs indicating that usual care dominated two (the Clearview Exact Strep A cassette and the Clearview Exact Strep A dipstick – test strip supplied by Abbott) of the thirteen tests considered in the economic modelling. Base-case probabilistic ICERs for the remaining eleven tests ranged from £402,019 per QALY gained for Strep A rapid test – test strip supplied by Biopanda Reagents to £7,687,979 per QALY gained for Cobas Strep A Assay on Liat system supplied by Roche Diagnostics. The probability for testing to be cost-effective was zero under the base-case assumptions and model inputs regardless of the point-of-care test used in comparison to usual care.

Table 31: Adult primary care model: Probabilistic sensitivity analysis results

TestID	Test Name	Mean costs / 1000 individuals	Mean QALYs / 1000 individuals	Inc. Costs / 1000 individuals	Inc. QALYS / 1000 individuals	ICER versus usual care	Probability of cost-effectiveness at £20,000 per QALY
	Usual care (Clinical scoring based on Centor ≥3 plus clinical assessment)	£49,248	863.385033	£0	0.0000000		1
1	Clearview Exact Strep A cassette (Abbott)	£51,231	863.3810729	£1,983	-0.0039601	Dominated	0
2	Clearview Exact Strep A dipstick – test strip (Abbott)	£51,025	863.3811219	£1,777	-0.0039111	Dominated	0
3	BD Veritor Plus system group A Strep Assay - cassette (Beckton Dickinson)						
4	Strep A rapid test - cassette (Biopanda Reagents)	£55,639	863.3881185	£6,391	0.0030855	£2,071,429	0
5	Strep A rapid test – test strip (Biopanda Reagents)*	£50,467	863.388065	£1,219	0.0030320	£402,019	0
6	NADAL Strep A - test strip (nal von minden GmbH)	£55,732	863.3885375	£6,484	0.0035045	£1,850,335	0
7	NADAL Strep A - cassette (nal von minden GmbH)	£55,795	863.3885208	£6,547	0.0034878	£1,877,146	0
8	NADAL Strep A plus - cassette (nal von minden GmbH)	£55,825	863.3885467	£6,576	0.0035137	£1,871,672	0
9	NADAL Strep A plus - test strip (nal von minden GmbH)	£55,765	863.3885251	£6,517	0.0034921	£1,866,277	0
10	NADAL Strep A scan test - cassette (nal von minden GmbH)	£55,943	863.3885931	£6,695	0.0035601	£1,880,491	0
11	OSOM Strep A test – test strip (Sekisui Diagnostics)						
12	QuikRead Go Strep A test kit (Orion Diagnostica)	£56,414	863.388524	£7,166	0.0034910	£2,052,580	0
13	Alere TestPack Plus Strep A - cassette (Abbott)	£51,055	863.3879231	£1,807	0.0028900	£625,287	0
14	Bionexia Strep A plus - cassette (Biomerieux)						
15	Bionexia Strep A dipstick – test strip (Biomerieux)						
16	Biosynex Strep A - cassette (Biosynex)						
17	Sofia Strep A FIA (Quidel)						
18	ALERE i Strep A (Abbott)						
19	ALERE i Strep A 2 (Abbott)						
20	Cobas Strep A Assay on Liat system (Roche Diagnostics)	£75,269	863.3884176	£26,020	0.0033846	£7,687,979	0
21	Xpert Xpress Strep A (Cepheid)	£63,728	863.3889669	£14,480	0.0039339	£3,680,754	0

[1]Test accuracy data for Biopanda Reagents Strep A rapid test – test strip not available, assumed the strip has same test accuracy as the cassette version of the test

4.2.13 Adult primary care model: exploratory sensitivity analyses

Exploratory analyses were conducted to test the robustness of the economic base-case estimates for adults presenting in primary care with suspected GAS. The base-case ICERs are highly sensitive to various modelling assumptions and input values. In the sections that follow, sensitivity analysis results are presented only for those tests where the ICER is sensitive to the alternative modelling assumptions and parameter inputs (as indicated by changes in the direction of incremental costs or incremental QALYs compared with usual care).

4.2.13.1 Adult primary care model - Prevalence of GAS and clinical score threshold for starting antibiotics (usual care arm) and testing (intervention arm)

In the base-case, a cut-off of three points on the Centor scale was used as threshold for starting antibiotic treatment with scores ≥3 indicating positive GAS infection. Changing this threshold to a score ≥2 had minimal impact on the base-case cost-effectiveness estimates. However, a threshold of ≥1 for initiating point-of-care testing in primary care (equivalent to a test all approach) changed the QALY difference from incremental QALY loss (-0.00396 per 1,000 individuals) to incremental QALY gain (0.00346 per 1,000 individuals) for Clearview Exact Strep A test cassette (Abbott) and Clearview Exact Strep A dipstick - test strip (Abbott) compared with usual care (Table 32). The corresponding ICERs changed from these two tests being dominated in the base-case to £1,890,627 and £2,087,056 per QALY gained for the dipstick and cassette versions respectively when compared with usual care.

The cost-effectiveness estimates were also sensitive to the prevalence GAS among adults presenting in primary care. Increasing the prevalence rate from 22.6% (base-case model) to 35.9% (upper estimate from studies included in systematic review of test accuracy studies) generally favoured usual care (results not shown here); however, whilst decreasing the prevalence to 10% (the value used in the Neuner 2003 study[75]) favoured the intervention arm (i.e. testing). In the majority of cases, the ICERs did not changed substantially to influence interpretation of cost-effectiveness, but the ICERs for Clearview Exact Strep A dipstick - test strip and Clearview Exact Strep A test - cassette (Abbott) changed from being dominated (less effective and more costly) to being more effective and more costly at 10% prevalence rate (Table 32).

Table 32: Adult primary care model: Deterministic sensitivity analyses - Centor threshold for starting antibiotic therapy and prevalence of GAS

Test	Base case			Sensitivity analysis		
	Inc. Costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. Costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#2 - changed Centor threshold for starting antibiotics from ≥3 to ≥1						
Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£7,220	0.00346	£2,087,056
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,757	-0.00396	Dominated	£6,540	0.00346	£1,890,627
SA#5 - Changed GAS prevalence from 22.6% (base-case) to 10% (Neuner 2003[75])						
Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£1,809	0.00131	£1,377,303
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,757	-0.00396	Dominated	£1,640	0.00131	£1,248,775

4.2.13.2 Adult primary care model - Complications rates in treated and untreated GAS infection

Only ICERs for Clearview Exact Strep A dipstick - test strip (Abbott) and Clearview Exact Strep A dipstick – cassette (Abbott) were sensitive to modelled rates of complications (peritoncillar abscess, quinsy and cellulitis as the probabilities used in the model represented all these complications as shown Table 33). In the base-case analysis, GAS related complications rates were set to 1.5% for untreated infection and 1.3% for treated GAS infection based on UK primary care data published by Little et al (2013b).[81] Halving and doubling the complications rates in the treated group did not influence ICERs substantially but doubling complications in the untreated infection to 3% favoured the intervention arm. The ICER for the Clearview Exact Strep A dipstick - test strip (Abbott) and Clearview Exact Strep A cassette (Abbott) changing from being dominated in the base-case by usual care to £712,813 and £839,805 per QALY gained compared with usual care respectively (Table 33).

Table 33: Adult primary care model: Deterministic sensitivity analyses – Complications following GAS infection

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#10 - Doubled complications in treated GAS to 2.6%						
Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£1,323	0.00158	£839,805
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,757	-0.00396	Dominated	£1,123	0.00158	£712,813

4.2.13.3 Adult primary care model - Side-effects of penicillin

Cost-effectiveness estimates were most sensitive to modelled rates of penicillin-induced anaphylaxis. In the base-case, penicillin-induced anaphylaxis were set to 0.01% probability (Table 26) and utility decrement of 9 quality-adjusted life-days lost (Table 27) based on figures reported in the Neuner et al. 2003 study[75] with £1,744 in treatment costs (Hex et al. 2017)[88] reflecting the rare but serious nature of this event. Changing the rate of penicillin-induced rash from 0.01% to 0.64% as reported in Van Howe and Kusnier (2006)[76] favoured testing – the ICER for Clearview Exact Strep A dipstick - test strip (Abbott) and Clearview Exact Strep A cassette (Abbott) changed from being dominated by usual care in the base-case to £19,668 and £30,270 per QALY gained compared with usual care. ICERs for Strep A rapid test - test strip (Biopanda Reagents) and Alere TestPack Plus Strep A - cassette (Abbott) also changed from £388,314 and £609,714 per QALY gained in the base-case to £1,466 and £30,581 per QALY gained compared with usual care, respectively. When the rate of mild penicillin rash was doubled from 2 to 4%, the ICER for Clearview Exact Strep A dipstick - test strip (Abbott) and Clearview Exact Strep A cassette (Abbott) changed from being dominated by usual care in the base-case to £1,524,891 and £1,711,314 per QALY gained compared with usual care, respectively (see Table 34).

Table 34: Adult primary care model: Deterministic sensitivity analyses – Exploring impact of complications of penicillin

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#16 - Doubled rates of mild penicillin reaction (rash) to 4%						

Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£1,836	0.00107	£1,711,314
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,757	-0.00396	Dominated	£1,636	0.00107	£1,524,891
SA#17 - Changed rates of anaphylaxis from 0.01% (Neuner 2003[75]) to 0.64% (Van Howe and Kusnier 2006[76])						
Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£571	0.01887	£30,270
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,757	-0.00396	Dominated	£371	0.01887	£19,668
Strep A rapid test - test strip (Biopanda Reagents)	£1,206	0.00311	£388,314	£33	0.02243	£1,466
Alere TestPack Plus Strep A - cassette (Abbott)	£1,785	0.00293	£609,714	£657	0.0215	£30,581

Note that of the tests with ICERs in the region of £30,000/QALY, only the Alere TestPack Plus used test accuracy data from published peer-reviewed studies. See Table 15 for more information.

4.2.13.4 **Adult primary care model - Cost of testing in primary care**

Excluding the cost of confirmatory throat-culture following a negative test result favoured testing but only the ICER for Strep A rapid test - test strip (Biopanda Reagents) decreased to below £100,000 per QALY gained to £22,428 per QALY gained in comparison to usual care (Table 35). ICERs for the remaining tests suggests that testing was either dominated by usual care (Clearview Exact Strep A cassette (Abbott) and Clearview Exact Strep A dipstick - test strip (Abbott)) or were substantially higher than £100,000 per QALY gained.

Table 35: Adult primary care model: Deterministic sensitivity analyses - Exclude cost of confirmatory throat culture given negative test result

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#20 - Assume no Swab culture in those with a negative test result						
Strep A rapid test - cassette (Biopanda Reagents)	£1,206	0.00311	£388,314	£70	0.00311	£22,428

4.2.13.5 Adult primary care model - Utility decrement, Strep A sore throat and related complications

The base-case estimates were sensitive to changes in disutility associated with GAS sore throat and related complications. Decreasing the utility decrement associated with untreated GAS by half, favoured testing, whilst doubling it favoured usual care (see Table 36). All other testing scenarios involving doubling the utility decrements for treated GAS infection and penicillin-induced rash produced ICERs favourable to testing (key result changes are presented below).

Table 36: Adult primary care model: Deterministic sensitivity analyses - Utility decrement, Strep A sore throat and related complications

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#27 - Halved the utility decrement, untreated GAS						
Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£1,957	0.00667	£293,426
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,757	-0.00396	Dominated	£1,757	0.00667	£263,430
SA#28 - Doubled utility decrement, untreated GAS						
Strep A rapid test - cassette (Biopanda Reagents)	£6,335	0.00311	£2,039,376	£6,335	-0.0002	Dominated
Strep A rapid test - test strip (Biopanda Reagents)	£1,206	0.00311	£388,314	£1,206	-0.0002	Dominated
Alere TestPack Plus Strep A - cassette (Abbott)	£1,785	0.00293	£609,714	£1,785	-0.0004	Dominated
SA#30 - Doubled utility decrement, treated GAS						
Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£1,957	0.00879	£222,505
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,757	-0.00396	Dominated	£1,757	0.00879	£199,759
SA#36 - Doubled utility decrement, penicillin-induced rash						
Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£1,957	0.00107	£1,823,596

Clearview Exact Strep A dipstick - test strip (Abbott)	£1,757	-0.00396	Dominated	£1,757	0.00107	£1,637,173

4.2.14 Adult secondary care model: base-case analysis results

The primary care adult model (section 4.2.2) was adapted to model adult patients presenting with suspected GAS infection in secondary care settings (urgent care/walk-in centres and emergency departments). The modelled pathways remain the same as the adult primary care model depicted in Figure 12 to Figure 14. Sensitivity and specificity of the clinical score at the specified Centor score ≥ 3 for a positive GAS infection were left unchanged as in the adult primary care model (Table 23) as were the modelled pathway probabilities (Table 26) and health-state utility values (Table 27). However, the two models differ in the way treatment and testing costs are calculated. The secondary care model assumes the care pathways associated with suspected cases of GAS infections are presenting for the first time in secondary care and have not received any treatment in primary care. The cost of the initial GP consultation included in the adult primary care model is therefore excluded from the cost. The model does however account for patients attending a GP consultation (and the associated costs) following hospital discharge at a rate equal to the proportion attending repeat GP consultations in the primary care model (14.2% based on figures reported in Little et al 2013b[81]). Additionally, we assume that point-of-care testing within secondary care settings can be performed within the standard allocated time for most hospital-based appointments, such that no additional time is required for administering and processing test results.

Test accuracy estimates were obtained from our systematic review and remained broadly the same as those used to inform the adults in primary care model (Table 24) except for three tests (OSOM Strep A test strip, QuikRead Go Strep A test kit and the Alere TestPack Plus Strep A – cassette). Table 37 presents test accuracy estimates used in adult secondary care model for these three point-of-care tests. Estimates of sensitivity changed from 92% in primary care to 94% in secondary care for OSOM Strep A test strip, from 100% in primary care to 87% in secondary care for QuikRead Go Strep A test kit and from 95% in primary to 90% in secondary care for the Alere TestPack Plus Strep A – cassette. Estimates of specificity for the three tests however remain broadly unchanged across primary and secondary care settings.

Table 37: Adult secondary care model: Test accuracy of point-of-care tests used in economic model*

Test Name Manufacturer	Sensitivity (95% CI)	Specificity (95% CI)	Assumed distribution	Data source
OSOM Strep A test - test strip (Sekisui Diagnostics)	0.94 (0.89, 0.98)	0.95 (0.91, 0.98)	Normal (logit)	5 studies (Bura 2017; Llor 2009; Llor 2011; Rogo 2011; Weinzierl 2018)
QuikRead Go Strep A test kit (Orion Diagnostica)	0.87 (0.78, 0.95)	0.78 (0.71, 0.85)	Normal (logit)	2 studies (Azrad 2019; Stefaniuk 2017)
Alere TestPack Plus Strep A – cassette (Abbott)	0.90 (0.86, 0.94)	0.95 (0.92, 0.96)	Normal (logit)	1 study (Rosenberg 2002) and 1 abstract (Valverde 2018)

*Only tests with secondary care accuracy estimates that are different from those used to inform the adult primary care model are presented in this table

Table 38 presents the cost-effectiveness results for adults in a secondary care setting. As with the adult primary care model, only 13 of the 21 tests that have test accuracy and costs data have been included in this analysis. The pattern and direction of cost-effectiveness in the secondary care adult model is similar to what has been observed in the adult primary care model.

Two tests (Abbotts's Clearview Exact Strep A cassette and Clearview Exact Strep A dipstick – test trip) generated fewer QALYs than usual care and produced ICERs indicating being dominated by usual care (i.e. were less effective and more costly). The remaining 11 tests all generated marginally more QALYs than usual care. The ICERs ranged from £346,005 per QALY gained for NADAL Strep A - test strip (nal von minden GmbH) to £13,785,774 per QALY gained for the QuikRead Go Strep A test kit supplied by Orion Diagnostica.

Table 38: Adult secondary care model: Base-case cost-effectiveness results

TestID	Test Name	Mean costs / 1000 individuals	Mean QALYs/ 1000 individuals	Inc. Costs / 1000 individuals	Inc. QALYS / 1000 individuals	ICER
	Usual care (Clinical scoring based on Centor ≥3 plus clinical assessment)	£49,147	859.82458955	£0	0.0000000	
1	Clearview Exact Strep A cassette (Abbott)	£51,103	859.82063008	£1,957	-0.0039595	Dominated
2	Clearview Exact Strep A dipstick - test strip (Abbott)	£50,903	859.82063008	£1,757	-0.0039595	Dominated
3	BD Veritor Plus system group A Strep Assay - cassette (Beckton Dickinson)					
4	Strep A rapid test - cassette (Biopanda Reagents)	£50,405	859.82769587	£1,259	0.0031063	£405,222
5	Strep A rapid test - test strip (Biopanda Reagents)*	£50,353	859.82769587	£1,206	0.0031063	£388,314
6	NADAL Strep A - test strip (nal von minden GmbH)	£50,488	859.82846603	£1,341	0.0038765	£346,005
7	NADAL Strep A - cassette (nal von minden GmbH)	£50,548	859.82846603	£1,401	0.0038765	£361,488
8	NADAL Strep A plus - cassette (nal von minden GmbH)	£50,578	859.82846603	£1,431	0.0038765	£369,227
9	NADAL Strep A plus - test strip (nal von minden GmbH)	£50,518	859.82846603	£1,371	0.0038765	£353,746
10	NADAL Strep A scan test - cassette (nal von minden GmbH)	£50,698	859.82846603	£1,551	0.0038765	£400,190
11	OSOM Strep A test - test strip (Sekisui Diagnostics)					
12	QuikRead Go Strep A test kit (Orion Diagnostica)	£51,306	859.82474622	£2,160	0.0001567	£13,785,774
13	Alere TestPack Plus Strep A - cassette (Abbott)	£50,995	859.82627789	£1,849	0.0016883	£1,094,955
14	Bionexia Strep A plus - cassette (Biomerieux)					
15	Bionexia Strep A dipstick - test strip (Biomerieux)					
16	Biosynex Strep A - cassette (Biosynex)					
17	Sofia Strep A FIA (Quidel)					
18	ALERE i Strep A (Abbott)					
19	ALERE i Strep A 2 (Abbott)					
20	Cobas Strep A Assay on Liat system (Roche Diagnostics)	£68,841	859.82824206	£19,694	0.0036525	£5,391,982
21	Xpert Xpress Strep A (Cepheid)	£51,308	859.82854357	£2,162	0.0039540	£546,659

[1]Test accuracy data for Biopanda Reagents Strep A rapid test – test strip not available, assumed the strip has same test accuracy as the cassette version of the test

Note: Missing cost-effectiveness estimates are for the tests where either test accuracy or cost data were not available

4.2.15 Adults secondary care model: probabilistic sensitivity analyses

Probabilistic results for the adult secondary care model mirrored the adult primary care PSA model. Results shown below in Table 39.

Table 39: Adult secondary care model: Probabilistic sensitivity analysis results

TestID	Test Name	Mean costs / 1000 individuals	Mean QALYs / 1000 individuals	Inc. Costs / 1000 individuals	Inc. QALYS / 1000 individuals	ICER versus usual care	Probability of cost-effectiveness at £20,000 per QALY
	Usual care (Clinical scoring based on Centor ≥3 plus clinical assessment)	£49,129	859.7355672	£0	0.0000000		1
1	Clearview Exact Strep A cassette (Abbott)	£51,122	859.7315202	£1,993	-0.0040471	Dominated	0
2	Clearview Exact Strep A dipstick – test strip (Abbott)	£50,917	859.731603	£1,788	-0.0039643	Dominated	0
3	BD Veritor Plus system group A Strep Assay - cassette (Beckton Dickinson)						
4	Strep A rapid test - cassette (Biopanda Reagents)	£50,419	859.7386446	£1,289	0.0030773	£418,995	0
5	Strep A rapid test – test strip (Biopanda Reagents)*	£50,363	859.738715	£1,234	0.0031477	£391,971	0
6	NADAL Strep A - test strip (nal von minden GmbH)	£50,517	859.739076	£1,388	0.0035087	£395,451	0
7	NADAL Strep A - cassette (nal von minden GmbH)	£50,575	859.7391528	£1,445	0.0035855	£403,108	0
8	NADAL Strep A plus - cassette (nal von minden GmbH)	£50,605	859.7391322	£1,476	0.0035650	£413,950	0
9	NADAL Strep A plus - test strip (nal von minden GmbH)	£50,543	859.7391416	£1,414	0.0035743	£395,606	0
10	NADAL Strep A scan test - cassette (nal von minden GmbH)	£50,729	859.7390965	£1,600	0.0035292	£453,269	0
11	OSOM Strep A test – test strip (Sekisui Diagnostics)						
12	QuikRead Go Strep A test kit (Orion Diagnostica)	£51,332	859.7355944	£2,203	0.0000272	£81,028,152	0
13	Alere TestPack Plus Strep A - cassette (Abbott)	£51,013	859.7373365	£1,884	0.0017693	£1,064,593	0
14	Bionexia Strep A plus - cassette (Biomerieux)						
15	Bionexia Strep A dipstick – test strip (Biomerieux)						
16	Biosynex Strep A - cassette (Biosynex)						
17	Sofia Strep A FIA (Quidel)						
18	ALERE i Strep A (Abbott)						
19	ALERE i Strep A 2 (Abbott)						
20	Cobas Strep A Assay on Liat system (Roche Diagnostics)	£69,131	859.7389475	£20,001	0.0033802	£5,917,173	0
21	Xpert Xpress Strep A (Cepheid)	£51,332	859.7395664	£2,203	0.0039991	£550,881	0

[1]Test accuracy data for Biopanda Reagents Strep A rapid test – test strip not available, assumed the strip has same test accuracy as the cassette version of the test

4.2.16 Adults secondary care model: exploratory sensitivity analyses

Exploratory analyses were conducted to test the robustness of the economic base-case estimates for adults presenting in secondary care with suspected GAS infection. The base-case ICERs are highly sensitive to various modelling assumptions and input values. In the sections that follow, sensitivity analysis results are presented only for those tests where the ICER is sensitive to the alternative modelling assumptions and parameter inputs (as indicated by changes in the direction of incremental costs or incremental QALYs compared with usual care).

4.2.16.1 Adults in secondary care - Centor threshold for starting antibiotics and testing

In the base-case secondary care model, Centor score ≥3 was used as an indication for starting antibiotic treatment in the usual care arm and to initiate testing using a point-of-care test in the intervention arm. Changing this threshold to Centor score ≥2 had minimal impact on the base-case cost-effectiveness of all tests included in the analysis. However, using a threshold of ≥1 changed the ICER for Clearview Exact Strep A cassette (Abbott) and Clearview Exact Strep A cassette (Abbott) from being dominated by usual care to £2,087,056 and £1,890,627 per QALY gained in comparison to usual care, respectively (Table 40).

Table 40: Adult secondary care model: Deterministic sensitivity analyses - Centor threshold for starting antibiotics

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£7,220	0.00346	£2,087,056
Clearview Exact Strep A cassette (Abbott)	£1,757	-0.00396	Dominated	£6,540	0.00346	£1,890,627

4.2.16.2 Adults in secondary care - Prevalence of GAS

Changing the prevalence of GAS infection in secondary care from 22.6% base-case value to 35.9% (upper value reported in studies included in the test accuracy systematic review) was

less favourable to testing with usual care dominating QuikRead Go Strep A test kit (Orion Diagnostica) and Alere TestPack Plus Strep A - cassette (Abbott) in comparison to base-case results (Table 41). In contrast, a lower prevalence of disease was more favourable to testing with ICERs for Clearview Exact Strep A cassette (Abbott) and Clearview Exact Strep A dipstick - test strip (Abbott) changing from being dominated by usual care to £1,377,303 and £1,248,775 per QALY gained, respectively in comparison to usual care (Table 41). ICERs for all other tests did not change substantially to suggest change in the direction of cost-effectiveness in comparison to usual care.

Table 41: Adult secondary care model: Deterministic sensitivity analyses - Prevalence of GAS

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#4 - Changed Strep A prevalence from 22.6% to 35.9%						
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,160	0.00016	£13,785,774	£2,318	-0.00241	Dominated
Alere TestPack Plus Strep A - cassette (Abbott)	£1,849	0.00169	£1,094,955	£1,886	-0.00055	Dominated
SA#4 - Changed Strep A prevalence from 22.6% to 10%						
Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£1,809	0.00131	£1,377,303
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,757	-0.00396	Dominated	£1,640	0.00131	£1,248,775

4.2.16.3 **Adults in secondary care - Complication rates**

In the base case analysis, GAS related complications rates were set to 1.5% for untreated infection and 1.3% for treated infection based on UK primary care data published by Little et al (2013b).[81] Halving complications in the treated group to 0.65% was less favourable to testing with usual care now dominating QuikRead Go Strep A test kit (Orion Diagnostica) in comparison to ICER produced under base-case assumptions. Doubling the complications rates in the treated group to 2.6% on the otherhand favoured testing, the ICER for the Clearview Exact Strep A dipstick - test strip (Abbott) and Clearview Exact Strep A cassette

(Abbott) changed from being dominated by usual care to £712,813 and £839,805 per QALY gained compared with usual care respectively (Table 42). ICERs for all other tests were much lower in comparison to the base-case estimates but still remained well above £100,000 per QALY gained in comparison to usual care. Doubling complications in the untreated group to 3% was less favourable to testing with usual care now dominating both QuikRead Go Strep A test kit (Orion Diagnostica) and the Alere TestPack Plus Strep A - cassette (Abbott) in comparison to ICERs produced under base-case assumptions.

Table 42: Adult secondary care model: Deterministic sensitivity analyses – complications following GAS sore throat

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#9 - Halved complications, treated infection to 0.65%						
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,160	0.00016	£13,785,774	£2,289	-0.00097	Dominated
SA#10 - Doubled complications in treated infection to 2.6%						
Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£1,323	0.00158	£839,805
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,757	-0.00396	Dominated	£1,123	0.00158	£712,813
SA#12 - Doubled complications, untreated infection to 3%						
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,160	0.00016	£13,785,774	£2,457	-0.00244	Dominated
Alere TestPack Plus Strep A - cassette (Abbott)	£1,849	0.00169	£1,094,955	£2,077	-0.00031	Dominated

4.2.16.4 **Adults in secondary care - Adverse effects of penicillin**

Cost-effectiveness estimates were most sensitive to adverse-effects of penicillin. Halving the mild/uncomplicated side-effects of penicillin (rash) to 1.0% favoured usual care, whilst doubling it favoured testing (Table 43). In the base-case, penicillin-induced anaphylaxis were set to 0.01% probability (Table 26) and utility decrement of 9 quality-adjusted life-days lost (Table 27) based on figures reported in Neuner et al. 2003 study[75] with £1,744 in treatment

costs (Hex et al. 2017)[88] reflecting the rare but serious nature of this event. Changing the rate of penicillin-induced rash from 0.01% to 0.64% as reported in Van Howe and Kusnier (2006)[76] favoured testing, generating ICERs ranging from £1,466 for Strep A rapid test - test strip (Biopanda Reagents) to £67,661 QuikRead Go Strep A test kit (Orion Diagnostica) (Table 43).

Table 43: Adult secondary care model: Deterministic sensitivity analyses – adverse effect of penicillin

	Base case			Sensitivity analysis		
Test	**Inc. costs per 1000 individuals**	**Inc. QALYs per 1000 individuals**	**ICER**	**Inc. costs per 1000 individuals**	**Inc. QALYs per 1000 individuals**	**ICER**
SA#15 - Halved prob mild penicillin reaction (rash) to 1%						
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,160	0.00016	£13,785,774	£2,204	-0.00169	Dominated
Alere TestPack Plus Strep A - cassette (Abbott)	£1,849	0.00169	£1,094,955	£1,900	-0.00046	Dominated
SA#16 - Doubled rates of mild penicillin reaction (rash) to 4%						
Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£1,836	0.00107	£1,711,314
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,757	-0.00396	Dominated	£1,636	0.00107	£1,524,891
SA#17 – changed penicillin-induced anaphylaxis from 0.01% (Neuner 2003[75]) to 0.64% (Van Howe and Kusnier 2006[76])						
Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£571	0.01887	£30,270
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,757	-0.00396	Dominated	£371	0.01887	£19,668
Strep A rapid test - cassette (Biopanda Reagents)	£1,259	0.00311	£405,222	£85	0.02243	£3,807
Strep A rapid test - test strip (Biopanda Reagents)	£1,206	0.00311	£388,314	£33	0.02243	£1,466
NADAL Strep A - test strip (nal von minden GmbH)	£1,341	0.00388	£346,005	£195	0.02275	£8,587
NADAL Strep A - cassette (nal von minden GmbH)	£1,401	0.00388	£361,488	£255	0.02275	£11,225

NADAL Strep A plus - cassette (nal von minden GmbH)	£1,431	0.00388	£369,227	£285	0.02275	£12,544
NADAL Strep A plus - test strip (nal von minden GmbH)	£1,371	0.00388	£353,746	£225	0.02275	£9,906
NADAL Strep A scan test - cassette (nal von minden GmbH)	£1,551	0.00388	£400,190	£405	0.02275	£17,819
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,160	0.00016	£13,785,774	£1,143	0.0169	£67,661
Alere TestPack Plus Strep A - cassette (Abbott)	£1,849	0.00169	£1,094,955	£664	0.0212	£31,324
Xpert Xpress Strep A (Cepheid)	£2,162	0.00395	£546,659	£1,071	0.02192	£48,845

Note that of the tests with ICERs in the region of £30,000/QALY, only the Alere TestPack Plus and QuikRead Go tests used test accuracy data from published peer-reviewed studies. See Table 15 for more information.

4.2.16.5 **Adults in secondary care - Cost of testing in secondary care**

Excluding the cost of confirmatory throat-culture following a negative test result favoured testing but only the ICERs for Strep A rapid test - test strip and cassette (Biopanda Reagents) and the different formulations of the NADAL Strep A test supplied by nal von minden GmbH decreased to below £100,000 per QALY gained in comparison to usual care (Table 44). ICERs for the remaining tests suggest testing was either dominated by usual care (Clearview Exact Strep A cassette (Abbott) and Clearview Exact Strep A dipstick - test strip (Abbott)) or were substantially higher than £100,000 per QALY gained.

Table 44: Adult secondary care model: Deterministic sensitivity analyses - Exclude cost of confirmatory throat culture given negative test result

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#20 - Assume no swab culture in those with a negative test result						
Strep A rapid test - cassette (Biopanda Reagents)	£1,259	0.00311	£405,222	£122	0.00311	£39,333

Strep A rapid test - test strip (Biopanda Reagents)	£1,206	0.00311	£388,314	£70	0.00311	£22,428
NADAL Strep A - test strip (nal von minden GmbH)	£1,341	0.00388	£346,005	£231	0.00388	£59,665
NADAL Strep A - cassette (nal von minden GmbH)	£1,401	0.00388	£361,488	£291	0.00388	£75,148
NADAL Strep A plus - cassette (nal von minden GmbH)	£1,431	0.00388	£369,227	£321	0.00388	£82,890
NADAL Strep A plus - test strip (nal von minden GmbH)	£1,371	0.00388	£353,746	£261	0.00388	£67,407

4.2.16.6 Adults in secondary care - Utility decrement, Strep A sore throat and related complications

The base-case estimates were sensitive to changes in disutility associated with GAS related complications (Table 45). Decreasing the utility decrement associated with treated infection and utility decrement for penicillin-induced rash by a half, doubling the decrement associated with untreated infection and doubling the utility decrement for abscess each favoured usual care and produced ICERs suggesting that usual care dominated testing (see Table 45 for specific tests) in comparison to the base-case assumptions. Halving the utility decrement for untreated infection and doubling the utility decrements for treated infection and penicillin-induced rash all favoured testing.

Table 45: Adult secondary care model: Deterministic sensitivity analyses - Utility decrement, Strep A sore throat and related complications

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#27 - Halved utility decrement, untreated infection						
Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£1,957	0.00667	£293,426
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,757	-0.00396	Dominated	£1,757	0.00667	£263,430
SA#28 - Doubled utility decrement, untreated infection						

Strep A rapid test - cassette (Biopanda Reagents)	£1,259	0.00311	£405,222	£1,259	-0.00022	Dominated
Strep A rapid test - test strip (Biopanda Reagents)	£1,206	0.00311	£388,314	£1,206	-0.00022	Dominated
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,160	0.00016	£13,785,774	£2,160	-0.00848	Dominated
Alere TestPack Plus Strep A - cassette (Abbott)	£1,849	0.00169	£1,094,955	£1,849	-0.00495	Dominated
SA#29-Halved utility decrement, treated infection						
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,160	0.00016	£13,785,774	£2,160	-0.00243	Dominated
Alere TestPack Plus Strep A - cassette (Abbott)	£1,849	0.00169	£1,094,955	£1,849	-0.0003	Dominated
SA#30 - Doubled utility decrement, treated infection						
Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£1,957	0.00879	£222,505
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,757	-0.00396	Dominated	£1,757	0.00879	£199,759
SA#32-Doubled utility decrement, abscess						
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,160	0.00016	£13,785,774	£2,160	-0.00019	Dominated
SA#35-Halved utility decrement, penicillin-induced rash						
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,160	0.00016	£13,785,774	£2,160	-0.00169	Dominated
Alere TestPack Plus Strep A - cassette (Abbott)	£1,849	0.00169	£1,094,955	£1,849	-0.00046	Dominated
SA#36 - Doubled utility decrement, penicillin-induced rash						
Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£1,957	0.00107	£1,823,596
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,757	-0.00396	Dominated	£1,757	0.00107	£1,637,173

4.2.17 Children primary care model: base-case results

The primary care adult model (section 4.2.2) was adapted to model children presenting with suspected GAS infection in a primary care setting. The modelled pathways remain the same as the adult primary care model as depicted in Figure 12 to Figure 14. The prevalence of

GAS changed from 22.6% in the adult primary care model to 30.2% - median prevalence in our systematic review of test accuracy studies among children in primary care settings (see Table 25). Sensitivity and specificity of the clinical score at the specified Centor score ≥3 for a positive GAS infection were left unchanged as in the adult primary care model (see estimates displayed in Table 23) as well as the modelled pathway probabilities (see Table 26) and health-state utility values (see Table 27). Test accuracy estimates were obtained from our systematic review and remained broadly the same as those used to inform the adults in primary care model (Table 24) except for five tests (BD Veritor Plus system group A Strep Assay - cassette supplied by Beckton Dickinson, OSOM Strep A test - test strip supplied by Sekisui Diagnostics, QuikRead Go Strep A test kit by Orion Diagnostica and Alere TestPack Plus Strep A - cassette and ALERE i Strep A both supplied by Abbott) – see Table 24 for further details.

Treatment costs for peritonsillar abscess and related complications of GAS infection in children were estimated at £1,420.50 (Tonsillectomy, 18 years and under with HRG code CA60B)[89], this is slightly lower than the estimate used in the adult primary care model for these complications (£1,571.28 for Tonsillectomy, 19 years and over with HRG code CA60A).[89] Treatment costs for penicillin-induced rash were left unchanged as in the adult models at £10 (assuming treatment switched to another antibiotic e.g. Erythromycin 500mg) and anaphylaxis at £1,744.64.[88]

Table 46: Children primary care model - Base-case cost-effectiveness results

TestID	Test Name	Mean costs / 1000 individuals	Mean QALYs / 10,000 individuals	Inc. Costs / 10,000 individuals	Inc. QALYS / 10,000 individuals	ICER versus usual care
	Usual care (Clinical scoring based on Centor ≥3 plus clinical assessment)	£50,185	939.77019917	£0	0.0000000	
1	Clearview Exact Strep A cassette (Abbott)	£52,219	939.76305927	£2,034	-0.0071399	Dominated
2	Clearview Exact Strep A dipstick – test strip (Abbott)	£52,000	939.76305927	£1,815	-0.0071399	Dominated
3	BD Veritor Plus system group A Strep Assay - cassette (Beckton Dickinson)					
4	Strep A rapid test - cassette (Biopanda Reagents)	£56,943	939.77244279	£6,758	0.0022436	£3,012,257
5	Strep A rapid test – test strip (Biopanda Reagents)*	£51,331	939.77244279	£1,146	0.0022436	£510,969
6	NADAL Strep A - test strip (nal von minden GmbH)	£57,027	939.77347194	£6,842	0.0032728	£2,090,579
7	NADAL Strep A - cassette (nal von minden GmbH)	£57,092	939.77347194	£6,908	0.0032728	£2,110,644
8	NADAL Strep A plus - cassette (nal von minden GmbH)	£57,125	939.77347194	£6,940	0.0032728	£2,120,676
9	NADAL Strep A plus - test strip (nal von minden GmbH)	£57,060	939.77347194	£6,875	0.0032728	£2,100,613
10	NADAL Strep A scan test - cassette (nal von minden GmbH)	£57,257	939.77347194	£7,072	0.0032728	£2,160,807
11	OSOM Strep A test – test strip (Sekisui Diagnostics)					
12	QuikRead Go Strep A test kit (Orion Diagnostica)	£58,198	939.76701428	£8,013	-0.0031849	Dominated
13	Alere TestPack Plus Strep A - cassette (Abbott)	£52,143	939.76939575	£1,958	-0.0008034	Dominated
14	Bionexia Strep A plus - cassette (Biomerieux)					
15	Bionexia Strep A dipstick – test strip (Biomerieux)					
16	Biosynex Strep A - cassette (Biosynex)					
17	Sofia Strep A FIA (Quidel)					
18	ALERE i Strep A (Abbott)					
19	ALERE i Strep A 2 (Abbott)					
20	Cobas Strep A Assay on Liat system (Roche Diagnostics)	£78,229	939.77326996	£28,044	0.0030708	£9,132,658
21	Xpert Xpress Strep A (Cepheid)	£65,704	939.77368771	£15,520	0.0034885	£4,448,757

[1]Test accuracy data for Biopanda Reagents Strep A rapid test – test strip not available, assumed the strip has same test accuracy as the cassette version of the test

Note: Missing cost-effectiveness estimates are for the tests where either test accuracy or cost data were not available

Overall, 13 of the 21 tests were included in the child primary care model. Cost-effectiveness estimates for these tests compared with usual care are presented in Table 46. Simulated mean costs and QALYs were multiplied by 1,000 to aid clarity in presentation because of the small amount of QALYs accrued over a one-year time horizon. The base-case cost-effectiveness for children presenting in primary care largely mirrored those for the adult population. However, because of the slightly higher prevalence of GAS in children (30.2%) compared with adults (22.6%), simulated costs over the one-year time horizon were generally higher in the children model than those in the adult primary care model.

The mean costs simulated under base-case assumptions were £50,185 (£49,147 in the adult primary care model) per 1,000 children treated in primary care under usual care practice and ranged from £51,331 (£50,353 in the adult primary care model) per 1,000 children for Biopanda Reagents's Strep A rapid test strip to £78,229 (£74,932 adult primary care model) per 1,000 children treated in primary care for Cobas Strep A Assay on Liat system supplied by Roche Diagnostics. Simulated QALYs were also higher for children treated in primary care than adults because of the higher baseline utility in children (0.94) compared with a utility norm of 0.863 for adults in the UK. Simulated mean QALYs were 939.7702 (859.8246 in the adult primary care model) for children treated in primary care under usual care practice and ranged from 939.7631 (859.8206 adult primary care model) for Abbott's Clearview Exact Strep A test cassette and strip to 939.7737 (859.8285 in the adult primary care model) for the other tests.

In terms of incremental cost-effectiveness, the base-case estimates suggest usual care was cheaper and generated marginally more QALYs than (and therefore dominated) the QuikRead Go Strep A test kit (Orion Diagnostica), the cassette and strip versions of the Clearview Exact Strep A test cassette supplied by Abbott and the Alere TestPack Plus Strep A - cassette also supplied by Abbott. Incremental cost-effectiveness ratios (ICERs) for the remaining nine tests suggest that testing for children in primary care under base-case assumptions produced ICERs ranging from £510,969 per QALY gained for Strep A rapid test – test strip supplied by Biopanda Reagents to £9,132,658 per QALY gained for the Xpert Xpress Strep A by Cepheid compared with usual care.

4.2.18 Children primary care model: probabilistic sensitivity analyses

Probabilistic results for the children primary care model are shown below and are in line with the deterministic results for the children primary care model (see Table 47).

Table 47: Children primary care model: Probabilistic sensitivity analysis results

TestID	Test Name	Mean costs / 1000 individuals	Mean QALYs / 1000 individuals	Inc. Costs / 1000 individuals	Inc. QALYS / 1000 individuals	ICER versus usual care	Probability of cost-effectiveness at £20,000 per QALY
	Usual care (Clinical scoring based on Centor ≥3 plus clinical assessment)	£50,155	939.4373891	£0	0.0000000		1
1	Clearview Exact Strep A cassette (Abbott)	£52,213	939.4302793	£2,058	-0.0071099	Dominated	0
2	Clearview Exact Strep A dipstick – test strip (Abbott)	£51,992	939.4302569	£1,836	-0.0071322	Dominated	0
3	BD Veritor Plus system group A Strep Assay - cassette (Beckton Dickinson)						
4	Strep A rapid test - cassette (Biopanda Reagents)	£56,974	939.4395964	£6,819	0.0022073	£3,089,328	0
5	Strep A rapid test – test strip (Biopanda Reagents)*	£51,324	939.4395809	£1,169	0.0021918	£533,492	0
6	NADAL Strep A - test strip (nal von minden GmbH)	£57,076	939.4401856	£6,921	0.0027965	£2,474,937	0
7	NADAL Strep A - cassette (nal von minden GmbH)	£57,142	939.4401949	£6,986	0.0028058	£2,489,969	0
8	NADAL Strep A plus - cassette (nal von minden GmbH)	£57,174	939.4402158	£7,019	0.0028266	£2,483,238	0
9	NADAL Strep A plus - test strip (nal von minden GmbH)	£57,108	939.4401998	£6,953	0.0028107	£2,473,868	0
10	NADAL Strep A scan test - cassette (nal von minden GmbH)	£57,306	939.4402599	£7,150	0.0028707	£2,490,798	0
11	OSOM Strep A test – test strip (Sekisui Diagnostics)						
12	QuikRead Go Strep A test kit (Orion Diagnostica)	£58,250	939.4337734	£8,095	-0.0036157	Dominated	0
13	Alere TestPack Plus Strep A - cassette (Abbott)	£52,139	939.4366259	£1,984	-0.0007632	Dominated	0
14	Bionexia Strep A plus - cassette (Biomerieux)						
15	Bionexia Strep A dipstick – test strip (Biomerieux)						
16	Biosynex Strep A - cassette (Biosynex)						
17	Sofia Strep A FIA (Quidel)						
18	ALERE i Strep A (Abbott)						
19	ALERE i Strep A 2 (Abbott)						
20	Cobas Strep A Assay on Liat system (Roche Diagnostics)	£78,466	939.4400902	£28,311	0.0027010	£10,481,369	0
21	Xpert Xpress Strep A (Cepheid)	£65,796	939.4408971	£15,641	0.0035080	£4,458,622	0

[1]Test accuracy data for Biopanda Reagents Strep A rapid test – test strip not available, assumed the strip has same test accuracy as the cassette version of the test

4.2.19 Children primary care model: exploratory sensitivity analyses

Exploratory analyses conducted to test the robustness of economic base-case estimates for children presenting in primary care with suspected GAS infection. The base-case ICERs are highly sensitive to various modelling assumptions and input values. In the sections that follow, sensitivity analysis results are presented only for those tests where the ICER is sensitive to the alternative modelling assumptions and parameter inputs (as indicated by changes in the direction of incremental costs or incremental QALYs compared with usual care).

4.2.19.1 Children primary care model - Centor threshold for starting antibiotics and testing

In the base-case children primary care model, Centor score ≥3 was used cut-off for starting antibiotic treatment in the usual care arm and to initiate testing in the intervention arm. Lowering the threshold to Centor score ≥1 favoured testing with the ICER for QuikRead Go Strep A test kit (Orion Diagnostica) and Alere TestPack Plus Strep A - cassette (Abbott) changing from being dominated in the base-case to £802,519 per QALY gained for Alere TestPack Plus Strep A - cassette (Abbott) and £7,537,017 per QALY gained for the QuikRead Go Strep A test kit (Orion Diagnostica) compared with usual care (see Table 48). Lowering the threshold to Centor score ≥2 favoured testing with the ICER for Alere TestPack Plus Strep A - cassette (Abbott) changing from being dominated in the base-case to £1,904,134 per QALY gained compared with usual care. ICERs for the other tests remain unchanged in comparison to base-case ICERs.

Table 48: Children primary care model: Deterministic sensitivity analyses - Centor threshold for starting antibiotic therapy

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#1 - Changed Centor threshold for starting antibiotics from ≥3 to ≥2						
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958.0	-0.0008	Dominated	£4,294	0.00226	£1,904,134

SA#2 - Changed Centor threshold for starting antibiotics from ≥3 to ≥1						
QuikRead Go Strep A test kit (Orion Diagnostica)	£8,013.0	-0.00318	Dominated	£25,963	0.00344	£7,537,017
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958.0	-0.0008	Dominated	£7,148	0.00891	£802,519

4.2.19.2 Children primary care model - Prevalence of GAS

Changing the prevalence of GAS infection among children presenting in primary care from 30.2% base-case value to 40.1% (upper value reported in studies included in the test accuracy systematic review) had minimal impact on base-case cost-effectiveness results. Changing the prevalence rate to 10% favoured testing but only the ICERs for Clearview Exact Strep A test - cassette (Abbott), the Clearview Exact Strep A dipstick - test strip (Abbott), QuikRead Go Strep A test kit (Orion Diagnostica) and Alere TestPack Plus Strep A - cassette (Abbott) changed from being dominated in the base-case to values between £526,042 per QALY gained for Alere TestPack Plus Strep A - cassette (Abbott) and £2,622,192 per QALY gained for QuikRead Go Strep A test kit (Orion Diagnostica) compared with usual care (Table 49).

Table 49: Children primary care model: Deterministic sensitivity analyses - prevalence of GAS

	Base case Sensitivity analysis			Sensitivity analysis		
Test	**Inc. costs per 1000 individuals**	**Inc. QALYs per 1000 individuals**	**ICER**	**Inc. costs per 1000 individuals**	**Inc. QALYs per 1000 individuals**	**ICER**
SA#5 - Changed GAS prevalence from 30.2% to 10% (Neuner et al 2003[75])						
Clearview Exact Strep A cassette (Abbott)	£2,034.0	-0.00714	Dominated	£1,805	0.00131	£1,374,151
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,815.0	-0.00714	Dominated	£1,636	0.00131	£1,245,623
QuikRead Go Strep A test kit (Orion Diagnostica)	£8,013.0	-0.00318	Dominated	£6,472	0.00247	£2,622,192
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958.0	-0.0008	Dominated	£1,876	0.00357	£526,042

4.2.19.3 Children primary care model - Complication rates in treated and untreated GAS infection

In the base-case analysis, GAS related complications rates were set to 1.5% for untreated infection and 1.3% for treated GAS infection based on UK primary care data published by Little et al (2013b).[81] Doubling the complications rate in treated group to 2.6% favoured testing and changed the ICERs for Clearview Exact Strep A cassette (Abbott), Clearview

Exact Strep A dipstick - test strip (Abbott), QuikRead Go Strep A test kit (Orion Diagnostica) and Alere TestPack Plus Strep A - cassette (Abbott) from being dominated to values between £667,160 per QALY gained for Alere TestPack Plus Strep A - cassette (Abbott) to £5,240,954 per QALY gained for QuikRead Go Strep A test kit (Orion Diagnostica) compared with usual care (Table 50). Decreasing complications in the untreated group to 0.75% favoured testing and changed the ICER for Alere TestPack Plus Strep A - cassette (Abbott) from being dominated in the base-case analysis to £1,659,379 per QALY gained compared with usual care. The ICERs for all other tests were much lower in comparison to the base-case estimates but remained well above £100,000 per QALY gained in comparison to usual care.

Table 50: Children primary care model: Deterministic sensitivity analyses – complications following GAS infection

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	**Inc. QALYs per 1000 individuals**	**ICER**	**Inc. costs per 1000 individuals**	**Inc. QALYs per 1000 individuals**	**ICER**
SA#10 - Doubled complications in treated GAS infection to 2.6%						
Clearview Exact Strep A cassette (Abbott)	£2,034.0	-0.00714	Dominated	£1,268	0.00026	£4,949,827
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,815.0	-0.00714	Dominated	£1,049	0.00026	£4,095,204
QuikRead Go Strep A test kit (Orion Diagnostica)	£8,013.0	-0.00318	Dominated	£7,535	0.00144	£5,240,954
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958.0	-0.0008	Dominated	£1,623	0.00243	£667,160
SA#11 - Halved complications, untreated GAS infection to 0.075%						
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958.0	-0.0008	Dominated	£1,765	0.00106	£1,659,379

4.2.19.4 **Children primary care model - Side-effects of penicillin**

Changing the rate of penicillin-induced rash from 0.01% to 0.64% as reported in Van Howe and Kusnier (2006)[76] favoured testing – the ICERs for Clearview Exact Strep A dipstick - test strip (Abbott), Clearview Exact Strep A cassette (Abbott), QuikRead Go Strep A test kit (Orion Diagnostica) and Alere TestPack Plus Strep A - cassette (Abbott) changed from being dominated by usual care in the base-case, ranging from £28,181 and £416,3321 per QALY gained compared with usual care. ICER for Strep A rapid test - test strip (Biopanda Reagents) changed from £510,969 per QALY gained in the base-case to £3,421 per gained

respectively compared with usual care when the rate of penicillin-induced anaphylaxis was set to 0.64% (Table 51). When the rate of mild penicillin rash was doubled from 2 to 4% favoured testing for two tests (QuikRead Go Strep A test kit (Orion Diagnostica) and Alere TestPack Plus Strep A - cassette (Abbott)).

Table 51: Children primary care model: Deterministic sensitivity analyses - complications of penicillin

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	**Inc. QALYs per 1000 individuals**	**ICER**	**Inc. costs per 1000 individuals**	**Inc. QALYs per 1000 individuals**	**ICER**
SA#16 - Doubled rates of mild penicillin reaction (rash) to 4%						
QuikRead Go Strep A test kit (Orion Diagnostica)	£8,013.0	-0.00318	Dominated	£7,910	0.00113	£6,987,678
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958.0	-0.0008	Dominated	£1,854	0.00355	£522,352
SA#17 - Changed rates of anaphylaxis from 0.01% (Neuner 2003[75]) to 0.64% (Van Howe and Kusnier 2006[76])						
Clearview Exact Strep A cassette (Abbott)	£2,034.0	-0.00714	Dominated	£657	0.01554	£42,266
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,815.0	-0.00714	Dominated	£438	0.01554	£28,181
Strep A rapid test - test strip (Biopanda Reagents)	£1,146.0	0.00224	£510,969	£68	0.02	£3,421
QuikRead Go Strep A test kit (Orion Diagnostica)	£8,013.0	-0.00318	Dominated	£6,825	0.0164	£416,221
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958.0	-0.0008	Dominated	£759	0.01894	£40,104

Note that of the tests with ICERs in the region of £30,000/QALY, only the Alere TestPack Plus used test accuracy data from published peer-reviewed studies. See Table 15 for more information.

4.2.19.5 Children primary care model - Cost of testing in primary care

Excluding the cost of confirmatory throat-culture following a negative test result favoured testing but only the ICER for Strep A rapid test - test strip (Biopanda Reagents) decreased to below £100,000 per QALY gained from £510,969 in the base-case to £45,556 per QALY gained in comparison to usual care (Table 52).

Table 52: Children primary care model: Deterministic sensitivity analyses – exclude cost of confirmatory negative test result

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#20 - Assume no Swab culture in those with a negative test result						
Strep A rapid test - test strip (Biopanda Reagents)	£1,146.0	0.00224	£510,969	£102	0.00224	£45,556

4.2.19.6 **Children primary care model - Utility decrement, Strep A sore throat and related complications**

As in the adult primary and secondary care models, decreasing the utility decrement associated with untreated GAS by half, doubling the utility treatment for treated GAS and doubling the utility decrement for penicillin-induced rash all favoured testing; whilst doubling the decrement associated with untreated infection favoured usual care (Table 53).

Table 53: Children primary care model: Deterministic sensitivity analyses - utility decrements associated with GAS related complications

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#27 - Halved the utility decrement, untreated GAS						
Clearview Exact Strep A cassette (Abbott)	£2,034.0	-0.00714	Dominated	£2,034	0.00706	£287,940
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,815.0	-0.00714	Dominated	£1,815	0.00706	£256,947
QuikRead Go Strep A test kit (Orion Diagnostica)	£8,013.0	-0.00318	Dominated	£8,013	0.00569	£1,407,832
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958.0	-8.00E-04	Dominated	£1,958	0.00541	£361,868

SA#28 - Doubled utility decrement, untreated GAS						
Strep A rapid test - cassette (Biopanda Reagents)	£6,758.0	0.00224	£3,012,257	£6,758	-0.00219	Dominated
Strep A rapid test - test strip (Biopanda Reagents)	£1,146.0	0.00224	£510,969	£1,146	-0.00219	Dominated
SA#30 - Doubled utility decrement, treated GAS						
Clearview Exact Strep A cassette (Abbott)	£2,034.0	-0.00714	Dominated	£2,034	0.0099	£205,352
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,815.0	-0.00714	Dominated	£1,815	0.0099	£183,248
QuikRead Go Strep A test kit (Orion Diagnostica)	£8,013.0	-0.00318	Dominated	£8,013	0.00747	£1,073,116
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958.0	-8.00E-04	Dominated	£1,958	0.00665	£294,273
SA#36 - Doubled utility decrement, penicillin-induced rash						
QuikRead Go Strep A test kit (Orion Diagnostica)	£8,013.0	-0.00318	Dominated	£8,013	0.00113	£7,078,988
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958.0	-0.0008	Dominated	£1,958	0.00355	£551,716

4.2.19.7 Children primary care model - Lower and upper estimates of the accuracy for the clinical score and test

Changing the test accuracy data from the central estimate of test sensitivity and specificity to the lower confidence limit for all test and the Centor score favoured testing but only the ICER for Alere TestPack Plus Strep A - cassette (Abbott) changed from being dominated by usual care under base-case assumption to £5,046,664 per QALY gained compared with usual care (Table 54). The upper limits of test sensitivity and specificity favoured testing (results not presented) but none of the ICERs changed substantially to suggest different interpretation of base-case cost-effectiveness results.

Table 54: Children primary care model: Deterministic sensitivity analyses - Lower and upper limits of confidence intervals for test accuracy data

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#39 - Lower confidence limits of test accuracy						
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958.0	-0.0008	Dominated	£2,567	0.00051	£5,046,664

4.2.20 Children in secondary care: base-case analysis results

The models for adults in secondary care (section 4.2.14) and children in primary care (section 4.2.17) were adapted to model suspected GAS infection among children in secondary care settings (urgent care/walk-in centres and emergency departments). The modelled pathways remain the same as depicted Figure 12 to Figure 14. The prevalence rate was maintained at 30.2% as in the children primary care model. Test accuracy estimates obtained from our systematic review remained broadly the same as those used to inform the primary care models except for six tests (BD Veritor Plus system group A Strep Assay - cassette (Beckton Dickinson), OSOM Strep A test - test strip (Sekisui Diagnostics), QuikRead Go Strep A test kit (Orion Diagnostica), Alere TestPack Plus Strep A - cassette (Abbott), ALERE i Strep A (Abbott) and Xpert Xpress Strep A (Cepheid)). Table 55 presents test accuracy estimates used in children secondary care model for these tests.

Table 55: Children secondary care model: Test accuracy of point-of-care tests used in economic model*

Test Name Manufacturer	Sensitivity (95% CI)	Specificity (95% CI)	Assumed distribution	Data source
OSOM Strep A test - test strip (Sekisui Diagnostics)	0.94 (0.89, 0.98)	0.97 (0.95, 0.99)	Normal (logit)	2 studies (Rogo 2011; Weinzierl 2018)
QuikRead Go Strep A test kit (Orion Diagnostica)	0.87 (0.78, 0.95)	0.78 (0.71, 0.85)	Normal (logit)	2 studies (Azrad 2019; Stefaniuk 2017)
Alere TestPack Plus Strep A - cassette (Abbott)	0.77 (0.73, 0.8)	0.97 (0.93, 0.99)	Normal (logit)	4 studies (Kurtz 2000; Lacroix 2018; Penney

Test Name Manufacturer	Sensitivity (95% CI)	Specificity (95% CI)	Assumed distribution	Data source
				2016; Santos 2003)

*Only tests with secondary care accuracy estimates that are different from those used to inform the children primary care model are presented here

Table 56 presents cost-effectiveness estimates for children treated in secondary care. As with the adult primary care model, only 13 of the 21 tests that have test accuracy and costs data are included in this analysis. The base-case estimates suggest usual care was cheaper and generated marginally more QALYs than (and therefore dominated) four tests (Clearview Exact Strep A cassette (Abbott), Clearview Exact Strep A dipstick – test strip (Abbott), QuikRead Go Strep A test kit (Orion Diagnostica) and Alere TestPack Plus Strep A - cassette (Abbott)). Incremental cost-effectiveness ratios (ICERs) for the remaining tests suggest testing was more costly and more effective than usual care with ICERs ranging from £393,398 per QALY gained for the NADAL Strep A - test strip (nal von minden GmbH) to £6,962,076 per QALY gained for Cobas Strep A Assay on Liat system (Roche Diagnostics) compared with usual care.

Table 56: Children secondary care model: Base-case cost-effectiveness results

TestID	Test Name	Mean costs / 1000 individuals	Mean QALYs / 1000 individuals	Inc. Costs / 1000 individuals	Inc. QALYS / 1000 individuals	ICER versus usual care
	Usual care (Clinical scoring based on Centor ≥3 plus clinical assessment)	£50,185	939.77019917	£0	0.0000000	
1	Clearview Exact Strep A cassette (Abbott)	£52,219	939.76305927	£2,034	-0.0071399	Dominated
2	Clearview Exact Strep A dipstick – test strip (Abbott)	£52,000	939.76305927	£1,815	-0.0071399	Dominated
3	BD Veritor Plus system group A Strep Assay - cassette (Beckton Dickinson)					
4	Strep A rapid test - cassette (Biopanda Reagents)	£51,389	939.77244279	£1,204	0.0022436	£536,579
5	Strep A rapid test – test strip (Biopanda Reagents)*	£51,331	939.77244279	£1,146	0.0022436	£510,969
6	NADAL Strep A - test strip (nal von minden GmbH)	£51,472	939.77347194	£1,288	0.0032728	£393,398
7	NADAL Strep A - cassette (nal von minden GmbH)	£51,538	939.77347194	£1,353	0.0032728	£413,460
8	NADAL Strep A plus - cassette (nal von minden GmbH)	£51,571	939.77347194	£1,386	0.0032728	£423,495
9	NADAL Strep A plus - test strip (nal von minden GmbH)	£51,505	939.77347194	£1,320	0.0032728	£403,429
10	NADAL Strep A scan test - cassette (nal von minden GmbH)	£51,702	939.77347194	£1,517	0.0032728	£463,626
11	OSOM Strep A test – test strip (Sekisui Diagnostics)					
12	QuikRead Go Strep A test kit (Orion Diagnostica)	£52,644	939.76701428	£2,459	-0.0031849	Dominated
13	Alere TestPack Plus Strep A - cassette (Abbott)	£52,143	939.76939575	£1,958	-0.0008034	Dominated
14	Bionexia Strep A plus - cassette (Biomerieux)					
15	Bionexia Strep A dipstick – test strip (Biomerieux)					
16	Biosynex Strep A - cassette (Biosynex)					
17	Sofia Strep A FIA (Quidel)					
18	ALERE i Strep A (Abbott)					
19	ALERE i Strep A 2 (Abbott)					
20	Cobas Strep A Assay on Liat system (Roche Diagnostics)	£71,564	939.77326996	£21,379	0.0030708	£6,962,076
21	Xpert Xpress Strep A (Cepheid)	£52,374	939.77368771	£2,189	0.0034885	£627,449

[1]Test accuracy data for Biopanda Reagents Strep A rapid test – test strip not available, assumed the strip has same test accuracy as the cassette version of the test

*Note: Missing cost-effectiveness estimates are for the tests where either test accuracy or cost data were not available.

4.2.21 Children secondary care model: probabilistic sensitivity analyses

Probabilistic results for the children secondary care model mirrored the children primary care PSA model. Results shown below in Table 57.

Table 57: Children secondary care model: Probabilistic sensitivity analysis results

TestID	Test Name	Mean costs / 1000 individuals	Mean QALYs / 1000 individuals	Inc. Costs / 1000 individuals	Inc. QALYS / 1000 individuals	ICER versus usual care	Probability of cost-effectiveness at £20,000 per QALY
	Usual care (Clinical scoring based on Centor ≥3 plus clinical assessment)	£50,171	938.7482213	£0	0.0000000		1
1	Clearview Exact Strep A cassette (Abbott)	£52,225	938.7409486	£2,054	-0.0072727	Dominated	0
2	Clearview Exact Strep A dipstick – test strip (Abbott)	£51,999	938.7411016	£1,828	-0.0071197	Dominated	0
3	BD Veritor Plus system group A Strep Assay - cassette (Beckton Dickinson)						
4	Strep A rapid test - cassette (Biopanda Reagents)	£51,391	938.7503898	£1,220	0.0021685	£562,702	0
5	Strep A rapid test – test strip (Biopanda Reagents)*	£51,336	938.7503186	£1,165	0.0020972	£555,354	0
6	NADAL Strep A - test strip (nal von minden GmbH)	£51,490	938.7510217	£1,319	0.0028003	£470,954	0
7	NADAL Strep A - cassette (nal von minden GmbH)	£51,558	938.750955	£1,387	0.0027337	£507,442	0
8	NADAL Strep A plus - cassette (nal von minden GmbH)	£51,592	938.7509441	£1,421	0.0027227	£521,782	0
9	NADAL Strep A plus - test strip (nal von minden GmbH)	£51,522	938.7510304	£1,351	0.0028091	£481,010	0
10	NADAL Strep A scan test - cassette (nal von minden GmbH)	£51,718	938.7510872	£1,547	0.0028659	£539,746	0
11	OSOM Strep A test – test strip (Sekisui Diagnostics)						
12	QuikRead Go Strep A test kit (Orion Diagnostica)	£52,664	938.7444727	£2,493	-0.0037486	Dominated	0
13	Alere TestPack Plus Strep A - cassette (Abbott)	£52,149	938.7473253	£1,978	-0.0008960	Dominated	0
14	Bionexia Strep A plus - cassette (Biomerieux)						
15	Bionexia Strep A dipstick – test strip (Biomerieux)						
16	Biosynex Strep A - cassette (Biosynex)						
17	Sofia Strep A FIA (Quidel)						
18	ALERE i Strep A (Abbott)						
19	ALERE i Strep A 2 (Abbott)						
20	Cobas Strep A Assay on Liat system (Roche Diagnostics)	£71,667	938.75085	£21,496	0.0026286	£8,177,637	0
21	Xpert Xpress Strep A (Cepheid)	£52,377	938.7517067	£2,205	0.0034854	£632,779	0

[1]Test accuracy data for Biopanda Reagents Strep A rapid test – test strip not available, assumed the strip has same test accuracy as the cassette version of the test

4.2.22 Children secondary care model: exploratory sensitivity analyses

Exploratory analyses conducted to test the robustness of economic base-case estimates for children presenting in secondary care with suspected GAS infection. The base-case ICERs are highly sensitive to various modelling assumptions and input values. In the sections that follow, sensitivity analysis results are presented only for those tests where the ICER is sensitive to the alternative modelling assumptions and parameter inputs (as indicated by changes in the direction of incremental costs or incremental QALYs compared with usual care).

4.2.22.1 Children secondary care model - Centor threshold for starting antibiotics and testing

In the base-case model for children treated in secondary care, a threshold of ≥3 on the Centor score plus clinical assessment was used as the basis for immediate antibiotic treatment in the usual care arm and to initiate testing in the intervention arm. Changing this threshold to Centor score ≥2 had minimal impact on the base-case cost-effectiveness of all tests included in the analysis (except the Alere TestPack Plus Strep A - cassette (Abbott)). Using a threshold of ≥1 on the Centor score, favoured testing and changed the ICERs for the Alere TestPack Plus Strep A - cassette (Abbott) and the QuikRead Go Strep A test kit (Orion Diagnostica) from being dominated in the base-case to £802,519 per QALY gained and £2,473,336 per QALY gained compared with usual care, respectively (Table 58).

Table 58: Children secondary care model: Deterministic sensitivity analyses - Centor threshold for starting antibiotics and testing

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#1 - Changed Centor threshold for starting antibiotics from ≥3 (base case) to ≥2						
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958	-0.0008	Dominated	£4,294	0.00226	£1,904,134
SA#1 - Changed Centor threshold for starting antibiotics from ≥3 (base case) to ≥1						
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,459	-0.00318	Dominated	£8,520	0.00344	£2,473,336
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958	-0.0008	Dominated	£7,148	0.00891	£802,519

4.2.22.2 **Children secondary care model - Prevalence of GAS**

Changing the prevalence of GAS infection among children presenting in secondary care from 30.2% to 40.1% (upper value reported in studies included in the test accuracy systematic review) had minimal impact on the base-case ICERs in the children secondary care model. In contrast (see Table 59), a lower prevalence of disease at 10% was more favourable to testing with ICERs ranging from £526,042 per QALY gained for Alere TestPack Plus Strep A - cassette (Abbott) to £1,374,151 per QALY gained for the Clearview Exact Strep A cassette (Abbott) compared with usual care. ICERs for all other tests did not change substantially to change the direction of the base-case cost-effectiveness estimates.

Table 59: Children secondary care model: Deterministic sensitivity analyses - prevalence of GAS infection among children presenting in secondary care

	Base case			Sensitivity analysis		
Test	**Inc. costs per 1000 individuals**	**Inc. QALYs per 1000 individuals**	**ICER**	**Inc. costs per 1000 individuals**	**Inc. QALYs per 1000 individuals**	**ICER**
SA#4 - Changed GAS prevalence from 22.6% to 10%						
Clearview Exact Strep A cassette (Abbott)	£2,034	-0.00714	Dominated	£1,805	0.00131	£1,374,151
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,815	-0.00714	Dominated	£1,636	0.00131	£1,245,623
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,459	-0.00318	Dominated	£2,189	0.00247	£886,724
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958	-8.00E-04	Dominated	£1,876	0.00357	£526,042

4.2.22.3 **Children secondary care model - Complication rates**

Halving complications in the treated group to 0.65% and doubling the rate in the untreated group was less favourable to testing with usual care now dominating both QuikRead Go Strep A test kit (Orion Diagnostica) and the Alere TestPack Plus Strep A - cassette (Abbott) in comparison to ICERs produced under base-case assumptions. Doubling the complications rate in the treated group to 2.6% on the otherhand favoured testing, the ICER for the Clearview Exact Strep A dipstick - test strip (Abbott) and Clearview Exact Strep A cassette (Abbott) changed from being dominated by usual care to £712,813 and £839,805 per QALY

gained compared with usual care respectively (Table 60). ICERs for all other tests were much lower in comparison to the base-case estimates but still remained well above £100,000 per QALY gained in comparison to usual care.

Table 60: Children secondary care model: Deterministic sensitivity analyses - complications of GAS infection

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	**Inc. QALYs per 1000 individuals**	**ICER**	**Inc. costs per 1000 individuals**	**Inc. QALYs per 1000 individuals**	**ICER**
SA#9 - Halved complications, treated infection to 0.65%						
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,160	0.00016	£13,785,774	£2,289	-0.00097	Dominated
SA#10 - Doubled complications, treated infection to 2.6%						
Clearview Exact Strep A cassette (Abbott)	£1,957	-0.00396	Dominated	£1,323	0.00158	£839,805
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,757	-0.00396	Dominated	£1,123	0.00158	£712,813
SA#12 - Doubled complications, untreated infection to 3%						
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,160	0.00016	£13,785,774	£2,457	-0.00244	Dominated
Alere TestPack Plus Strep A - cassette (Abbott)	£1,849	0.00169	£1,094,955	£2,077	-0.00031	Dominated

4.2.22.4 Children secondary care model - Adverse effects of penicillin

Cost-effectiveness estimates were most sensitive to adverse effects of penicillin. Halving the mild/uncomplicated side-effects of penicillin (rash) to 1.0% favoured usual care (results not shown here), whilst doubling it favoured testing (see Table 61). Changing the rate of penicillin-induced rash from 0.01% to the 0.65% favoured testing and generated ICERs ranging from £3,421 for Strep A rapid test - test strip (Biopanda Reagents) to £77,459 per QALY gained for the QuikRead Go Strep A test kit (Orion Diagnostica) compared with usual care (Table 61).

Table 61: Children secondary care model: Deterministic sensitivity analyses – Adverse effects of penicillin

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#16 - Doubled rates of mild penicillin reaction (rash) to 4%						
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,459	-0.00318	Dominated	£2,355	0.00113	£2,080,760
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958	-0.0008	Dominated	£1,854	0.00355	£522,352
SA#17 – Changed penicillin-induced anaphylaxis from 0.01% (Neuner 2003[75]) to 0.64% (Van Howe and Kusnier 2006[76])						
Clearview Exact Strep A cassette (Abbott)	£2,034	-0.00714	Dominated	£657	0.01554	£42,266
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,815	-0.00714	Dominated	£438	0.01554	£28,181
Strep A rapid test - cassette (Biopanda Reagents)	£1,204	0.00224	£536,579	£126	0.02	£6,294
Strep A rapid test - test strip (Biopanda Reagents)	£1,146	0.00224	£510,969	£68	0.02	£3,421
NADAL Strep A - test strip (nal von minden GmbH)	£1,288	0.00327	£393,398	£246	0.02043	£12,051
NADAL Strep A - cassette (nal von minden GmbH)	£1,353	0.00327	£413,460	£312	0.02043	£15,266
NADAL Strep A plus - cassette (nal von minden GmbH)	£1,386	0.00327	£423,495	£345	0.02043	£16,874
NADAL Strep A plus - test strip (nal von minden GmbH)	£1,320	0.00327	£403,429	£279	0.02043	£13,659
NADAL Strep A scan test - cassette (nal von minden GmbH)	£1,517	0.00327	£463,626	£476	0.02043	£23,303
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,459	-0.00318	Dominated	£1,270	0.0164	£77,459
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958	-0.0008	Dominated	£759	0.01894	£40,104
Xpert Xpress Strep A (Cepheid)	£2,189	0.00349	£627,449	£1,203	0.01974	£60,926

Note that of the tests with ICERs in the region of £30,000/QALY, only the Alere TestPack Plus used test accuracy data from published peer-reviewed studies. See Table 15 for more information.

4.2.22.5 Children secondary care model - Cost of testing in secondary care

Excluding the cost of confirmatory throat-culture following a negative test result favoured testing but only the ICERs for Strep A rapid test - test strip and cassette (Biopanda Reagents), NADAL Strep A - test strip (nal von minden GmbH) and NADAL Strep A plus - test strip (nal von minden GmbH) decreased to below £100,000 per QALY gained compared with usual care (Table 62).

Table 62: Children secondary care model: Deterministic sensitivity analyses - exclude cost of confirmatory culture given negative test result

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#20 - Assume no Swab culture in those with a negative test result						
Strep A rapid test - cassette (Biopanda Reagents)	£1,204	0.00224	£536,579	£160	0.00224	£71,166
Strep A rapid test - test strip (Biopanda Reagents)	£1,146	0.00224	£510,969	£102	0.00224	£45,556
NADAL Strep A - test strip (nal von minden GmbH)	£1,288	0.00327	£393,398	£279	0.00327	£85,188
NADAL Strep A plus - test strip (nal von minden GmbH)	£1,320	0.00327	£403,429	£312	0.00327	£95,222

4.2.22.6 Children secondary care model - Utility decrement, GAS related complications

The base-case estimates were sensitive to changes in disutility associated with GAS related complications (Table 63). Doubling disutility associated with untreated infection and decreasing disutility associated with treated infection produced ICERs in which usual care dominated the Strep A rapid test – cassette, Strep A rapid test - test strip supplied by Biopanda Reagents, QuikRead Go Strep A test kit (Orion Diagnostica) and Alere TestPack Plus Strep A - cassette (Abbott) in comparison to the base-case estimates. Decreasing the disutility for untreated infection and doubling the decrements for treated infection and penicillin-induced rash all favoured testing.

Table 63: Children secondary care model: Deterministic sensitivity analyses - Utility decrement, Strep A sore throat and related complications – children secondary care model

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#27 - Halved utility decrement, untreated infection						
Clearview Exact Strep A cassette (Abbott)	£2,034	-0.00714	Dominated	£2,034	0.00706	£287,940
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,815	-0.00714	Dominated	£1,815	0.00706	£256,947
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,459	-0.00318	Dominated	£2,459	0.00569	£431,970
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958	-0.0008	Dominated	£1,958	0.00541	£361,868
SA#28 - Doubled utility decrement, untreated infection						
Strep A rapid test - cassette (Biopanda Reagents)	£1,204	0.00224	£536,579	£1,204	-0.00219	Dominated
Strep A rapid test - test strip (Biopanda Reagents)	£1,146	0.00224	£510,969	£1,146	-0.00219	Dominated
SA#29-Halved utility decrement, treated infection						
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,160	0.00016	£13,785,774	£2,160	-0.00243	Dominated
Alere TestPack Plus Strep A - cassette (Abbott)	£1,849	0.00169	£1,094,955	£1,849	-0.0003	Dominated
SA#30 - Doubled utility decrement, treated infection						
Clearview Exact Strep A cassette (Abbott)	£2,034	-0.00714	Dominated	£2,034	0.0099	£205,352
Clearview Exact Strep A dipstick - test strip (Abbott)	£1,815	-0.00714	Dominated	£1,815	0.0099	£183,248
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,459	-0.00318	Dominated	£2,459	0.00747	£329,268
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958	-0.0008	Dominated	£1,958	0.00665	£294,273
SA#36 - Doubled utility decrement, penicillin-induced rash						
QuikRead Go Strep A test kit (Orion Diagnostica)	£2,459	-0.00318	Dominated	£2,459	0.00113	£2,172,070
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958	-0.0008	Dominated	£1,958	0.00355	£551,716

4.2.22.7 Children secondary care model - Lower and upper estimates of the accuracy for the clinical score and test

Changing the test accuracy data from the central estimate of test sensitivity and specificity to the lower confidence limit for all test and the Centor score favoured testing but only the ICER for Alere TestPack Plus Strep A - cassette (Abbott) changed from being dominated by usual care under base-case assumption to £5,046,664 per QALY gained compared with usual care (Table 64). Note that these results are exactly the same as the children primary care model because the two models only differed in terms of cost of additional clinician time required for procession of the test which is 0. The upper limits of test sensitivity and specificity favoured testing (results not presented) but none of the ICERs changed substantially to suggest different interpretation of base-case cost-effectiveness results.

Table 64: Children secondary care model: Deterministic sensitivity analyses - Lower and upper limits of confidence intervals for test accuracy data

Test	Base case			Sensitivity analysis		
	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER	Inc. costs per 1000 individuals	Inc. QALYs per 1000 individuals	ICER
SA#39 - Lower confidence limits of test accuracy						
Alere TestPack Plus Strep A - cassette (Abbott)	£1,958.0	-0.0008	Dominated	£2,567	0.00051	£5,046,664

4.2.23 Additional sensitivity analyses

Table 65 in Appendix 7 displays the list of the 39 deterministic sensitivity analyses conducted to explore the impact of alternative modelling assumptions and parameter inputs on base-case ICERs. In the majority of cases, the ICERs were robust to the implemented changes in the majority of the analyses implemented and the base-case cost-effectiveness conclusions remain unchanged. In particular, assuming a shorter 14-day time horizon (SA#3) which is consistent with typical duration and resolution of symptoms of GAS sore throat infection favoured usual care but the ICERs did not changed substantially to suggest a different interpretation of the base-case cost-effectiveness. Assuming that the treating primary care healthcare professional in both the intervention and usual care arms is a nurse or a pharmacist (SA#19) rather than a GP doctor favoured testing, only if the test cannot be done with the

allocated consultation time. In this instance, the costs associated with the additional clinician time taken to administer and process test results is much lower if seen by a nurse or pharmacist in comparison to if the treating clinician is a GP doctor. Similarly, excluding the cost of the additional clinician time required for process test results (SA#22) favoured testing only where testing cannot be done within allocated primary care consultation time.

4.3 Summary of economic modelling

We undertook a systematic search for economic evaluation studies of the use of point-of-care tests as listed in the NICE scope for patients with suspected GAS infection. We did not identify any relevant economic models that could be adapted. Hence, a *de novo* decision tree model was built to compare point-of-care testing in conjunction with clinical scoring tools with clinical scoring tools alone for children and adults presenting with GAS infection in primary and secondary care settings.

The model took account of the presenting prevalence of disease in the modelled population, accuracy of clinical scoring and testing, the prescribing behaviour of treating clinicians and complications of the infection and treatment. In the base-case analysis, costs were calculated from a UK NHS/PSS perspective over the one-year time horizon. The health impact of intervention was expressed in QALYs captured through application of disutilities associated with treated and untreated infection and related complications over the modelled time horizon.

The scope of the appraisal had called for 21 tests to be evaluated in comparison to usual care practice, however, difficulties in obtaining reliable test accuracy and cost data for all tests, meant that we were only able to include 13 of the 21 tests for which relevant data were available in final economic modelling. Under the base-case model assumptions for adults presenting with suspected GAS in primary care, the incremental cost-effectiveness ratio (ICER) suggest usual care dominated two tests (Clearview Exact Strep A cassette and Clearview Exact Strep A dipstick - test strip both supplied by Abbott). For the remaining 11 tests, testing was marginally more effective and more costly than usual care with ICERs ranging from £388,314 per QALY gained for Biopanda's Strep A rapid test strip to

£7,059,731 per QALY gained for Roche Diagnostics's Cobas Strep A Assay on Liat system compared with usual care.

Probabilistic analyses based on 1,000 Monte-Carlo simulations of the ICER assessed parameter uncertainty and generated probability statements about the cost-effectiveness of point-of-care testing across a range of willingness-to-pay thresholds. Probabilistic ICERs produced results similar to the deterministic base-case ICERs, and suggested that testing was associated with zero probability of cost-effectiveness at willingness to pay thresholds of £0 to £100,000 per QALY gained under base-case assumptions. Similar cost-effectiveness results were obtained in the base-case models for adults presenting in secondary care, and primary and secondary care models for children.

Extensive exploratory deterministic sensitivity analyses of the base-case inputs and assumptions were conducted to understand key model drivers. The findings suggest that the ICER is highly sensitive to parameter inputs and assumptions that (i) increase the cost of testing (acquisition cost of test, additional clinician time for administering and processing test results, and cost of confirmatory throat culture for those testing negative) and ii) the penalty for antibiotic over-prescription/unnecessary antibiotic use (acquisition cost of antibiotic and probabilities for penicillin-induced anaphylaxis and rash). Factoring in costs associated with additional clinician time (at £4 per minute of GP time) for administering tests and £8 for a confirmatory throat culture given a negative test in the base-case both favour usual care as these costs can be substantially higher than the actual cost of the test and are applied only to the intervention arm. On the other-hand, the model predicts lower antibiotic use with testing compared with usual care; however, the cost of antibiotic treatment at £0.91 per course of penicillin, the treatment of choice for GAS infection, is considerably cheaper (than the acquisition costs for majority of the test kits) such that the penalty for supplying antibiotics to those who don't need it.

The base-case incorporates serious adverse-effects of penicillin such as penicillin-induced anaphylaxis with associated high treatment costs and disutility but the modelled rate of 0.01%[75] used in the base-case suggests anaphylaxis is very rare and its impact is therefore minimal on the cost-effectiveness of testing. Sensitivity analyses increasing the rate of

anaphylaxis to 0.64% based on another economic evaluation of GAS pharyngitis[76] produced substantially lower ICERs ranging from £1,466 for per QALY gained for Strep A rapid test – test strip (Biopanda Reagents) to £30,581 per QALY gained for Alere TestPack Plus Strep A - cassette (Abbott) compared with usual care. However, of the tests with ICERs in the region of £30,000 per QALY, only the Alere TestPack Plus and QuikRead Go used test accuracy data from at least one published peer-reviewed study. Other tests with comparable ICERs used data either from abstracts or from information provided by manufacturers which may overestimate test accuracy, and the estimates are not reliable enough to underpin policy decisions.

Cost-effectiveness estimates were also sensitive to the prevalence of GAS infection (higher prevalence favouring usual care and lower prevalence favouring testing), the disutility for untreated infection (lower values favour testing whilst doubling the decrement associated with untreated infection favoured usual care), and disutility for treated GAS infection (doubling the disutility favours testing).

However, testing was also seen to be cost-effective, if the cost of confirmatory throat-culture following a negative test result was excluded as shown in Table 35 for the adult primary care model (the ICER for Strep A rapid test - test strip (Biopanda Reagents) decreased to below £100,000 per QALY gained to £22,428 per QALY gained in comparison to usual care).

4.3.1 Points for discussion regarding the economic modelling

A number of limitations apply to the economic model:

- Although the economic model represented the clinical care pathway in the NHS, practice and management will vary from site to site (within and across both primary care and secondary care settings).
- We could only compare point-of-care testing for 13 of the 21 tests listed in the NICE scope as we did not have test accuracy and/or cost data for the other 8 point-of-care tests.
- There was not enough information on test accuracy data to model Strep A infection in the pharmacy setting or for the elderly population.

- Inputs (except for the sensitivity and specificity data from our effectiveness review) were generally available as point estimates without associated measures of uncertainty such as confidence intervals and standard errors required for probabilistic modelling. Thus, we have had to follow the common practice of assuming a ±10% around the central estimate to incorporate uncertainty in our modelling. This approach to probabilistic analysis is itself associated with degree of uncertainty as it may underestimate or overestimate the true uncertainty in the evidence.
- Our protocol had specified a time horizon of 14-days as the evidence suggest GAS infection is a self-limiting illness with majority of patients making a complete recovery within two-weeks of the infection.[81] However, we extended the time horizon to one-year in the base-case model to accommodate the impact of rare complications of GAS such as acute rheumatic fever where we found evidence to suggest that these complications could be associated with as much as 75 quality-adjusted-life days lost.[75-77] This longer time horizon however required further assumptions to keep the modelling feasible and supported by appropriate evidence. In particular, we assumed only one-episode of GAS infection (the initial index episode) per patient with no possibility for a recurrent infection within the one-year time horizon which is unlikely to represent true reflection of sore throat infections in the community (these point-of-care tests would not be used in people with recurrent sore throats and this was excluded from the scope of work). Extending the time horizon to one-year may also not adequately to capture all cost and consequences associated with infection. For example, there is evidence to suggesting increased risk of death from rheumatic heart disease associated with complications of GAS,[82,79] but this cannot be fully incorporated within the one-year time horizon considered in our base-case.
- We did not explore the impact of a life-time horizon on the cost-effectiveness point-of-care testing because: (i) GAS infection is self-limiting illness (see point above) and (ii) the decision tree structure is not suitable for economic models with lifetime modelling. The model does however account for rare but serious complications of the infection such as acute rheumatic fever and anaphylactic reactions to penicillin, both of which can have long lasting impact. For these, we assume that the one-year time horizon considered in the base-case is sufficient to capture the costs and consequences associated with such complications.

- Sensitivity analyses assuming a shorter time horizon of two weeks (14-days) corresponding to the expected time for symptom resolution, did not alter the conclusions of the base-case cost-effectiveness results.
- Although the model captures the unwanted effects of antibiotic treatment such as penicillin-induced rash and anaphylaxis through incorporating appropriate costs and disutility for these events, resistance to antimicrobial therapy was not explicitly modelled because of evidence suggesting that GAS is highly susceptible to penicillin,[79,90] the treatment of choice for GAS infection.
- The model captures suppurative and non-suppurative complications of GAS infection and the unwanted effects of penicillin use. The probability estimates for suppurative complications were derived from combining data for all such complications (quinsy, sinusitis, otitis media and cellulitis) reported in a large UK cohort study by Little and colleagues (2013b).[81] As the Little et al. (2013b) data does include figures for non-suppurative complications of the infection such as the acute rheumatic fever, all such complications were included in the modelling under the assumption that majority of complications of GAS infection were suppurative with no more than 0.01% being non-suppurative. Sensitivity analysis suggest this assumption has minimal impact on base-case cost-effectiveness results. Other complications of the GAS infection not included in the economic modelling include mortality outcomes and scarlet fever in the children models due to lack of data informing probability estimates for these events; hence, the costs may be underestimated and outcomes overestimated in the models.
- Transmissions between infected and susceptible individuals are not modelled due to lack of evidence to inform transmission rates in dynamic disease modelling. There is also evidence to suggest seasonality effect, e.g. a more increased presentation of GAS infection during the winter months and around Easter time, but this was not explicitly modelled. However, we carried out exploratory analysis in which we varied the prevalence of disease which can be taken as proxy for seasonality effect. These exploratory analyses suggest increasing prevalence of disease among adults and children in primary care generally favoured usual care but the ICERs did not change substantially to suggest a different conclusion from the base-case cost-effectiveness results. On the other hand, lowering the prevalence favoured testing but again the ICERs did not change substantially to alter conclusions of the base-case analyses.

- The modelling may have underestimated the costs as we did not take into account the contribution to antimicrobial stewardship, due to the lack of evidence.
- The model has not accounted for certain high-risk populations such as immunosuppressed patients or pregnant women, as these patients would all be offered antibiotics.
- We have not taken into account that some of these point-of-care tests may also detect other strains of Strep infections such as Strep C and Strep G in addition to Strep A.
- The modelling may have underestimated the costs as we did not take into account the different strains of GAS which may have influenced test performance and disease characteristics, potentially altering the profile of complications.
- We did not consider the impact that introducing routine point-of-care testing might have on patient presentation with sore throat, which could influence the cost-effectiveness results.
- We did not place any monetary value on the impact a point-of-care test might have in including the patient in the treatment decision making process.
- We have not taken into account any broader societal costs such as lost productivity or time off work, due to suspected Strep A infection.
- Finally, modelled changes in costs and QALYs are simulations and have not been observed. Findings should be verified through properly designed and conducted research.

5 Discussion

5.1 Decision problem and objectives

The overall objective was to undertake a clinical and cost-effectiveness analysis of rapid antigen detection and molecular tests in those with high clinical scores, compared to the use of clinical scoring tools alone, for increasing the diagnostic confidence of suspected Group A Streptococcal infection in people who present with an acute sore throat in primary and secondary care. The literature informing clinical effectiveness and cost-effectiveness was systematically reviewed and summarised. A *de novo* economic model was developed to assess the cost-effectiveness of rapid antigen detection and molecular tests in conjunction with clinical scoring tools compared to clinical scoring tools alone in England and Wales.

5.2 Summary of Methods and Findings

5.2.1 Clinical Effectiveness

We searched a number of databases including MEDLINE, EMBASE, Web of Science and the Cochrane library. We found 3,309 unique records, of which 38 were included (26 full text articles, 3 abstracts, 5 manufacturer's submissions (submitted to NICE in response to a request for information) and 4 FDA documents). There were 26 studies which reported on test accuracy data. In general, the methodological quality of the included studies was poor. In particular, in 65.4% (17/26) studies it was unclear whether the sample was consecutive or convenience. Convenience samples may not provide a true representation of the prevalence of GAS. There was judged to be a high level of bias surrounding the subjective reading of some of the point-of-care tests and through lack of adherence to manufacturer's guidance by using the same swab to streak the microbiological culture and then perform the point-of-care test. Additionally, microbiological culture is unlikely to be 100% accurate and may vary with different culture media.

Overall the findings reveal wide variations in the point estimates for the sensitivity (67.9% to 100%) and specificity (73.3% to 100%) of the different point-of-care tests. These estimates were 82.9% to 94.6% for sensitivity and 84.9% to 99.1% for specificity in high-risk populations, including patients with Centor/McIsaac scores > 2, representing the population

of interest. These estimates do not account of any of the unpublished manufacturer submissions.

Clinical scoring tools (FeverPAIN and Centor) have been proposed as a method by which clinicians can identify which patients are most likely to benefit from antibiotic use for sore throat.[8] These tools were developed to predict Strep A (Centor, FeverPAIN), C (FeverPAIN) and G (FeverPAIN). Direct comparison to sore-throat clinical scoring tools and point-of-care tests indicated that specificity estimates were higher for the point-of-care tests, and that sensitivity was generally comparable between the two approaches. Direct comparison to sore-throat clinical scoring tools revealed that point-of-care tests were generally more specific. However, one methodological limitation concerns the varying way clinical scoring tools have been implemented across the included studies. For instance, different studies apply different clinical score cut-offs when recruiting patients. No studies were identified which matched the proposed pathway of care and treatment for patients with acute streptococcal pharyngitis, which would entail evaluating the test accuracy of a combined strategy of sore throat clinical scores at the recommended NICE thresholds (Centor / McIsaac ≥3 or FeverPAIN ≥4) and point-of-care tests. No evidence was identified for the elderly population or in a pharmacy setting. Likewise, data for test accuracy were sparse for each combination of test, population and setting. There were very few head-to-head (direct) comparison studies between index tests.

It was not possible to identify which test is the most accurate due to the lack of evidence. The large degree of heterogeneity among results for studies performing the same rapid test suggests that is unlikely any single study will accurately capture a tests true performance. The apparent accuracy of a test may be penalised for having more studies, compared to tests with a single study, particularly those where the manufacturer has conducted that study. The heterogeneity introduced by the differing characteristics of the studies, further confounded attempts to produce meaningful estimates of test performance, such as care setting, age group, throat score restriction and disease prevalence. Due to the potential heterogeneity, estimates for the sensitivity and specificity of each test were stratified by age group, throat score and care setting, though a lack of evidence meant generalisations had to be made for the majority of estimates.

There is some RCT evidence to suggest the use of rapid antigen detection tests may help reduce antibiotic prescribing rates, but there was no evidence on the effect of using molecular technologies. If a test was proven to be extremely accurate, then it is plausible that clinical staff would trust the outcomes. There was no evidence found on time to antimicrobial prescribing decision, number of appointments required per episode, and onward transmission of infection.

5.2.2 Cost-effectiveness

The systematic review of cost-effectiveness studies identified three studies that used the rapid antigen detection tests as identified in the NICE scope and were classed as economic evaluations. Two studies had some notable limitations and could not be fully data extracted. The one study that allowed a full data extraction, was classed as a high quality economic evaluation when checked against the CHEERS reporting tool.

Thirteen of the twenty-one tests listed in the NICE scope had relevant data on test accuracy and costs, to be included in the final economic modelling. In the base-case analysis, which included adult patients seen in primary care with suspected GAS infection, the economic model found considerable uncertainty about the cost-effectiveness of the different point-of-care tests for suspected GAS infection. This finding was also seen in the other economic models which were adapted for the different patient groups and settings (adult patients seen in the hospital, children seen in primary care and children seen in the hospital). Important uncertainties in the model include parameter inputs and assumptions that increase (i) the cost of testing (acquisition cost of test, additional clinician time for administering and processing test results, and cost of confirmatory throat culture for those testing negative) and ii) the penalty for antibiotic over-prescription/unnecessary antibiotic use (acquisition cost of antibiotic and probabilities for penicillin-induced anaphylaxis and rash).

5.3 Strengths and Limitations

We used a rigorous and exhaustive search to conduct a comprehensive systematic review (literature search, data extraction and analysis) and locate primary studies. All relevant

studies were systematically reviewed and agreement between the two reviewers was very high. We also built a *de novo* decision tree model to assess cost-effectiveness of point-of-care testing. The economic model provides a representation of the clinical care pathway in primary and secondary care settings. The decision tree with populated with probabilities and test accuracy values from the clinical evidence review, published studies and clinical expert opinion.

No studies on point-of-care test use in a pharmacy setting or in the elderly population were retrieved. Additionally, no study matched the proposed pathway of care and treatment for patients with acute streptococcal pharyngitis, which would entail evaluating the test accuracy of a combined strategy of sore throat clinical scores at the recommended NICE thresholds (Centor/McIsaac ≥3 or FeverPAIN ≥4) and point-of-care tests in the age groups defined in the scope.

Children under five years of age were not explicitly considered in this review. Whilst they may benefit from a point-of-care test, following advice from healthcare professionals, we understood that their diagnostic pathway is likely to differ from older age groups, and were considered beyond the scope of this review.

For the purpose of this review we classified GP surgeries, healthcare centres, family practices and primary care clinics as primary care. Secondary care included emergency departments, private pediatric clinics, outpatient clinics, urgent care clinics and walk-in centres. In practice, other countries may define primary and secondary care differently. For example, pediatric clinics could be part of primary care. However, given that it is unclear if test accuracy differs by setting we do not know the impact this could have on the cost-effectiveness estimates.

We only included only English Language studies and studies directly matching the test name, unless we had confirmation that a test had been taken over by another manufacturer (for example, in the case of IMI testpack becoming Alere). We did not include studies where it was unclear whether later iterations of the test were different. During the EAG write up of the final report, Abbott notified NICE that the Alere i Strep A test is no longer available. Also,

the Alere i Strep A 2 has been rebranded as the ID NOW Strep A 2. Our previously excluded studies were re-screened by test name, and there were none that used the ID NOW Strep A 2 test. It is the EAGs understanding that the information in this report relevant to the Alere i Strep A 2 is transferrable to the ID NOW Strep A 2. Additionally, the Clearview Exact Strep A tests (both cassette and dipstick) were replaced with new Clearview Exact Strep A 2 editions. The manufacturers supplied NICE with information that there are procedural differences between previous Clearview tests and the new A 2 editions. Therefore, the results for the Clearview Exact Strep A tests in this review may therefore not be generalizable to the current Clearview products on the market. Furthermore, our previously excluded studies were re-screened by test name, and there were none that used the Clearview Exact Strep A 2 editions.

We did not explore the effect of culture medium on test accuracy within the review. One of the included test accuracy studies found that using different culture media was showing Strep A positivity on samples that were initially negative.[39] This could indicate possible differences in accuracy of different culture media.

Test accuracy may also vary greatly based on the quality of the swabbing. It is unclear how the level of training of clinical staff involved in these studies compares to routine care, which could limit the generalisability of these results.

The evidence informing the test accuracy estimates was not sufficient to produce reliable of robust estimates that we could be confident actually reflected the tests true performance in any particular patient group. This concern extends to the economic modelling which used the estimates for each test.

The studies within this review determined antibiotic appropriateness to be based upon Strep A positivity in the culture. However, culture may detect Strep A carriage as opposed to disease. PCR was a potential alternative reference standard, but was less widely used, and encounters the same issue of carriage detection.

Although the economic model represented the clinical care pathway in the NHS, practice and management will vary from site to site (within and across both primary care and secondary care settings). There was not enough information on test accuracy data to model GAS for the pharmacy setting or for the elderly population. Furthermore, we could only compare 13 of the 21 point-of-care tests as listed in the NICE scope as we did not have test accuracy and/or cost data for the other 8 point-of-care tests. The modelling may have underestimated the costs as we did not take into account the different strains of GAS which may have influenced test performance and alter the profile of complications, seasonality of GAS infection, resistance to antimicrobial therapy, the onward transmission of the infection and the broader societal costs.

6 Conclusions

The systematic review and cost-effectiveness model identify uncertainties around the adoption of point-of-care tests within primary and secondary care settings in England and Wales. The available evidence is heterogeneous in populations studied, design, methods and analysis. Although sensitivity and specificity estimates are promising, we have little information on the best point-of-care test to use. While there is potential for the point-of-care tests to be cost-effectiveness in both primary and secondary care settings, key parameter inputs and modelling assumptions need to be confirmed and model findings remain uncertain.

6.1 Recommendations for future research

Further research is needed to understand the test accuracy of point-of-care tests within the proposed NHS pathway and within comparable settings and patient groups. Future work which considers head-to-head test accuracy studies or randomised controlled trials using multiple point-of-care tests in relevant patient populations and healthcare settings considered in the NICE scope would provide relevant comparator information and help determine the value of point-of-care testing.

7 References

1. Hannaford PC, Simpson JA, Bisset AF, Davis A, McKerrow W, Mills R. The prevalence of ear, nose and throat problems in the community: results from a national cross-sectional postal survey in Scotland. *Fam Pract* 2005;**22**:227-33.
2. Maxwell S, Smith S, Smith GE, Heatlie H, Bashford JNR, Ashcroft DM*, et al.* Reducing variation in antibacterial prescribing rates for 'cough/cold' and sore throat between 1993 and 2001: Regional analyses using the general practice research database. *Public Health* 2006;**120**:752-9.
3. Clinical Knowledge Summaries. *Sore throat - acute: Background information*. National Institute for Health and Care Excellence; 2018. URL: https://cks.nice.org.uk/sore-throat-acute#!background (Accessed 5 April 2019).
4. Martin E. *Concise medical dictionary*: Oxford University Press, USA; 2015.
5. Ashworth M, Charlton J, Ballard K, Latinovic R, Gulliford M. Variations in antibiotic prescribing and consultation rates for acute respiratory infection in UK general practices 1995-2000. *Br J Gen Pract* 2005;**55**:603-8.
6. Little P, Hobbs FD, Moore M, Mant D, Williamson I, McNulty C*, et al.* Clinical score and rapid antigen detection test to guide antibiotic use for sore throats: randomised controlled trial of PRISM (primary care streptococcal management). *BMJ* 2013;**347**:f5806.
7. Centor RM, Witherspoon JM, Dalton HP, Brody CE, Link K. The diagnosis of strep throat in adults in the emergency room. *Med Decis Making* 1981;**1**:239-46. http://dx.doi.org/10.1177/0272989x8100100304
8. National Institute for Health and Care Excellence. *Sore throat (acute): antimicrobial prescribing: NICE guideline [NG84]*. 2018. URL: https://www.nice.org.uk/guidance/ng84 (Accessed 2 April 2019).
9. Linder JA, Stafford RS. Antibiotic treatment of adults with sore throat by community primary care physicians: a national survey, 1989-1999. *JAMA* 2001;**286**:1181-6.
10. Gulliford MC, Dregan A, Moore MV, Ashworth M, van Staa T, McCann G*, et al.* Continued high rates of antibiotic prescribing to adults with respiratory tract infection: survey of 568 UK general practices. *BMJ open* 2014;**4**:e006245.
11. National Institute for Health and Care Excellence. *Rapid Tests for Group A Streptococcal infections in people with a sore throat: Final Scope*. 2018. URL: https://www.nice.org.uk/guidance/gid-dg10025/documents/final-scope (Accessed 10 April 2019).
12. Petersen I, Johnson A, Islam A, Duckworth G, Livermore D, Hayward A. Protective effect of antibiotics against serious complications of common respiratory tract infections: retrospective cohort study with the UK General Practice Research Database. *BMJ* 2007;**335**:982.
13. Public Health England. *Invasive group A streptococcal disease: managing close contacts*. 2008. URL: https://www.gov.uk/government/publications/invasive-group-a-streptococcal-disease-managing-community-contacts (Accessed 5 April 2019).
14. Corner M. *ONS: Sickness absence in the UK labour market: 2016*. 2017. URL: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/labourproductivity/articles/sicknessabsenceinthelabourmarket/2016 (Accessed 10 April 2019).
15. Newcastle and York External Assessment Centre. *Point-of-care diagnostic testing in primary care for strep A infection in sore throat*. National Institute for Health and Care Excellence; 2018. URL: https://www.nice.org.uk/advice/mib145 (Accessed 21 September 2018, 2018).
16. McCormick A, Fleming D, Charlton J. *Morbidity statistics from general practice: fourth national study 1991-1992*. London HMSO; 1995.
17. Fine AM, Nizet V, Mandl KD. Large-scale validation of the Centor and McIsaac scores to predict group A streptococcal pharyngitis. *Arch Intern Med* 2012;**172**:847-52.

18. Bryant AE, Stevens DL. 199: Streptococcus pyogenes. In: Bennett JE, Dolin R, Blaser MJ, eds. Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases. 8 ed. Philadelphia, PA: Elsevier/Saunders; 2015.
19. Berry GJ, Miller CR, Prats MM, Marquez C, Oladipo OO, Loeffelholz MJ*, et al.* Comparison of the Alere i Strep A Test and the BD Veritor System in the Detection of Group A Streptococcus and the Hypothetical Impact of Results on Antibiotic Utilization. *J Clin Microbiol* 2018;**56**:e01310-17.
20. Uhl JR, Adamson SC, Vetter EA, Schleck CD, Harmsen WS, Iverson LK*, et al.* Comparison of LightCycler PCR, rapid antigen immunoassay, and culture for detection of group A streptococci from throat swabs. *J Clin Microbiol* 2003;**41**:242-9.
21. Kocoglu E, Karabay O, Yilmaz F, Ekerbicer H. The impact of incubating the throat culture for 72 h on the diagnosis of group A beta-hemolytic streptococci. *Auris Nasus Larynx* 2006;**33**:311-3.
22. Lacroix L, Cherkaoui A, Schaller D, Manzano S, Galetto-Lacour A, Pfeifer U*, et al.* Improved Diagnostic Performance of an Immunofluorescence-based Rapid Antigen Detection Test for Group A Streptococci in Children With Pharyngitis. *Pediatr Infect Dis J* 2018;**37**:206-11.
23. Wang F, Tian Y, Chen L, Luo R, Sickler J, Liesenfeld O*, et al.* Accurate Detection of Streptococcus pyogenes at the Point of Care Using the cobas Liat Strep A Nucleic Acid Test. *Clin Pediatr (Phila)* 2017;**56**:1128-34.
24. National Institute for Health and Care Excellence. *Fever in under 5s: assessment and initial management: Clinical guideline [CG160]*. 2013. URL: nice.org.uk/guidance/cg160 (Accessed 27 March 2019).
25. Roper SM, Edwards R, Mpwo M, Mutandiro C, Devaraj S. Reducing Errors in an Emergency Center Setting Using an Automated Fluorescence Immunoassay for Group A Streptococcus Identification. *Clin Pediatr (Phila)* 2017;**56**:675-7.
26. Public Health England. *UK Standards for Microbiology Investigations: Investigation of Throat Related Specimens*. London: Standards Unit, Microbiology Services, PHE; 2015. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/423204/B_9i9.pdf (Accessed 27 March 2019).
27. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD*, et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;**343**:d5928.
28. *The Joanna Briggs Institute Critical Appraisal tools for use in JBI Systematic Reviews: Checklist for Analytical Cross Sectional Studies*. Adelaide, Australia: Joanna Briggs Institute; 2017. URL: http://joannabriggs.org/research/critical-appraisal-tools.html (Accessed 1 May 2019).
29. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10 Analysing and Presenting Results. In: Deeks JJ, Bossuyt PM, Gatsonis C, editors. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. Version 1.0.*: Cochrane Collaboration; 2010.
30. Takwoingi Y, Guo B, Riley RD, Deeks JJ. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. *Stat Methods Med Res* 2017;**26**:1896-911.
31. Azrad M, Danilov E, Goshen S, Nitzan O, Peretz A. Detection of group a Streptococcus in pharyngitis by two rapid tests: comparison of the BD VeritorTM and the QuikRead go Strep A. *Eur J Clin Microbiol Infect Dis* 2019 1-7. [Epub ahead of print].
32. Bird C, Winzor G, Lemon K, Moffat A, Newton T, Gray J. A Pragmatic Study to Evaluate the Use of a Rapid Diagnostic Test to Detect Group A Streptococcal Pharyngitis in Children With the Aim of Reducing Antibiotic Use in a UK Emergency Department. *Pediatr Emerg Care* 2018 [Epub ahead of print].
33. Bura M, Michalak M, Chojnicki M, Padzik M, Mozer-Lisewska I. Moderate and severe pharyngitis in young adult inhabitants of poznan, Western Poland. *Fam Med Prim Care Rev* 2017;**19**:12-7.
34. Cohen DM, Russo ME, Jaggi P, Kline J, Gluckman W, Parekh A. Multicenter Clinical Evaluation of the Novel Alere i Strep A Isothermal Nucleic Acid Amplification Test. *J Clin Microbiol* 2015;**53**:2258-61.

35. Dimatteo LA, Lowenstein SR, Brimhall B, Reiquam W, Gonzales R. The relationship between the clinical features of pharyngitis and the sensitivity of a rapid antigen test: evidence of spectrum bias. *Ann Emerg Med* 2001;**38**:648-52.
36. Humair J-P, Revaz SA, Bovier P, Stalder H. Management of Acute Pharyngitis in Adults. *Arch Intern Med* 2006;**166**:640.
37. Johansson L, Mansson NO. Rapid test, throat culture and clinical assessment in the diagnosis of tonsillitis. *Fam Pract* 2003;**20**:108-11.
38. Johnson DR, Kaplan EL. False-positive rapid antigen detection test results: reduced specificity in the absence of group A streptococci in the upper respiratory tract. *J Infect Dis* 2001;**183**:1135-7.
39. Kurtz B, Kurtz M, Roe M, Todd J. Importance of inoculum size and sampling effect in rapid antigen detection for diagnosis of Streptococcus pyogenes pharyngitis. *J Clin Microbiol* 2000;**38**:279-81.
40. Lindbaek M, Hoiby EA, Lermark G, Steinsholt IM, Hjortdahl P. Which is the best method to trace group A streptococci in sore throat patients: culture or GAS antigen test? *Scand J Prim Health Care* 2004;**22**:233-8.
41. Llor C, Calvino O, Hernandez S, Crispi S, Perez-Bauer M, Fernandez Y*, et al.* Repetition of the rapid antigen test in initially negative supposed streptococcal pharyngitis is not necessary in adults. *Int J Clin Pract* 2009;**63**:1340-4.
42. Llor C, Madurell J, Balague-Corbella M, Gomez M, Cots JM. Impact on antibiotic prescription of rapid antigen detection testing in acute pharyngitis in adults: a randomised clinical trial. *Br J Gen Pract* 2011;**61**:e244-51.
43. McIsaac WJ, Kellner JD, Aufricht P, Vanjaka A, Low DE. Empirical validation of guidelines for the management of pharyngitis in children and adults. *JAMA* 2004;**291**:1587-95.
44. Nerbrand C, Jasir A, Schalen C. Are current rapid detection tests for Group A Streptococci sensitive enough? Evaluation of 2 commercial kits. *Scand J Infect Dis* 2002;**34**:797-9.
45. Penney C, Porter R, O'Brien M, Daley P. Operator Influence on Blinded Diagnostic Accuracy of Point-of-Care Antigen Testing for Group A Streptococcal Pharyngitis. *Can J Infect Dis Med Microbiol* 2016;**2016**:Article ID 1710561.
46. Rogo T, Schwartz RH, Ascher DP. Comparison of the Inverness Medical Acceava Strep A test with the Genzyme OSOM and Quidel QuickVue Strep A tests. *Clin Pediatr (Phila)* 2011;**50**:294-6.
47. Rosenberg P, McIsaac W, Macintosh D, Kroll M. Diagnosing streptococcal pharyngitis in the emergency department: Is a sore throat score approach better than rapid streptococcal antigen testing? *CJEM* 2002;**4**:178-84.
48. Santos O, Weckx LL, Pignatari AC, Pignatari SS. Detection of Group A beta-hemolytic Streptococcus employing three different detection methods: culture, rapid antigen detecting test, and molecular assay. *Braz J Infect Dis* 2003;**7**:297-300.
49. Stefaniuk E, Bosacka K, Wanke-Rytt M, Hryniewicz W. The use of rapid test QuikRead go Strep A in bacterial pharyngotonsillitis diagnosing and therapeutic decisions. *Eur J Clin Microbiol Infect Dis* 2017;**36**:1733-8.
50. Thornley T, Marshall G, Howard P, Wilson AP. A feasibility service evaluation of screening and treatment of group A streptococcal pharyngitis in community pharmacies. *J Antimicrob Chemother* 2016;**71**:3293-9.
51. Weinzierl EP, Jerris RC, Gonzalez MD, Piccini JA, Rogers BB. Comparison of Alere i Strep A Rapid Molecular Assay With Rapid Antigen Testing and Culture in a Pediatric Outpatient Setting. *Am J Clin Pathol* 2018;**19**:19.
52. Worrall G, Hutchinson J, Sherman G, Griffiths J. Diagnosing streptococcal sore throat in adults: randomized controlled trial of in-office aids.[Erratum appears in Can Fam Physician. 2007 Jun;53(6):1006]. *Can Fam Physician* 2007;**53**:666-71.
53. Andersen JB, Dahm TL, Nielsen CT, Frimodt-Moller N. [Diagnosis of streptococcal tonsillitis in the pediatric department with the help of antigen detection test]. *Ugeskr Laeger* 2003;**165**:2291-5.

54. Pauchard JY, Verga ME, Bersier J, Durusell C, Gehri M, Vaudaux B. Performance of a rapid antigen detection test in group A beta-haemolytic streptococcal pharyngitis in comparison with three clinical decision rule in a tertiary paediatric emergency department. *Swiss Med Wkly* 2013;**197**:6S [meeting abstract].
55. Valverde ED, Colmenarejo C, llescas S, Gonzalez JC. P0820: Evaluation of a rapid streptococcal Group A antigen test in different age groups. Paper presented at: ECCMID; Madrid, Spain.
56. *FDA Decision Summary: Substantial Equivalence determination for the BD Veritor™ System for Rapid Detection of Group A Streptococcus (Group A Strep). K122718*. Silver Spring, MD, USA: U.S. Food and Drug Administration; 2013. URL: http://www.accessdata.fda.gov/cdrh_docs/reviews/K122718.pdf (Accessed 6 March 2019).
57. *FDA Decision Summary: Substantial Equivalence determination for the Sophia Strep A FIA assay for use with the Sophia Analyzer. K123793*. Silver Spring, MD, USA: U.S. Food and Drug Administration; 2013. URL: https://www.accessdata.fda.gov/cdrh_docs/reviews/K123793.pdf (Accessed 6 March 2019).
58. *FDA Decision Summary: Substantial Equivalence determination for the Alere i Strep A 2 performed on the Alere i Analyzer for the detection of Streptococcus pyogenes (Group A Streptococcus). K173653*. Silver Spring, MD, USA: U.S. Food and Drug Administration; 2018. URL: http://www.accessdata.fda.gov/cdrh_docs/reviews/K173653.pdf (Accessed 6 March 2019).
59. *FDA Decision Summary: Substantial Equivalence determination for Xpert Xpress Strep A test performed on the GeneXpert Xpress System. K173398*. Silver Spring, MD, USA: U.S. Food and Drug Administration; 2018. URL: http://www.accessdata.fda.gov/cdrh_docs/reviews/K173398.pdf (Accessed 6 March 2019).
60. Esposito S, Blasi F, Bosis S, Droghetti R, Faelli N, Lastrico A*, et al.* Aetiology of acute pharyngitis: the role of atypical bacteria. *J Med Microbiol* 2004;**53**:645-51.
61. Gieseker KE, Roe MH, MacKenzie T, Todd JK. Evaluating the American Academy of Pediatrics diagnostic standard for Streptococcus pyogenes pharyngitis: backup culture versus repeat rapid antigen testing. *Pediatrics* 2003;**111**:e666-70.
62. Felsenstein S, Faddoul D, Sposto R, Batoon K, Polanco CM, Dien Bard J. Molecular and clinical diagnosis of group A streptococcal pharyngitis in children. *J Clin Microbiol* 2014;**52**:3884-9.
63. Thamlikitkul V, Rachata T, Popum S, Chinswangwatanakul P, Srisomnuek A, Seenama C*, et al.* Accuracy and utility of rapid antigen detection tests for group A beta-hemolytic Streptococcus on ambulatory adult patients with sore throat associated with acute respiratory infections at Siriraj hospital. *J Med Assoc Thai* 2018;**101**:441-9.
64. Ramos JL, Fraile MT, Chanza M, Tormo N, Lurbe A, Gimeno C. Rapid detection of Streptococcus pyogenes in peripheral medical centres. A pilot custody assay. *Clin Microbiol Infect* 2011;**17**:S250 [meeting abstract].
65. Hoffmann K, Reichardt B, Zehetmayer S, Maier M. Evaluation of the implementation of a rapid streptococcal antigen test in a routine primary health care setting: from recommendations to practice. *Wien Klin Wochenschr* 2012;**124**:633-8.
66. Lean WL, Arnup S, Danchin M, Steer AC. Rapid diagnostic tests for group A streptococcal pharyngitis: a meta-analysis. *Pediatrics* 2014;**134**:771-81.
67. Roveta S, Marchese A, Debbia EA. Evaluation of the Uro-Quick, a new rapid automated system, for the detection of well-characterized antibiotic-resistant bacteria. *J Chemother* 2004;**16**:107-18.
68. Ebrahimi S, Mohabatkar H, Behbahani M. Predicting Promiscuous T Cell Epitopes for Designing a Vaccine Against Streptococcus pyogenes. *Appl Biochem Biotechnol* 2018;**11**:11. http://dx.doi.org/https://dx.doi.org/10.1007/s12010-018-2804-5
69. Leeflang MM, Bossuyt PM, Irwig L. Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis. *J Clin Epidemiol* 2009;**62**:5-12.

70. Gerber MA, Shulman ST. Rapid diagnosis of pharyngitis caused by group A streptococci. *Clin Microbiol Rev* 2004;**17**:571-80.
71. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D*, et al.* Consolidated health economic evaluation reporting standards (CHEERS) statement. *Cost Eff Resour Alloc* 2013;**11**:6.
72. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R*, et al.* Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;**8**:1-158.
73. Little P, Hobbs FD, Moore M, Mant D, Williamson I, McNulty C*, et al.* PRImary care Streptococcal Management (PRISM) study: in vitro study, diagnostic cohorts and a pragmatic adaptive randomised controlled trial with nested qualitative study and cost-effectiveness study. *Health Technol Assess* 2014;**18**:vii-xxv, 1-101.
74. National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal*. London: NICE; 2013. URL: https://www.nice.org.uk/process/pmg9 (Accessed 10 March 2019).
75. Neuner JM, Hamel MB, Phillips RS, Bona K, Aronson MD. Diagnosis and management of adults with pharyngitis. A cost-effectiveness analysis. *Ann Intern Med* 2003;**139**:113-22.
76. Van Howe RS, Kusnier LP, 2nd. Diagnosis and management of pharyngitis in a pediatric population based on cost-effectiveness and projected health outcomes. *Pediatrics* 2006;**117**:609-19.
77. Klepser DG, Bisanz SE, Klepser ME. Cost-effectiveness of pharmacist-provided treatment of adult pharyngitis. *Am J Manag Care* 2012;**18**:e145-54.
78. National Center for Immunization and Respiratory Diseases, Division of Bacterial Diseases. *Pharyngitis (Strep Throat)*. Centers for Disease Control and Prevention; 2018. URL: https://www.cdc.gov/groupastrep/diseases-hcp/strep-throat.html (Accessed 17 April 2019).
79. Wessels MR. *Streptococcus pyogenes : Basic Biology to Clinical Manifestations*. Bethesda MD, USA: National Center for Biotechnology Information; 2016. URL: https://www.ncbi.nlm.nih.gov/books/NBK333418/ (Accessed 25 April 2019).
80. Aalbers J, O'Brien KK, Chan WS, Falk GA, Teljeur C, Dimitrov BD*, et al.* Predicting streptococcal pharyngitis in adults in primary care: a systematic review of the diagnostic accuracy of symptoms and signs and validation of the Centor score. *BMC Med* 2011;**9**:67. http://dx.doi.org/10.1186/1741-7015-9-67
81. Little P, Stuart B, Hobbs FD, Butler CC, Hay AD, Campbell J*, et al.* Predictors of suppurative complications for acute sore throat in primary care: prospective clinical cohort study. *BMJ* 2013;**347**:f6867. http://dx.doi.org/https://dx.doi.org/10.1136/bmj.f6867
82. *Point-of-care tests for group A streptococcus: horizon scanning report*. Oxford: NIHR Community Healthcare MedTech and In vitro Diagnostics Co-operative; 2015. URL: https://www.community.healthcare.mic.nihr.ac.uk/reports-and-resources/horizon-scanning-reports/point-of-care-tests-for-group-a-streptococcus (Accessed 25 April 2019).
83. Kind P, Dolan P, Gudex C, Williams A. Variations in population health status: results from a United Kingdom national questionnaire survey. *BMJ* 1998;**316**:736-41.
84. *NHS Supply chain catalogue*. London: NHS Supply Chain; 2019. URL: https://my.supplychain.nhs.uk/catalogue (Accessed 25 April 2019).
85. *Clearview Exact Strep A Dipstick x 25*. Medisave UK Ltd. URL: https://www.medisave.co.uk/clearview-exact-strep-a-dipstick-x-25-test-kit-p-7660.html (Accessed 17 April 2019).
86. Curtis L, Burns A. *Unit Costs of Health and Social Care* Canterbury, UK: Personal Social Services Research Unit, University of Kent; 2017. http://dx.doi.org/10.22024/UniKent/01.02/65559
87. Joint Formulary Committee. *BNF 74: September 2017*. London: Pharmaceutical Press; 2017.
88. Hex N, Retzler J, Bartlett C, Arber M. *The Cost of Sepsis Care in the UK: Final Report [YHEC]*. North Grey Literature Collection; 2017. URL: http://allcatsrgrey.org.uk/wp/wpfb-file/yhec-sepsis-report-17-02-17-final-pdf/ (Accessed 17 April 2019).

89. *NHS Reference Costs: 2017/18 reference cost data*. NHS Improvement; 2018. URL: https://improvement.nhs.uk/resources/reference-costs/#rc1718 (Accessed 17 April 2019).
90. Centers for Disease CaP. *Group A Streptococcal Disease (GAS) Disease - Pharyngitis (Strep Throat)*. 2018. URL: https://www.cdc.gov/groupastrep/diseases-hcp/strep-throat.html (Accessed 20 May, 2019).
91. Banerjee S, Ford C. Rapid Tests for the Diagnosis of Group A Streptococcal Infection: A Review of Diagnostic Test Accuracy, Clinical Utility, Safety, and Cost-Effectiveness. *Canadian Agency for Drugs and Technologies in Health CADTH Rapid Response Reports* 2018;**05**:31.
92. Kose E, Sirin Kose S, Akca D, Yildiz K, Elmas C, Baris M*, et al.* The Effect of Rapid Antigen Detection Test on Antibiotic Prescription Decision of Clinicians and Reducing Antibiotic Costs in Children with Acute Pharyngitis. *J Trop Pediatr* 2016;**62**:308-15.
93. Benjamin JT. The costs of testing for streptococcal pharyngitis in the office laboratory. *Arch Pediatr Adolesc Med* 2000;**154**:93-4.
94. Tsevat J, Kotagal UR. Management of sore throats in children: a cost-effectiveness analysis. *Arch Pediatr Adolesc Med* 1999;**153**:681-8.
95. Boyler PA, Humair J, Revaz SA, Stalder H. A cost-effectiveness analysis of recommended strategies for acute pharyngitis. *J Gen Intern Med* 2002;**17**:135-6.
96. Ehrlich JE, Demopoulos BP, Daniel KR, Jr., Ricarte MC, Glied S. Cost-effectiveness of treatment options for prevention of rheumatic heart disease from Group A streptococcal pharyngitis in a pediatric population. *Prev Med* 2002;**35**:250-7.
97. Giraldez-Garcia C, Rubio B, Gallegos-Braun JF, Imaz I, Gonzalez-Enriquez J, Sarria-Santamera A. Diagnosis and management of acute pharyngitis in a paediatric population: a cost-effectiveness analysis. *Eur J Pediatr* 2011;**170**:1059-67.
98. Klepser D, Grismer SE, Klepser ME. Cost-effectiveness of pharmacist provided care for the treatment of adult pharyngitis. *J Manag Care Pharm* 2011;**17**:241 [meeting abstract].
99. Komaroff AL, Pass TM, Pappius EM. A cost-effectiveness analysis of alternate strategies for management of sore throat. *Clin Res* 1983;**31**:A299 [meeting abstract].
100. Lathia N, Sullivan K, Tam K, Brna M, MacNeil P, Saltmarche D*, et al.* Cost-minimization analysis of community pharmacy-based point-of-care testing for strep throat in 5 Canadian provinces. *Can Pharm J (Ott)* 2018;**151**:322-31.
101. Maizia A, Letrilliart L, Colin C. [Diagnostic strategies for acute tonsillitis in France: a cost-effectiveness study]. *Presse Med* 2012;**41**:e195-203.
102. Malecki M, Mazur A, Sobolewski M, Binkowska-Bury M, Marc M, Januszewicz P. Rapid strip tests as a decision-making tool about antibiotic treatment in children - A prospective study. *Pediatr Pol* 2017;**92**:149-55.
103. Meier FA, Howland J, Johnson J, Poisson R. Effects of a rapid antigen test for group A streptococcal pharyngitis on physician prescribing and antibiotic costs. *Arch Intern Med* 1990;**150**:1696-700.
104. Mlejnek JR, Almulhem K, Spadafore S. Utility and cost effectiveness of throat culture in the treatment of patients with negative rapid strep screens. *Acad Emerg Med* 2014;**21**:S51 [meeting abstract].
105. Polisena J, Spry C. *Point of care testing for streptococcal sore throat: a review of diagnostic accuracy, cost-effectiveness, and guidelines*. Ottawa: Canadian Agency for Drugs and Technologies in Health (CADTH); 2009. URL: http://www.cadth.ca/media/pdf/htis/L0088%20Point%20of%20Care%20Testing%20for%20Streptococcal%20Sore%20Throat%20final.pdf (Accessed 23 January 2019).

Appendix

Appendix 1: Record of searches - Clinical Effectiveness

Bibliographic databases

Summary of bibliographic database searches

Database	Date of search	Number of records (+ number from update search)
MEDLINE (Ovid)	26/11/2018 (updated 07/03/2019)	1646 (+33)
Embase (Ovid)	27/11/2018 (updated 12/03/2019)	2546 (+177)
Cochrane Library	29/11/2018 (updated 12/03/2019)	118 (+1)
Science Citation Index and Conference Proceedings Science (Web of Science)	03/12/2018 (updated 12/03/2019)	1275 (+67)
DARE	22/01/2019 (updated 12/03/2019)	30 (+0)
HTA	22/01/2019 (updated 12/03/2019)	2 (+0)

Total from database searches: 5617 (+278 from 2019 update search) = 5895

Total after deduplication: 3240 (+45 from 2019 update search) = 3285

MEDLINE (Ovid)

Searched on 26/11/2018 (updated on 07/03/2019, see at the end of this search record)

Databases: Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions(R) <1946 to November 21, 2018>

Original search, 26/11/2018

Search Strategy:

---

1 exp Pharyngitis/ (15049)

2 pharyngit*.ti,ab,kf. (5455)

3 (nasophyryngit* or rhinopharyngit* or epipharyngit*).ti,ab,kf. (177)

4 (tonsillit* or tonsilit*).ti,ab,kf. (5589)

5 ((sore or pain* or ache* or aching or inflam* or infect*) adj3 (pharyn* or throat* or tonsil* or nasopharyn* or rhinopharyn* or epipharyn*)).ti,ab,kf. (9903)

6 or/1-5 (25137)

7 Streptococcal Infections/di, mi (13347)

8 Streptococcus pyogenes/im, ip (5444)

9 7 or 8 (16609)

10 ((strep or streptococcal or group) adj2 A).ti,ab,kf. (558959)

11 9 and 10 (4831)

12 (strep* adj5 (throat* or pharyn* or tonsil*)).ti,ab,kf. (3397)

13 streptoco* A.ti,ab,kf. (475)

14 (group A adj5 streptoco*).ti,ab,kf. (9481)

15 ((streptococcus or strep) adj1 (pyogenes or pyogenic)).ti,ab,kf. (7683)

16 ((streptococcus or strep) adj1 (epidemicus or erysipelatis or erysipelatos or hemolyticus or haemolyticus or scarlatinae or lancefield)).ti,ab,kf. (237)

17 (s pyogenes or pyogenes s or micrococcus scarlatinae).ti,ab,kf. (2485)

18 lancefield group.ti,ab,kf. (475)

19 gabhs.ti,ab,kf. (392)

20 or/11-19 (18796)

21 Point-of-Care Systems/ (11122)

22 exp Reagent Kits, Diagnostic/ (19326)

23 Antigens, Bacterial/an (7619)

24 (point-of-care or poc or poct or pocts).ti,ab,kf. (17665)

25 ((rapid* or bedside*1 or bed-side*1 or near-patient or nearpatient or extra-laboratory or extralaboratory or office*1) adj6 (test or tests or testing or tested or detect* or diagnos* or screen* or kit or kits or assay* or immunoassay* or determin* or identif* or antigen*1)).ti,ab,kf. (136637)

26 (radt or radts or rdt or rdts).ti,ab,kf. (1813)

27 (antigen*1 adj6 (test or tests or testing or tested or detect* or diagnos* or screen* or kit or kits or assay* or immunoassay* or determin* or identif*)).ti,ab,kf. (100724)

28 (clearview exact* or BD veritor* or strep A rapid test* or quikread go* or alere i* or cobas liat* or genexpert* or ((alere* or testpack* or test-pack* or bionexia* or bio-nexia* or biosynex* or veritor* or cobas* or quikread* or quik-read* or NADAL* or OSOM* or sofia* or xpert*) and (strep A or point of care or point-of-care or POC))).ti,ab,kf. (804)

29 ((abbott or beckton dickinson or biopanda or nal von minden or sekisui or orion diagnostica or roche or cepheid or biomerieux or quidel) and (strep A or point of care or POC or rapid test* or rapid antigen or antigen test*)).ti,ab,kf,in. (618)

30 or/21-29 (269698)

31 (6 or 20) and 30 (1759)

32 exp animals/ not humans/ (4517568)

33 31 not 32 (1646)

34 31 use medp,prem,mesx (114)

35 33 or 34 (1646)

Updated search, 07/03/2019

Re-ran above search with following date limits:

36 limit 35 to ed=20181126-20190307 (17)

37 limit 35 to ep=20181126-20190307 (14)

38 (2018 11* or 2018 12* or 2019*).dt,ez. (243915)

39 35 and 38 (13)

40 36 or 37 or 39 (33)

Total after removing duplicates with previous search: 16

Embase (Ovid)

Searched on 27/11/2018 (updated on 12/03/2019, see at the end of this search record)

Database: Embase Classic+Embase <1947 to 2018 November 21>

Original search, 27/11/2018

Search Strategy:

---

1 streptococcal pharyngitis/ or pharyngitis/ or rhinopharyngitis/ or sore throat/ or tonsillitis/ or chronic tonsillitis/ or palatine tonsillitis/ (51206)

2 pharyngit*.ti,ab,kw. (7851)

3 (nasophyryngit* or rhinopharyngit* or epipharyngit*).ti,ab,kw. (379)

4 (tonsillit* or tonsilit*).ti,ab,kw. (8320)

5 ((sore or pain* or ache* or aching or inflam* or infect*) adj3 (pharyn* or throat* or tonsil* or nasopharyn* or rhinopharyn* or epipharyn*)).ti,ab,kw. (15900)

6 or/1-5 (59836)

7 Streptococcus infection/di (3821)

8 Streptococcus pyogenes/ or streptococcus group a/ or group A streptococcal infection/ (23921)

9 7 or 8 (26865)

10 ((strep or streptococcal or group) adj2 A).ti,ab,kw. (792961)

11 9 and 10 (9617)

12 (strep* adj5 (throat* or pharyn* or tonsil*)).ti,ab,kw. (4842)

13 streptoco* A.ti,ab,kw. (636)

14 (group A adj5 streptoco*).ti,ab,kw. (12213)

15 ((streptococcus or strep) adj1 (pyogenes or pyogenic)).ti,ab,kw. (9259)

16 ((streptococcus or strep) adj1 (epidemicus or erysipelatis or erysipelatos or hemolyticus or haemolyticus or scarlatinae or lancefield)).ti,ab,kw. (388)

17 (s pyogenes or pyogenes s or micrococcus scarlatinae).ti,ab,kw. (3223)

18 lancefield group.ti,ab,kw. (567)

19 gabhs.ti,ab,kw. (504)

20 or/11-19 (24055)

21 point of care system/ or point of care testing/ (11966)

22 rapid test/ or diagnostic kit/ (8892)

23 antigen detection/ or bacterial antigen/an or Streptococcus antigen/ (24501)

24 (point-of-care or poc or poct or pocts).ti,ab,kw. (25553)

25 ((rapid* or bedside*1 or bed-side*1 or near-patient or nearpatient or extra-laboratory or extralaboratory or office*1) adj6 (test or tests or testing or tested or detect* or diagnos* or screen* or kit or kits or assay* or immunoassay* or determin* or identif* or antigen*1)).ti,ab,kw. (177813)

26 (radt or radts or rdt or rdts).ti,ab,kw. (2974)

27 (antigen*1 adj6 (test or tests or testing or tested or detect* or diagnos* or screen* or kit or kits or assay* or immunoassay* or determin* or identif*)).ti,ab,kw. (130835)

28 (clearview exact* or BD veritor* or strep A rapid test* or quikread go* or alere i* or cobas liat* or genexpert* or ((alere* or testpack* or test-pack* or bionexia* or bio-nexia* or biosynex* or veritor* or cobas* or quikread* or quik-read* or NADAL* or OSOM* or sofia* or xpert*) and (strep A or point of care or point-of-care or POC))).ti,ab,kw. (1633)

29 ((abbott or beckton dickinson or biopanda or nal von minden or sekisui or orion diagnostica or roche or cepheid or biomerieux or quidel) and (strep A or point of care or POC or rapid test* or rapid antigen or antigen test*)).ti,ab,kw,in. (1404)

30 or/21-29 (345022)

31 (6 or 20) and 30 (2856)

32 (exp animal/ or nonhuman/) not exp human/ (6749742)

33 31 not 32 (2546)

Updated search, 12/03/2019

Re-ran above search with following date limits:

34 limit 33 to dd=20181127-20190312 (18)

35 limit 33 to em=201811-201903 (152)

36 34 or 35 (159)

37 limit 33 to dc=20181127-20190312 (41)

38 36 or 37 (177)

Total after removing duplicates with other update and previous searches: 25

Cochrane Library (including Cochrane Database of Systematic Reviews, CENTRAL)

Searched on 29/11/2018 (updated on 12/03/2019, see at the end of this search record)

Original search, 29/11/2018

ID Search Hits

#1 MeSH descriptor: [Pharyngitis] explode all trees 1138

#2 pharyngit*:ti,ab,kw 1916

#3 (nasophyryngit* or rhinopharyngit* or epipharyngit*):ti,ab,kw 2597

#4 (tonsillit* or tonsilit*):ti,ab,kw 826

#5 ((sore or pain* or ache* or aching or inflam* or infect*) near/3 (pharyn* or throat* or tonsil* or nasopharyn* or rhinopharyn* or epipharyn*)):ti,ab,kw 3198

#6 #1 or #2 or #3 or #4 or #5 7030

#7 MeSH descriptor: [Streptococcal Infections] explode all trees and with qualifier(s): [diagnosis - DI, microbiology - MI] 306

#8 MeSH descriptor: [Streptococcus pyogenes] explode all trees and with qualifier(s): [immunology - IM, isolation & purification - IP] 89

#9 #7 or #8 351

#10 ((strep or streptococcal or group) near/2 A):ti,ab,kw 109570

#11 #9 and #10 126

#12 (strep* near/5 (throat* or pharyn* or tonsil*)):ti,ab,kw 499

#13 streptoco* next A:ti,ab,kw 26

#14 (group A near/5 streptoco*):ti,ab,kw 689

#15 ((streptococcus or strep) near/1 (pyogenes or pyogenic)):ti,ab,kw 423

#16 ((streptococcus or strep) near/1 (epidemicus or erysipelatis or erysipelatos or hemolyticus or haemolyticus or scarlatinae or lancefield)):ti,ab,kw 1

#17 ("s pyogenes" or "pyogenes s" or "micrococcus scarlatinae"):ti,ab,kw 60

#18 "lancefield group":ti,ab,kw 6

#19 gabhs:ti,ab,kw 109

#20 #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 1123

#21 MeSH descriptor: [Point-of-Care Systems] explode all trees 424

#22 MeSH descriptor: [Reagent Kits, Diagnostic] explode all trees 267

#23 MeSH descriptor: [Antigens, Bacterial] explode all trees and with qualifier(s): [analysis - AN] 63

#24 (point-of-care or poc or poct or pocts):ti,ab,kw 2560

#25 ((rapid* or bedside*1 or bed-side*1 or near-patient or nearpatient or extra-laboratory or extralaboratory or office*1) near/6 (test or tests or testing or tested or detect* or diagnos* or screen* or kit or kits or assay* or immunoassay* or determin* or identif* or antigen*1)):ti,ab,kw 3506

#26 (radt or radts or rdt or rdts):ti,ab,kw 302

#27 (antigen*1 near/6 (test or tests or testing or tested or detect* or diagnos* or screen* or kit or kits or assay* or immunoassay* or determin* or identif*)):ti,ab,kw 0

#28 (clearview next exact* or BD next veritor* or "strep A rapid" next test* or quikread next go* or alere next i* or cobas next liat* or genexpert* or ((alere* or testpack* or test-pack* or bionexia* or bio-nexia* or biosynex* or veritor* or cobas* or quikread* or quik-read* or NADAL* or OSOM* or sofia* or xpert*) and ("strep A" or "point of care" or point-of-care or POC))):ti,ab,kw 114

#29 ((abbott or "beckton dickinson" or biopanda or "nal von minden" or sekisui or "orion diagnostica" or roche or cepheid or biomerieux or quidel) and ("strep A" or "point of care" or POC or rapid next test* or rapid next antigen* or antigen next test*)):ti,ab,kw 47

#30 #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 6235

#31 (#6 or #20) and #30 118

Total: 118

CDSR – Reviews: 15

CDSR – Protocols: 1

CENTRAL: 102

Updated search, 12/03/2019

Re-ran above search and sorted by date, newest first.

New since 29/11/2018:

CDSR – Reviews: 0

CDSR – Protocols: 0

CENTRAL: 1

Total after removing duplicates with other update and previous searches: 0

Science Citation Index and Conference Proceedings (Web of Science)

Searched on 03/12/2018 (updated on 12/03/2019, see at the end of this search record)

Original search, 03/12/2018

Note, search record reads from bottom to top

Set	Results	History
# 23	1,275	(#5 OR #14) AND #22 *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 22	265,727	#15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 21	487	TS=((abbott OR "beckton dickinson" OR biopanda OR "nal von minden" OR sekisui OR "orion diagnostica" OR roche OR cepheid OR biomerieux OR quidel) AND ("strep A" OR "point* of care" OR poc OR poct OR pocts OR "rapid test*" OR "rapid antigen" OR "antigen test*")) *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 20	849	TS=("clearview exact*" OR "BD veritor*" OR "strep A rapid test*" OR "quikread go*" OR "alere i*" OR "cobas liat*" OR genexpert* OR ((alere* OR testpack* OR test-pack* OR bionexia* OR bio-nexia* OR biosynex* OR veritor* OR cobas* OR quikread* OR quik-read* OR NADAL* OR OSOM* OR sofia* OR xpert*) AND ("strep A" OR "point* of care" OR poc OR poct OR pocts))) *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 19	86,024	TS=(antigen* NEAR/5 (test OR tests OR testing OR tested OR detect* OR diagnos* OR screen* OR kit OR kits OR assay* OR immunoassay* OR determin* OR identif*)) *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 18	2,261	TS=(radt OR radts OR rdt OR rdts) *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 17	165,166	TS=((rapid* OR bedside* OR bed-side* OR near-patient OR nearpatient OR extra-laboratory OR extralaboratory OR office*) NEAR/5 (test OR tests OR testing OR tested OR detect* OR diagnos* OR screen* OR kit OR kits OR assay* OR immunoassay* OR determin* OR identif* OR antigen*)) *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 16	22,883	TS=("point* of care" OR poc OR poct OR pocts)

		Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 15	219	TS=(diagnostic AND (reagent NEAR/0 (kit* OR strip*))) *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 14	17,280	#6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 13	308	TS=gabhs *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 12	444	TS="lancefield group" *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 11	2,042	TS=("s pyogenes" OR "pyogenes s" OR "micrococcus scarlatinae") *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 10	59	TS=((strep*) NEAR/0 (epidemicus OR erysipelatis OR erysipelatos OR hemolyticus OR haemolyticus OR scarlatinae OR lancefield)) *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 9	7,107	TS=((strep*) NEAR/0 (pyogenes OR pyogenic)) *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 8	9,638	TS=("group A" NEAR/4 strep*) *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 7	1,156	TS="strep* A" *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 6	2,875	TS=(strep* NEAR/4 (throat* OR pharyn* OR tonsil*)) *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 5	12,426	#1 OR #2 OR #3 OR #4 *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 4	6,980	TS=((sore OR pain* OR ache* OR aching OR inflam* OR infect*) NEAR/2 (pharyn* OR throat* OR tonsil* OR nasopharyn* OR rhinopharyn* OR epipharyn*)) *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 3	2,703	TS=(tonsillit* OR tonsilit*) *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 2	96	TS=(nasophyryngit* OR rhinopharyngit* OR epipharyngit*) *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*
# 1	4,651	TS=pharyngit* *Indexes=SCI-EXPANDED, CPCI-S Timespan=All years*

Updated search, 12/03/2019

Re-ran above search with following date limits:

# 23	67	(#5 OR #14) AND #22 *Indexes=SCI-EXPANDED, CPCI-S Timespan=2018-2019*

Total after removing duplicates with other update and previous searches: 4

DARE (CRD) and HTA Database (CRD)

Searched on 22/01/2019 (Not updated because no new records have been added to DARE since 31st March 2015 or to the HTA database since 31 March 2018. The INAHTA website was checked in March 2019 to see if a new platform for the HTA database was available).

Original search, 22/01/2019

1 MeSH DESCRIPTOR Pharyngitis EXPLODE ALL TREES IN DARE,NHSEED,HTA 73

2 (pharyngit*) 85

3 (nasophyryngit*) OR (rhinopharyngit*) OR (epipharyngit*) 5

4 (tonsillit* or tonsilit*) 43

5 (((sore or pain* or ache* or aching or inflam* or infect*) adj3 (pharyn* or throat* or tonsil* or nasopharyn* or rhinopharyn* or epipharyn*))) 91

6 #1 OR #2 OR #3 OR #4 OR #5 163

7 MeSH DESCRIPTOR Streptococcal Infections WITH QUALIFIERS DI, MI IN DARE,NHSEED,HTA 31

8 MeSH DESCRIPTOR Streptococcus pyogenes WITH QUALIFIERS IM, IP IN DARE,NHSEED,HTA 13

9 #7 OR #8 36

10 (((strep or streptococcal or group) adj2 A)) 2025

11 #9 AND #10 17

12 ((strep* adj5 (throat* or pharyn* or tonsil*))) 39

13 (streptoco* adj1 A) 10

14 ((group A adj5 streptoco*)) 27

15 (((streptococcus or strep or staphylococcus) adj1 (pyogenes or pyogenic))) 25

16 (((streptococcus or strep) adj1 (epidemicus or erysipelatis or erysipelatos or hemolyticus or haemolyticus or scarlatinae or lancefield))) 0

17 ((s pyogenes or pyogenes s or micrococcus scarlatinae)) 1

18 (lancefield group) 0

19 (gabhs) 8

20 #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 51

21 #6 AND #20 43

22 (#21) IN DARE 30

22 (#21) IN HTA 2

PROSPERO (International Prospective Register of Systematic Reviews)

Searched on 20/02/2019

#1 MeSH DESCRIPTOR Pharyngitis EXPLODE ALL TREES (29)

#2 pharyngit* (48)

#3 nasophyryngit* OR rhinopharyngit* OR epipharyngit* (3)

#4 tonsillit* OR tonsilit* (35)

#5 (sore OR pain* OR ache* OR aching OR inflam* OR infect*) ADJ3 (pharyn* or throat* OR tonsil* OR nasopharyn* OR rhinopharyn* OR epipharyn*) (105)

#6 #1 OR #2 OR #3 OR #4 OR #5 (125)

#7 strep* ADJ5 (throat* or pharyn* or tonsil*) (8)

#8 #6 OR #7 (125)

Status of Review: Completed or Published (17)

Browsed online by information specialist, none relevant

Trials registers

ClinicalTrials.gov,

Searched on 20/02/2019

33 Studies found for: Active, not recruiting, Completed, Suspended, Terminated, Withdrawn, Unknown status Studies | "strep throat" OR (strep OR streptococcus OR streptococcal OR "group a" OR gabhs) AND (throat OR pharynx OR tonsils) OR pharyngitis OR rhinopharyngitis OR epipharyngitis OR tonsillitis OR tonsilitis OR "sore throat" | rapid OR antigen OR radt OR radts OR rdt OR rdts OR "point of care" OR poc OR poct OR pocts OR bedside OR bed-side OR near-patient OR nearpatient OR diagnostic OR diagnosis OR test OR tests OR testing OR kit OR kits OR clearview OR veritor OR quikread OR quik-read OR alere OR cobas OR genexpert OR testpack OR test-pack OR bionexia OR bio-nexia OR biosynex OR nadal OR osom OR sofia OR xpert OR abbott OR "beckton dickinson" OR biopanda OR "nal von minden" OR sekisui OR "orion diagnostica" OR roche OR cepheid OR biomerieux OR quidel

Downloaded to Excel and screened by information specialist against inclusion criteria and with reference to included studies from database searches. No new studies identified.

Conferences and professional organisations

Selected with advice from several advisors (Noel McCarthy and NICE specialist committee members)

Federation of Infection Societies conference

Searched 06/03/2019

2019 November (not available yet)

2018 https://fis2018.co.uk/ (browsed Abstracts > Diagnostics) – 0 relevant

2017 http://event.federationinfectionsocieties.com/ (browsed Abstracts) – 0 relevant

2016 https://www.journalofhospitalinfection.com/issue/S0195-6701(16)X0012-6 (searched Abstracts (Poster and Oral presentations and Invited Speaker Abstracts) one term at a time. Terms used: strep or group a or throat or pharyn or tonsil (search looked for these within words as well as whole words)) – 0 relevant

2015 – Abstracts appear not to be available online

2014 (searched Abstracts one term at a time. Terms used: strep or group a or throat or pharyn or tonsil (search looked for these within words as well as whole words)) – 0 relevant

The European Congress of Clinical Microbiology and Infectious (ECCMID)

Searched 05/03/2019

https://www.eccmid.org/

Embase indexes up to 22nd European Congress of Clinical Microbiology and Infectious Diseases. London United Kingdom. 2012.

Older and more recent years available via ESCMID eLibrary
https://www.escmid.org/escmid_publications/escmid_elibrary/

Searched ESCMID eLibrary on 05/03/2019 for the following terms, with no date limit:

strep – 25 results (3 sent to reviewers)

"group a" AND streptococcus, limited to 'Topics: Diagnostic Bacteriology & General Microbiology' – 10 (1 sent to reviewers)

"group a" AND streptococcal, limited to 'Topics: Diagnostic Bacteriology & General Microbiology' – 8 (1 sent to reviewers)

American Society of Microbiology

ASM Microbes

https://www.asm.org/ (website restructured, past meeting abstracts unavailable)

British Society for Antimicrobial Chemotherapy

Searched 06/03/2019

http://www.bsac.org.uk

BSAC Spring meeting abstracts 2016 – 2018 and general website search one term at a time. Terms used: strep or group a or throat or pharyn or tonsil (search looked for these within words as well as whole words)) – Screened online, none relevant

British Infection Association

Searched 06/03/2019

https://www.britishinfection.org/

None relevant

n.b. BSAC Spring conference - Thursday 21st - Friday 22nd March

Public Health England Annual Conference and Public Health Research and Science Annual Conference

Searched 12/03/2019

2016 - 2018 https://phe.multilearning.com/phe/#!*menu=6*browseby=3*sortby=2
Searched one term at a time. Terms used: strep or streptococcal or streptococcus of group a or throat or pharyngitis or pharynx or tonsillitis or tonsilitis (search looked for whole words) – Screened online, none relevant

Streptococcal biology conference

Searched 12/03/2019

https://www.grc.org/streptococcal-biology-conference/2018/

Searched one term at a time. Terms used: throat or pharyn or tonsil or rapid or point or diagnos (search looked for these within words as well as whole words)) – 0

Lancefield International Symposium on Streptococci and Streptococcal Diseases

Searched 12/03/2019

Not indexed in Embase. Some abstracts indexed in Web of Science, but only up to 2009

2017 http://lisssd2017.org/abstracts/ website unavailable

Not able to find a list of full abstracts for most recent five years online

Microbiology Society Conference

Searched 12/03/2019

2019 (abstract book for April 2019 available) https://microbiologysociety.org/event/annual-conference/annual-conference.html

2018 https://microbiologysociety.org/event/annual-conference/annual-conference-2018.html#tab-2

2017 https://microbiologysociety.org/events/annual-conference.html?eventYear=2017

2016 https://microbiologysociety.org/event/annual-conference/annual-conference-2016.html

2015 https://microbiologysociety.org/event/annual-conference/annual-conference-2015.html

Searched one term at a time. Terms used: streptococcal or streptococcus of group a or throat or pharyn or tonsil (search looked for these within words as well as whole words)) – Screened online, none relevant

Association of Clinical Biochemistry and Laboratory Medicine

Searched 12/03/2019

2018 and 2019 searched

http://www.acb.org.uk/whatwedo/events/national_meetings.aspx

http://www.acb.org.uk/whatwedo/events/national_meetings/focus-2018/abstracts/posterabstracts

Searched one term at a time. Terms used: streptococcal or streptococcus of group a or throat or pharyn or tonsil (search looked for these within words as well as whole words)) – Screened online, none relevant.

Also searched website using Google Advanced Search with http://www.acb.org.uk in the domain

strep* OR throat OR tonsil* OR pharyn* OR "group a" site:www.acb.org.uk: 7, screened online. None relevant

Royal College of Pathologists

Searched 12/03/2019

Searched website using Google Advanced Search with http://www.acb.org.uk in the domain

strep* OR throat OR tonsil* OR pharyn* OR "group a" site:www.rcpath.org: 55, screened online. None relevant

Included studies in relevant reviews

Reviews found in searches

Vachhani R, Patel T, Centor RM, Estrada CA. Sensitivity for Diagnosing Group A Streptococcal Pharyngitis from Manufacturers is 10% Higher than Reported in Peer-Reviewed Publications. Southern Medical Journal 2017;110(1):59-64.

Vachhani et al., 2017 focusses on manufacturers' package inserts. Refers to the following systematic reviews in the background:

- Lean WL, Arnup S, Danchin M, Steer AC. Rapid diagnostic tests for group A streptococcal pharyngitis: a meta-analysis. Pediatrics 2014;134(4):771-81. *Cross-checked 14 articles that mention a test name within our scope, out of the 48 total included studies. All 14 have already been picked up and sifted.*

- Stewart EH, Davis B, Clemans-Taylor BL, Littenberg B, Estrada CA, Centor RM. Rapid antigen group A streptococcus test to diagnose pharyngitis: a systematic review and meta-analysis. PLoS ONE [Electronic Resource] 2014;9(11):e111727. *Cross-checked 58 included studies with database search results. 57 of the 58 have already been picked up and sifted. The one remaining (Parviainen M, Koskela M, Ikäheimo I, Kelo E, Sirola H, et al. (2011) A novel strep A test for a rapid test reader compared with standard culture method and a commercial antigen assay. Eur Infect Dis 5: 143–145.) includes tests not within scope (ReaScan Strep A test (Reagena International Ltd, Toivala, Finland) vs standard culture vs TestPack® Strep A test (Inverness Medical, Cranfield, UK)).*
- Ruiz-Aragon J, Rodriguez Lopez R, Molina Linde JM. [Evaluation of rapid methods for detecting Streptococcus pyogenes. Systematic review and meta-analysis]. Anales de Pediatria 2010;72(6):391-402. *(Not in English and older (published 2010))*

Cohen JF, Bertille N, Cohen R, Chalumeau M. Rapid antigen detection test for group A streptococcus in children with pharyngitis. Cochrane Database of Systematic Reviews 2016;7:CD010502. *Included studies scanned for test names in scope and cross-checked against results of database searches*

Found some that we excluded due to not having the test name in the meeting abstract, where Cochrane reviewers contacted authors and were given more information. Sent to reviewers:

Mlejnek 2014

Go to characteristics of included studies:

https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD010502.pub2/references#characteristicStudies

And search using ctrl+f for: Mlejnek 2014

Pauchard 2013

Pauchard JY, Verga ME, Bersier J, Durusell C, Gehri M, Vaudaux B. Performance of rapid antigen detection test in group A β-haemolytic streptococcal pharyngitis in comparison with three clinical decision rule in a tertiary paediatric emergency department. Swiss Medical Weekly 2013;143(Suppl 197):6S.

Go to the same link above and search using ctrl+f for 'Pauchard 2013'

Pauchard 2012

Pauchard JY, Verga ME, Bersier J, Prod'Hom G, Gehri M, Vaudaux B. Performance of rapid antigen diagnostic test for group A β-haemolytic streptococcal pharyngitis in a tertiary paediatric emergency department. Swiss Medical Weekly 2012;142(Suppl 192):35S.

Go to the same link above and search using ctrl+f for 'Pauchard 2012'.

Not found in our searches:

Schwartz RH. Evaluation of rapid streptococcal detection tests. Pediatric Infectious Disease Journal 1997;16(11):1099-100. (OSOM Strep A (Wyntek))

Letter

Schwartz 1997a

Sedki M, Salama H, Salama E, Abdalla N, Ezz H. Rapid diagnostic test for streptococcal throat infection in Egyptian children. Medical Journal of Cairo University 2010;78(2):177-82. *Checked full-text - not a test in our scope (Streptatest ®, Dektra Pharm, Strasbourg, France)*

Sedki 2010

Also checked:

Banerjee 2018 CADTH rapid response review "Rapid Tests for the Diagnosis of Group A Streptococcal Infection: A Review of Diagnostic Test Accuracy, Clinical Utility, Safety, and Cost-Effectiveness"

References of our included studies

Searched in March 2019 for studies.

Manufacturers' websites

Searched in January 2019 for studies and data.

Rachel also checked manufacturers' submissions for mention of studies. Forwarded relevant details, abstracts, posters and package inserts to reviewers.

Regulatory bodies

FDA, CLIA - Clinical Laboratory Improvement Amendments database

Targeted searches undertaken on 06/03/2019

Test System / Manufacturer: NADAL. None found

Test System / Manufacturer: nal von minden. None found

Test System / Manufacturer: Cepheid AND Analyte Name: Streptococcus, group A. Two found, most relevant one sent to reviewers.

Test System / Manufacturer: Cobas AND Analyte Name: Streptococcus, group A. Two found, earliest one sent to reviewers.

Test System / Manufacturer: Biopanda. None found

Test System / Manufacturer: Biomerieux AND Analyte Name: Streptococcus, group A. None found

Test System / Manufacturer: Bionexia AND Analyte Name: Streptococcus, group A. None found

Test System / Manufacturer: Alere i AND Analyte Name: Streptococcus, group A One found (Alere i) – sent to reviewers.

Test System / Manufacturer: Abbott AND Analyte Name: Streptococcus, group A. One found (TestPack Plus) – sent to reviewers.

Test System / Manufacturer: Clearview AND Analyte Name: Streptococcus, group A. None found for Abbott

Test System / Manufacturer: BD Veritor AND Analyte Name: Streptococcus, group A. One found – sent to reviewers.

Test System / Manufacturer: OSOM AND Analyte Name: Streptococcus, group A. Several found, but no details of studies in summaries/statements.

Test System / Manufacturer: Orion Diagnostica AND Analyte Name: Streptococcus, group A. None found for QuikRead Go

Test System / Manufacturer: QuikRead Go AND Analyte Name: Streptococcus, group A. None found

Test System / Manufacturer: Biosynex. None found

Test System / Manufacturer: Quidel AND Analyte Name: Streptococcus, group A. several found, earliest one sent to reviewers.

European commission medical devices

Checked on 06/03/2019 N.b. Eudamed database not yet publically available.

Health services research agencies

INAHTA

Searched on 08/03/19. HTA database also searched (see Bibliographic database searches above).

Google

Searched on 27/02/19

Targeted search for NADAL Strep A scan using various terms. No study data found.

Appendix 2: Data extraction form for primary studies

Name of first reviewer: *Name of second reviewer:*

Study details	
Study ID (Endnote ref)	
First author surname	
Year of publication	
Country	
Study design	
Study setting	
Number of centres	
Study duration	
Question type	
Test accuracy (1) Other outcomes only (2)	
Aim of the study	
Patient selection	
Inclusion criteria:	
Exclusion criteria:	

Baseline characteristics/Population	
Number at baseline	
Women (%)	
Age years, median (range)	
Number with index test results included in the final analysis	
Adults (n)	
Adults aged 15 -75 y, n (%)	
Adults aged > 75 y, n (%)	
Children (n)	
Children aged 5 - 14 y, n (%)	
Number with centor score ≥3	
Number with centore score <3	
Number with FeverPAIN score ≥4	
Number with FeverPAIN score 2-3	
Number with FeverPAIN score <2	
Number with McIsaac score > 2	
Co-morbidities	
Recent Abx Tx prior to study enrolment, n(%)	

Strep A prevalence in entire or target population (%)	

Index test and comparator for Strep A	
Name of index test	
Proportion of index test results validated with microbiological culture, n (%)	
Comparison with Centor score (Yes or No)	
Proportion with a centor score, n (%)	
Comparison with FeverPAIN score (Yes or No)	
Proportion with a FeverPAIN score, n (%)	
Comparison with McIsaac score (Yes or No)	
Proportion with McIsaac score, n(%)	
Reference standard	
Notes / Comments:	

Diagnostic accuracy of clinical score	
Clinical score groupings	
True positives	
False positives	
True negatives	
False negatives	
Total number	
Sensitivity (%) (95% CI	
Specificity (%) (95% CI)	
PPV (%) (95% CI)	
NPV (%) (95% CI)	

Diagnostic accuracy of point-of-care test	
True positives	
False positives	
True negatives	
False negatives	
Total number	
Sensitivity (%) (95% CI	
Specificity (%) (95% CI)	
PPV (%) (95% CI)	
NPV (%) (95% CI)	

Diagnostic accuracy of point-of-care test with PCR adjudication	
True positives	
False positives	
True negatives	
False negatives	
Total number	
Sensitivity (%) (95% CI	
Specificity (%) (95% CI)	
PPV (%) (95% CI)	
NPV (%) (95% CI)	
Other outcomes	
Discordant results with standard microbiology tests	
Number of delayed or immediate Abx prescriptions used	
Contribution to antimicrobial stewardship	
Time to test results	
Test failure rate	
Time to Abx prescribing decision	
Morbidity	
Mortality	
HLQoL	
Patient or carer satisfaction with POC tests	
Healthcare professional satisfaction with test	
Any other comments:	

Appendix 3: Excluded studies with reasons

Reference	Reason for exclusion
1. (2016). " Multicenter evaluation of the solana group a streptococcus assay: comparison with culture." Journal of clinical microbiology 54(9): 2388-2390.	Wrong test
2. Abd El-Ghany, S.M., Abdelmaksoud, A.A., Saber, S.M., Abd El Hamid, D.H. (2015). "Group A beta-hemolytic streptococcal pharyngitis and carriage rate among Egyptian children: a case-control study." Annals of Saudi Medicine 35(5): 377-82.	Wrong test
3. Abu-Sabaah, A.H., Ghazi, H.O. (2006). " Better diagnosis and treatment of throat infections caused by group A beta-haemolytic streptococci." British Journal of Biomedical Science 63(4): 155-8.	Wrong test
4. Agarwal, M., Raghuwanshi, S.K., Asati, D.P. (2015). "Antibiotic Use in Sore Throat: Are We Judicious?" Indian Journal of Otolaryngology & Head & Neck Surgery 67(3): 267-70.	Wrong test
5. Alper, Z., Uncu, Y., Akalin, H., Ercan, I., Sinirtas, M., Bilgel, N.G (2013). " Diagnosis of acute tonsillopharyngitis in primary care: a new approach for low-resource settings." Journal of Chemotherapy 25(3): 148-55.	Wrong test
6. Al-Tawfiq, J.A., Alawami, A.H (2017). "A multifaceted approach to decrease inappropriate antibiotic use in a pediatric outpatient clinic." Annals of Thoracic Medicine 12(1):51-54.	No specific RADT mentioned
7. Amorim, R., Filho, A.F., Abath, A., Hatem, T., Mourato, F., Gomes, R., Mattos, S (2017). "Prevalence of positive rapid antigen group a streptococcus test in children and adolescents in a state from Northeast Brazil." Cardiology in the Young 27(4):s484.	Wrong test
8. Anderson, K. B., Simasathien, S., Watanaveeradej, V., Weg, A. L., Ellison, D. W., Suwanpakdee, D., Jarman, R. G. (2018). "Clinical and laboratory predictors of influenza infection among individuals with influenza-like illness presenting to an urban Thai hospital over a five-year period." PLoS ONE 13(3): e0193050.	Wrong test
9. Anderson, N. W., Buchan, B. W., Mayne, D., Mortensen, J. E., Mackey, T. L., Ledeboer, N. A. (2013). "Multicenter clinical evaluation of the illumigene group A Streptococcus DNA amplification assay for detection of group A Streptococcus from pharyngeal swabs." Journal of Clinical Microbiology 51(5): 1474-1477	Wrong test
10. Andre, M., Eriksson, M., Molstad, S., Stalsbylundborg, C., Jacobsson, A., Odenholt, I., Swedish Study Group on Antibiotic, U. (2005). "The management of infections in children in general practice in Sweden: a repeated 1-week diagnosis-prescribing study in 5 counties in 2000 and 2002." Scandinavian Journal of Infectious Diseases 37(11-12): 863-869.	No specific RADT mentioned
11. Andrews, D., Chetty, Y., Cooper, B. S., Virk, M., Glass, S. K., Letters, A., Jeyaratnam, D. (2017). "Multiplex PCR point of care testing versus routine, laboratory-based testing in the treatment of adults with respiratory tract infections: a quasi-randomised study assessing impact on length of stay and antimicrobial use." BMC Infectious Diseases 17(1).	Wrong test
12. Anonymous. (2003). Group A streptococcal pharyngitis: Diagnosis and management. Drug Benefit Trends, 15(12), 29-32.	Wrong test

13. Aoki, A., Ashizawa, T., Ebata, A., Nasu, Y., & Fujii, T. (2014). Group A Streptococcus pharyngitis outbreak among university students in a judo club. *Journal of Infection & Chemotherapy, 20*(3), 190-193. doi:https://dx.doi.org/10.1016/j.jiac.2013.10.004	Wrong test
14. Araujo Filho, B. C., Imamura, R., Sennes, L. U., Sakae, F. A. (2005). "Role of rapid antigen detection test for the diagnosis of group A beta-hemolytic streptococcus in patients with pharyngotonsillitis." Revista Brasileira de Otorrinolaringologia 71(2): 168-171.	Wrong test
15. Araujo Filho, B. C., Imamura, R., Sennes, L. U., & Sakae, F. A. (2006). "Role of rapid antigen detection test for the diagnosis of group-A beta-hemolytic streptococcus in patients with pharyngotonsillitis." Revista Brasileira de Otorrinolaringologia 72(1): 12-15.	Wrong test
16. Arbefeville, S., Nelson, K., Thonen-Kerr, E., & Ferrieri, P. (2018). "Prospective Postimplementation Study of Solana Group A Streptococcal Nucleic Acid Amplification Test vs Conventional Throat Culture." American Journal of Clinical Pathology 150(4): 333-337.	Wrong test
17. Armengol, C. E., Hendley, J. O., & Schlager, T. A. (2004). Could repetition of the rapid antigen detection test for group a streptococci on a second swab replace the backup throat culture? Pediatric Research, 55(4), 341A-341A.	Meeting abstract couldn't be located
18. Armengol, C. E., Schlager, T. A., Hendley, J. O. (2004). "Sensitivity of a rapid antigen detection test for group A streptococci in a private pediatric office setting: answering the Red Book's request for validation." Pediatrics 113(4): 924-926.	Wrong test
19. Atlas, S. J., McDermott, S. M., Mannone, C., Barry, M. J. (2005). "The role of point of care testing for patients with acute pharyngitis." Journal of General Internal Medicine 20(8): 759-761.	Wrong test
20. Ayanruoh, S., Waseem, M., Quee, F., Humphrey, A., & Reynolds, T. (2009). Impact of rapid streptococcal test on antibiotic use in a pediatric emergency department. Pediatric Emergency Care, 25(11), 748-750. doi:https://dx.doi.org/10.1097/PEC.0b013e3181bec88c	Wrong test
21. Balasubramanian, S., Amperayani, S., Dhanalakshmi, K., Senthilnathan, S., Chandramohan, V. (2018). "Rapid antigen diagnostic testing for the diagnosis of group A beta-haemolytic streptococci pharyngitis." National Medical Journal of India 31(1): 8-10.	Wrong test
22. Ba-Saddik, I. A., Munibari, A. A., Alhilali, A. M., Ismail, S. M., Murshed, F. M., Coulter, J. B., Cuevas, L.E., Hart, C.A., Brabin, B.J., Parry, C. M. (2014). "Prevalence of Group A beta-haemolytic Streptococcus isolated from children with acute pharyngotonsillitis in Aden, Yemen." Tropical Medicine & International Health 19(4): 431-439.	Wrong test
23. Bergmark, R., Bergmark, B., Blander, J., Fataki, M., Janabi, M. (2010). "Burden of disease and barriers to the diagnosis and treatment of group a beta-hemolytic streptococcal pharyngitis for the prevention of rheumatic heart disease in Dar Es Salaam, Tanzania." Pediatric Infectious Disease Journal 29(12): 1135-1137.	No specific RADT mentioned
24. Bjerrum, L., Cots, J. M., Llor, C., Molist, N., & Munck, A. (2006). Effect of intervention promoting a reduction in antibiotic prescribing by improvement of diagnostic procedures: A	No comparison with biological

	prospective, before and after study in general practice. European Journal of Clinical Pharmacology, 62(11), 913-918. doi:http://dx.doi.org/10.1007/s00228-006-0187-y	culture or clinical scores
25.	Brennan-Krohn, T., Ozonoff, A., Sandora, T. J. (2018). "Adherence to guidelines for testing and treatment of children with pharyngitis: a retrospective study." BMC Pediatrics,18(1): 43.	No specific RADT mentioned
26.	Briel, M., Young, J., Tschudi, P., Hersberger, K. E., Hugenschmidt, C., Langewitz, W., & Bucher, H. C. (2006). Prevalence and influence of diagnostic tests for acute respiratory tract infections in primary care. Swiss Medical Weekly, 136(15-16), 248-253.	Wrong population
27.	Brittain-Long, R., Westin, J., Olofsson, S., Lindh, M., Andersson, L. (2011). "Access to a polymerase chain reaction assay method targeting 13 respiratory viruses can reduce antibiotics: A randomised, controlled trial." BMC Medicine 9(44).	No specific RADT mentioned
28.	Brook, I., Gober, A. E. (2008). "Concurrent influenza A and group A beta-hemolytic streptococcal pharyngotonsillitis." Annals of Otology, Rhinology & Laryngology 117(4): 310-312.	Wrong test
29.	Bursle, E., Robson, J. (2016). "Non-culture methods for detecting infection." Australian Prescriber 39(5):171-175.	Review
30.	Camurdan, A. D., Camurdan, O. M., Ok, I., Sahin, F., Ilhan, M. N., Beyazova, U. (2008). "Diagnostic value of rapid antigen detection test for streptococcal pharyngitis in a pediatric population." International Journal of Pediatric Otorhinolaryngology 72(8): 1203-1206.	Wrong test
31.	Cao, C., Zhang, F., Ji, M., Pei, F., Fan, X., Shen, H., Wang, Q., Yang, W., Wang, Y. (2016). "Development of a loop-mediated isothermal amplification method for rapid detection of streptococcal pyrogenic exotoxin B." Toxicon 117: 53-58.	No specific RADT mentioned
32.	Cardoso, D. M., Gilio, A. E., Hsin, S. H., Machado, B. M., de Paulis, M., Lotufo, J. P., Martinez, M.B., Grisi, S. J. (2013). Impact of the rapid antigen detection test in diagnosis and treatment of acute pharyngotonsillitis in a pediatric emergency room. Revista Paulista de Pediatria, 31(1), 4-9.	No comparison with biological culture or clinical scores
33.	Chen, F. M. (2000). Culture confirmation of negative rapid strep test results. Journal of Family Practice, 49(4), 371-372.	Wrong test
34.	Cheng, C., Han, B., Smoot, B., Chen, Y., & Exner, M. M. (2010). Real-time PCR detection of group a streptococcus using the 3M integrated cycler. Journal of Molecular Diagnostics, 12 (6), 883.	Wrong test
35.	Cohen, J. F., Chalumeau, M., Levy, C., Bidet, P., Benani, M., Koskas, M., Bingen, E., Cohen, R. (2013). "Effect of clinical spectrum, inoculum size and physician characteristics on sensitivity of a rapid antigen detection test for group A streptococcal pharyngitis." European Journal of Clinical Microbiology & Infectious Diseases 32(6): 787-793.	Wrong test
36.	Cohen, J. F., Chalumeau, M., Levy, C., Bidet, P., Thollot, F., Wollner, A., Bingen, E., Cohen, R. (2012). "Spectrum and inoculum size effect of a rapid antigen detection test for group A streptococcus in children with pharyngitis." PLoS ONE 7(6): e39085.	Wrong test
37.	Cohen, J. F., Cohen, R., Bidet, P., Elbez, A., Levy, C., Bossuyt, P. M., Chalumeau, M. (2017). "Efficiency of a clinical prediction model for selective rapid testing in children with pharyngitis: A prospective, multicenter study." PLoS ONE 12(2): e0172871.	Wrong test

38. Cohen, J. F., Cohen, R., Bidet, P., Levy, C., Deberdt, P., d'Humieres, C., Ligouri, A., Corrard, F., Thollot, F., Mariani-Kurkdijan, P., Chalumeau, M., Bingen, E. (2013). "Rapid-antigen detection tests for group a streptococcal pharyngitis: revisiting false-positive results using polymerase chain reaction testing." Journal of Pediatrics 162(6): 1282-1284, 1284.e1281.	Wrong test
39. Cohen, R., Levy, C., Ovetchkine, P., Boucherat, M., Weil-Olivier, C., Gaudelus, J., de la Rocque, F., Bingen, E. (2004). "Evaluation of streptococcal clinical scores, rapid antigen detection tests and cultures for childhood pharyngitis." European Journal of Pediatrics 163(4-5): 281-282.	Wrong test
40. Dagnelie, C. F., Bartelink, M. L., van der Graaf, Y., Goessens, W., de Melker, R. A. (1998). "Towards a better diagnosis of throat infections (with group A beta-haemolytic streptococcus) in general practice." British Journal of General Practice 48(427): 959-962.	Wrong test
41. Dale, J. C., Novak, R., Higgens, P., Wahl, E. (2002). "Testing for group a streptococci." Archives of Pathology & Laboratory Medicine 126(12): 1467-1470.	No specific RADT mentioned
42. Dawson, E. D., Taylor, A. W., Smagala, J. A., Rowlen, K. L. (2009). "Molecular detection of Streptococcus pyogenes and Streptococcus dysgalactiae subsp. equisimilis." Molecular Biotechnology 42(1), 117-127.	Wrong test
43. Demore, B., Tebano, G., Gravoulet, J., Wilcke, C., Ruspini, E., Birge, J., Boivin, J.M., Henard, S., Dieterling, A., Munerol, L., Husson, J., Rabaud, C., Pulcini, C., Malblanc, S. (2018). "Rapid antigen test use for the management of group A streptococcal pharyngitis in community pharmacies." European Journal of Clinical Microbiology & Infectious Diseases 06: 06.	Wrong test
44. Deniz, R., Aktas, E., Baris, A., Bayraktar, B. (2018). "The use of rapid antigen testing and matrix-assisted laser desorption/ionization-time of flight mass spectrometry in the diagnosis of group A beta-hemolytic streptococci in throat swab samples." Turkish Journal of Medical Sciences 48(5): 939-944.	Wrong test
45. Dodd, M., Adolphe, A., Parada, A., Brett, M., Culbreath, K., Mercier, R. C. (2018). "Clinical Impact of a Rapid Streptococcal Antigen Test on Antibiotic Use in Adult Patients." Diagnostic Microbiology & Infectious Disease 91(4): 339-344.	Wrong test
46. Donato LJ, Myhre NK, Murray MA, McDonah MR, Myers JF, Maxson JA, et al. Assessment of Test Performance and Potential for Environmental Contamination Associated with a Point-of-Care Molecular Assay for Group A Streptococcus in an End User Setting. Journal of Clinical Microbiology 2019;57(2). http://dx.doi.org/https://dx.doi.org/10.1128JCM.01629-18	No comparison with culture or clinical scores
47. Dulaney, K., Hohmeier, K., Fisher, C., Cardosi, L., & Wasson, M. (2018). Exploring pharmacists' perceptions regarding influenza and streptococcal testing within a chain pharmacy. Journal of the American Pharmacists Association: JAPhA, 58(4), 438-441.e431. doi:https://dx.doi.org/10.1016/j.japh.2018.03.004	No specific RADT mentioned. No comparison with culture or clinical scores
48. Dut, R., & Kocagoz, S. (2016). Use of Streptococcal Tonsillopharyngitis Diagnostic Tests in Children. Journal of Pediatric Infectious Diseases, 11(4), 126-130. doi:http://dx.doi.org/10.1055/s-0036-1597697	Wrong test

49. Dut, R., Kocagoz, S. (2016). "Clinical Signs and Diagnostic Tests in Acute Respiratory Infections." Indian Journal of Pediatrics 83(5): 380-385.	No specific RADT mentioned
50. Edin, A., Granholm, S., Koskiniemi, S., Allard, A., Sjostedt, A., Johansson, A. (2015). "Development and laboratory evaluation of a real-time PCR assay for detecting viruses and bacteria of relevance for community-acquired pneumonia." Journal of Molecular Diagnostics 17(3): 315-324.	No specific RADT mentioned
51. Edmonson, M. B., Farwell, K. R. (2005). "Relationship between the clinical likelihood of group a streptococcal pharyngitis and the sensitivity of a rapid antigen-detection test in a pediatric practice." Pediatrics 115(2): 280-285.	Wrong type of test
52. Edmonson, M. B., & Weix, K. R. (2003). Relationship of pre-test likelihood of group A streptococcal (GAS) pharyngitis and sensitivity of a rapid antigen detection test (RADT) in pediatric practice. Pediatric Research, 53(4), 180A-180A.	Abstract couldn't be located
53. Ehrlich, J. E., Demopulos, B. P., Daniel, K. R., Ricarte, M. C., Glied, S. (2002). "Cost-effectiveness of treatment options for rheumatic heart disease from group A streptococcal pharyngitis in a pediatric population." Preventive Medicine 35: 250-257	Wrong test name
54. Ehsanipour, F., Mirghorbani, M., Masoumi Asl, H., Harandi, N. V., & Khanaliha, K. (2016). "Comparison of clinical findings and rapid streptococcal antigen detection test in the diagnosis of group a Streptococcal (GAS) pharyngitis." Archives of Clinical Infectious Diseases 11 (2).	Wrong test
55. Elf, S., Olli, J., Hirvonen, S., Auvinen, P., Eboigbodin, K. E. (2018). "Molecular Detection of Streptococcus pyogenes by Strand Invasion Based Amplification Assay." Molecular Diagnosis & Therapy 22(5): 595-602.	Wrong type of test
56. Elmas, B., Koroglu, M., Terzi, H. A., Aslan, F. G., Menekse, E., Kosecik, M., Altindis, M. (2017). "Performance of Clinical Features, Acute Phase Reactants and Group A Streptococcus Rapid Test in Evaluation of the Etiologic Agents for Tonsillopharyngitis in Children." Clinical Laboratory 63(7): 1223-1231.	Wrong test
57. Engstrom, S., Molstad, S., Lindstrom, K., Nilsson, G., & Borgquist, L. (2004). Excessive use of rapid tests in respiratory tract infections in Swedish primary health care. Scandinavian Journal of Infectious Diseases, 36(3), 213-218.	No comparison with clinical score or throat culture
58. Enright, K., Taheri, S., Beattie, T. (2009). "Emergency department testing for Streptococcus in children with sore throats." Emergency Medicine Journal 26(4): 310.	No specific RADT mentioned
59. Fakih, M. G., Berschback, J., Juzych, N. S., Massanari, R. M. (2006). "Compliance of resident and staff physicians with IDSA guidelines for the diagnosis and treatment of streptococcal pharyngitis." Infectious Diseases in Clinical Practice 14(2): 84-88.	Wrong test
60. FDA Decision Summary: Substantial Equivalence determination for the Sofia® Strep A+ FIA. K141775. Silver Spring, MD, USA: U.S. Food and Drug Administration; 2014. URL: http://www.accessdata.fda.gov/cdrh_docs/reviews/K141775.pdf (Accessed 6 March 2019).	Wrong test
61. FDA Decision Summary: Substantial Equivalence determination for Cobas Liat Strep A. K141338. Silver Spring, MD, USA: U.S. Food and Drug Administration; 2015. URL:	Data identical to that provided by the

http://www.accessdata.fda.gov/cdrh_docs/reviews/K141338.pdf (Accessed 6 March 2019).	manufacturer in response to a request for information by NICE
62. FDA Decision Summary: Substantial Equivalence determination for TestPack Plus Strep A. K971522. Silver Spring, MD, USA: U.S. Food and Drug Administration; 1997. URL: http://www.accessdata.fda.gov/cdrh_docs/pdf/K971522.pdf (Accessed 6 March 2019).	Insufficient information
63. Felsenstein, S., Faddoul, D., Sposto, R., Batoon, K., Polanco, C. M., & Dien Bard, J. (2014). Molecular and clinical diagnosis of group A streptococcal pharyngitis in children. Journal of Clinical Microbiology, 52(11), 3884-3889. doi:https://dx.doi.org/10.1128/JCM.01489-14	Wrong test
64. Fierro, J. L., Prasad, P. A., Russell Localio, A., Grundmeier, R. W., Wasserman, R. C., Zaoutis, T. E., & Gerber, J. S. (2014). "Variability in the diagnosis and treatment of group a streptococcal pharyngitis by primary care pediatricians." Infection Control and Hospital Epidemiology 35: S79-S85.	No specific RADT mentioned
65. Fontes, M. J., Bottrel, F. B., Fonseca, M. T., Lasmar, L. B., Diamante, R., Camargos, P. A. (2007). "Early diagnosis of streptococcal pharyngotonsillitis: assessment by latex particle agglutination test." Jornal de Pediatria 83(5): 465-470.	Wrong type of test
66. Forward, K. (2002). "Just the Berries: Diagnosing and managing Group A streptococcus pharyngitis." Canadian Family Physician 48(JAN.): 47-48.	Review
67. Forward, K. R., Haldane, D., Webster, D., Mills, C., Brine, C., Aylward, D. (2006). "A comparison between the Strep A Rapid Test Device and conventional culture for the diagnosis of streptococcal pharyngitis." The Canadian Journal of Infectious Diseases & Medical Microbiology 17(4): 221-223.	Wrong test
68. Fox, J. W., Cohen, D. M., Marcon, M. J., Cotton, W. H., Bonsu, B. K. (2006). "Performance of rapid streptococcal antigen testing varies by personnel." Journal of Clinical Microbiology 44(11): 3918-3922.	Wrong test
69. Fox, J. W., Marcon, M. J., Bonsu, B. K. (2006). "Diagnosis of streptococcal pharyngitis by detection of Streptococcus pyogenes in posterior pharyngeal versus oral cavity specimens." Journal of Clinical Microbiology 44(7): 2593-2594.	Wrong test
70. Gazewood, J. (2003). Negative antigen test misses <5% of strep pharyngitis. Journal of Family Practice, 52(10), 761-762.	Review
71. Gazzano, V., Berger, A., Benito, Y., Freydiere, A. M., Tristan, A., Boisset, S., Carricajo, A., Poyart, C., Vandenesch, F., Descours, G. (2016). Reassessment of the Role of Rapid Antigen Detection Tests in Diagnosis of Invasive Group A Streptococcal Infections. Journal of Clinical Microbiology, 54(4), 994-999. doi:https://dx.doi.org/10.1128/JCM.02516-15	Wrong population
72. Gieseker, K. E., Roe, M. H., MacKenzie, T., & Todd, J. K. (2003). Evaluating the American Academy of Pediatrics diagnostic standard for Streptococcus pyogenes pharyngitis: backup culture versus repeat rapid antigen testing. Pediatrics, 111(6 Pt 1), e666-670.	Wrong test

73. Gieseker, K. E., Mackenzie, T., Roe, M. H., & Todd, J. K. (2002). Comparison of two rapid Streptococcus pyogenes diagnostic tests with a rigorous culture standard. Pediatric Infectious Disease Journal, 21(10), 922-927.	Wrong test
74. Giraldez-Garcia, C., Rubio, B., Gallegos-Braun, J. F., Imaz, I., Gonzalez-Enriquez, J., & Sarria-Santamera, A. (2011). Diagnosis and management of acute pharyngitis in a paediatric population: a cost-effectiveness analysis. European Journal of Pediatrics, 170(8), 1059-1067. doi:https://dx.doi.org/10.1007/s00431-011-1410-0	No specific RADT mentioned
75. Gonsu, H. K., Bomki, C. M., Djomou, F., Toukam, M., Ndze, V. N., Lyonga, E. E., Mbakop, C.D., Koulla-Shiro, S. (2015). A comparative study of the diagnostic methods for Group A streptococcal sore throat in two reference hospitals in Yaounde, Cameroon. The Pan African medical journal, 20, 139.	Wrong test
76. Greiver, M. (1999). "Practice tips. Incorporating a rapid group A streptococcus assay with the sore throat score." Canadian Family Physician 45: 1181-1182.	Review
77. Grondal, H., Hedin, K., Strandberg, E. L., Andre, M., Brorsson, A. (2015). "Near-patient tests and the clinical gaze in decision-making of Swedish GPs not following current guidelines for sore throat - a qualitative interview study." BMC Family Practice 16: 81.	No specific RADT mentioned
78. Gurol, Y., Akan, H., Izbirak, G., Tekkanat, Z. T., Gunduz, T. S., Hayran, O., Yilmaz, G. (2010). "The sensitivity and the specifity of rapid antigen test in streptococcal upper respiratory tract infections." International Journal of Pediatric Otorhinolaryngology 74(6): 591-593.	Wrong test
79. Haldrup, S., Thomsen, R. W., Bro, F., Skov, R., Bjerrum, L., Sogaard, M. (2017). "Microbiological point of care testing before antibiotic prescribing in primary care: considerable variations between practices." BMC Family Practice 18(1): 9.	No specific RADT mentioned
80. Hall, M. C., Kieke, B., Gonzales, R., Belongia, E. A. (2004). "Spectrum bias of a rapid antigen detection test for group A beta-hemolytic streptococcal pharyngitis in a pediatric population." Pediatrics 114(1): 182-186.	Wrong test
81. Hammond-Collins, K., Strauss, B., Barnes, K., Demczuk, W., Domingo, M. C., Lamontagne, M. C., Lu, D., Martin, I., Tepper, M. (2018). "Group A Streptococcus Outbreak in a Canadian Armed Forces Training Facility." Military Medicine 21: 21.	Wrong test
82. Herranz, B., Rodriguez-Salinas, E., & Orden, B. (2007). From laboratory to clinic: Usefulness of rapid diagnostic techniques for the diagnostic techniques of Streptococcus pyogenes. [Spanish, English]. [Utilidad de las tecnicas de diagnostico rapido para la deteccion de streptococcus pyogenes.]. Anales de Pediatria Continuada, 5(2), 92-95. doi:http://dx.doi.org/10.1016/S1696-2818%2807%2974115-X	Foreign paper
83. Hinfey, P., Nicholls, B. H., Garcia, F., Ripper, J., Cameron, Y., Joshi, S. (2010). "Sensitivity of a Rapid Antigen Detection Test for the Diagnosis of Group a Streptoccal Pharyngitis In the Emergency Department." Annals of Emergency Medicine 56(3): S132-S132.	Wrong test
84. Hoffmann, K., Reichardt, B., Zehetmayer, S., Maier, M. (2012). "Evaluation of the implementation of a rapid streptococcal antigen test in a routine primary health care setting: from recommendations to practice." Wiener Klinische Wochenschrift 124(17-18): 633-638.	No specific RADT mentioned

85. Homme, J. H., Greenwood, C. S., Cronk, L. B., Nyre, L. M., Uhl, J. R., Weaver, A. L., Patel, R. (2018). "Duration of Group A Streptococcus PCR positivity following antibiotic treatment of pharyngitis." Diagnostic Microbiology & Infectious Disease 90(2): 105-108.	Wrong test
86. Honkanen, P. O., Rautakorpi, U. M., Huovinen, P., Klaukka, T., Palva, E., Roine, R., Sarkkinen, H., Varonen, H., Makela, M., Group, M. C. S. (2002). "Diagnostic tools in respiratory tract infections: use and comparison with Finnish guidelines." Scandinavian Journal of Infectious Diseases 34(11): 827-830.	No specific RADT mentioned
87. Humair, J. P., Revaz, S. A., & Stalder, H. (2002). Antibiotic prescription in strategies using a clinical score and a rapid streptococcal test for acute pharyngitis. Journal of General Internal Medicine, 17, 125-125.	Wrong test
88. Igarashi, H., Nago, N., Kiyokawa, H., Fukushi, M. (2017). "Abdominal pain and nausea in the diagnosis of streptococcal pharyngitis in boys." International journal of general medicine 10: 311-318.	Wrong test
89. Jayaratne, P., & Rutherford, C. (2015). Detection of Group A streptococci (GAS) by Loop-Mediated Isothermal Amplification (LAMP) directly from specimens: A rapid, simple and cost-effective alternative to culture. Canadian Journal of Infectious Diseases and Medical Microbiology, 26 (2), e19.	Wrong test
90. Joachim, L., Campos, D., Jr., Smeesters, P. R. (2010). "Pragmatic scoring system for pharyngitis in low-resource settings." Pediatrics 126(3): e608-614.	Wrong test
91. Kato, Y., Suzuki, K., Shibai, Y., & Iwamoto, H. (2015). Development of simultaneous detection lateral flow immunoassay kit for GAS and ADV. Clinical Chemistry, 1), S150.	Wrong test
92. Keahey, L., Bulloch, B., Jacobson, R., Tenenbein, M., Kabani, A. (2002). "Diagnostic accuracy of a rapid antigen test for GABHS performed by nurses in a pediatric ED." American Journal of Emergency Medicine 20(2): 128-130.	Wrong test
93. Khattak, M. H., Khan, M. A., Shafiullah, Orakzi, U. K. (2015). "Incidence of acute streptococcal pharyngitis." Journal of Medical Sciences (Peshawar) 23(3): 118-120.	No specific RADT mentioned
94. Kivi, N., Vanhanen, A. R., Nissinen, A. (2016). "Assessment of strep a point-of-care testing performance through external quality assurance (EQA) scheme: Results of a 6-year study period (2009-2015)." Clinical Chemistry and Laboratory Medicine 54 (10): eA277.	No specific RADT mentioned
95. Klepser, D. G., Klepser, M. E., Dering-Anderson, A. M., Morse, J. A., Smith, J. K., & Klepser, S. A. (2016). Community pharmacist-physician collaborative streptococcal pharyngitis management program. Journal of the American Pharmacists Association: JAPhA, 56(3), 323-329.e321. doi:https://dx.doi.org/10.1016/j.japh.2015.11.013	Wrong test
96. Klepser, D. G., Klepser, M. E., Smith, J. K., Dering-Anderson, A. M., Nelson, M., Pohren, L. E. (2018). "Utilization of influenza and streptococcal pharyngitis point-of-care testing in the community pharmacy practice setting." Research In Social & Administrative Pharmacy 14(4): 356-359.	No specific RADT mentioned

97. Klepser, D., Grismer, S. E., & Klepser, M. E. (2011). Cost-effectiveness of pharmacist provided care for the treatment of adult pharyngitis. Journal of Managed Care Pharmacy, 17 (3), 241.	No specific RADT mentioned
98. Klepser, D. G., Bisanz, S. E., & Klepser, M. E. (2012). Cost-effectiveness of pharmacist-provided treatment of adult pharyngitis. American Journal of Managed Care, 18(4), e145-154.	No specific RADT mentioned
99. Kose, E., Sirin Kose, S., Akca, D., Yildiz, K., Elmas, C., Baris, M., Anil, M. (2016). "The Effect of Rapid Antigen Detection Test on Antibiotic Prescription Decision of Clinicians and Reducing Antibiotic Costs in Children with Acute Pharyngitis." Journal of Tropical Pediatrics 62(4): 308-315.	Wrong test
100. Kreher, N. E., Hickner, J. M., Barry, H. C., & Messimer, S. R. (1998). Do gastrointestinal symptoms accompanying sore throat predict streptococcal pharyngitis? An UPRnet study. Journal of Family Practice, 46(2), 159-164.	Wrong test
101. Kucuk, O., Bicer, S., Giray, T., Col, D., Erdag, G. C., Gurol, Y., Kaspar, C.E., Vitrinel, A. (2014). "Validity of rapid antigen detection testing in group A beta-hemolytic streptococcal tonsillopharyngitis." Indian Journal of Pediatrics,81(2): 138-142.	Wrong test
102. Kulkarni, T., Aikawa, C., Nozawa, T., Murase, K., Maruyama, F., Nakagawa, I. (2016). "DNA-based culture-independent analysis detects the presence of group a streptococcus in throat samples from healthy adults in Japan." BMC Microbiology 16(1): 237.	Wrong type of test
103. Kurtz, B., Kurtz, M., Roe, M., & Todd, J. (1999). Importance of inoculum size and sampling effect in rapid antigen detection of S-pyogenes pharyngitis. Pediatric Research, 45(4), 166A-166A. doi:10.1203/00006450-199904020-00985	Not enough information
104. Lasseter, G., McNulty, C., Hobbs, F. D. R., Mant, D., & Investigators, P. (2014). In vitro evaluation of five rapid antigen detection tests for group A beta-haemolytic streptococcal sore throat infections. Health Technology Assessment, 18(6), 1-+. doi:10.3310/hta18060	Wrong population
105. Lasseter, G. M., McNulty, C. A., Hobbs, F. D., Mant, D., Little, P., & Group, P. R. c. S. M. I. (2011). Effect of swab type on the analytical sensitivity of five point-of-care tests for group A streptococci. British Journal of Biomedical Science, 68(2), 91-94.	Wrong population
106. Lasseter, G. M., McNulty, C. A., Richard Hobbs, F. D., Mant, D., Little, P., & Investigators, P. (2009). In vitro evaluation of five rapid antigen detection tests for group A beta-haemolytic streptococcal sore throat infections. Family Practice, 26(6), 437-444. doi:https://dx.doi.org/10.1093/fampra/cmp054	Wrong population
107. Lathia, N., Sullivan, K., Tam, K., Brna, M., MacNeil, P., Saltmarche, D., & Agro, K. (2018). Cost-minimization analysis of community pharmacy-based point-of-care testing for strep throat in 5 Canadian provinces. Canadian Pharmacists Journal, 151(5), 322-331. doi:http://dx.doi.org/10.1177/1715163518790993	No specific RADT mentioned
108. Leydon, G., McDermott, L., Moore, M., Williamson, I., Hobbs, F. D. R., Little, P., & Investigators, P. (2014). A qualitative study of general practitioner, nurse practitioner and patient views about the use of rapid streptococcus antigen detection tests in primary care: 'swamped with sore throats?'. Health Technology Assessment, 18(6), 41-50.	Duplicate

109. Leydon, G. M., McDermott, L., Moore, M., Williamson, I., Hobbs, F. D., Lambton, T., Cooper, R., Henderson, H., Little, P., Investigators, P. (2013). "A qualitative study of GP, NP and patient views about the use of rapid streptococcal antigen detection tests (RADTs) in primary care: 'swamped with sore throats?'." BMJ Open 3(4).	No specific RADT mentioned
110. Li, Y., Kim, H. J., Kong, H., Ranalli, T. A., Olivo, P. D., & Stenzel, T. T. (2013). Detection of group a streptococcus with a rapid, non-instrumented, isothermal molecular assay on throat swab specimens. Journal of Molecular Diagnostics, 15 (6), 890. doi:http://dx.doi.org/10.1016/S1525-1578%2813%2900181-5	No specific RADT mentioned
111. Little, P., Hobbs, F. D., Moore, M., Mant, D., Williamson, I., McNulty, C., Cheng, Y.E., Leydon, G., McManus, R., Kelly, J., Barnett, J., Glasziou, P., Mullee, M., investigators, P. (2014). PRImary care Streptococcal Management (PRISM) study: in vitro study, diagnostic cohorts and a pragmatic adaptive randomised controlled trial with nested qualitative study and cost-effectiveness study.[Erratum appears in Health Technol Assess. 2018 Oct;18(6):103-104; PMID: 30295591]. Health Technology Assessment (Winchester, England), 18(6), vii-xxv, 1-101. doi:https://dx.doi.org/10.3310/hta18060	Wrong population in study 1. Wrong test in study 3
112. Little, P., Moore, M., Hobbs, F. D. R., Mant, D., McNulty, C., Williamson, I., Cheng, M.Y.E., Glasziou, P., Mullee, M., Investigators, P. (2014). "Randomised controlled trial of a clinical score and rapid antigen detection test for sore throats." Health Technology Assessment, 18(6): 29-39.	Duplicate
113. Llor, C., Bjerrum, L., Munck, A., Cots, J. M., Hernandez, S., Moragas, A. (2014). "Access to point-of-care tests reduces the prescription of antibiotics among antibiotic-requesting subjects with respiratory tract infections." Respiratory Care 59(12):1918-1923.	No specific RADT mentioned
114. Llor, C., Hernandez, S., Sierra, N., Moragas, A., Hernandez, M., & Bayona, C. (2010). Association between use of rapid antigen detection tests and adherence to antibiotics in suspected streptococcal pharyngitis. Scandinavian Journal of Primary Health Care, 28(1), 12-17. doi:https://dx.doi.org/10.3109/02813431003669301	Wrong population and outcome
115. Llor, C., Moragas, A., Cots, J. M., Lopez-Valcarcel, B. G., Happy Audit Study, G. (2017). "Estimated saving of antibiotics in pharyngitis and lower respiratory tract infections if general practitioners used rapid tests and followed guidelines." Atencion Primaria 49(6): 319-325.	No specific RADT mentioned
116. Lotufo, J. P. B., Cardoso, D. M., Gilio, A. E., Hsin, S. H., Machado, B. M., De Paulis, M., Martinez, M.B., Grisi, S. J. E. (2012). "Impact of the use of rapid antigen detection test in the diagnosis and treatment of acute pharyngitis in pediatric emergency room." Paediatric Respiratory Reviews 13 (Supplement 1): S70.	No specific RADT mentioned
117. Luo R, Sickler J, Vahidnia F, Lee YC, Frogner B, Thompson M. Diagnosis and Management of Group a Streptococcal Pharyngitis in the United States, 2011-2015. BMC Infectious Diseases 2019;19(1):193. http://dx.doi.org/https://dx.doi.org/10.1186s12879-019-3835-4	No specific RADT mentioned
118. Madurell, J., Balague, M., Gomez, M., Cots, J. M., & Llor, C. (2010). Impact of rapid antigen detection testing on antibiotic prescription in acute pharyngitis in adults. FARINGOCAT	Protocol only. No results

STUDY: a multicentric randomized controlled trial. BMC Family Practice, 11, 25. doi:https://dx.doi.org/10.1186/1471-2296-11-25	
119. Makri, A., Tzanakaki, G., Kalimeratzi, S., Iliadou, H., Xiroyianni, A., Kremastinou, J., Voyatzi, A. (2010). "Evaluation of polymerase chain reaction as rapid diagnostic tool compared to culture in respiratory tract infections." Clinical Microbiology and Infection 2: S513.	No specific RADT mentioned
120. Malecki, M., Mazur, A., Sobolewski, M., Binkowska-Bury, M., Marc, M., & Januszewicz, P. (2017). Rapid strip tests as a decision-making tool about antibiotic treatment in children - A prospective study. Pediatria Polska, 92(2), 149-155. doi:http://dx.doi.org/10.1016/j.pepo.2017.01.006	No comparator
121. Maltezou, H. C., Tsagris, V., Antoniadou, A., Galani, L., Douros, C., Katsarolis, I., Maragos, A., Raftopopoulou, V., Biskini, P., Kanellakopoulou, K., Fretzayas, A., Papadimitriou, T., Nicolaidou, p., Giamarellou, H. (2008). Evaluation of a rapid antigen detection test in the diagnosis of streptococcal pharyngitis in children and its impact on antibiotic prescription. Journal of Antimicrobial Chemotherapy, 62(6), 1407-1412. doi:https://dx.doi.org/10.1093/jac/dkn376	Wrong test
122. Mayes, T., & Pichichero, M. E. (2001). Are follow-up throat cultures necessary when rapid antigen detection tests are negative for group A streptococci? Clinical Pediatrics, 40(4), 191-195.	Wrong test
123. Mazur, E., Bochynska, E., Juda, M., Koziol-Montewka, M. (2014). "Empirical validation of Polish guidelines for the management of acute streptococcal pharyngitis in children." International Journal of Pediatric Otorhinolaryngology 78(1): 102-106.	Wrong test
124. Messina, A., Bottaro, G., & Morselli, I. (2010). Utility of rapid antigen detection test for group a streptococci in a family paediatrician office setting. Acta Medica Mediterranea, 26(2), 101-105.	No outcomes by test
125. Michel-Lepage, A., Ventelou, B., Nebout, A., Verger, P., Pulcini, C. (2013). "Cross-sectional survey: risk-averse French GPs use more rapid-antigen diagnostic tests in tonsillitis in children." BMJ Open 3(10): e003540.	No specific RADT mentioned
126. Michel-Lepage, A., Ventelou, B., Verger, P., & Pulcini, C. (2014). Factors associated with the use of rapid antigen diagnostic tests in children presenting with acute pharyngitis among French general practitioners. European Journal of Clinical Microbiology & Infectious Diseases 33(5): 723-728.	No specific RADT mentioned
127. Mirjat, K. A., Fatima, I., & Mustafa, F. (2012). Prevalence of pharyngitis and tonsilitis among children. Medical Forum Monthly, 23(1), 64-67.	Not enough information
128. Mirjat, K. A., ValiRam, P., & Fatima, I. (2012). Role of rapid antigen detection test (RADT) and throat culture in the diagnosis of streptococcal pharyngotonsillitis. Medical Forum Monthly, 23(3), 60-63.	Not enough information
129. Mirza, A., Wludyka, P., Chiu, T. T., & Rathore, M. H. (2007). Throat culture is necessary after negative rapid antigen detection tests. Clinical Pediatrics, 46(3), 241-246.	Wrong test
130. Miyashita, N., Kawai, Y., Kato, T., Tanaka, T., Akaike, H., Teranishi, H., Nakano, T., Ouchi, K., Okimoto, N. (2016). "Rapid diagnostic method for the identification of Mycoplasma	Wrong test

pneumoniae respiratory tract infection." Journal of Infection & Chemotherapy 22(5): 327-330.	
131. Mlejnek, J. R., Almulhem, K., Spadafore, S. (2014). "Utility and cost effectiveness of throat culture in the treatment of patients with negative rapid strep screens." Academic Emergency Medicine 1: S51.	Unclear population
132. Moore, N. (2017). Rapid and Sensitive Isothermal Molecular Amplification of Group A Streptococcus (GAS) with Alere i Molecular Platform. Journal of Molecular Diagnostics, 19(6), 987-987.	Wrong outcome
133. Morandi, P. A., Deom, A., Mauris, A., & Rohner, P. (2003). External quality control of direct antigen tests to detect group A streptococcal antigen. European Journal of Clinical Microbiology & Infectious Diseases, 22(11), 670-674.	Wrong population
134. Murphy, M. L., Pichichero, M. E. (2002). "Prospective identification and treatment of children with pediatric autoimmune neuropsychiatric disorder associated with group A streptococcal infection (PANDAS)." Archives of Pediatrics & Adolescent Medicine 156(4): 356-361.	Wrong test
135. Nakhoul, G. N., Hickner, J. (2013). "Management of adults with acute streptococcal pharyngitis: minimal value for backup strep testing and overuse of antibiotics." Journal of General Internal Medicine 28(6): 830-834.	Wrong test
136. Nazgul, O., Yoshihisa, Y., Guli, S., Toshihiro, N., Mayramkan, A. (2010). 2Prevalence of group A b-hemolytic Streptococcus among children with tonsillopharyngitis in Kyrgyzstan: The difficulty of diagnostics and therapy." International Journal of Rheumatic Diseases 1: 212-213.	No specific RADT mentioned
137. Nct. (2017). Performance of Ellume·Lab Group A Strep Test Versus Culture for the Rapid Detection of Group A Streptococcus in Patients With Acute Pharyngitis. Https://clinicaltrials.gov/show/nct03171350.	No study publication. No results
138. Nct. (2017). Comparison of Two Rapid Antigen Detection Tests for the Detection of Group-A Streptococcal Pharyngitis in Children. Https://clinicaltrials.gov/show/nct03099018.	No outcome data
139. Neumark, T., Brudin, L., Molstad, S. (2010). "Use of rapid diagnostic tests and choice of antibiotics in respiratory tract infections in primary healthcare--a 6-y follow-up study." Scandinavian Journal of Infectious Diseases 42(2): 90-96.	No specific RADT mentioned
140. Nissinen, A., Stranden, P., Myllys, R., Takkinen, J., Bjorkman, Y., Leinikki, P., & Siitonen, A. (2009). Point-of-care testing of group A streptococcal antigen: performance evaluated by external quality assessment. European Journal of Clinical Microbiology & Infectious Diseases, 28(1), 17-20. doi:https://dx.doi.org/10.1007/s10096-008-0580-9	No comparison with culture or clinical score
141. Noorbakhsh, S., Tabatabaei, A., Farhadi, M., Ebrahimi, T. F. (2011). "Immunoasssay chromatographic antigen test for rapid diagnosis of Group A beta hemolytic Streptococcus pharyngitis in children: A cross/ sectional study." Iranian Journal of Microbiology 3(2): 99-103.	Wrong test
142. Orda, U., Mitra, B., Orda, S., Fitzgerald, M., Gunnarsson, R., Rofe, G., & Dargan, A. (2016). Point of care testing for group A streptococci in patients presenting with pharyngitis will improve appropriate antibiotic prescription. Emergency Medicine	Wrong reference standard

Australasia, 28(2), 199-204. doi:https://dx.doi.org/10.1111/1742-6723.12567	
143. Orda, U., Gunnarsson, R., Orda, S., Fitzgerald, M., Rofe, G., & Dargan, A. (2016). Etiologic predictive value of a rapid immunoassay for the detection of group A Streptococcus antigen from throat swabs in patients presenting with a sore throat. International Journal of Infectious Diseases, 45, 32-35. doi:https://dx.doi.org/10.1016/j.ijid.2016.02.002	No comparison with culture or clinical score
144. Ouchi, K., Hasegawa, K., Nonaka, Y., Matsushima, H., Komura, H., Maki, T., Nakazawa, T. (1999). "Rapid diagnosis of adenovirus respiratory tract infections by immunochromatography." Journal of Infection and Chemotherapy 5(4): 220-222.	Wrong type of test
145. Papastergiou, J., Diamantouros, A., Davidson, S., Saltmarche, D. (2017). "Community pharmacist-directed point-of-care group A strep testing: Results of a Canadian pilot program." International Journal of Clinical Pharmacy 39 (1): 208.	No specific RADT mentioned
146. Papastergiou, J., Trieu, C. R., Saltmarche, D., & Diamantouros, A. (2018). Community pharmacist-directed point-of-care group A Streptococcus testing: Evaluation of a Canadian program. Journal of the American Pharmacists Association: JAPhA, 58(4), 450-456. doi:https://dx.doi.org/10.1016/j.japh.2018.03.003	No comparison with culture or clinical score
147. Park, S. Y., Gerber, M. A., Tanz, R. R., Hickner, J. M., Galliher, J. M., Chuang, I., Besser, R. E. (2006). C"linicians' management of children and adolescents with acute pharyngitis." Pediatrics 117(6): 1871-1878.	No specific RADT mentioned
148. Pauchard, J. Y., Verga, M. E., Bersier, J., Prod'Hom, G., Gehri, M., & Vaudaux, B. (2012). Spectrum Bias of Rapid Antigen Diagnostic Test for Group A beta-Haemolytic Streptococcal Pharyngitis in a tertiary paediatric emergency department. Swiss Medical Weekly, 142, 9S-10S.	No specific RADT mentioned
149. Pauchard, J. Y., Verga, M. E., Bersier, J., Prod'Hom, Gehri, M., & Vaudaux, B. (2012). Performance of Rapid Antigen Diagnostic Test for Group A beta-Haemolytic Streptococcal Pharyngitis in a tertiary paediatric emergency department. Swiss Medical Weekly, 142, 35S-35S.	No specific RADT mentioned
150. Peralta, N. V., & Alcaraz, L. E. (2018). Frequency of isolates of streptococcus pyogenes in patients with clinical diagnosis of acute pharyngotonsillitis in a private laboratory in the city of San Luis. Biocell, 3), 26-27.	No specific RADT mentioned
151. Phung, E., Mirzaian, E., & Arouchanova, D. (2018). Utilization of pharmacist-performed rapid influenza and group a streptococcus testing and treatment in the community pharmacy setting: Economic value and patient satisfaction. Journal of the American Pharmacists Association, 58 (3), e129. doi:http://dx.doi.org/10.1016/j.japh.2018.04.004	Abstract only. No extractable data
152. Pitetti, R. D., Drenning, S. D., & Wald, E. R. (1998). Evaluation of a new rapid antigen detection kit for group A beta-hemolytic streptococci. Pediatric Emergency Care, 14(6), 396-398.	Wrong test
153. Plainvert, C., Duquesne, I., Touak, G., Dmytruk, N., & Poyart, C. (2015). In vitro evaluation and comparison of 5 rapid antigen detection tests for the diagnosis of beta-hemolytic group A streptococcal pharyngitis. Diagnostic Microbiology & Infectious Disease, 83(2), 105-111. doi:https://dx.doi.org/10.1016/j.diagmicrobio.2015.06.012	Wrong population

154. Pulcini, C., Pauvif, L., Paraponaris, A., Verger, P., Ventelou, B. (2012). "Perceptions and attitudes of French general practitioners towards rapid antigen diagnostic tests in acute pharyngitis using a randomised case-vignette study: A cross-sectional study." Clinical Microbiology and Infection 3: 494.	No specific RADT mentioned
155. Pulcini, C., Pauvif, L., Paraponaris, A., Verger, P., Ventelou, B. (2012). "Perceptions and attitudes of French general practitioners towards rapid antigen diagnostic tests in acute pharyngitis using a randomized case vignette study." Journal of Antimicrobial Chemotherapy 67(6): 1540-1546.	No specific RADT mentioned
156. Ramos, J. L., Fraile, M. T., Chanza, M., Tormo, N., Lurbe, A., & Gimeno, C. (2011). Rapid detection of Streptococcus pyogenes in peripheral medical centres. A pilot custody assay. Clinical Microbiology and Infection, 4), S250. doi:http://dx.doi.org/10.1111/j.1469-0691.2011.03558.x	Wrong test
157. Rao A, Berg B, Quezada T, Fader R, Walker K, Tang S, et al. Diagnosis and antibiotic treatment of group a streptococcal pharyngitis in children in a primary care setting: impact of point-of-care polymerase chain reaction. BMC Pediatrics 2019;19(1):24. https://dx.doi.org/10.1186/s12887-019-1393-y	Wrong reference standard and no comparison to clinical score
158. Rathi, S. K., & Ahmed, R. (2014). Pakistan prevalence survey in acute pharyngitis. JPMA - Journal of the Pakistan Medical Association, 64(8), 928-931.	Wrong test
159. Rimoin, A. W., Vince, A., Hamza, H., da Cunha, A. L. A., Chitale, R., Oazi, S., & Steinhoff, M. C. (2004). Evaluation of a rapid test for streptococcal pharyngitis in children in 3 countries. Pediatric Research, 55(4), 279A-279A.	Meeting abstract couldn't be located
160. Rimoin, A. W., Walker, C. L., Hamza, H. S., Elminawi, N., Ghafar, H. A., Vince, A., da Cunha, A.L., Qazi, S., Gardovska, D., Steinhoff, M. C. (2010). "The utility of rapid antigen detection testing for the diagnosis of streptococcal pharyngitis in low-resource settings." International Journal of Infectious Diseases 14(12): e1048-1053.	Wrong test
161. Russo, M. E., Kline, J., Jaggi, P., Leber, A. L., & Cohen, D. M. (2017). The Challenge of Patient Notification and the Work of Follow-Up Generated by a 2-Step Testing Protocol for Group A Streptococcal Pharyngitis in the Pediatric Emergency Department. *Pediatric Emergency Care, 30*, 30. doi:https://dx.doi.org/10.1097/PEC.0000000000001144	No comparison with throat score or culture
162. Sancho, A., Diaz-Almiron, M., Yebra, J., Hawkins, M. (2014). "S. pyogenes reviewed in a paediatric population: Age and predictive models." Archives of Disease in Childhood 2: A325.	No specific RADT mentioned
163. Sarikaya, S., Aktas, C., Ay, D., Cetin, A., & Celikmen, F. (2010). Sensitivity and specificity of rapid antigen detection testing for diagnosing pharyngitis in the emergency department. Ear, Nose, & Throat Journal, 89(4), 180-182.	Wrong test
164. Sayyahfar, S., Fahimzad, A., Naddaf, A., & Tavassoli, S. (2015). Antibiotic Susceptibility Evaluation of Group A Streptococcus Isolated from Children with Pharyngitis: A Study from Iran. Infection & Chemotherapy, 47(4), 225-230.	Wrong type of test
165. Scheel, A., DeWyer, A., Sarnacki, R., Kamarembo, J., Okello, E., & Beaton, A. (2018). The Utility of Existing Clinical Decision Rules For Streptococcal Pharyngitis In Ugandan School Children. Global Heart, 13 (4), 508-509.	No specific RADT mentioned

166. Schwartz, R. H., Kim, D., Martin, M., & Pichichero, M. E. (2015). A Reappraisal of the Minimum Duration of Antibiotic Treatment Before Approval of Return to School for Children With Streptococcal Pharyngitis. Pediatric Infectious Disease Journal, 34(12), 1302-1304. doi:https://dx.doi.org/10.1097/INF.0000000000000883	Wrong test
167. Schwartz, K., Monsur, J., Northrup, J., West, P., & Neale, A. V. (2004). Pharyngitis clinical prediction rules: effect of interobserver agreement: a MetroNet study. Journal of Clinical Epidemiology, 57(2), 142-146.	Wrong test
168. Shapiro, D. J., Lindgren, C. E., Neuman, M. I., & Fine, A. M. (2017). Viral Features and Testing for Streptococcal Pharyngitis. Pediatrics, 139(5).	No specific RADT mentioned
169. Sheeler, R. D., Houston, M. S., Radke, S., Dale, J. C., & Adamson, S. C. (2002). Accuracy of rapid strep testing in patients who have had recent streptococcal pharyngitis. Journal of the American Board of Family Practice, 15(4), 261-265.	Wrong population
170. Singh, S., Dolan, J. G., & Centor, R. M. (2006). Optimal management of adults with pharyngitis--a multi-criteria decision analysis. BMC Medical Informatics & Decision Making, 6, 14.	No specific RADT mentioned
171. Skoog, G., Edlund, C., Giske, C. G., Molstad, S., Norman, C., Sundvall, P. D., & Hedin, K. (2016). A randomized controlled study of 5 and 10 days treatment with phenoxymethylpenicillin for pharyngotonsillitis caused by streptococcus group A - a protocol study. BMC Infectious Diseases, 16, 484.	No specific RADT mentioned
172. Slinger, R., Goldfarb, D., Rajakumar, D., Moldovan, I., Barrowman, N., Tam, R., & Chan, F. (2011). Rapid PCR detection of group A Streptococcus from flocked throat swabs: a retrospective clinical study. Annals of Clinical Microbiology & Antimicrobials, 10, 33.	Wrong type of test
173. St Sauver, J. L., Weaver, A. L., Orvidas, L. J., Jacobson, R. M., & Jacobsen, S. J. (2006). Population-based prevalence of repeated group A beta-hemolytic streptococcal pharyngitis episodes. Mayo Clinic Proceedings, 81(9), 1172-1176.	Wrong test
174. Subashini, B., Anandan, S., & Balaji, V. (2015). Evaluation of a rapid antigen detection test for the diagnosis of group-A beta-hemolytic Streptococcus in pharyngotonsillitis. Journal of global infectious diseases, 7(2), 91-92. doi:https://dx.doi.org/10.4103/0974-777X.154447	Letter
175. Sultan, A. M., & Seliem, W. A. (2018). Evaluating the use of dedicated swab for rapid antigen detection testing in group a streptococcal pharyngitis in children. African Journal of Clinical and Experimental Microbiology, 19(1), 24-29.	Wrong test
176. Supon, P. A., Tunnell, S., Greene, M., & Ostroff, R. M. (1998). Rapid detection of group A streptococcal antigen with a new optical immunoassay. Pediatric Infectious Disease Journal, 17(4), 349-351.	Wrong test
177. Syriopoulou, T., Konstantelos, D., Papoula, M., Karachanidi, E., Maggana, I., Straka, K., Michael, E., Fourlani, E., Kapogli, A. (2011). Laboratory methods for diagnosing streptococcal pharyngitis: Predictive value, usefulness. Clinical Biochemistry, 44 (7), 534-535. doi:http://dx.doi.org/10.1016/j.clinbiochem.2011.03.080	No specific RADT mentioned
178. Tanz, R. R., Gerber, M. A., Kabat, W., Rippe, J., Seshadri, R., & Shulman, S. T. (2009). Performance of a rapid antigen-detection	Wrong test

test and throat culture in community pediatric offices: implications for management of pharyngitis.[Erratum appears in Pediatrics. 2009 Aug;124(2):846]. Pediatrics, 123(2), 437-444.	
179. Tanz, R. R., Zheng, X. T., Carter, D. M., Steele, M. C., & Shulman, S. T. (2018). Caution Needed: Molecular Diagnosis of Pediatric Group A Streptococcal Pharyngitis. Journal of the Pediatric Infectious Diseases Societ, 7(3), e145-e147.	Wrong test
180. Teratani, Y., Hagiya, H., Koyama, T., Ohshima, A., Zamami, Y., Tatebe, Y., ... & Hinotsu, S. (2019). Association between rapid antigen detection tests and antibiotics for acute pharyngitis in Japan: A retrospective observational study. Journal of Infection and Chemotherapy, 25(4), 267-272.	No specific RADT mentioned
181. Thamlikitkul V, Rachata T, Popum S, Chinswangwatanakul P, Srisomnuek A, Seenama C, et al. Accuracy and utility of rapid antigen detection tests for group A beta-hemolytic Streptococcus on ambulatory adult patients with sore throat associated with acute respiratory infections at Siriraj hospital. Journal of the Medical Association of Thailand 2018;101(4).	Wrong test
182. Toepfner, N., Henneke, P., Berner, R., & Hufnagel, M. (2013). Impact of technical training on rapid antigen detection tests (RADT) in group A streptococcal tonsillopharyngitis. European Journal of Clinical Microbiology & Infectious Diseases, 32(5), 609-611.	Wrong test
183. Tsevat, J., & Kotagal, U. R. (1999). Management of sore throats in children: a cost-effectiveness analysis. Archives of Pediatrics & Adolescent Medicine, 153(7), 681-688.	Wrong test
184. Tsung, L. Y., Choi, K. C., Nelson, E. A. S., Chan, P. K. S., & Sung, R. Y. T. (2010). Factors associated with length of hospital stay in children with respiratory disease. Hong Kong Medical Journal, 16(6), 440-446.	Wrong type of test
185. Tsutsumi, H., Ouchi, K., Ohsaki, M., Yamanaka, T., Kuniya, Y., Takeuchi, Y., Nakai, C., Meguro, H., Chiba, S. (1999). Immunochromatography test for rapid diagnosis of adenovirus respiratory tract infections: Comparison with virus isolation in tissue culture. Journal of Clinical Microbiology, 37(6), 2007-2009.	Wrong type of test
186. Upton, A., Lowe, C., Stewart, J., Taylor, S., & Lennon, D. (2014). In vitro comparison of four rapid antigen tests for group A streptococcus detection. New Zealand Medical Journal, 127(1398), 77-83.	Wrong population. No comparison with culture or clinical score
187. Vachhani, R., Patel, T., Centor, R. M., & Estrada, C. A. (2017). Sensitivity for Diagnosing Group A Streptococcal Pharyngitis from Manufacturers is 10% Higher than Reported in Peer-Reviewed Publications. Southern Medical Journal, 110(1), 59-64. doi:https://dx.doi.org/10.14423/SMJ.0000000000000597	Review
188. Van Howe, R. S., & Kusnier, L. P., 2nd. (2006). Diagnosis and management of pharyngitis in a pediatric population based on cost-effectiveness and projected health outcomes. Pediatrics, 117(3), 609-619.	No specific RADT mentioned
189. Van Limbergen, J., Kalima, P., Taheri, S., & Beattie, T. F. (2006). Streptococcus A in paediatric accident and emergency: are rapid streptococcal tests and clinical examination of any help? Emergency Medicine Journal, 23(1), 32-34.	Wrong test

190. Vedia, C., Garcia, J. A., Valles, R., Franzi, A., Morales, C., & Prat, N. (2016). Is it possible to decrease antibiotic prescription in pediatrics? Basic and Clinical Pharmacology and Toxicology, 119 (Supplement 1), 47.	No specific RADT mentioned
191. Waseem, M., Ayanruoh, S., Humphrey, A., & Reynolds, T. (2009). Impact of rapid streptococcal test on antibiotic use in a pediatric emergency department. Annals of Emergency Medicine, 1), S41.	No specific RADT mentioned
192. Webb, K. H., Needham, C. A., & Kurtz, S. R. (2000). Use of a high-sensitivity rapid strep test without culture confirmation of negative results: 2 years' experience.[Erratum appears in J Fam Prac 2000 Apr;49(4):378]. Journal of Family Practice, 49(1), 34-38.	Wrong test
193. Williams, K. M., Jackson, M. A., & Hamilton, M. (2002). Rapid diagnostic testing for URIs in children: Impact on physician decision making and costs. Infections in Medicine, 19(3), 109-117.	No empirical data
194. Wong, M. C., & Chung, C. H. (2002). Group A streptococcal infection in patients presenting with a sore throat at an accident and emergency department: prospective observational study. Hong Kong Medical Journal, 8(2), 92-98.	Wrong test
195. Woodburn, J. D., Smith, K. L., & Nelson, G. D. (2007). Quality of care in the retail health care setting using national clinical guidelines for acute pharyngitis. American Journal of Medical Quality, 22(6), 457-462.	Wrong test
196. Wright, M., Williams, G., & Ludeman, L. (2007). Comparison of two rapid tests for detecting group A streptococcal pharyngitis in the pediatric population at wright-patterson air force base. Military Medicine, 172(6), 644-646.	Wrong test
197. Xu, J., Schwartz, K., Monsur, J., Northrup, J., & Neale, A. V. (2004). Patient-clinician agreement on signs and symptoms of 'strep throat': a MetroNet study. Family Practice, 21(6), 599-604.	Wrong test
198. Yang, J. H., Huang, P. Y., Shie, S. S., Yang, S., Tsao, K. C., Wu, T. L., . . . Huang, C. T. (2018). Diagnostic performance of the Sofia influenza A+B fluorescent immunoassay in adult outpatients in Northern Taiwan. Journal of Medical Virology, 90(6), 1010-1018. doi:https://dx.doi.org/10.1002/jmv.25043	Wrong test
199. Yoon, J., Yun, S. G., Nam, J., Choi, S. H., & Lim, C. S. (2017). The use of saliva specimens for detection of influenza A and B viruses by rapid influenza diagnostic tests. Journal of Virological Methods, 243, 15-19. doi:https://dx.doi.org/10.1016/j.jviromet.2017.01.013	Wrong test

Appendix 4: QUADAS-2 Tailored guidance notes and form

Modified QUADAS-2 and guidance notes for Strep A

Risk of bias should only be classed as low for each domain if all questions could be answered with 'yes'. If one or more signaling question is answered with 'no' the risk of bias should be classed as 'high' and equally if at least one question is answered with 'unclear' the risk of bias should be judged 'unclear'.

Domain 1: Patient selection

Test measurement ratings will differ depending on whether antibiotics have been previously prescribed.

A. Risk of bias

Guidance:

Was a consecutive or random sample of patients enrolled?

This question should only be answered with 'yes' if the study clearly states that children/adults were recruited consecutively or randomly. Case-control or two gate studies should be answered no

Was a case-control design avoided?

There is increased bias in a case control (2 gate) study as compared to a cohort (1 gate) study.

Were selection criteria clearly described (age limits and Centor/FeverPAIN scores)?

All inclusion criteria should be clearly specified. Lack of clear selection criteria, or different selection criteria introduce bias through unclear adherence to consecutive or random sampling, and because there is a recognized bias with the reference standard detecting GAS carriage (rather than GAS detection) which is exacerbated if a greater proportion of less symptomatic patients are introduced.

Did the study avoid inappropriate exclusions?

Patients who meet the inclusion criteria should be given the index test. If more than 5% meet the inclusion criteria but are not given the test, this is an inappropriate exclusion. If <5% and no reasons are provided, this is also an inappropriate exclusion.

All patients' who received the index test should have their results reported. If more than 5%

are not reported, this is an inappropriate exclusion. If <5% are reported but no reasons are provided, this is also an inappropriate exclusion.
We would expect the whole cohort to receive a rapid test(s) (from one of our included list: clearview exact, BD Veritor Plus, Strep A rapid test, NADAL Strep A, OSOM Strep A trest, QuikRead Go Strep A test kit, Alere TestPack Plus Strep A, Bionexia Strep A, Biosynex, Sofia Strep A FIA) or a molecular tests (from one of our included list: Alere I, Cobas Strep A Assay on Liat system or Xpert Xpress Strep A). Also a comparator (Centor [modified Centor or McIsaac] or FeverPAIN) where included in the study design and a biological culture as the reference standard. Very small numbers of exclusions (<5%) may be acceptable, if accompanied by reasonable explanations.

Were patients seen in an ambulatory care setting?
Patients seen as inpatients may vary in severity and have co-morbidities affecting their diagnosis.

B. Concerns regarding applicability
Guidance:
Patients aged under 5 years do not meet our inclusion criteria. If more than 10% of the sample are under 5 this should be rated as high.
In the UK the test would be given following an assessment using Centor or FeverPAIN. The rapid test would only be given in people with centor scores >2 and FeverPAIN scores >1. If the study does not mention these tests or no assessment test was undertaken it should be rated high concern. If the study included people with scores 2 or lower on Centor, or 1 on FeverPAIN this can only be classed as low risk of bias if the test accuracy is reported separately for with centor scores >2 and FeverPAIN scores >1. If the test accuracy for low and high rated centor/FeverPAIN groups are ONLY reported together this should be reported as a high concern for applicability.

Domain 2: Index test
The main sources of bias introduced by conducting and interpreting the index test are blinding and defining the threshold. If the reference standard is carried out before the index test (e.g. in case control studies) it is important to blind personnel to the results of the reference standard.

The QUADAS-2 tool requires a threshold to be pre-specified in the methods in order to avoid adjustment of the threshold according to the test outcome. In manufactured tests the threshold has been pre-determined. There is some subjectivity in how the RADT tests are read. If the operator claimed to follow the product insert then the subjectivity has been reduced, however a bias still exists. There is no subjectivity in the molecular tests which tell you on the screen whether Strep A is present or not. In studies of test development, the threshold must be reported and must be pre-specified.

A. Risk of bias

Were the index test results interpreted without knowledge of the results of the reference standard?

In cohort designs where the reference standard was given after or at the same time as the index test answer yes. This is because the reference standard is read after a longer time period than the rapid test. If timing is unclear, or the study has a case-control design then this is only a yes if blinding is specifically mentioned or the index test is fully automated with no human interpretation.

Was a separate swab undertaken for the index test?

Manufacturer's specifications require separate swabs be taken for the index and reference standard. Using one swab for multiple purposes may reduce the amount of the sample and affect the accuracy of the test.

Was a threshold explicitly pre-specified?

All manufactured rapid tests and have an inbuilt threshold therefore the answer should be low. If the threshold is not pre-specified then it must be rated high risk of bias. In test development studies it must explicitly state that the threshold has been pre-specified and what the threshold is.

Is the test reading objective?

Molecular tests provide the result on the screen so should always be answered yes (low risk of bias). All rapid tests are subjectively read based on the internal inbuilt threshold bar the BD Veritor plus system, NADAL Strep A scan test, QuikRead Go Strep A test and Sofia Strep A FIA which use analysers/readers to digitally display results. Any test where a

subjective reading is taken will have a high risk of bias and should be answered no.

B. Concerns about applicability

If the study does not specify that the test was carried out to the manufacturer's specification the rating should be noted as unclear. Previous versions of included tests should be rated as high.

Domain 3: Reference standard

The reference standard should be throat culture. FeverPAIN or Centor are appropriate comparator screening tests but not a reference standard.

The reference standard should be undertaken using Staph or Strep agar plate or simple blood agar. Cultures using a blood agar should be incubated in an anaerobic atmosphere at 35-37 degrees for 18-24 hours, with cultures read after more than 18 hours. Alternatively, blood agar could be incubated in 5-10% CO_2 at 35-37 degrees for 18-24 hours. Cultures using staph or strep selective agar should be incubated at 35-37 degrees in aerobic conditions for 18-48 hours and read after more than 24 hours. Current guidance advises to re-examine plates at 48 hours that yield negative results at 24 hours.[26] If the culture is not incubated in the correct manner then there will be a high risk of bias

Investigators won't be blinded to the clinical scoring tool but should be blinded to the reference standard.

A. Risk of bias

Was a separate swab taken for throat culture testing?

AAP recommends a separate swabs be taken for the index and reference standard testing [Mitul Patel, personal communication]. Using one swab for multiple purposes may reduce the amount of the sample and affect the accuracy of the test.

Is the reference standard likely to correctly classify the target condition?

If the reference standard used was throat culture and this was done appropriately then the answer should be yes. This should be a laboratory culture on a Staph, Strep or blood agar plate during 48 hours. Were the culture medium, atmosphere, duration of incubation and GAD-confirmation technique described?

Were the reference standard results interpreted without knowledge of the results of the index test?

This can be rated as low providing the operator in the lab is competency assessed and follows the Standard Operating Procedure. This is applicable to all types of lab cultures.

B. Concerns about applicability

The concern of applicability of the reference standard will be 'high' if any measure other than a throat culture is used. The culture should be done using a Staph or Strep or simple blood agar plate, incubated as described above and then serotyped. If any of these measures differ then there is a high risk of bias. If it is not reported then this should be noted as unclear.

Domain 4: Flow and Timing

The index test should be carried out prior to the reference standard and to antibiotic prescribing.

A. Risk of bias

Was there an appropriate interval between index test(s) and reference standard?

The swab for throat culture should be taken at the same time as the swab for the RADT and should be processed within 48 hours. Consider the following:

Were both index test(s) and reference standard (and comparator where included) all carried out at the same appointment?

Were all swabs processed within 48 hours?

If the answer to any of these is no then this is high risk of bias.

Were both index test(s) and reference standard (and comparator where included) all carried out prior to commencement of antibiotics?

Patients should not have been treated with antibiotics prior to receiving the index test(s) and/or reference standard.

Did all patients receive a reference standard?

All should receive both the index test and reference standard. Very small numbers of

exclusions (<5%) may be acceptable, if accompanied by reasonable explanations.

Did all patients receive the same reference standard?

This question should be answered with 'no' if patients received different reference standards or if positive cases on the index test received a different reference standard to negative subjects.

Were all patients included in the analysis?

All patients should be included in the analysis. If inconclusive or intermediate results are not considered in the analysis the question should be answered with 'no'. Very small numbers of exclusions (<5%) may be acceptable, if accompanied by reasonable explanations. If patients lost to follow up were not included in the analysis or >5% of patients were lost to follow up (even if considered in the analysis) the question should be answered with 'no'. (The actual proportion of patients lost to follow up needs to be recorded for each study.) In both cases the risk of bias should be classed as 'high'.

QUADAS-2 **(unadjusted)**

First author surname and year of publication:

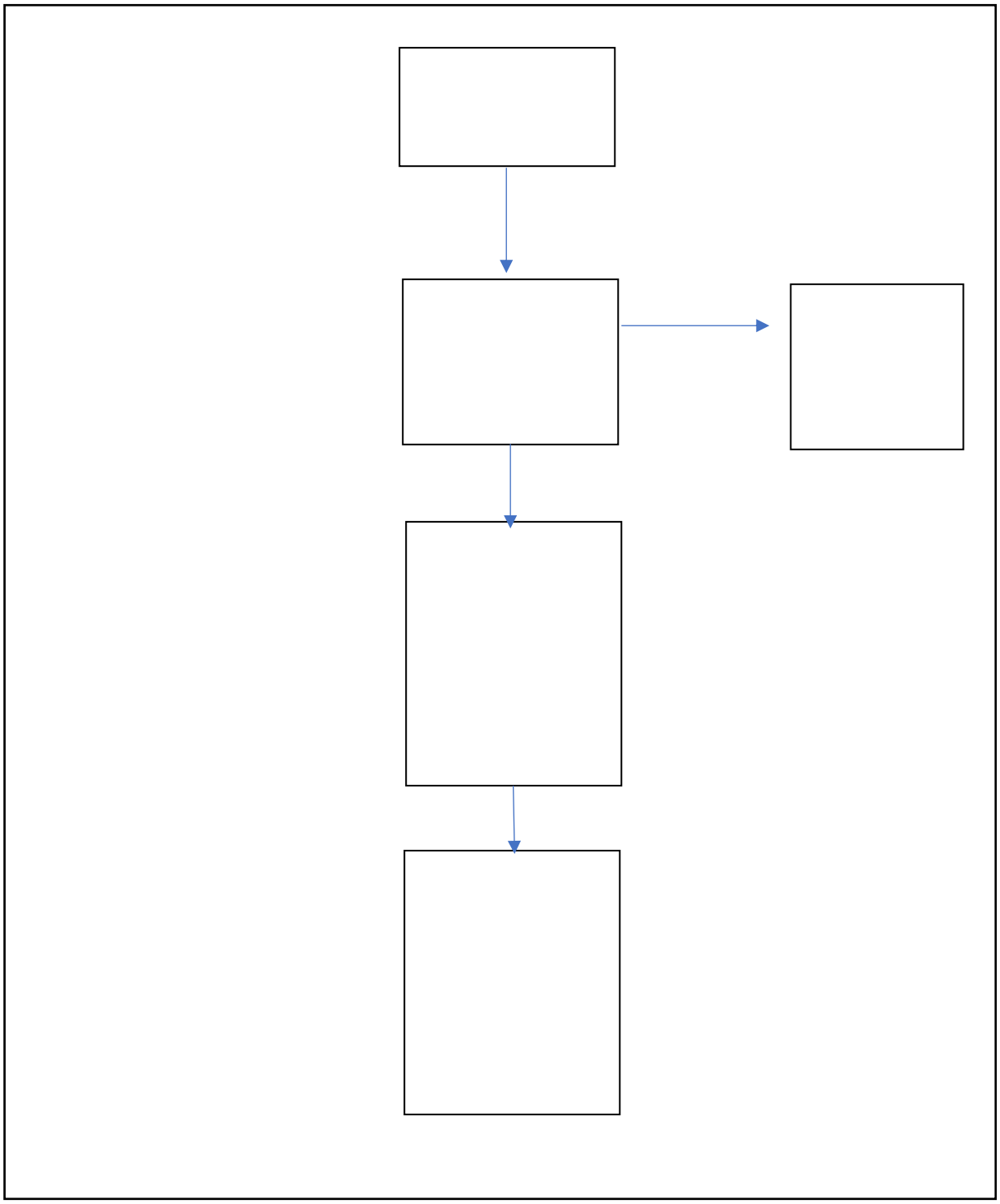
Name of first reviewer: HF *Name of second reviewer:*

Phase 1: State the review question:

Rapid Tests for Group A Streptococcal infections in people with a sore throat

Patients (setting, intended use of index test, presentation, prior testing):
Index test(s):
Comparator(s):
Reference standard and target condition: Culture. Strep A

Phase 2: Draw a flow diagram for the primary study



Phase 3: Risk of bias and applicability judgments

QUADAS-2 is structured so that 4 key domains are each rated in terms of the risk of bias and the concern regarding applicability to the research question (as defined above). Each key domain has a set of signalling questions to help reach the judgments regarding bias and applicability.

DOMAIN 1: PATIENT SELECTION **A. Risk of Bias**	
Describe methods of patient selection:	
+ Was a consecutive or random sample of patients enrolled?	Yes/No/Unclear
+ Was a case-control design avoided?	Yes/No/Unclear
+ Were selection criteria clearly described?	Yes/No/Unclear
+ Did the study avoid inappropriate exclusions?	Yes/No/Unclear
+ Were patients seen in ambulatory care setting?	Yes/No/Unclear
Could the selection of patients have introduced bias?	**RISK: LOW/HIGH/UNCLEAR**
B. Concerns regarding applicability	
Describe included patients (prior testing, presentation, intended use of index test and setting)**:**	
Is there concern that the included patients do not match the review question?	**CONCERN: LOW/HIGH/UNCLEAR**

DOMAIN 2: INDEX TEST(S)

If more than one index test was used, please complete for each test.

A. Risk of Bias

Describe the index test and how it was conducted and interpreted:	
+ Were the index test results interpreted without knowledge of the results of the reference standard?	Yes/No/Unclear
+ Was a separate swab undertaken for the index test?	Yes/No/Unclear
+ If a threshold was used, was it pre-specified?	Yes/No/Unclear
+ Is the test reading objective?	Yes/No/Unclear
Could the conduct or interpretation of the index test have introduced bias?	**RISK: LOW/HIGH/UNCLEAR**

B. Concerns regarding applicability

Is there concern that the index test, its conduct, or interpretation differ from the review question?	**CONCERN: LOW/HIGH/UNCLEAR**

DOMAIN 3: REFERENCE STANDARD

A. Risk of Bias

Describe the reference standard and how it was conducted and interpreted:

+ Was a separate swab taken for throat culture testing?	Yes/No/Unclear
+ Is the reference standard likely to correctly classify the target condition?	Yes/No/Unclear
+ Were the reference standard results interpreted without knowledge of the results of the index test?	Yes/No/Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	**RISK: LOW/HIGH/UNCLEAR**

B. Concerns regarding applicability

Is there concern that the target condition as defined by the reference standard does not match the review question?	**CONCERN: LOW/HIGH/UNCLEAR**

DOMAIN 4: FLOW AND TIMING

A. Risk of Bias

Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram):	
Describe the time interval and any intervention between index tests(s) and reference standard:	
+ Was there an appropriate interval (same appointment) between index test(s) and reference standard?	Yes/No/Unclear
+ Did all patients receive a reference standard?	Yes/No/Unclear
+ Did all patients receive the same reference standard?	Yes/No/Unclear
+ Were both index test(s) and reference standard (and comparator where included) all carried out prior to the commencement of antibiotics?	Yes/No/Unclear
+ Were all patients included in the analysis?	Yes/No/Unclear
Could the patient flow have introduced bias?	**RISK: LOW/HIGH/UNCLEAR**

Appendix 5: Record of searches - Cost-effectiveness

Sore throat / Strep A with economic evaluations / QoL / cost and resource use

Bibliographic databases

Summary of bibliographic database searches

Database	Date of search	Number of records from targeted search results (to screen first) + Other results picked up by broader search = Total number of records (+ update search results)
MEDLINE (Ovid)	22/01/2019 (updated 13/03/2019)	304 + 1728 = 2032 (+ 36)
Embase (Ovid)	22/01/2019 (updated 13/03/2019)	434 + 2673 = 3107 (+67)
NHS EED and HTA database (CRD)	22/01/2019 (not updated as no new records added)	13 + 42 = 55
Science Citation Index and Conference Proceedings Science (Web of Science)	29/01/2019 (updated 13/03/2019)	260 + 1,397 = 1,657 (+17)
CEA Registry	29/01/2019 (updated 13/03/2019)	3 (+0)
EconPapers (RePEc)	29/01/2019 (updated 13/03/2019)	6 (+0)
ScHARRHUD	29/01/2019 (updated 13/03/2019)	0 (+0)

Total from database searches: (1011 + 5849 = 6860) + 120 from 2019 update search = 6980

Total after deduplication: (522 + 2175 = 2697) + 58 from 2019 update search = 2755

MEDLINE (Ovid)

Searched on 22/01/2019 (updated on 13/03/2019, see at the end of this search record)

Databases: Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions(R) <1946 to January 21, 2019>

Original search, 22/01/2019

Search Strategy:

---

1 exp Pharyngitis/ (15095)

2 pharyngit*.ti,ab,kf. (5487)

3 (nasophyryngit* or rhinopharyngit* or epipharyngit*).ti,ab,kf. (178)

4 (tonsillit* or tonsilit*).ti,ab,kf. (5615)

5 ((sore or pain* or ache* or aching or inflam* or infect*) adj3 (pharyn* or throat* or tonsil* or nasopharyn* or rhinopharyn* or epipharyn*)).ti,ab,kf. (9975)

6 or/1-5 (25268)

7 Streptococcal Infections/di, mi (13421)

8 Streptococcus pyogenes/im, ip (5463)

9 7 or 8 (16691)

10 ((strep or streptococcal or group) adj2 A).ti,ab,kf. (564113)

11 9 and 10 (4859)

12 (strep* adj5 (throat* or pharyn* or tonsil*)).ti,ab,kf. (3410)

13 streptoco* A.ti,ab,kf. (480)

14 (group A adj5 streptoco*).ti,ab,kf. (9515)

15 ((streptococcus or strep or staphylococcus) adj1 (pyogenes or pyogenic)).ti,ab,kf. (7726)

16 ((streptococcus or strep) adj1 (epidemicus or erysipelatis or erysipelatos or hemolyticus or haemolyticus or scarlatinae or lancefield)).ti,ab,kf. (240)

17 (s pyogenes or pyogenes s or micrococcus scarlatinae).ti,ab,kf. (2497)

18 lancefield group.ti,ab,kf. (476)

19 gabhs.ti,ab,kf. (394)

20 or/11-19 (18885)

21 exp Economics/ (571394)

22 exp "Costs and Cost Analysis"/ (221362)

23 Health Status/ (75366)

24 exp "Quality of Life"/ (171033)

25 exp Quality-Adjusted Life Years/ (10672)

26 (pharmacoeconomic* or pharmaco-economic* or economic* or cost* or price or prices or pricing).ti,ab,kf. (752907)

27 (expenditure$ not energy).ti,ab,kf. (27109)

28 (value adj1 money).ti,ab,kf. (32)

29 budget*.ti,ab,kf. (26932)

30 (health state* or health status).ti,ab,kf. (57854)

31 (qaly* or ICER or utilit* or EQ5D or EQ-5D or euroqol or euro-qol or short-form 36 or shortform 36 or SF-36 or SF36 or SF-6D or SF6D or SF-12 or SF12 or health utilities index or HUI).ti,ab,kf. (224115)

32 (markov or time trade off or TTO or standard gamble or SG or hrql or hrqol or disabilit* or disutilit* or net benefit or contingent valuation).ti,ab,kf. (215735)

33 (quality adj2 life).ti,ab,kf. (248124)

34 (decision adj2 model).ti,ab,kf. (6096)

35 (visual analog* scale* or discrete choice experiment* or health* year* equivalen* or (willing* adj2 pay)).ti,ab,kf. (54743)

36 resource*.ti,ab,kf. (294615)

37 (well-being or wellbeing).ti,ab,kf. (77269)

38 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 (2072673)

39 6 and 38 (1622)

40 20 and 38 (714)

41 39 and 40 (304)

42 39 or 40 (2032)

43 42 not 41 (1728)

Updated search, 13/03/2019

Re-ran above search with following date limits:

44 limit 42 to ed=20190122-20190313 (8)

45 limit 42 to ep=20190122-20190313 (17)

46 2019*.dt,ez. (265815)

47 42 and 46 (29)

48 44 or 45 or 47 (36)

Total after removing duplicates with previous search: 27

Embase (Ovid)

Searched on 22/01/2019 (updated on 13/03/2019, see at the end of this search record)

Database: Embase Classic+Embase <1947 to 2019 Week 03>

Original search, 22/01/2019

Search Strategy:

---

1 *streptococcal pharyngitis/ or *pharyngitis/ or *rhinopharyngitis/ or *sore throat/ or *tonsillitis/ or *chronic tonsillitis/ or *palatine tonsillitis/ (12255)

2 pharyngit*.ti,ab,kw. (7907)

3 (nasophyryngit* or rhinopharyngit* or epipharyngit*).ti,ab,kw. (381)

4 (tonsillit* or tonsilit*).ti,ab,kw. (8351)

5 ((sore or pain* or ache* or aching or inflam* or infect*) adj3 (pharyn* or throat* or tonsil* or nasopharyn* or rhinopharyn* or epipharyn*)).ti,ab,kw. (15999)

6 or/1-5 (32848)

7 Streptococcus infection/di (3828)

8 Streptococcus pyogenes/ or streptococcus group a/ or group A streptococcal infection/ (24060)

9 7 or 8 (27010)

10 ((strep or streptococcal or group) adj2 A).ti,ab,kw. (799616)

11 9 and 10 (9653)

12 (strep* adj5 (throat* or pharyn* or tonsil*)).ti,ab,kw. (4855)

13 streptoco* A.ti,ab,kw. (636)

14 (group A adj5 streptoco*).ti,ab,kw. (12259)

15 ((streptococcus or strep or staphylococcus) adj1 (pyogenes or pyogenic)).ti,ab,kw. (9749)

16 ((streptococcus or strep) adj1 (epidemicus or erysipelatis or erysipelatos or hemolyticus or haemolyticus or scarlatinae or lancefield)).ti,ab,kw. (391)

17 (s pyogenes or pyogenes s or micrococcus scarlatinae).ti,ab,kw. (3246)

18 lancefield group.ti,ab,kw. (566)

19 gabhs.ti,ab,kw. (507)

20 or/11-19 (24568)

21 exp health economics/ (803214)

22 exp health status/ (219256)

23 exp "quality of life"/ (447670)

24 exp quality adjusted life year/ (23005)

25 (pharmacoeconomic* or pharmaco-economic* or economic* or cost* or price or prices or pricing).ti,ab,kw. (986866)

26 (expenditure* not energy).ti,ab,kw. (37545)

27 (value adj2 money).ti,ab,kw. (2246)

28 budget*.ti,ab,kw. (35940)

29 (health state* or health status).tw. (75069)

30 (qaly* or ICER or utilit* or EQ5D or EQ-5D or euroqol or euro-qol or short-form 36 or shortform 36 or SF-36 or SF36 or SF-6D or SF6D or SF-12 or SF12 or health utilities index or HUI).ti,ab,kw. (321459)

31 (markov or time trade off or TTO or standard gamble or SG or hrql or hrqol or disabilit* or disutilit* or net benefit or contingent valuation).ti,ab,kw. (311593)

32 (quality adj2 life).tw. (384281)

33 (decision adj2 model).tw. (9229)

34 (visual analog* scale* or discrete choice experiment* or health* year* equivalen* or (willing* adj2 pay)).tw. (78125)

35 resource*.ti,ab,kw. (375642)

36 (well-being or wellbeing).tw. (99946)

37 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 (2880444)

38 6 and 37 (2459)

39 20 and 37 (1082)

40 38 and 39 (434)

41 38 or 39 (3107)

42 41 not 40 (2673)

Updated search, 13/03/2019

Re-ran above search with following date limits:

43 limit 41 to dd=20190122-20190313 (16)

44 limit 41 to em=201901-201903 (25)

45 43 or 44 (41)

46 limit 41 to dc=20190122-20190313 (42)

47 45 or 46 (67)

Total after removing duplicates with other update and previous searches: 25

NHS EED and HTA database (CRD)

Searched on 22/01/2019 (Not updated because no new records have been added to NHS EED since 31st March 2015 or to the HTA database since 31 March 2018. The INAHTA website was checked in March 2019 to see if a new platform for the HTA database was available).

Original search, 22/01/2019

1 MeSH DESCRIPTOR Pharyngitis EXPLODE ALL TREES IN DARE,NHSEED,HTA 73

2 (pharyngit*) 85

3 (nasophyryngit*) OR (rhinopharyngit*) OR (epipharyngit*) 5

4 (tonsillit* or tonsilit*) 43

5 (((sore or pain* or ache* or aching or inflam* or infect*) adj3 (pharyn* or throat* or tonsil* or nasopharyn* or rhinopharyn* or epipharyn*))) 91

6 #1 OR #2 OR #3 OR #4 OR #5 163

7 MeSH DESCRIPTOR Streptococcal Infections WITH QUALIFIERS DI, MI IN DARE,NHSEED,HTA 31

8 MeSH DESCRIPTOR Streptococcus pyogenes WITH QUALIFIERS IM, IP IN DARE,NHSEED,HTA 13

9 #7 OR #8 36

10 (((strep or streptococcal or group) adj2 A)) 2025

11 #9 AND #10 17

12 ((strep* adj5 (throat* or pharyn* or tonsil*))) 39

13 (streptoco* adj1 A) 10

14 ((group A adj5 streptoco*)) 27

15 (((streptococcus or strep or staphylococcus) adj1 (pyogenes or pyogenic))) 25

16 (((streptococcus or strep) adj1 (epidemicus or erysipelatis or erysipelatos or hemolyticus or haemolyticus or scarlatinae or lancefield))) 0

17 ((s pyogenes or pyogenes s or micrococcus scarlatinae)) 1

18 (lancefield group) 0

19 (gabhs) 8

20 #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 51

21 #6 AND #20 43

22 (#21) IN NHSEED, HTA 13

23 #6 OR #20 171

24 (#23) IN NHSEED, HTA 55

25 (#24 NOT #22) IN NHSEED, HTA 42

Science Citation Index and Conference Proceedings Science (Web of Science)

Searched on 29/01/2019 (updated on 13/03/2019, see at the end of this search record)

Original search, 29/01/2019

Note, search record reads from bottom to top.

# 20	1,397	#19 not #18 *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 19	1,657	#17 OR #16 *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 18	260	#17 AND #16 *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 17	709	#15 AND #14 *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 16	1,208	#15 AND #5 *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 15	3,164,661	TS=("quality of life" or qol or hrql or hrqol or ("quality adjusted life" NEAR/0 year*) or qaly* or icer or cost* or economic* or pharmacoeconomic* or pharmaco-economic* or price or prices or pricing or (expenditure* not energy) or (value NEAR/1 money) or budget* or euro-qol or utilit* or disutilit* or (net NEAR/0 benefit*) or (contingent NEAR/0 valuation*) or euroqol or "euro qol" or eq5d or eq-5d or "short-form 36" or "shortform 36" or sf-36 or sf36 or sf-6d or sf6d or sf-12 or sf12 or "health utilities index" or hui or (time NEAR/0 trade*) or tto or "standard gamble" or sg or markov or (decision NEAR/1 model*) or (visual NEAR/0 analog*) or "discrete choice" or ((health* NEAR/0 year*) NEAR/0 equivalen*) or (health NEAR/0 stat*) or (willing* NEAR/1 pay) or resource* or wellbeing or well-being) *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 14	17,381	#6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 13	308	TS=gabhs *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 12	445	TS="lancefield group" *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 11	2,059	TS=("s pyogenes" OR "pyogenes s" OR "micrococcus scarlatinae") *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 10	60	TS=((strep*) NEAR/0 (epidemicus OR erysipelatis OR erysipelatos OR hemolyticus OR haemolyticus OR scarlatinae OR lancefield)) *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 9	7,163	TS=((strep*) NEAR/0 (pyogenes OR pyogenic)) *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 8	9,682	TS=("group A" NEAR/4 strep*) *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 7	1,165	TS="strep* A" *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 6	2,877	TS=(strep* NEAR/4 (throat* OR pharyn* OR tonsil*)) *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 5	12,508	#1 OR #2 OR #3 OR #4 *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 4	7,034	TS=((sore OR pain* OR ache* OR aching OR inflam* OR infect*) NEAR/2 (pharyn* OR throat* OR tonsil* OR nasopharyn* OR rhinopharyn* OR epipharyn*)) *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 3	2,716	TS=(tonsillit* OR tonsilit*) *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*

# 2	96	TS=(nasophyryngit* OR rhinopharyngit* OR epipharyngit*) *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*
# 1	4,667	TS=pharyngit* *Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2019*

Updated search, 13/03/2019

Re-ran above search with following date limits:

# 19	17	#17 or #16 *Indexes=SCI-EXPANDED, CPCI-S Timespan=2019-2019*

Total after removing duplicates with other update and previous searches: 6

CEA Registry

Searched on 29/01/2019 (updated on 13/03/2019, see at the end of this search record)

Original search, 29/01/2019

Single term searches, de-duplicated and screened online. Results (number selected):

pharyngitis	6 (3 selected)
pharynx	4 (0)
nasopharyngitis	0 (0)
nasopharynx	0 (0)
rhinopharyngitis	0 (0)
rhinopharynx	0 (0)
epipharyngitis	0 (0)
epipharynx	0 (0)
tonsillitis	0 (0)
tonsilitis	0 (0)
tonsil	3 (1, already got from search on pharyngitis above)
throat	6 (2, both already got from search on pharyngitis above)
streptococcus	22 (2, both already got from search on pharyngitis above)
streptococcal	7 (2, both already got from search on pharyngitis above)
strep	30 (3, both already got from search on pharyngitis above)

Potentially relevant downloaded to EndNote: 3

Updated search, 13/03/2019

Re-ran above searches on 13/03/2019. No further records added

EconPapers (RePEc)

Searched on 30/01/2019 (updated on 13/03/2019, see at the end of this search record)

Original search, 30/01/2019

Advanced Search

#1

((pharyn* | nasopharyn* | rhinopharyn* | epipharyn* | tonsil* | throat) + (strep* | "lancefield group" | pyogenes | micrococcus)) 6

#2

((sore | pain* | ache* | aching | inflam* | infect*) + (throat | pharyn* | tonsil* | nasopharyn* | rhinopharyn* | epipharyn*)) 31

#3

pharyngitis | nasopharyngitis | rhinopharyngitis | epipharyngitis | tonsilitis | tonsillitis 28

Above three searches combined with OR (|):

#4

((pharyn* | nasopharyn* | rhinopharyn* | epipharyn* | tonsil* | throat) + (strep* | "lancefield group" | pyogenes | micrococcus)) | ((sore | pain* | ache* | aching | inflam* | infect*) + (throat | pharyn* | tonsil* | nasopharyn* | rhinopharyn* | epipharyn*)) | (pharyngitis | nasopharyngitis | rhinopharyngitis | epipharyngitis | tonsilitis | tonsillitis) 52

De-duplicated and screened online, selecting all potentially relevant.

Potentially relevant downloaded to EndNote: 6

Updated search, 13/03/2019

Re-ran above combination search on 13/03/2019. One further record added, but this was not relevant.

ScHARRHUD

Searched on 30/01/2019 (updated on 13/03/2019, see at the end of this search record)

Original search, 30/01/2019

Single term searches, de-duplicated and screened online. Results (number selected):

pharynx*	0 (0)
nasopharyn*	0 (0)
rhinopharyn*	0 (0)
epipharyn*	0 (0)
tonsil*	0 (0)
throat*	1 (0)
strep*	0 (0)
gapbs	0 (0)
pyogene*	0 (0)

Potentially relevant downloaded to EndNote: 0

Updated search, 13/03/2019

Re-ran above searches on 13/03/2019. No further records added.

Other sources

In addition to these searches, any relevant cost-effectiveness studies identified during the clinical effectiveness review were brought to the attention of the reviewers.

Search Engine

Google, searched 22/03/2019

(HTA OR "health technology assessment") AND (pharyngitis OR strep OR streptococcus OR streptococcal)

Checked first 20 records

Appendix 6: Excluded studies after full text papers received for Strep A economics search

Author (Year) Journal	Title	Reason for exclusion
Banerjee and Ford (2018) Canadian Agency for Drugs and Technologies in Health. CADTH Rapid Response Reports 05: 31.[91]	Rapid Tests for the Diagnosis of Group A Streptococcal Infection: A Review of Diagnostic Test Accuracy, Clinical Utility, Safety, and Cost-Effectiveness	The review provides information on two cost-effectiveness studies. One study has been included [73] and the other study was excluded as it is not an economic evaluation* and the test is outside the NICE scope [92]
Benjamin (2000) Archives of Pediatrics & Adolescent Medicine. 154(1): 93-94. [93]	The costs of testing for streptococcal pharyngitis in the office laboratory	Letter to editor commenting on Tsevat and Kotagal [94] Not an economic evaluation*
Bovier et al (2002) Journal of General Internal Medicine 17: 135-136. [95]	A cost-effectiveness analysis of recommended strategies for acute pharyngitis	Abstract Test outside NICE scope
Ehrlich et al (2002) Preventive Medicine 35(3): 250-257.[96]	Cost-effectiveness of treatment options for prevention of rheumatic heart disease from Group A streptococcal pharyngitis in a pediatric population	No specific test stated
Giraldez-Garcia et al (2011) European Journal of Pediatrics 170(8): 1059-1067 [97]	Diagnosis and management of acute pharyngitis in a paediatric population: a cost-effectiveness analysis	No specific test stated
Klepser et al (2011) Journal of Managed Care Pharmacy 17 (3): 241 [98]	Cost-effectiveness of pharmacist provided care for the treatment of adult pharyngitis	Abstract No specific test stated
Klepser et al (2012) American Journal of Managed Care 18 (4): 145-154 [77]	Cost-effectiveness of pharmacist provided care for the treatment of adult pharyngitis	No specific test stated.
Komaroff et al (1983) Clinical Research 31(2): A299-A299 [99]	A cost-effectiveness analysis of alternate strategies for management of sore throat	Abstract No specific test stated
Lathi et al (2018) Canadian Pharmacists Journal 151(5): 322-331 [100]	Cost-minimization analysis of community pharmacy-based point-of-care testing for strep throat in 5 Canadian provinces	No specific test stated

Author (Year) Journal	Title	Reason for exclusion
Maizia et al (2012) Presse Medicale 41(4): e195-203 [101]	Diagnostic strategies for acute tonsillitis in France: A cost-effectiveness study	Not English (in French) No specific test stated.
Malecki et al (2017) Pediatria Polska 92(2): 149-155 [102]	Rapid strip tests as a decision-making tool about antibiotic treatment in children - A prospective study	Not an economic evaluation* No comparator
Meier et al (1990) Archives of Internal Medicine 150(8): 1696-1700 [103]	Effects of a rapid antigen test for group A streptococcal pharyngitis on physician prescribing and antibiotic costs	No specific test stated
Mlejnek et al (2014) Academic Emergency Medicine 1): S51 [104]	Utility and cost effectiveness of throat culture in the treatment of patients with negative rapid strep screens	No specific test stated
Neuner et al (2003) Annals of Internal Medicine 139(2): 113-122 [75]	Diagnosis and management of adults with pharyngitis. A cost-effectiveness analysis	Test outside NICE scope
Polisena and Spry (2009) CADTH Health Technology Inquiry Service [105]	Point of care testing for streptococcal sore throat: a review of diagnostic accuracy, cost-effectiveness, and guidelines	The review provides information on one cost-effectiveness study which was excluded as it did not mention a specific test [76]
Tsevat and Kotagal (1999). Archives of Pediatrics & Adolescent Medicine 153(7): 681-688 [94]	Management of sore throats in children: a cost-effectiveness analysis	No specific test stated
Van Howe and Kusnier (2006) Pediatrics 117(3): 609-619 [76]	Diagnosis and management of pharyngitis in a pediatric population based on cost-effectiveness and projected health outcomes	No specific test stated

*Not looking at incremental costs and incremental benefits

Appendix 7: Additional sensitivity analyses

Table 65: List of additional sensitivity analyses

SA	Description of sensitivity analysis	Model worksheet and cell	Updated input parameter
0	Base case	None	
1	Changed Centor threshold for starting antibiotics from ≥3 (base case) to ≥2	Settings, B6	2
2	Changed Centor threshold for starting antibiotics from ≥3 (base case) to ≥1	Settings, B6	1
3	**Changed time horizon 14 days**	**Settings, B7**	**14**
4	Changed strep A prevalence (adults) from 22.6% to 35.9% (upper value reported in studies included in the test accuracy systematic review)	Settings, B21	0.359
5	Changed strep A prevalence (adults) from 22.6% to 10% (Neuner et al 2003)	Settings, B21	0.1
6	Delayed prescription rate set to 27.3% in both arms (RADT group, Little 2013a)	Settings, B25	0.273
7	Delayed prescription rate set to 51% in both arms (clinical score group, Little 2013a)	Settings, B26	0.51
8	Doubled proportion who use delayed antibiotics to 92%	Settings, B27	0.92
9	Halved probability of strep A complications given antibiotics from 0.013 (Little 2012) to 0.0065 (analyst assumption)	Settings, B30	0.0065
10	Doubled probability strep A complication given antibiotics from 0.013 (Little 2012) to 0.026 (analyst assumption)	Settings, B30	0.026
11	Halved probability of strep A complications given no antibiotics from 0.015 (Little 2012) to 0.0075 (analyst assumption)	Settings, B31	0.0075
12	Doubled probability of strep A complications given no antibiotics 0.015 (Little 2012) to 0.03 (analyst assumption)	Settings, B31	0.03
13	Halved prob. rheumatic fever to 0.00005	Settings, B32	0.00005
14	Increase prob. rheumatic fever 10-fold to 0.001	Settings, B32	0.001
15	Halved mild penicillin reaction (rash) to 0.01	Settings, B36	0.01
16	Doubled mild penicillin reaction (rash) to 0.04	Settings, B36	0.04
17	Changed prob. anaphylaxis from 0.0001 (Neuner 2003) to 0.0064 (Van Howe Kusnier 2006)	Settings, B37	0.0064
18	Changed cost of antibiotics from £0.74 (BNF 2017, 15 capsules amoxy 500mg) to £6.11 (NG51 costing report)	Settings, B47	6.11
19	Assume patient is seen by practice nurse(£62/hr PSSRU UC 2017, section 10.1) instead of doctor	Settings, B49	1.03
20	Assume no Swab culture in those with a negative test result	Settings, B50	0
21	Double the cost of alternative antibiotic in those with penicillin-induced rash to £20	Settings, B51	20
22	Assume testing within standard GP time	Settings, B9	Yes
23	Doubled cost of anaphylaxis to £3,489.28	Settings, B53	£3,489.28
24	Doubled cost of abscess to 3142.56	Settings, B54	3142.56
25	Doubled cost of acute rheumatic fever to £3,544.88	Settings, B55	£3,544.88
26	Changed baseline utility from 0.863 (UK norm) to 0.6305 (PRISM study, Table 17)	Settings, B58	0.6305
27	Halved utility decrement, untreated strep A	Settings, B59	0.125

28	Doubled utility decrement, untreated strep A	Settings, B59	0.5
29	Halved utility decrement, treated strep A	Settings, B60	0.075
30	Doubled utility decrement, treated strep A	Settings, B60	0.3
31	Halved utility decrement, strep A related abscess	Settings, B61	2.5
32	Doubled utility decrement, strep A related abscess	Settings, B61	10
33	Halved utility decrement, acute rheumatic fever	Settings, B62	38.25
34	Doubled utility decrement, acute rheumatic fever	Settings, B62	153
35	Halved utility decrement, penicillin-induced rash	Settings, B63	0.3125
36	Doubled utility decrement, penicillin-induced rash	Settings, B63	1.25
37	Halved utility decrement, strep A related sepsis	Settings, B64	4.5
38	Doubled utility decrement, strep A related sepsis	Settings, B64	18
39	Lower confidence limits of test accuracy	TestAccuracy	
40	Upper confidence limits of test accuracy	TestAccuracy	

Appendix 8: Summary of manufacturer information

Biopanda

1. Checklist of confidential information
2. Product Insert: Strep A Rapid Test RAPG-STRA-001
3. Declaration of Conformity DOCSTRA1826
4. Response to Request for Information

Cepheid

1. Package Insert: Xpert Xpress Strep A XPRSTREPA-CE-10
2. CE Declaration of Conformity
3. The GeneXpert System. CE-IVD Test Menu 2
4. The GeneXpert System. CE-IVD Test Menu
5. Patricia Ferrieri, Kari Nelson, Elizabeth Thonen-Kerr, Sophie Arbefeville; Prospective Evaluation of Xpert Xpress Strep A Automated PCR Assay vs Solana Group A Streptococcal NAAT vs Conventional Throat Culture, *American Journal of Clinical Pathology*, Volume 150, Issue suppl_1, 21 September 2018, Pages S157, https://doi.org/10.1093/ajcp/aqy112.367
6. Matthys J, De Meyere M, van Driel ML, De Sutter A. Differences among international pharyngitis guidelines: not just academic. *Ann Fam Med*. 2007;5(5):436-43.
7. Response to Request for Information
8. Xpert Xpress Strep A brochure CEIVD 3106-01.A
9. Xpert Xpress StrepA Datasheet CEIVD 3105-01

Nal von minden

1. Gazzano V, Berger A, Benito Y, Freydiere A-M, Tristan A, Boisset S, et al. Reassessment of the Role of Rapid Antigen Detection Tests in Diagnosis of Invasive Group A Streptococcal Infections. Journal of Clinical Microbiology 2016;54:994. http://dx.doi.org/10.1128/JCM.02516-15

NADAL Strep A

1. EC-Declaration of Conformity for product number 221002A – signed 30.01.2017
2. EC-Declaration of Conformity for product number 221002A – signed 09.02.2017
3. EC-Declaration of Conformity for product number 222008 – signed 28.07.2017
4. Instructions for use for NADAL Strep A Test (test strip), Ref 221001A, version 2.2, 2017-08-11
5. Instructions for use for NADAL Strep A Test (test cassette), Ref 222001A, version 2.3, 2017-10-24
6. Checklist of confidential information. For Test Strip
7. Checklist of confidential information. For Cassette.
8. Response to Request for Information NADAL strep A cassette
9. Response to Request for Information NADAL strep A test strip

NADAL Strep A plus

1. EC-Declaration of Conformity. Product number 221050N-50
2. Instructions for use for NADAL Strep A plus Test (test strip) 221050N-50
3. Instructions for use for NADAL Strep A plus Test (test cassette) 222007
4. Instructions for use for NADAL Strep A plus Test (test cassette) 222008
5. Checklist of confidential information. For test strip.

6. Checklist of confidential information. For test cassette.
7. Response to Request for Information NADAL strep A+ cassette
8. Response to Request for Information NADAL strep A+ test strip

NADAL Strep A Scan

1. EC-Declaration of Conformity. Product number 222049NBUL-20
2. Instructions for use for NADAL Strep A Scan Test (test cassette) 222049NBUL-20
3. Checklist of confidential information. For NADAL Strep A scan (cassette)
4. Response to Request for Information NADAL strep A Scan (cassette)

Orion Diagnostica

1. Shallcross, Laura J and Dame Sally C Davies. "Antibiotic overuse: a key driver of antimicrobial resistance" *British journal of general practice : the journal of the Royal College of General Practitioners* vol. 64,629 (2014): 604-5.
2. Checklist of confidential information. 10122018
3. Clinical impact of rapid POC test for acute sore throat Poster ECCMID 2016 http://www.oriondiagnostica.com/globalassets/documents-and-materials/quikread-go/quikread-go-strep-a/9031_clinical_impact_of_rapid_poc_tests_for_accute_sore_throat_eccmid_2016_a3_web.pdf
4. Response to Request for Information
5. Declaration of Conformity for QuikRead go Strep A System and QuiRead go Strep A Cat no 135883
6. Instructions for use QuikRead go Strep A. 136262-3
7. Poster ESPID 2013
8. QuikRead go Strep A – An evaluation of performance in comparison to Alere TestPack+Plus with OBC, by Oulun Työterveys laboratory
9. Evaluation of QuikRead go Strep A test regarding the detection level of Streptococcus pyogenes, by Pia Karlsson at Microbiology laboratory of Medicinsk Diagnostik, Jönköping, Sweden
10. Stefaniuk E, Bosacka K, Wanke-Rytt M, Hryniewicz W. The use of rapid test QuikRead go(R) Strep A in bacterial pharyngotonsillitis diagnosing and therapeutic decisions. *Eur J Clin Microbiol Infect Dis* 2017;**36**:1733-8. http://dx.doi.org/10.1007/s10096-017-2986-8
11. The report from Scandinavian evaluation of laboratory equipment for primary health care (SKUP) on QuikRead go Strep A

Roche

1. Declaration of Conformity DOC-2017-38
2. Cobas Strep A – Nucleic acid test for use on the cobas Liat System - Package Insert
3. Response to Request for Information
4. Checklist of confidential information