Rapid antigen detection and molecular tests for group A streptococcal infections for acute sore throat: systematic reviews and economic evaluation

Hannah Fraser, Daniel Gallacher, Felix Achana, Rachel Court, Sian Taylor-Phillips, Chidozie Nduka, Chris Stinton, Rebecca Willans, Paramjit Gill and Hema Mistry*

Warwick Medical School, University of Warwick, Coventry, UK

*Corresponding author Hema.Mistry@warwick.ac.uk

Declared competing interests of authors: Sian Taylor-Phillips, Chris 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 University Hospitals Birmingham NHS Foundation Trust. Sian Taylor-Phillips is funded by a NIHR Career Development fellowship (reference CDF-2016-09-018).

Published June 2020
DOI: 10.3310/hta24310

Scientific summary

Antigen detection and molecular tests for strep A infection

Health Technology Assessment 2020; Vol. 24: No. 31
DOI: 10.3310/hta24310

NIHR Journals Library www.journalslibrary.nihr.ac.uk

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, which are both infections of the upper respiratory airway affecting the mucosa. Most cases are viral; however, a small number of these infections may be caused by the group A *Streptococcus* bacterium. Most sore throats resolve spontaneously within a few weeks. An analysis of UK primary care use 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 use.

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 streptococcal infection, to 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 of ≥ 3 points, FeverPAIN scores of ≥ 4 points); these include primary and secondary care, and community pharmacies.

Decision question

The decision problem for this assessment is what is the clinical effectiveness and cost-effectiveness of rapid antigen detection and molecular tests in patients with high clinical scores (Centor scores of ≥ 3 points, FeverPAIN scores of ≥ 4 points), compared with 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?

Objectives

To systematically review the evidence for the clinical effectiveness of 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 with 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 reviews and the cost-effectiveness reviews.

The following databases were searched in November and December 2018 and searches were updated in March 2019: MEDLINE [via OvidSP (Health First, Rockledge, FL, USA)], MEDLINE In-Process & Other Non-Indexed Citations (via OvidSP), MEDLINE Epub Ahead of Print (via OvidSP), MEDLINE Daily Update (via OvidSP), EMBASE (via OvidSP), Cochrane Database of Systematic Reviews [via Wiley Online Library (John Wiley & Sons, Inc., Hoboken, NJ, USA)], Cochrane Central Register of Controlled

Trials (CENTRAL) (via Wiley Online Library), Database of Abstracts of Reviews of Effects (DARE) (via the Centre for Reviews and Dissemination), Health Technology Assessment database (via the Centre for Reviews and Dissemination), Science Citation Index and Conference Proceedings [via the Web of Science™ (Clarivate Analytics, Philadelphia, PA, USA)] and the PROSPERO International Prospective Register of Systematic Reviews (via the Centre for Reviews and Dissemination).

Supplementary searches were used to identify additional published and unpublished studies. Reference lists of the 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 presenting with symptoms of an acute sore throat.
- Intervention – point-of-care tests for group A *Streptococcus* (including rapid antigen detection tests and molecular tests), preferably in those identified as being at high risk.
- Comparator – antibiotic-prescribing decisions using clinical scoring tools for group A *Streptococcus*, such as FeverPAIN or Centor/modified Centor (McIsaac) alone.
- Reference standard – microbiological culture.
- Outcomes – any patient-related outcome, test accuracy (the ability of a test to correctly differentiate between people who do and people who do not have a disease) or performance, prescribing behaviour and cost-effectiveness estimates.
- Study design – clinical test accuracy studies that compare the index tests and/or FeverPAIN/Centor/McIsaac scores with 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 that compared the index test and/or clinical scores (i.e. FeverPAIN/Centor/McIsaac) with culture was eligible. For cost-effectiveness, any full economic evaluations (or economic models) reporting both cost and outcome estimates were eligible.
- Health-care setting – primary care (general practice 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 was synthesised using narrative review and statistical methods where appropriate. Meta-analyses were undertaken in Stata® version 15 (StataCorp LP, College Station, TX, USA).

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, Joanna Briggs Institute 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® (Microsoft Corporation, Redmond, WA, USA) to assess the cost-effectiveness of rapid antigen detection and molecular tests in conjunction with clinical scoring tools, compared with the use of clinical scoring tools alone. The base-case economic model included adult patients who were seen in primary care with suspected group A streptococcal infection. The base-case model was adapted to look at the following subgroups with suspected group A streptococcal 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 quality-adjusted life-years for each rapid antigen detection test and molecular test over a 1-year time horizon, and adopted an NHS and Personal Social Services perspective. Costs were in 2017/18 prices. No discounting of costs and outcomes was undertaken. 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 were also undertaken (1000 model runs).

Results

The searches identified 5919 records of clinical effectiveness and 6980 records 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 scope and/or clinical scores with biological cultures as a reference standard. These comprised 26 full-text articles, three abstracts, five manufacturer's submissions (submitted directly to the National Institute for Health and Care Excellence in response to a request for information) and four US Food and Drug Administration documents. There were 26 studies (23 full-text articles and three abstracts) that reported test accuracy data. The methodological quality of the included studies was poor. In particular, in 65.4% (17/26) of studies it was unclear whether the sample was consecutive or convenience. Convenience samples may not provide a true representation of the prevalence of group A streptococci. There was judged to be a high level of bias concerning methods of patient selection. Overall, the findings reveal variations in the sensitivity (0.679 to 1.00) and specificity (0.733 to 1.00) estimates of point-of-care tests. These point estimates were 0.829 to 0.946 for sensitivity and 0.849 to 0.991 for specificity in high-risk populations, including patients with Centor/McIsaac scores of > 2 points, representing the population of interest. These estimates do not account for any of the unpublished manufacturer submissions.

Direct comparison with sore throat clinical scoring tools revealed that point-of-care tests were generally more specific. However, one methodological limitation concerns the varying way that clinical scoring tools have been implemented across the included studies. For instance, different studies apply different clinical score cut-off points 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 National Institute for Health and Care Excellence thresholds (Centor/McIsaac score of ≥ 3 points or FeverPAIN score of ≥ 4 points) 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 over-representation of the Alere™ TestPack +Plus (Abbott Laboratories, Abbott Park, IL, USA) group A streptococcal 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 owing to the lack of available evidence. There was a large degree of heterogeneity among results for studies using the same rapid test. Where a test is reported in several studies, its accuracy may appear lower than that of tests reported in only a single study, particularly those at high risk of bias or with unpublished methods, so we report on number 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. Owing to the potential heterogeneity, estimates for the sensitivity and specificity of each test were stratified by age group, throat score and care setting, although a lack of evidence meant that generalisations had to be made for the majority of estimates.

Of the scoped secondary outcomes, our search identified only studies discussing antibiotic-prescribing rates and appropriateness ($n = 12$). There is some randomised controlled trial evidence to suggest

that the use of rapid antigen detection tests may help to 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 the number of appointments required per episode, morbidity, mortality, onward transmission of infection, health-related quality of life, patient satisfaction with the test or health-care 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 National Institute for Health and Care Excellence scope, and were classed as economic evaluations. Two studies had some notable limitations and data could not be fully 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.

Fourteen of the 21 tests listed in the National Institute for Health and Care Excellence's scope had data on test accuracy and costs that were relevant to be included in the final economic modelling. In the base-case analysis, which included adult patients seen in primary care with suspected group A streptococcal infection, the economic model found considerable uncertainty about the cost-effectiveness of the different point-of-care tests for suspected group A streptococcal infection. This finding was also seen in the other economic models that 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 (1) 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 (2) penalty for antibiotic overprescription/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 in 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. Although 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 strategy (literature search, data extraction and analysis) 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. In addition, 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 National Institute for Health and Care Excellence thresholds (i.e. Centor/McIsaac score of ≥ 3 points or FeverPAIN score of ≥ 4 points) 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 settings and secondary care settings). The modelling may have underestimated the costs as we did not take into account the different strains of group A *Streptococcus*, which may have influenced test performance and altered the

profile of complications, seasonality of group A streptococcal infection and the onward transmission of the infection.

Implications for health care

Our findings indicate that point-of-care testing was 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 that was available to us. Health-care 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 that 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.

Study registration

The protocol of the review is registered as PROSPERO CRD42018118653.

Funding

This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 24, No. 31. See the NIHR Journals Library website for further project information.

Health Technology Assessment

NICE TAR and DAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 3.819

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, the Cochrane Library and Clarivate Analytics Science Citation Index.

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

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

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

Criteria for inclusion in the *Health Technology Assessment* journal

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

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

HTA programme

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

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

This report

The research reported in this issue of the journal was funded by the HTA programme as project number NIHR127666. The contractual start date was in November 2018. The draft report began editorial review in June 2019 and was accepted for publication in November 2019. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

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

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