Cost-effectiveness of non-invasive methods for assessment and monitoring of liver fibrosis and cirrhosis in patients with chronic liver disease: systematic review and economic evaluation

Catriona Crossan,[1] Emmanuel A Tsochatzis,[2] Louise Longworth,[1]* Kurinchi Gurusamy,[3] Brian Davidson,[3] Manuel Rodríguez-Perálvarez,[2] Konstantinos Mantzoukis,[2] Julia O'Brien,[2] Evangelos Thalassinos,[2] Vassilios Papastergiou[2] and Andrew Burroughs[2]

[1]Health Economics Research Group, Brunel University London, Uxbridge, UK
[2]Sheila Sherlock Liver Centre, Royal Free Hospital and UCL Institute for Liver and Digestive Health, Royal Free Hospital, London, UK
[3]Royal Free Campus, UCL Medical School, London, UK

*Corresponding author

Declared competing interests of authors: none

Published January 2015
DOI: 10.3310/hta19090

Scientific summary

Methods for assessment and monitoring of liver fibrosis and cirrhosis

Health Technology Assessment 2015; Vol. 19: No. 9
DOI: 10.3310/hta19090

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

Scientific summary

Background

In 2011, it was estimated that 2.3 million people, or approximately 5% of the population of England, had liver disease. Currently, liver biopsy is used in patients with suspected liver disease to determine the extent of liver fibrosis and to help inform treatment decisions. However, biopsy is an invasive procedure associated with morbidity and mortality risks. Alternatives to liver biopsy include non-invasive liver tests (NILTs) which can be serum tests or imaging modalities and have in many cases replaced liver biopsy in clinical practice. As liver biopsy is high risk and costly, NILTs may offer cost-effective alternatives.

Objectives

There were two related objectives for the study:

1. to determine the diagnostic accuracy of different NILTs in the diagnosis and monitoring of liver fibrosis and cirrhosis in patients with various aetiologies for chronic liver disease; and
2. to estimate the incremental cost-effectiveness of the NILTs.

Methods

A systematic review was conducted to identify studies which reported the diagnostic accuracy of NILTs used for the identification of liver fibrosis in patients with various causes of liver disease. The causes of liver disease included hepatitis C (HCV), hepatitis B (HBV), alcoholic liver disease (ALD) and non-alcoholic fatty liver disease (NAFLD).

The following databases were searched: MEDLINE (PubMed), EMBASE, Science Citation Index Expanded, Bioscience Information Service (BIOSIS), Cochrane Central Register of Controlled Trials (CENTRAL), Latin American and Caribbean Health Sciences Literature (LILACS) and Cumulative Index to Nursing and Allied Health Literature (CINAHL). The search was conducted from 1998 to April 2012 for all databases. Reference lists of identified studies and reviews, and conference proceedings from recent conferences, were hand-searched to identify further studies.

Data from relevant studies were reviewed by two independent reviewers using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool.

Studies were included if they reported the diagnostic accuracy of non-invasive tests using liver biopsy as the reference standard. Studies were excluded if the time difference between the tests was greater than 6 months.

Decision-analytic models were developed to assess the cost-effectiveness of the NILTs. Health outcomes were measured using quality-adjusted life-years (QALYs) and took into account the long-term consequences of test results where possible. Costs were estimated from a NHS perspective. Fully incremental analyses were conducted. Separate models were constructed for each of the four causes of liver disease included in the systematic review. Two models were constructed for HBV representing the different disease progression and epidemiology for patients with hepatitis B e antigen (HBeAg)-positive and HBeAg-negative. Additionally, an analysis of the NILTs in the diagnosis of cirrhosis (irrespective of cause) was conducted.

For HBV and HCV, the analysis reflected a use of the tests to determine when patients should receive antiviral treatment. The NILTs were compared with each other as single tests; combinations of NILTs using four alternative strategies; biopsy; a strategy of treating all with suspected fibrosis; and no testing or treatment. Markov models were developed to estimate the long-term outcomes of each test result.

For ALD, there are no specific treatments initiated as result of the diagnosis of liver fibrosis. These patients are advised to abstain from alcohol intake. For our cost-effectiveness analysis of the NILTs, we hypothesised that people diagnosed with cirrhosis would be more likely to abstain from drinking and estimated the benefit to long-term health outcomes and associated costs.

It was not possible to conduct the same analysis for NAFLD as there are currently no specific therapeutic interventions which depend on the degree of fibrosis. Instead, we conducted a cost per correct diagnosis which allowed us to identify the incremental cost associated with an additional correct diagnosis for each test. The results are presented separately for correct positive and negative diagnoses. We also conducted an exploratory analysis for NAFLD to assess the impact of using the non-invasive tests to determine referral of patients to tertiary care for treatment and monitoring.

Results

Results of the systematic review

During the search, 114,071 studies were found and, after review, 302 of these papers were deemed suitable. The highest number of studies identified was for HCV, and ALD had the lowest number.

Data from tests that converged using the bivariate model are more robust; however, despite the vast amount of literature, very few tests' results converged. In HBV, only five tests had a robust evidence base. There were no tests in ALD where the bivariate model converged; therefore, the use of NILTs in such patients for treatment decision is not yet proven. In patients with NAFLD, the evidence base is slightly larger than for ALD; for the diagnosis of F3 (using Kleiner score), five tests converged. HCV has the highest number of tests where the bivariate model converged (14 NILTs). The findings show that the evidence base for many NILTs is not yet proven and further studies are required.

Diagnostic threshold cut-offs of NILTs to determine specific fibrosis stages were not always predetermined or sufficiently validated; this represents a significant limitation in the interpretation of their results. Among all NILTs, aspartate aminotransferase (AST) to platelet ratio index (APRI) (low and high cut-offs) had established cut-offs which were almost universally used in published studies.

Fibroscan (Echosens, Paris) was the NILT assessed in most studies across diseases aetiologies (37 studies in HCV, 13 in HBV, eight in NAFLD and six in ALD). APRI was also widely assessed in HBV and HCV but not in NAFLD or ALD.

The methodological quality of included studies as assessed by the QUADAS-2 tool was poor; only 5 of the 302 studies (1.6%) were of high methodological quality. The most common causes were that diagnostic threshold cut-offs were not predetermined and liver biopsy samples were not of adequate length or did not have a sufficient number of portal tracts for reliable staging. Therefore, all reported results are likely biased.

Results of the economic evaluation

Using a cost-effectiveness threshold of £20,000, the results from the analysis for HBV suggests that for people with HBeAg-positive disease, using two non-invasive tests together [first NILT-hyaluronic acid, with magnetic resonance (MR) elastography to confirm positive results] is cost-effective with an incremental cost-effectiveness ratio (ICER) of £19,612. There was, however, a substantial amount of uncertainty around

© Queen's Printer and Controller of HMSO 2015. This work was produced by Crossan *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

this result and the probability that it would have the highest expected net benefit is < 5% and several combinations of tests had similar costs and outcomes.

The results for the HBeAg-negative analysis differed from those for HBeAg-positive disease and found that treating all patients suspected of fibrosis without prior testing for the extent of fibrosis was the most cost-effective option only if the upper bound of the standard UK cost-effectiveness threshold range of £30,000 is considered acceptable (mean ICER: £28,137), with a probability having the highest expected net benefit of 38%. The reasons for the difference in results between HBeAg-positive and HBeAg-negative are due to the underlying characteristics for both groups; the HBeAg-negative population tend to be older with a higher proportion of males who have a higher all-cause mortality risk than females.

Treating patients with HCV regardless of the degree of fibrosis was the most cost-effective option given a cost-effectiveness threshold value of £20,000. The results imply that there is no necessity for a diagnostic test in patients with HCV to determine fibrosis stage. This concurs with current National Institute for Health and Care Excellence guidance for HCV, which recommends early treatment in all patients with mild chronic HCV rather than waiting for disease progression to fibrosis. This finding was robust to a number of sensitivity analyses. It was sensitive to the size of treatment effect in people with very mild disease, but remained the cost-effective option provided that the benefit to these patients is at least approximately 75% of those with more severe fibrosis.

In patients with ALD, abstinence is usually recommended. The analysis of the tests for ALD was limited as there are few data available on whether or not abstinence rates are influenced by the diagnosis of liver fibrosis. It has been theorised that liver biopsy, due to its invasive nature, may encourage abstinence in more people than non-invasive tests. We incorporated this assumption into our analysis and the base-case results indicated that liver biopsy was the most effective test to use in patients with ALD; however, the conclusions were sensitive to some assumptions including differential abstinence rates, which led to non-invasive testing becoming the cost-effective option.

The analysis of the incremental cost per correct positive diagnosis for NAFLD found that most of the tests were dominated or extendedly dominated by liver biopsy; however, hyaluronic acid had an ICER of £1.27 and NAFIC (ferritin, fasting insulin, type IV collagen) (low cut-off) had an ICER of £1.29. The analysis found that it costs an additional £112.30 to obtain an additional correctly diagnosed positive result from biopsy, compared with these NILTs. The analysis of the incremental cost per correct negative diagnosis found that FIB-4 (high cut-off) and non-alcoholic fatty liver disease fibrosis score (NFS) (high cut-off) had ICERs of below £1, the ICERs for NFS enhanced liver fibrosis test was £5.72 and for biopsy was £145.39. Whether or not the ICERs for the biopsy represent good value for money is difficult to judge as there are no established cost-effectiveness thresholds for this measure.

The analysis of the NILTs in people with cirrhosis found that the most cost-effective NILT to select patients for intensive hepatocellular cancer surveillance and monitoring was Forns index. This test has an ICER of £2032 per additional QALY gained and, if the cost-effectiveness threshold is set at £20,000, is 50% likely to be the optimal test.

Discussion

We have comprehensively assessed the evidence on the accuracy of the non-invasive tests for liver fibrosis and the economic implications of using them routinely within a NHS setting for a variety of aetiologies. In some cases, such as for HCV, the results suggest that early treatment without the need for fibrosis staging is cost-effective. In other cases, such as for HBeAg-positive disease, the NILTs (single or in combination) may be more appropriately used to determine treatment; however, several of these tests have very similar long-term expected mean health and cost outcomes.

Given the robustness of the data, the results must be approached conservatively. Most studies had a high risk of bias; therefore, reported results might be biased. Moreover, reported cut-offs for specific fibrosis stages were seldom predetermined and in most cases insufficiently validated. In addition, a considerable number of NILTs were tested in only one or a very few studies in specific disease aetiologies, most notably HBV and ALD; therefore, further studies are needed to assess their diagnostic accuracy.

Furthermore, as some NILTs from the direct tests and imaging modalities categories are not yet widely available, they cannot be universally applied. Further high-quality research is required on the diagnostic accuracy of the tests.

As the diagnostic accuracy of most tests was based on studies conducted within tertiary care settings, the population analysed in the studies may have had more advanced disease than the general population. This could overestimate the prevalence of the disease, leading to an overestimation of the diagnostic accuracy for each test.

All reported non-invasive tests were developed and compared with reference to liver biopsy, which is a reference standard with limitations, most notably misclassifications due to sample variability and intra- and interobserver variability. A potential solution would be to develop and validate non-invasive tests with reference to clinical outcomes; however, this would take time.

The findings of the cost-effectiveness study imply that for HCV the best option is to treat all patients regardless of stage of liver disease. For HBeAg-negative chronic HBV, this is also the case if the higher bound of the standard cost-effectiveness threshold is considered acceptable. These findings would be applicable in settings similar to the UK; however, in resource-poor settings, a treat-all strategy may not be possible. In this case, from our findings, a non-invasive test may be a better diagnostic option than liver biopsy.

Conclusion

The evidence suggests that, for HCV, treating all patients without prior diagnostic testing is the most cost-effective option. This analysis has not included the recently approved, more costly, interferon-free regimes. For HBV, the results differed for patients with HBeAg-positive and HBeAg-negative. The results suggest that, if the upper band of the standard UK cost-effectiveness threshold is accepted for patients with HBeAg-negative disease, the strategy of treating all patients regardless of fibrosis level is cost-effective. For similar patients with HBeAg-positive disease, at standard UK cost-effectiveness thresholds the results are highly uncertain, with several test strategies having similar expected outcomes and costs.

Abstinence is recommended for patients with ALD. There was a lack of data to allow robust modelling of the impact of testing on abstinence rates and whether or not these are affected by the degree of invasiveness of the tests. If abstinence is likely to increase following diagnosis of fibrosis or cirrhosis, and if it is likely to be higher following an invasive test, then biopsy will be cost-effective.

For NAFLD, most interventions are aimed at behavioural change rather than treatment and are not specifically recommended to reduce or halt fibrosis progression (e.g. weight-loss programmes for obesity); therefore, it is not possible to robustly determine the long-term costs and health consequences of fibrosis testing.

Suggested research priorities

Research on treatment effectiveness for patients with NAFLD is required, such as on the impact of fibrosis diagnosis on weight loss and other behaviour changes, and the relative effectiveness of primary care interventions versus secondary referrals.

© Queen's Printer and Controller of HMSO 2015. This work was produced by Crossan *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

High-quality studies with a low risk of bias for NILTs are required to allow for sufficient validation of specific cut-offs to stage fibrosis in different disease aetiologies. These require the use of predetermined cut-offs for the NILTs, adequate biopsy samples, selection of consecutive patients with no inappropriate exclusions and adequate reporting of patient flow and indeterminate results.

The potential use of NILTs to predict liver-related complications rather than to stage fibrosis should be further explored. This would provide a hard end point and overcome the need for liver biopsy.

Currently-available NILTs cannot differentiate simple steatosis from steatohepatitis. Therefore, there is a need to develop reliable non-invasive tests for this, as simple steatosis is usually non-progressive, whereas steatohepatitis could potentially progress to significant fibrosis and cirrhosis.

Further research on abstinence rates following diagnosis with either a NILT or liver biopsy is required.

The applicability of the findings for HBV and HCV to different countries and settings would benefit from future research.

The impact of new therapies on cost-effectiveness (higher costs but fewer side effects and better efficacy) for HCV also warrant further investigation.

Study registration

This study is registered as PROSPERO CRD42011001561.

Funding

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

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 5.116

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index and is assessed for inclusion in the Database of Abstracts of Reviews of Effects.

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

Editorial contact: nihredit@southampton.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

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

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

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

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 09/114/02. The contractual start date was in November 2011. The draft report began editorial review in August 2013 and was accepted for publication in February 2014. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

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

© Queen's Printer and Controller of HMSO 2015. This work was produced by Crossan *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

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

Editor-in-Chief of *Health Technology Assessment* and NIHR Journals Library

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

NIHR Journals Library Editors

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

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

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

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

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

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Peter Davidson Director of NETSCC, HTA, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

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

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

Professor Geoffrey Meads Professor of Health Sciences Research, Faculty of Education, University of Winchester, UK

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

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

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

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

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

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

Editorial contact: nihredit@southampton.ac.uk