HEALTH TECHNOLOGY ASSESSMENT

VOLUME 19 ISSUE 9 JANUARY 2015
ISSN 1366-5278

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, Emmanuel A Tsochatzis, Louise Longworth, Kurinchi Gurusamy, Brian Davidson, Manuel Rodríguez-Perálvarez, Konstantinos Mantzoukis, Julia O'Brien, Evangelos Thalassinos, Vassilios Papastergiou and Andrew Burroughs

NHS
National Institute for Health Research

DOI 10.3310/hta19090

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

This report should be referenced as follows:

Crossan C, Tsochatzis EA, Longworth L, Gurusamy K, Davidson B, Rodríguez-Perálvarez M, *et al*. 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. *Health Technol Assess* 2015;**19**(9).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents®*/ Clinical Medicine.

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.

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

Abstract

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 Louise.Longworth@brunel.ac.uk

Background: Liver biopsy is the reference standard for diagnosing the extent of fibrosis in chronic liver disease; however, it is invasive, with the potential for serious complications. Alternatives to biopsy include non-invasive liver tests (NILTs); however, the cost-effectiveness of these needs to be established.

Objective: To assess the diagnostic accuracy and cost-effectiveness of NILTs in patients with chronic liver disease.

Data sources: We searched various databases from 1998 to April 2012, recent conference proceedings and reference lists.

Methods: We included studies that assessed the diagnostic accuracy of NILTs using liver biopsy as the reference standard. Diagnostic studies were assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool. Meta-analysis was conducted using the bivariate random-effects model with correlation between sensitivity and specificity (whenever possible). Decision models were used to evaluate the cost-effectiveness of the NILTs. Expected costs were estimated using a NHS perspective and health outcomes were measured as quality-adjusted life-years (QALYs). Markov models were developed to estimate long-term costs and QALYs following testing, and antiviral treatment where indicated, for chronic hepatitis B (HBV) and chronic hepatitis C (HCV). NILTs were compared with each other, sequential testing strategies, biopsy and strategies including no testing. For alcoholic liver disease (ALD), we assessed the cost-effectiveness of NILTs in the context of potentially increasing abstinence from alcohol. Owing to a lack of data and treatments specifically for fibrosis in patients with non-alcoholic fatty liver disease (NAFLD), the analysis was limited to an incremental cost per correct diagnosis. An analysis of NILTs to identify patients with cirrhosis for increased monitoring was also conducted.

Results: Given a cost-effectiveness threshold of £20,000 per QALY, treating everyone with HCV without prior testing was cost-effective with an incremental cost-effectiveness ratio (ICER) of £9204. This was robust in most sensitivity analyses but sensitive to the extent of treatment benefit for patients with mild fibrosis. For HBV [hepatitis B e antigen (HBeAg)-negative)] this strategy had an ICER of £28,137, which was cost-effective only if the upper bound of the standard UK cost-effectiveness threshold range (£30,000) is acceptable. For HBeAg-positive disease, two NILTs applied sequentially (hyaluronic acid and magnetic resonance elastography) were cost-effective at a £20,000 threshold (ICER: £19,612); however, the results were highly uncertain, with several test strategies having similar expected outcomes and costs. For patients with ALD, liver biopsy was the cost-effective strategy, with an ICER of £822.

Limitations: A substantial number of tests had only one study from which diagnostic accuracy was derived; therefore, there is a high risk of bias. Most NILTs did not have validated cut-offs for diagnosis of specific fibrosis stages. The findings of the ALD model were dependent on assuptions about abstinence rates assumptions and the modelling approach for NAFLD was hindered by the lack of evidence on clinically effective treatments.

Conclusions: Treating everyone without NILTs is cost-effective for patients with HCV, but only for HBeAg-negative if the higher cost-effectiveness threshold is appropriate. For HBeAg-positive, two NILTs applied sequentially were cost-effective but highly uncertain. Further evidence for treatment effectiveness is required for ALD and NAFLD.

Study registration: This study is registered as PROSPERO CRD42011001561.

Funding: The National Institute for Health Research Health Technology Assessment programme.

Contents

List of tables

List of figures

Glossary

Clinical terms

Alcoholic fatty liver (hepatic steatosis) Accumulation of excess fat in the liver cells as a result of excess alcohol consumption.

Alcoholic hepatitis An acute inflammatory condition in patients who abuse alcohol involving liver damage, which is associated with high mortality.

Cirrhosis A consequence of chronic liver disease, most commonly alcoholism, hepatitis B and C, or fatty liver disease. It is characterised by diffuse nodular regeneration surrounded by dense fibrotic septa with subsequent parenchymal extinction and collapse of liver structures, together causing a marked distortion of hepatic vascular architecture, leading to loss of liver function.

Compensated cirrhosis Cirrhosis with no associated complications.

Hepatitis B A hepatotropic virus that causes chronic infection and is associated with progressive liver damage that can lead to cirrhosis and hepatocellular cancer.

Hepatitis C A hepatotropic virus that causes chronic infection and is associated with progressive liver damage that can lead to cirrhosis and hepatocellular cancer.

Hepatocellular carcinoma The most common type of liver cancer, usually secondary to scarring of the liver (cirrhosis) or hepatitide viral infection (hepatitis B or C).

Histological Related to the microscopic structure of tissue.

Liver biopsy Removal of a small sample of liver tissue using a hollow needle for examination in the laboratory.

Liver fibrosis Formation of excessive fibrous scar tissue in the liver.

Non-alcoholic fatty liver disease Liver disease characterised by the accumulation of fat in the liver cells of people who do not drink alcohol excessively but are obese and/or have other features of the metabolic syndrome.

Non-alcoholic steatohepatitis The progressive form of non-alcoholic fatty liver disease involving inflammation in and around the fatty liver cells. It may cause scarring of the liver and lead to cirrhosis.

Obese Having a body mass index ≥ 30 kg/m^2.

Oesophageal varices Varicose veins in the lower end of the oesophagus which develop as a complication of cirrhosis.

Percutaneous Performed through a needle puncture of the skin.

Steatosis A condition characterised by the accumulation of excess fat within the liver cells.

Transjugular Via the jugular vein in the neck.

Variceal bleeding Gastrointestinal bleeding caused by rupture of the oesophageal/gastric varices.

Varices varicose Veins usually in the stomach and lower end of the oesophagus (gullet) which develop as a complication of cirrhosis.

YKL-40 The name of a direct marker of liver fibrosis.

Diagnostic accuracy terms

Combined cut-off This refers to the use of a dual cut-off for classifying patients with a non-invasive fibrosis test. If a patient has a value below the low or above the high cut-off, then he or she is classified into a fibrosis stage; otherwise, he or she falls in the grey zone or indeterminate range and requires further non-invasive testing or a liver biopsy.

False negative A test result which indicates that a person does not have the disease when that person actually does have the disease.

False positive A test result which indicates that a person does have the disease when that person actually does not have the disease.

High cut-off This refers to the use of a cut-off for staging of liver fibrosis by a non-invasive fibrosis test that aims to maximise specificity. Usually the high cut-off is pre-set using the receiver operating characteristic curve to allow for a specificity of 90–95%.

Index test The test whose performance is being evaluated.

Low cut-off This refers to the use of a cut-off for staging of liver fibrosis by a non-invasive fibrosis test that aims to maximise sensitivity. Usually the low cut-off is pre-set using the receiver operating characteristic curve to allow for a sensitivity of 90–95%.

METADAS The name of a macro in the SAS software.

METAVIR score A histological scoring system developed by a French working group to assess and stage necro-inflammation and fibrosis.

QUADAS-2 A tool developed for quality assessment of diagnostic accuracy studies.

Receiver operating characteristic curve A receiver operating characteristic curve represents the relationship between the 'true-positive fraction' (sensitivity) and the 'false-positive fraction' (specificity). It displays the trade-offs between sensitivity and specificity as a result of varying the cut-off value for positivity in case of a continuous test result.

Reference standard Established test(s) against which the accuracy of a new test for detecting a particular condition can be evaluated.

Sensitivity (true-positive rate) The proportion of individuals with the target condition who are correctly identified by the index test.

Specificity (true-negative rate) The proportion of individuals free of the target condition who are correctly identified by the index test.

True negative A person without the disease correctly identified as negative by the index test.

True positive A person with the disease correctly identified as positive by the index test.

Non-invasive liver tests

Computed tomography A medical imaging technique using tomography created by computer processing to generate a three-dimensional internal image from a series of two-dimensional radiographic images.

Contrast-enhanced ultrasound The application of a contrast agent to conventional ultrasonography. Ultrasound contrast agents rely on the different ways that sound waves are reflected from interfaces between substances, for example microbubbles and human tissue. The difference in echogenicity (ability to reflect ultrasound waves) between microbubbles and surrounding tissues is very high and intravenous contrast injection can be used to visualise blood perfusion and to distinguish between benign and malignant tissue.

Magnetic resonance imaging A medical imaging technique that uses nuclear magnetic resonance to image the nuclei of atoms inside the body. It provides good contrast between the different tissues of the body and can be useful in distinguishing malignant from benign tumours.

Economic and statistical terms

Cost-effectiveness acceptability curve A graph that plots a range of possible cost-effectiveness thresholds on the horizontal axis (*x*-axis) against the probability that the intervention will be cost-effective on the vertical axis (*y*-axis) in order to give a representation of the decision uncertainty.

Cost-effectiveness acceptability frontier A graph that shows the probability that the technology with the highest expected net benefit is cost-effective.

Dominance An intervention is dominated if it has higher costs and worse outcomes than an alternative intervention. An intervention is dominant if it has lower costs and better outcomes than all other alternatives.

EQ-5D A generic preference-based instrument for measuring health-related quality of life.

Incremental cost-effectiveness ratio The ratio of the difference in the mean costs of two interventions divided by the difference in the mean outcomes.

Markov model An analytical method particularly suited to modelling repeated events or the progression of a chronic disease over time.

Meta-analysis Statistical techniques used to combine the results of two or more studies and obtain a combined estimate of effect.

Net benefit The overall benefit of the technology net of costs. Net (monetary) benefit is the difference between the monetary value of total expected quality-adjusted life-years (expected quality-adjusted life-years multiplied by the threshold value, e.g. £20,000) and total expected costs.

One-way sensitivity analysis An analysis that varies each parameter individually to assess the impact of its change on the results of the study.

Parameter A measurable or quantifiable characteristic.

Probabilistic sensitivity analysis An analysis in which probability distributions are assigned to the uncertain parameters in order to give a representation of decision uncertainty.

Quality-adjusted life-year A measure of length of survival that takes account of the patient's health-related quality of life during this time.

Randomised controlled trial A comparative study in which people are randomly allocated to intervention and control groups and followed up to examine differences in outcomes between groups.

Relative risk The number of times more or less likely an event is to happen in one group compared with another (calculated as the risk of the event in group A divided by the risk of the event in group B).

Sensitivity analysis An analysis to give an indication of the uncertainty in, or robustness of, the results of an analysis.

Threshold sensitivity analysis A sensitivity analysis in which the values of a parameter a systematically changed to identify the point above or below which the conclusions of the study will change.

Utility A measure of the strength of a person's preference for a specific health state in relation to alternative health states measured on a scale where 0 represents a state 'as bad as being dead' and 1 represents 'full health'.

List of abbreviations

AASLD	American Association for the Study of the Liver
ACE	angiotensin-converting enzyme
ALD	alcoholic liver disease
ALT	alanine aminotransferase
APRI	AST to platelet ratio index
ARFI	acoustic radiation force impulse
AST	aspartate aminotransferase
AUROC	area under the receiver operator curve
BARD	BMI, AST–ALT ratio, diabetes
BMI	body mass index
BNF	*British National Formulary*
CEAC	cost-effectiveness acceptability curve
CEAF	cost-effectiveness acceptability frontier
CELT	Cost-Effectiveness of Liver Transplantation
CI	confidence interval
CT	computed tomography
DNA	deoxyribonucleic acid
EASL	European Association for the Study of the Liver
ELF	enhanced liver fibrosis test
GGT	gamma-glutamyl transpeptidase
GP	general practitioner
GUCI	Göteborg University Cirrhosis Index
HBeAg	hepatitis B e antigen
HBV	hepatitis B
HCC	hepatocellular cancer
HCV	hepatitis C
HRQoL	health-related quality of life
HTA	Health Technology Assessment
ICER	incremental cost-effectiveness ratio
MR	magnetic resonance
MRI	magnetic resonance imaging
MTC	mixed-treatment comparison
NAFIC	ferritin, fasting insulin, type IV collagen
NAFLD	non-alcoholic fatty liver disease
NASH	non-alcoholic steatohepatitis
NFS	non-alcoholic fatty liver disease fibrosis score
NICE	National Institute for Health and Care Excellence
NILT	non-invasive liver test
PGAA	prothrombin time, GGT, apolipoprotein A1, α2-macroglobulin
PIIINP	amino-terminal propeptide of type III procollagen
PSA	probabilistic sensitivity analysis
PSSRU	Personal Social Services Research Unit
QALY	quality-adjusted life-year
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
RCT	randomised controlled trial
RR	relative risk

SAFE	sequential algorithm for fibrosis evaluation
SAPI	splenic artery pulsatility index
SHTAC	Southampton Health Technology Assessment Centre
SROC	summary receiver operating characteristic
SVR	sustained virological response
TE	transient elastography (Fibroscan)
TIMP	tissue metalloproteinase
ULN	upper limit of normal

Plain English summary

Liver has various functions including elimination of drugs and toxins. However, repeated insults to the liver including alcohol and viral infections can lead to loss of its structure and function. Fibrosis (slight) and cirrhosis (severe) are different degrees of loss of structure and function of the liver.

At present, there is no curative treatment for cirrhosis other than liver transplantation. Early diagnosis and treatment of fibrosis will improve survival and quality of life and reduce the need for liver transplantation. Traditionally, fibrosis is diagnosed by taking a bit of the liver using a wide-bore needle (biopsy); however, this is invasive and can cause serious complications. This study assessed alternative tests which are not invasive to determine whether or not they can replace liver biopsy and offer value for money.

A review of literature on how accurately the non-invasive tests are for the diagnosis of fibrosis and cirrhosis was carried out. We used this information to conduct an economic analysis to estimate the implications for NHS resources (including costs of testing and long-term costs of treating patients) and the health outcomes associated with each non-invasive test.

We found that the economic benefits vary according to the cause of the liver disease. In some cases, the non-invasive tests appeared best value for money (some types of hepatitis B); in others, biopsy was best (alcoholic cirrhosis); and in some cases, starting treatment early, without the need for testing, produced the highest health gain for a given cost (hepatitis C and some circumstances for hepatitis B).

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

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.

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.

Chapter 1 Background

Liver fibrosis is scarring of the liver.[1] Subsequently, areas of regenerating hepatocytes surrounded by fibrosis tissue develop, resulting in the development of liver cirrhosis.[1] Fibrosis and cirrhosis form chronic liver disease. Every year, around 6000 to 7000 people in the UK die from chronic liver disease[2,3] and about 600 adults have to have a liver transplant to survive.[4] In 2000, cirrhosis accounted for nearly 500 deaths in men aged 25 to 44 years and nearly 300 deaths in women of this age group, a seven- to eightfold increase in the deaths compared with the rate in 1970.[2] The age-standardised death rates from cirrhosis tripled from 2 per 100,000 population to 6 per 100,000 population between 1970 and 2000 in England,[2] and doubled from 9 per 100,000 population to 19 per 100,000 population between 1979 and 2007 in Scotland.[5]

Diagnostic testing for fibrosis and cirrhosis

Liver biopsy

Currently, histological examination of a tiny piece of liver tissue (liver biopsy) is considered the reference standard for the diagnosis and monitoring of liver fibrosis and cirrhosis. This is usually performed through the skin under the guidance of ultrasound[6–8] and involves taking a small section of the lesion using a sharp hollow needle. This can usually be performed under local anaesthesia.[6–8] The main risks of percutaneous biopsy are clinically significant bleeding (1.1–1.6%),[6,7] which can be fatal.[7]

Histological examination provides a spectrum of information, including liver architecture, presence and extent of steatosis, presence and grade of necroinflammation and presence and extent of liver fibrosis. It can also provide a diagnosis in cases of unexplained liver function test abnormalities. This amount of information is not provided by any non-invasive test, as they are mainly confined in the assessment of liver fibrosis. Therefore, liver biopsy will remain essential in many cases, whereas non-invasive liver function tests will be used in cases where the aetiology of liver disease is known and the clinical question is the extent of fibrosis.

Liver fibrosis is assessed in liver histological scoring systems using various staging systems that assess liver architecture and fibrosis. Such systems include Ishak, Knodell, Sheuer and METAVIR.[9,10] The METAVIR scoring system stages fibrosis in five categories, from 0 to 4, while the Ishak system stages fibrosis in seven categories, 0 to 6. Cirrhosis always represents the end stage of the spectrum and is characterised by bridging fibrosis and regenerative nodules.

It should be stressed that histological stages are descriptive semiquantitative categories that assess both liver architecture and liver fibrosis and do not provide a quantitative assessment of liver fibrosis.[10,11] The numbers that have been assigned to histological stages have no quantitative relationship between them, i.e. METAVIR fibrosis stage 2 does not mean twice the amount of fibrosis of stage 1.[12] Therefore, non-invasive fibrosis markers, which assess fibrosis quantitatively, should be ideally developed and validated with reference to a histological quantitative assessment of liver fibrosis.[13,14] Such histological methods have indeed been developed and quantify fibrosis by measuring liver collagen using digital image analysis.[15–17]

As liver biopsy assesses only a tiny amount of the liver, sample variability could potentially misclassify the extent of fibrosis. In addition, histological staging is also prone to intra- and interobserver variability, even when senior liver histopathologists are involved. A French study found that, in patients with chronic hepatitis C (HCV), 35% of biopsies 15 mm in length were not categorised correctly.[18] The study suggested that a sample at least 25 mm in length is necessary to evaluate fibrosis accurately with a semiquantitative score, with the possible exception of cirrhosis. Biopsies of such length are not always feasible with one needle pass in a percutaneous biopsy and, therefore, the patient's discomfort and also the complication

rate might increase. The misclassification rate (percentage of incorrect staging of fibrosis) of liver biopsy is the source of the myth that non-invasive fibrosis tests cannot achieve a high concordance with histological stages. This is only true for non-invasive tests for which their development was independent from liver histology, such as transient elastography (TE), although the diagnostic test accuracies of such tests are also evaluated using histology; it could be argued that in certain cases the false positive or false negative of such a test compared with the result of a liver biopsy is a fault of the biopsy rather than the test itself, i.e. the test diagnosed correctly what was missed by the biopsy. However, serum non-invasive fibrosis markers have been developed and calibrated with direct reference to a set of liver biopsies. Therefore, the perfect serum marker in this case would replicate the 'golden' histological standard and could theoretically reach an area under the receiver operator curve (AUROC) of 1, replicating even the misclassifications of a liver biopsy.[19]

Non-invasive fibrosis tests

During the last few years, there has been an explosive development and use of non-invasive fibrosis tests.[20,21] These tests in many cases have replaced liver biopsy in clinical practice in the staging of fibrosis and follow-up of patients with established chronic liver disease, especially in patients with chronic HCV. The non-invasive liver tests (NILTs) can be broadly divided into three categories: simple or indirect serum markers, direct serum markers and imaging modalities.

Indirect serum markers or class II biomarkers consist of the combination of routine biochemical or haematological tests, such as transaminases, platelet count and albumin, and patient demographics that are associated with fibrosis, such as age or the presence of diabetes.[20] These tests usually have dual cut-offs: a high cut-off with high specificity and a low cut-off with high sensitivity. Depending on the clinical scenario and the disease prevalence, the low or high cut-off is used at the expense of increased false positives and false negatives, respectively. If these cut-offs are combined, then the number of false positive and false negatives are minimised; however, a number of patients will fall in the indeterminate range of fibrosis (i.e. their score will be between the low and the high cut-off) and will need either further non-invasive testing or a liver biopsy. Commonly used indirect serum markers are FIB-4, aspartate aminotransferase (AST)-to-alanine aminotransferase (ALT) ratio and APRI (AST to platelet ratio index).

Direct serum non-invasive tests or class I biomarkers are intended to detect extracellular matrix turnover and/or fibrogenic cell changes. The most common markers used in current assays involve measuring products of extracellular matrix synthesis or degradation, and the enzymes that regulate their production or modification, such as hyaluronic acid, serum collagenases and their inhibitors and profibrogenic cytokines. It should be noted that these markers are not exclusively found in liver tissue; therefore, they reflect fibrogenic processes in various other organs. For instance, the enhanced liver fibrosis (ELF) biomarker is influenced by age and sex.[22] Moreover, their sensitivity is low in the initial stages of fibrosis.

Various direct and indirect tests have been combined in patented commercial algorithms that improve the diagnostic accuracy of tests when used singly. These are ELF, Fibrotest, Fibrospect, Fibroindex and Fibrometer. Of the tests, Fibrotest (Fibrosure in the USA) is the most widely validated panel; it consists of five parameters, namely total bilirubin, haptoglobin, gamma-glutamyl-transpeptidase, α2-macroglobulin and apolipoprotein A1, and has been studied in viral hepatitis, non-alcoholic fatty liver disease (NAFLD) and alcoholic liver disease (ALD).[23] The Fibroindex was developed for patients with chronic HCV and uses platelet count, AST and g-globulin levels.[24] Fibrospect includes hyaluronate, tissue metalloproteinase (TIMP)-1 and α2-macroglobulin, and is validated in chronic HCV.[25] Fibrometers are a family of six blood tests: one for staging and one for quantifying liver fibrosis in each of the three main causes of liver disease (chronic viral hepatitis, ALD and NAFLD).[26] The ELF biomarker is a panel of direct noninvasive markers that includes hyaluronic acid, type III collagen and TIMP-1.[27] It has been used in patients with chronic HCV and NAFLD.

New imaging modalities offer better sensitivity and specificity than conventional techniques, such as ultrasound, computed tomography (CT) and magnetic resonance imaging (MRI). The last of these can only identify cirrhosis, based on imaging findings of coarse echo-texture, collaterals suggestive of portal

hypertension and nodularity. These new modalities measure liver elasticity or liver stiffness based on ultrasound or magnetic resonance (MR) techniques. The most widely used imaging modality is TE or Fibroscan (Echosens, Paris).[28] Briefly, vibrations of mild amplitude and low frequency are transmitted by an ultrasound transducer, inducing an elastic shear wave that propagates within the liver. Pulse-echo ultrasonic acquisitions are performed to follow the shear wave and measure its speed, which is directly related to the tissue stiffness (the harder the tissue, the faster the shear propagates). Results are expressed in kilopascals (kPa) and correspond to the median value of 10 validated measurements ranging from 2.5 to 75 kPa, with 5.5 kPa reported to define normality. The volume of liver tissue evaluated by TE approximates a cylinder 4 × 1 cm which is at least 100 times bigger than a liver biopsy. Moreover, TE is painless and rapid (< 5 minutes) and thus highly acceptable for patients.

Other modalities include acoustic radiation force impulse (ARFI)[29] and MR elastography.[30] ARFI allows the evaluation of liver stiffness in a region of interest (ROI) involving mechanical excitation of tissue by the use of short-duration (≈262 μs) acoustic pulses while performing a real-time B-mode conventional hepatic ultrasound. Results are expressed in m/s. Although the volume of liver explored is smaller than that for TE (10 mm long × 6 mm wide), a critical advantage is the possibility to choose the representative area of interest, thereby avoiding large vessels and ribs. An advantage over TE is that it can be easily incorporated into a modified ultrasound machine. MR elastography uses a modified phase-contrast method to evaluate the propagation of the shear waves within the liver. It is a very promising technique but is not yet widely available and cost might be an important limiting factor.

Finally, algorithms of sequential or contemporary use of NILTs have been used mainly in chronic HCV, to improve the diagnostic accuracy of single tests.[31] These are typically based on an agree–disagree scenario or the sequential use of a second test if the result of the first test falls in the grey zone of an indeterminate result.

A major limitation of all the above NILTs is the absence of uniformly established and validated cut-offs for specific aetiologies of liver disease and fibrosis stages and the poor methodological quality of many of the published studies. In a recent meta-analysis on TE, only 6 of 41 included studies had both histological evaluation and Fibroscan measurements optimally performed, while all studies had a high risk of bias based on quality assessment by the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool.[32]

Aetiologies assessed

The study population comprised all patients with chronic liver disease (irrespective of the aetiology for chronic liver disease, age and clinical presentation). The aetiologies modelled and assessed were hepatitis B (HBV), HCV, ALD and NAFLD. We modelled and analysed these four aetiologies of liver disease, as the staging of fibrosis is pertinent in their prognosis and management. In all other causes of chronic liver disease, only the diagnosis of cirrhosis is important and liver biopsy is seldom performed for staging of fibrosis. Therefore, patients with primary biliary cirrhosis, autoimmune hepatitis, haemochromatosis and Wilson's disease are treated irrespective of fibrosis stage, whereas fibrosis evaluation is not pertinent in primary sclerosing cholangitis.[33–35] We also modelled liver cirrhosis irrespective of aetiology, as this diagnosis is important for every patient with chronic liver disease and heralds screening for oesophageal varices and hepatocellular carcinoma.[36]

Chronic HCV is a major cause of liver-related morbidity and mortality worldwide, and it is estimated that around 200 million people are infected worldwide.[37] The diagnosis of chronic HCV is based on serological testing and does not require a liver biopsy. The natural history of chronic HCV is variable; it is estimated that one-third of the infected patients will progress to cirrhosis.[38] Factors that are associated with fibrosis progression are age at infection > 40 years, male sex, obesity, alcohol abuse, presence of diabetes and coinfection with human immunodeficiency virus (HIV).[37,38] Current therapeutic options include dual therapy with pegylated interferon and ribavirin in patients with genotypes 2, 3 and 4, and triple therapy with the

addition of boceprevir or telaprevir in patients with genotype 1. The sustained virological response (SVR) rate in previously untreated patients is approximately 70% in genotype 4 and 80% in genotypes 1, 2 and 3.[37,39,40] Treatment is less successful in patients who were previously unsuccessfully treated, in obesity and in more advanced fibrosis.[37] Currently, antiviral treatment is recommended for all patients with chronic HCV irrespective of stage of fibrosis.

However, the decision to treat or not to treat is not always straightforward.[41,42] Many patients cannot tolerate the side effects of antiviral treatment, have unfavourable treatment response factors and in fact might never progress to severe fibrosis. Moreover, new treatment options with better efficacy and fewer side effects are rapidly emerging.[43,44] Non-invasive fibrosis tests offer the option not only of baseline fibrosis staging but also of follow-up measurements to determine the rate of fibrosis progression. Therefore, an alternative option would be to use an effective non-invasive fibrosis test for staging and treat only those patients with F2 and above, which represents clinically significant fibrosis. The test could be repeated in order to capture the false negatives and also potential fibrosis progression.

Chronic HBV is highly prevalent worldwide and it estimated that 350–400 million people are HBsAg carriers.[45] The natural history of the disease is variable; the virus itself is hepatotropic but not hepatotoxic, and liver damage is caused when the immune system attacks the hepatocytes that are infected by the virus. The natural history of the disease can be divided in four distinct phases.[46]

(1) The 'immune tolerant' phase is characterised by hepatitis B e antigen (HBeAg) positivity and high levels of virus replication but normal transaminases and no or minimal necroinflammation and progression of fibrosis. The virus, although in high concentrations, is not recognised by the immune system at that phase. This phase usually occurs in patients with perinatal infection in the first years of their lives.
(2) The 'immune reactive HBeAg(+)' phase is characterised by immune reaction, which leads to decreased HBV replication but also to destruction of hepatocyte, elevated transaminases, necroinflammation and fibrosis. This phase may last for several years and leads to HBeAg seroconversion to anti-HBe.
(3) The 'inactive HBC carrier state' phase is characterised by low or undetectable HBV deoxyribonucleic acid (DNA) and normal transaminases. This phase is characterised by immunological control of the infection and is associated with low risk of cirrhosis.
(4) The 'HBeAg(–) chronic hepatitis B' phase may follow immediately after the HBeAg sero-conversion or after years in the inactive carrier state. It is characterised by periodic virus reactivation with a pattern of fluctuating levels of HBV DNA and aminotransferases, active necroinflammation and progression of fibrosis.

Patients in the 'immune tolerant' and the 'inactive carrier' phase do not need antiviral treatment as they are not at imminent risk of fibrosis progression, but require regular follow-up with determination of viral load and transaminases.[45] Available treatment options include nucleoside or nucleotides analogues indefinitely or pegylated interferon alfa-2b for a finite period of 12 months. Treatment indications are based on the combination of criteria that take into account the HBV DNA levels, ALT levels and severity of liver disease based on histology. Current treatment guidelines advocate liver biopsy before initiating treatment in the majority of cases. The only exception is patients with obviously active chronic HBV, i.e. those with ALT > 2 upper limit of normal (ULN) and HBV DNA > 20,000 IU/ml, who may start treatment without a biopsy.[45] Patients with abnormal transaminases, HBV DNA > 2000 IU/ml and a biopsy showing moderate to severe active necroinflammation and/or at least moderate fibrosis using a standardised scoring system should be started on antiviral treatment.[45] Non-invasive fibrosis tests could potentially substitute liver biopsy in such patients, i.e. those with a non-invasive diagnosis of ≥ F2. A minority of patients with moderate necroinflammation but < F2 fibrosis, who would need treatment according to guidelines, would not be captured with a non-invasive fibrosis test, and would only be treated once they progressed to F2.

Non-alcoholic fatty liver disease affects approximately 20% of the general population and encompasses a wide range of liver disease, from simple steatosis to necroinflammation, fibrosis and cirrhosis.[47]

It is associated with obesity and is considered the hepatic manifestation of metabolic syndrome.[48] Non-alcoholic steatohepatitis (NASH) is the progressive form of NAFLD and affects 15–20% of patients with NAFLD.[49] Only patients with steatohepatitis have increased liver-related mortality.[47]

Data on natural history of NAFLD are still scarce; in a meta-analysis of 10 studies comprising 221 patients, 37.6% had progressive fibrosis, 41.6% had no change and 20.8% had improvement in fibrosis over a mean follow-up of 5.3 years.[50] Age and initial necroinflammation grade were the only factors associated with progression of fibrosis.[50] Even in patients with NASH, the primary cause of death was cardiovascular disease, with liver disease being only the third cause.[51] Compensated cirrhosis due to NASH is associated with a lower mortality rate than that due to HCV, and also with lower rates of development of ascites, hyperbilirubinemia and hepatocellular carcinoma.[52]

Treatment strategies for NAFLD/NASH are mainly based on lifestyle changes, including weight loss and exercise, and treatment of the individual components of the metabolic syndrome, such as diabetes, hypertension and hyperlipidaemia.[47] Vitamin E in non-diabetic patients and pioglitazone may improve steatosis and necroinflammation but not fibrosis, as shown in randomised controlled trials (RCTs).[47]

Currently, no validated non-invasive tests are available to differentiate NAFLD from NASH.[47] Diagnosis of patients with advanced fibrosis ($\geq$ F3) is of significance, as such patients could benefit from multidisciplinary treatment of metabolic syndrome components, targeted intervention for weight loss and specific treatment (vitamin E or pioglitazone) in selected cases.

Alcoholic liver disease encompasses a spectrum of injury that ranges from simple steatosis to cirrhosis.[53] The amount of ingested alcohol is the most important risk factor for the development of ALD.[54] Suggested safe limits are 21 units per week in men and 14 units per week in women.[53] Development of ALD is not dose dependent, as ALD is found in only a subset of patients. Women are more susceptible to alcohol-mediated liver injury than men.[55] Binge drinking and consumption of alcohol outside meal times are both associated with a higher risk of ALD.[55] The risk of developing cirrhosis is increased with ingestion of > 60–80 g/day of alcohol for > 10 years in men and > 20 g/day in women.[53]

The only effective treatment in patients with ALD is abstinence.[53] Prognosis is determined both by the degree of liver fibrosis and by the subsequent drinking behaviour. Interestingly, 5-year mortality in patients with well-compensated ALD cirrhosis was 10% in those who abstained and 30% in those who continued drinking.[56] Abstinence improves the histological features of ALD and may reverse fibrosis or decompensated cirrhosis to compensated cirrhosis. Diagnosis of patients with advanced fibrosis ($\geq$ F3) is of significance, as it will allow the timely provision of interventions to induce and maintain abstinence before cirrhosis occurs.

Decision problem to be addressed

As liver biopsy is an invasive procedure and is associated with morbidity and mortality risk, it is important (1) to assess the diagnostic accuracy of the different non-invasive fibrosis tests available and (2) to determine the most cost-effective approach in the clinical management of patients with chronic liver disease using either biopsy or non-invasive fibrosis tests for clinical decisions.

A range of non-invasive tests have become available and offer potential alternatives to liver biopsy. In order to assess the most appropriate use of the tests within a NHS setting, the relative accuracy and cost-effectiveness of the tests need to be evaluated. Furthermore, as liver biopsy is costly, and associated with morbidity and a small risk of mortality, the non-invasive tests may offer cost-effective alternatives.

Our analysis aims to assess the diagnostic accuracy and cost-effectiveness of the non-invasive tests in people with suspected liver fibrosis or cirrhosis. The tests are compared with each other, liver biopsy and strategies without testing. Fully incremental analyses are conducted wherever possible.

When assessing the cost-effectiveness of a test, it is important to consisder the consequences of the test result. A positive test result is likely to lead to a different course of treatment or action than a negative result; therefore, the consequences of an incorrect positive diagnosis are likely to differ from the consequences of an incorrect negative diagnosis. In order to reflect this, and a range of mobidity outcomes and mortality, our analyses are conducted using the quality-adjusted life-year (QALY) as the measure of outcome where possible. Where this has not been possible, analyses have been conducted to reflect potential differences between positive and negative diagnoses.

Structure of report

The rest of this report is structured as follows. The methods of the systematic review and overall methodological approach to the cost-effectiveness analysis are described in *Chapter 3*. *Chapter 4* presents results of the systematic review and meta-analysis. *Chapters 5–9* present the aetiology-specific methods and results of the cost-effectiveness analyses for HBV, HCV, ALD, NAFLD and cirrhosis, respectively. *Chapter 10* is a discussion of the findings from the study is provided and *Chapter 11* presents our conclusions.

Chapter 2 Objectives

There were two related objectives for the study:

1. To compare the diagnostic accuracy of different non-invasive tests in the diagnosis and monitoring of liver fibrosis and cirrhosis in patients with various aetiologies for chronic liver disease.
2. To estimate the incremental cost-effectiveness of the non-invasive tests in patients with various aetiologies for chronic liver disease.

Chapter 3 Methods of systematic review and economic evaluation modelling

Section 1 outlines the systematic review and meta-analysis methodology used in the study. Section 2 outlines the modelling methodology employed for the five aetiologies; HBV, HCV, NAFLD, ALD and cirrhosis.

Section 1: overview of systematic review methodology

Criteria for considering studies for review

The aim of the systematic review was to identify papers comparing the diagnostic accuracy of different non-invasive tests in the diagnosis and monitoring of liver fibrosis and cirrhosis with liver biopsy, and to synthesise the outcomes where possible. We included studies providing cross-sectional information of the index test(s) and reference test. In other words, we included all studies that reported staging of fibrosis by index test(s) and reference standard so that it is possible to know how many patients had a certain stage of fibrosis by index test and reference test (true positive), how many had that stage by index test but not on the reference test (false positive), how many did not have that particular stage by index test but were found to have that stage by reference test (false negative), and how many patients did not have a certain stage of fibrosis by index test or reference test (true negative) in the appropriate patient population, irrespective of language or publication status, or whether data were collected prospectively or retrospectively.

We also included comparative studies in which the different index tests were performed in the same study population, or studies in which different individuals in the study population received different index tests, and the choice of tests that the different individuals received were determined in a random manner or if all the patients underwent both the index tests that were assessed. We excluded diagnostic case–control studies from the analysis if there were at least four cross-sectional or comparative studies for that test. We also excluded studies where the maximum interval between the reference standard (liver biopsy) and the non-invasive fibrosis test (index test) was > 6 months.

Participants

Adult patients with chronic liver disease (irrespective of the aetiology and clinical presentation). Studies reporting on paediatric patients were excluded.

Index tests

Ultrasound, CT scan, MRI, elastography (TE by ultrasound or MR elastography), and direct and indirect serum markers (such as AST–ALT ratio, APRI, ELF test, Fibrotest, etc.).

Target condition

Liver fibrosis and cirrhosis.

Reference standards

Histopathological examination of liver tissue (percutaneous or transjugular or laparoscopic biopsy). The staging and grading of liver biopsy can be performed by various histological scoring systems such as Ishak, METAVIR, Knodell and others.[57] We included studies irrespective of the histological scoring system used. For data synthesis and analysis we transformed the histological scores used in individual studies to METAVIR for HBV, HCV and alcohol and to Kleiner for NAFLD/NASH as these are the most commonly used histological scores. Conversion of various histological stages to METAVIR is shown in *Table 1*.

TABLE 1 Conversion of various histological systems to METAVIR

Ishak	Knodell	Scheuer	METAVIR
0	0	0	0
1	1	1	1
2, 3	1	2	2
4, 5	3	3	3
6	4	4	4

Search methods for identification of studies

Electronic searches

The following databases were searched from 1988 until April 2012: 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).[58,59]

The search strategies for the different databases are provided in *Appendix 1*.

Initially, we did not use any filter; however, this yielded 200,000 references and a compromise had to be arranged, as it would not be possible to complete the analysis within the time scale allowed for this study. Therefore, a methodological filter is included but does not act as a filter for all search results (see *Appendix 1*). This represents a potential limitation in our search strategy.

Searching other sources

Reference lists of identified studies and reviews, and conference proceedings from the recent hepatobiliary conferences (last 2 years), were hand-searched to identify further studies.

Data collection and analysis

Selection of studies

The references were searched by two researchers independently for identification of relevant studies. No restrictions were placed on the language or the publication status (full text vs. abstract from conference proceedings). However, studies which reported on a total of fewer than 10 patients with fibrosis or cirrhosis were excluded. Full texts were obtained for the references that at least one of the reviewers considered relevant. Full-text articles were then used to include or exclude studies for the review.

Data extraction and management

Data were extracted by two reviewers independently. Any differences in the data extraction were resolved by the lead applicant, Professor Burroughs, and Dr Gurusamy. Data necessary to calculate the true positive, false positive, true negative and false negative diagnostic test results were extracted using the reference standard of liver biopsy. If the information on true positive, false positive, false negative and true negative diagnostic test results were not available directly, these were calculated from information available in the study. Data were entered into a Microsoft Excel (Microsoft Corporation, Redmond, WA, USA) file created for the purpose.

The following data were extracted:

1. year of publication
2. country/ethnicity of included patients
3. inclusion criteria
4. exclusion criteria
5. total number of patients
6. patients included in the analysis
7. mean age
8. mean body mass index (BMI)
9. sex
10. mean ALT
11. aetiology of liver disease
12. technical failure in undertaking liver biopsy or non-invasive tests
13. non-invasive test used
14. fibrosis histological scoring system used
15. non-invasive test cut-off for diagnosing specific fibrosis stages
16. distribution of patients across histological stages
17. sensitivity, specificity, true positive, false positive, false negative, true negative of non-invasive test for diagnosing different histological stages
18. number of patients with uninterpretable liver biopsies or index tests
19. number of patients with indeterminate non-invasive test for a specific fibrosis stage
20. methodological quality using the QUADAS-2 assessment tool.

Assessment of methodological quality

The quality of the studies was assessed independently by two reviewers using the QUADAS-2 assessment tool.[60–62] This tool comprises four domains: patient selection, index test, reference standard, and flow and timing. Each domain is assessed in terms of risk of bias, and the first three domains are also assessed in terms of concerns regarding applicability. Signalling questions are included to help judge the risk of bias. The quality criteria that were derived from the QUADAS-2 tool and were assessed are shown in *Table 2*.

Statistical analysis and data synthesis

The data obtained from the various studies are presented in the form of summary sensitivity and specificity with their corresponding 95% confidence intervals (CIs). The data were combined using the bivariate random-effects model with correlation between sensitivity and specificity[63] using the METADAS macro developed by the Systematic Review Diagnostic Test Accuracy Group[64] in the SAS 9.2 statistical software (SAS Institute Inc., Cary, NC, USA). We calculated the summary sensitivity and specificity at specific thresholds for tests with explicit thresholds such as serum markers and calculated the overall summary sensitivity and specificity for tests that do not have an explicit threshold (such as ultrasound).

The bivariate model allows for meta-analysis of diagnostic test accuracy studies to be conducted in which both the index test under study and the reference test (gold standard) are dichotomous. Bivariate analysis involves statistical distributions at two levels. At the lower level, it models the cell counts in the 2×2 tables extracted from each study using binomial distributions and logistic (log-odds) transformations of proportions. At the higher level, random study effects are assumed to account for heterogeneity in diagnostic test accuracy between studies beyond that accounted for by sampling variability at the lower level.[65]

If the results did not converge using the above random-effects model with correlation between sensitivity and specificity, we performed the meta-analysis with variations of bivariate analysis. The variations included different assumptions such as no correlation between the sensitivity and specificity in the studies; random-effects model for sensitivity but fixed-effect model for specificity; fixed-effect model for sensitivity but random-effects model for specificity; and fixed-effect models for both sensitivity and specificity (Takwoingi, University of Birmingham, March 2013, personal communication).

TABLE 2 Assessment of methodological quality using the QUADAS-2 tool

Quality assessed	Description	Choice	Comment
Domain 1: patient sampling	Was a consecutive or random sample of patients enrolled?	Yes/no/unclear	
	Was a case–control design avoided?	Yes/no/unclear	
	Did the study avoid inappropriate exclusions?	Yes/no/unclear	For example exclusion of patients with severe or low fibrosis, obese, etc.
Risk of bias	Could the selection of patients have introduced bias?	Low risk/high risk/unclear	Summarises previous questions: if any has no as answer then high risk, if any has unclear then unclear
Concerns about applicability	Are there concerns that the included patients and setting do not match the review question?	High/low concern/ unclear	Tertiary centres, selected difficult cases
Domain 2: index test	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes/no/unclear	Relevant only in US, CT, MRI
	If a threshold was used, was it prespecified?	Yes/no/unclear	
Risk of bias	Could the conduct or interpretation of the index test have introduced bias?	Low risk/high risk/unclear	Summarises previous questions: if any has no as answer then high risk, if any has unclear then unclear
Concerns about applicability	Are there concerns that the index test, its conduct, or interpretation differs from the review question?	High/low concern/ unclear	Index test not conducted according to manufacturer recommendations
Domain 3: reference standard	Is the reference standard likely to correctly classify the target condition?	Yes/no/unclear	Yes if biopsy length > 15 mm and/or > 6 portal tracts
	Was the reference standard results interpreted without knowledge of the results of the index tests?	Yes/no/unclear	
Risk of bias	Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk/high risk/unclear	Summarises previous questions: if any has no as answer then high risk, if any has unclear then unclear
Concerns about applicability	Are there concerns that the target condition as defined by the reference standard does not match the question?	High/low concern/unclear	Always low concern
Domain 4: flow and timing	Was there an appropriate interval between index test and reference standard?	Yes/no/unclear	Yes if interval between biopsy and index test < 3 months, no if interval > 3 but < 6 months, excluded study if interval > 6 months
	Did all patients receive the same reference standard?	Yes/no/unclear	
	Were all patients included in the analysis?	Yes/no/unclear	No if patients with uninterpretable results were not included in the analysis or if there were patients with indeterminate results
Risk of bias	Could the patient flow have introduced bias?	Low risk/high risk/unclear	Summarises previous questions: if any has no as answer then high risk, if any has unclear then unclear

US, ultrasound.

It must, however, be pointed out that the assumptions used to perform the above analysis (e.g. if one assumes that there is no correlation between the sensitivity and specificity, one has to ensure this from a scatterplot and correlation coefficient, and when one assumes a fixed-effect model, the values should be relatively close to each other) were not always met and the summary values of a model that converged was used. This could have resulted in a biased effect estimate. The alternative was not to conduct a meta-analysis for those tests which would have meant that the information could not be used in the cost-effectiveness analysis.

We also calculated the median, the lowest and the highest prevalence for the specific stages of fibrosis in the studies included.

Investigations of heterogeneity

The following sources of heterogeneity were explored.

1. Studies of high methodological quality versus low methodological quality.
2. Different stages of fibrosis (different scoring systems were converted to comparable stages in METAVIR in viral diseases and alcohol, and to Kleiner scoring system in NAFLD).
3. Different reference histological scoring systems (e.g. Ishak scoring, METAVIR, Knodell score, etc.).[57]
4. Different aetiological diagnosis (e.g. ALD, HCV infection, etc.).
5. Different threshold levels for classification of positive and negative results. We performed a meta-analysis for every possible cut-off in each fibrosis stage of the reference standard.
6. Studies not published in full text were compared with studies published in full text.
7. Different ranges of transaminases (normal, between normal and up to three times the normal level, and more than three times the normal level).

Section 2: overview of economic modelling methodology

The population of interest is patients who are suspected of having liver fibrosis or cirrhosis (patients who a hepatologist would wish to biopsy to inform treatment decisions). Owing to differences in treatment and natural history of disease, the analysis is conducted separately for subgroups defined according to aetiology. Five subgroups are defined for the analysis: patients with HBV, HCV, ALD, NAFLD and cirrhosis. More details are given in the dedicated chapters according to disease aetiology (see *Chapter 5* for HBV, *Chapter 6* for HCV, *Chapter 7* for ALD and *Chapter 8* for NAFLD).

The overall aim of the health economic analysis was to assess the incremental cost-effectiveness of the NILTs. Wherever possible, the analyses take a lifetime perspective. Health outcomes were measured using QALYs. A NHS perspective was taken for the estimation of costs. Both costs and QALYs were discounted at 3.5% in accordance with current National Institute for Health and Care Excellence (NICE) guidelines.[66]

The consequences following diagnosis are estimated and included in the analyses. In most cases the test diagnoses are expected to potentially affect decisions about future treatments (HBV, HCV and cirrhosis) or behaviour change (ALD). The long-term costs and health outcomes as a result of these treatments/behaviour changes are taken into account in the analysis (including the potential impacts of correct and incorrect diagnoses). Where this has not been possible, due to insufficient evidence or lack of treatments specifically aimed at fibrosis, the analysis has been restricted to an incremental cost per correct diagnoses, supplemented by exploratory analyses (NAFLD). In the cost per correct diagnoses, correct positive diagnoses have been presented separately from correct negative diagnoses as the consequences of each are likely to be very different.

Comparators

Where a large number of applicable NILTs were located by the systematic literature review (HBV and HCV), a two-stage approach to the analysis was conducted. The first stage compared each NILT identified from the systematic review (per aetiology) with each other and with liver biopsy. Where analyses involved treatment (HBV, HCV and cirrhosis), two additional testing approaches were included: a 'treat all' approach, where everyone is treated, and a 'no treatment' approach, where no diagnostic tests or treatments are administered.

The second stage of the analysis evaluated comparisons of sequential testing strategies, again compared with each other, biopsy and the treat-all and treat-no-one strategy. For this, combinations of the two most cost-effective tests within each category were chosen based on an incremental analysis using a cost-effectiveness threshold value of £20,000.[66] We assumed a decision rule whereby the two most cost-effective tests from each category were combined with tests from the other categories (reflecting combinations which would happen in actual current practice or potential future practice). Some of the NILTs evaluated have defined 'low' or 'high' cut-off thresholds and were analysed as separate test options. Combinations of tests considered to be clinically implausible were excluded; for example, a NILT with a low cut-off diagnostic threshold would not be followed by a second NILT with a low cut-off diagnostic threshold in practice but could be followed by a test with a high cut-off threshold. The following assumptions were made when combining the tests:

- If the first NILT used was an indirect serum marker, a patented or direct serum marker or an imaging modality or liver biopsy could be administered as a second test.
- If the first NILT used was a direct or patented serum marker, an imaging modality or liver biopsy could be administered as a second test.
- If the first test used was an imaging modality, a liver biopsy could be administered as a second test.

The analysis also assumed that the sensitivity and specificity of each test were independent of each other, i.e. there was no correlation of sensitivities and specificities of the tests used in the first stage and the second stage. The combinations were assumed to take four possible sequential testing strategies (*Table 3*).

The probabilities of having each of the four possible diagnoses (true negative, true positive, false negative, false positive) for the four sequential testing strategies were determined by multiplying the probabilities (i.e. using decision tree calculation methodology: multiplying probabilities along pathways from left to right to estimate the probability of each pathway).

Each of the sequential tests were compared with each other; liver biopsy alone; 'treat all' and 'no treatment' approaches; each cost-effective test singly; reported tests which used a combined cut-off; and any reported tests whose efficacy was estimated using a published algorithm derived from two or more tests used sequentially.

TABLE 3 Sequential testing strategies

	First NILT result		Second NILT result	
Strategy number	**Positive**	**Negative**	**Positive**	**Negative**
Strategy 1	Treat patients	Liver biopsy		
Strategy 2	Do second test	Watchful waiting	Treat patients	Liver biopsy
Strategy 3	Do second test	Liver biopsy	Treat patients	Liver biopsy
Perform two NILTs regardless of test outcome				
Strategy 4	Agree (+): treat		Disagree: liver biopsy	
	Agree (–): watchful waiting		Positive: treatment	
			Negative: watchful waiting	

Synthesis of economic evidence

A decision tree model was constructed to estimate the cost-effectiveness of all comparators. Sensitivity and specificity data included in the decision tree were extracted from the meta-analysis (see *Chapter 4*).

Long-term costs and QALYs were taken from the literature if estimates specifically matching the decision tree pathways were available. Where this was not possible, long-term costs and outcomes were estimated using a series of Markov models. *Figure 1* depicts the flow of data between the different modelling elements for the models estimating incremental cost per QALY.

The watchful waiting strategy incorporated a retest every 2 years. We assumed that the retest would have perfect sensitivity and specificity in the base case for modelling practicality due to the large number of applicable NILTs identified.

Literature review

Literature searches were undertaken to identify incremental-cost-per-QALY analyses of the non-invasive tests for each aetiology. Titles and abstracts were reviewed and full papers were retrieved if deemed relevant. If existing systematic reviews were available, these were reviewed and the searches updated and/or amended as required.

Studies were excluded if not in the English language due to resource limitations. We gave preference to UK-based studies for cost data as there may be transferability issues using data from other populations due to underlying differences between the populations.

Literature searching was undertaken to populate input parameters for the models (for natural history, costs and QALY inputs). Titles and abstracts were reviewed and full papers were retrieved if deemed relevant. We started by identifying existing recent reviews. The papers identified in these were reviewed. The searches were updated, amended if needed, and rerun.

For data on natural history, inclusion criteria related to the population of interest. Judgements about the relevance of studies also took into account the country of origin (preference for UK data), high-quality and recent studies. For cost studies, those reporting data from a NHS perspective were preferred. For studies reporting health-related utility inclusion criteria requiring data from the population of interest (depending on aetiology), information on health had to be collected directly from patients and the method of preference elicitation had to be a choice-based method (e.g. time trade-off) in a UK population. As per standard NICE methods guidance,[66] data obtained through the EQ-5D measure were preferred.

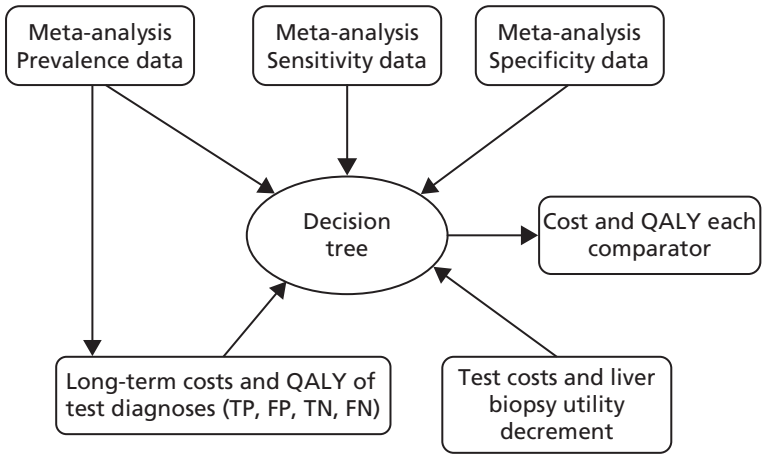

FIGURE 1 Illustration of data flow (input into decision tree from other modelling elements). FN, false negative; FP, false positive; TN, true negative; TP, true positive.

Further details of the search results are described in the cost-effectiveness chapters (see *Chapters 5–9*) and the individual search strategies are listed in *Appendix 2*.

Costs

All unit costs reported in the analysis for health states and liver biopsy are priced for the year 2012. Where required, costs were inflated to 2012 prices using NHS inflation indices.[67] Test costs for the NILTs are costed for the year 2012–13 as costs for some of the components for the NILTs were sourced in early January 2013.

Incremental-cost-per-quality-adjusted-life-year analyses

All analyses were fully incremental. In the incremental analyses, test strategies were ordered according to the least effective and test strategies which were found to be more costly and less effective ('dominated') than another strategy were ruled out of the analysis. Incremental cost-effectiveness ratios (ICERs) were calculated for the tests that were not dominated and test strategies with an ICER greater than that of a more effective intervention ('extendedly dominated') were also ruled out; the ICER was calculated using the formula

$$\text{ICER} = [(C_1 - C_0)/(E_1 - E_0)] \quad (1)$$

where C_1 equals the cost of strategy 1, C_0 equals the cost of (the next best) strategy 0, E_1 equals QALYs from strategy 1 and E_0 equals QALYs from (the next best) strategy.

The cost-effectiveness results for the remaining strategies which were not ruled out (not 'dominated' or 'extendedly dominated') were presented as ICERs.[68]

Probabilistic sensitivity analysis (PSA) was conducted. With a PSA, rather than using the average values for each parameter input, the value is instead drawn from a distribution. The probability distribution for each input variable (natural history data, mortality rates, costs, QALYs, treatment effectiveness and test effectiveness) was constructed using estimates of the mean value and standard error (if required for probability distribution) and Monte Carlo simulation was used to randomly sample from each input distribution simultaneously for 1000 runs of the models. For each of the decision tree model outputs (1000 simulation runs), an average total lifetime cost and QALY was calculated for each testing strategy.

To summarise the uncertainty around the cost-effectiveness result, we constructed cost-effectiveness acceptability curves (CEACs) which are derived from a joint distribution of the costs and effects (QALYs) to represent the probability that a testing option is cost-effective (had the highest net monetary benefit) at different levels of a cost-effectiveness threshold (varied from £0 to £60,000 in analysis). Net benefit is calculated using the formula

$$\text{Net benefit} = E \times CR - C \quad (2)$$

where E is equivalent to the health outcome for a testing strategy, CR equals the ceiling ratio which is the cost-effectiveness threshold (range between £0 to £60,000) and C equals the cost of the testing strategy.[69]

The CEAC represents the probability that a testing option has the highest probability of being cost-effective over a range of threshold values. However, as Fenwick *et al.*[69] have shown, the testing option with the highest probability of being cost-effective may not necessarily have the highest expected net benefit. In this case, the CEAC should not be used to identify the optimal option; instead, the cost-effectiveness acceptability frontier (CEAF) which plots the uncertainty associated with the optimal testing option (option with highest expected net benefit) for different cost-effectiveness threshold values may be more applicable.

We also present the CEAFs to illustrate the probability of any testing strategy being optimal (has the highest expected net benefit) compared with each other over a range of different cost-effectiveness thresholds (threshold value range varied from £0 to £60,000).[69,70]

Cost per correct diagnosis (alcoholic liver diseases and non-alcoholic fatty liver disease)

The cost-per-correct-diagnoses analyses are presented incrementally. We carried out a probabilistic analysis where we estimated the number of correct true responses for each tests (positive and negative responses). We then compared the results of each test incrementally using the cost for each test to rule out tests which were more costly and provided less correct results. Liver biopsy was included as a comparator in the cost-per-correct-diagnosis analyses.

Chapter 4 Results of systematic review and meta-analysis

Systematic review results

Description of studies

The search strategy initially retrieved 114,071 studies, or after duplicate exclusion, 91,097 studies. The flow chart is shown in *Figure 2*. Finally, data from 302 studies were analysed (HCV $n = 162$, HBV $n = 52$, NAFLD $n = 48$, radiology $n = 60$, ALD $n = 12$).[23–31,71–363] All but five of the included studies were captured by the search strategy[79,256,276,332,334] These five studies were retrieved by manually searching the reference lists of included studies and published meta-analyses.

Meta-analysis results

Data analysis was performed separately according to disease aetiology (HCV, HBV, NAFLD and ALD) as there are distinct patterns of fibrosis development in different aetiologies of chronic liver disease. For example, fibrosis in chronic viral hepatitis is characterised by portal-central septa and interface hepatitis, whereas capillarisation of sinusoids and intercellular fibrosis (chicken-wire fibrosis) are typical of alcoholic and non-alcoholic steatohepatitis.[364] This results in a statistically different amount of fibrosis as measured by liver collagen in patients with different aetiologies of liver disease but the same histological stage.[365] This is reflected in disease-specific cut-offs of non-invasive markers for the same histological stage, for example the cut-offs using Fibroscan for F2 fibrosis differ in HBV and HCV,[32] but also in differences in diagnostic accuracy depending on the aetiology of liver disease.[20]

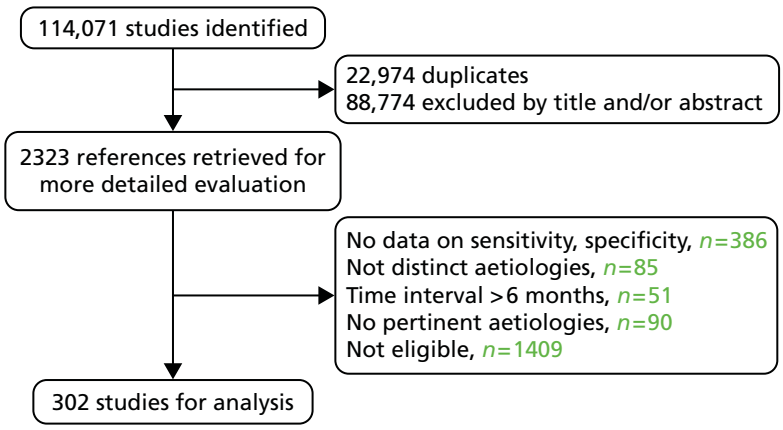

FIGURE 2 Flow chart of literature review search.

Data from radiological methods of fibrosis assessment were pooled and analysed together irrespective of aetiology, as these methods are based on size and contour of the liver, echotexture and signs of portal hypertension rather than on disease-specific fibrotic patterns. Data on Fibroscan, ARFI and real-time elastography were analysed according to the aetiology of liver disease.

Non-invasive test cut-offs for the diagnosis of specific histological stages were not always predetermined, and consequently, varied in the included studies. This probably resulted in higher diagnostic accuracies of the non-invasive tests assessed when the cut-off was not predetermined, as such cut-offs were statistically determined to correlate in the best way with the biopsy results. We opted not to perform a separate meta-analysis for each stage-specific cut-off of a non-invasive test, but to group together cut-offs if the range was reasonable. Therefore, all reported sensitivities and specificities of a non-invasive test, when a range of cut-offs is mentioned in the results tables, are probably overestimated.

A number of NILTs, mainly indirect non-invasive fibrosis tests, report sensitivities and specificities at dual cut-offs, a high cut-off with high specificity and a low cut-off with high sensitivity. The low and high cut-off is usually set at 90–95% of sensitivity and specificity, respectively. Depending on the clinical scenario and the disease prevalence, the low or high cut-off is used at the expense of increased false positives and false negatives respectively. We performed separate meta-analyses for low and high cut-offs whenever such cut-offs were reported and were similar across studies. Patients who have test results greater than the higher cut-off are considered to be test positive and those with test results lower than the lower cut-off are considered to be test negative. If these cut-offs are combined, then false positives and false negatives are minimised but a number of patients will fall in the indeterminate range of fibrosis (i.e. their score will be between the low and the high cut-off) and will need either further non-invasive testing or a liver biopsy. Such patients with intermediate results were considered to have undergone a second test.

Table 4 provides a list of NILTs found, applicable aetiologies and a list of the components.

Results: hepatitis C virus

Data on patients with HCV were extracted from 162 studies.[23–29,31,71–224] Meta-analysis was performed separately for each non-invasive test which had been assessed at each METAVIR stage (F1–F4). Summary sensitivity and specificity for F2 and F4 are shown in *Tables 5* and *6*, while the sensitivity and specificity estimates for F1 and F3 are reported in *Appendix 3*. Individual study characteristics are shown in *Appendix 4*. The median prevalence (minimum–maximum) of fibrosis stages F1–F4 in included studies was for F1 0.875 (0.157–0.968), F2 0.522 (0.063–0.893), F3 0.291 (0.051–0.778) and F4 0.17 (0.026–0.681). Forest plots and summary receiver operating characteristic (SROC) plots of different NILTs across fibrosis stages are presented in *Appendices 5* and *6*, respectively.

For the diagnosis of fibrosis stage ≥ F2, which was the one mainly used in economic modelling, 19 non-invasive tests were evaluated in single studies. Of 47 different evaluated tests, only 18 converged with the bivariate random-effects model [APRI low and high cut-offs, AST–ALT ratio, FIB-4 low and high cut offs, Forns index low and high cut-off, Göteborg University Cirrhosis Index (GUCI), Lok's index, platelet count, hyaluronic acid, Hepascore, Fibrometer, Fibrotest standard, low and high cut-offs, platelet-to-spleen-diameter ratio and Fibroscan]. The most commonly evaluated non-invasive tests were APRI (low cut-off), which was evaluated in 47 studies,[24,31,73,74,79,81,84,85,89,90,91,94,97,98,103,107,109,121,126,127,130,131,134,137,140,143,144,146,150,152–154,156–158,163,164,168,182,185,187,189,194,195,210,218,220,223] followed by TE in 37 studies[28,29,75,76,86–88,91,95,98–100,102,105,106,110,116,119,130,137,141,147,153,155,159,161,164,170,172,173,194,199–201,211,223,224] and APRI (high cut-off) in 37 studies.[24,31,72–74,79,81,89,90,91,94,97,100,103,121,123,126,131,134,140,143,146,150,152–154,156–158,182,187,195,209,210,218,220,223]

TABLE 4 Components of NILTs and applicable aetiology

Test	Components	Comments
Indirect serum non-invasive fibrosis tests		
APGA	AST, platelet count, GGT, α-fetoprotein	HBV
APRI	AST, platelet count	HBV, HCV, NAFLD, ALD
Age–Platelet Index	Age, platelet count	HBV, HCV, NAFLD
AST–ALT ratio	AST, ALT	HBV, HCV, NAFLD
BARD	BMI, AST, ALT, presence of diabetes	
CDS	AST, ALT, platelet count, INR	HBV, HCV, NAFLD
FIB-4	Age, AST, ALT, platelet count	HBV, HCV, NAFLD
Forns index	Age, γ-GT, cholesterol, platelet count	HBV, HCV, ALD
FibroQ	Age, AST, ALT, INR, platelet count	HCV
Fibrosis probability index	Age, past alcohol intake, AST, cholesterol, HOMA-IR	HCV
GUCI	AST, ALT, platelet count	HBV, HCV
Hui index	BMI, total bilirubin, platelet count, albumin	HBV
King's	Age, AST, INR, platelet count	HCV
Lok's index	AST, ALT, platelet count, INR	HBV, HCV, NAFLD
NAFLD fibrosis score	Age, BMI, presence of diabetes or IFG, AST, ALT, platelet count, albumin	NAFLD
NIHCED	Age, prothrombin time, platelet count, AST, ALT, splenomegaly, caudate lobe hypertrophy, right liver lobe atrophy	HCV
PAPAS	Platelet count, age, ALP, α-fetoprotein, AST	HBV
PGAA	Prothrombin time, GGT, apolipoprotein A1, α2-macroglobulin	ALD
Platelet count	Platelets count	HCV, NAFLD
Pohl index	AST, ALT, platelets	HCV
Direct non-invasive fibrosis tests		
^{13}C-caffeine breath test		HBV, HCV, NAFLD
Amino-breath test	Aminopyrine breath test	HCV
CTGF	Connective tissue growth factor	HBV
Fontana	Hyaluronic acid, TIMP-1, platelet count	HCV
Hyaluronic acid	Hyaluronic acid	HBV, HCV, NAFLD
Hepascore	Age, sex, α2-macroglobulin, hyaluronate, bilirubin, γ-GT	HBV, HCV, NAFLD
NAFIC	Ferritin, fasting insulin, type IV collagen	NAFLD
NAFLD diagnostic panel – advanced fibrosis	Presence of diabetes, AST, triglycerides, TIMP-1	NAFLD
NAFLD diagnostic panel – any fibrosis	Presence of diabetes, sex, BMI, triglycerides, M30, M65-M30	NAFLD
PIIINP	Amino-terminal propeptide of type III procollagen	HCV
PIIINP/MMP1 index	PIIINP, MMP1	HCV

continued

TABLE 4 Components of NILTs and applicable aetiology (*continued*)

Test	Components	Comments
Type IV collagen	Type IV collagen	HBV, HCV, NAFLD
YKL-40	YKL-40	HCV, ALD
Commercial non-invasive serum fibrosis tests		
ELF	PIIINP, hyaluronate, TIMP-1	HCV, NAFLD
Fibroindex	Platelet count, AST, γ-globulin	HCV
Fibrometer	Platelets, prothrombin time, macroglobulin, AST, hyaluronate, age, urea	HCV
FibrospectII	α2-macroglobulin, hyaluronate and TIMP-1	HCV
Fibrotest	γ-GT, haptoglobin, bilirubin, A1 apolipopotein, α2-macroglobulin	HBV, HCV, NAFLD, ALD
Imaging modalities		
ARFI	Acoustic radiation force impulse imaging	HBV, HCV, NAFLD
Platelet–spleen ratio	Platelet count, spleen diameter	HCV
Real-time elastography	Real-time elastography	HBV, HCV, NAFLD
Fibroscan	Transient elastography	HBV, HCV, NAFLD, ALD
CT	Computed tomography scan	All aetiologies
MRI	Magnetic resonance imaging	All aetiologies
DW-MRI	Diffusion-weighted magnetic resonance imaging	All aetiologies
MR elastography	Liver stiffness measured with MRI	All aetiologies
US	Conventional ultrasound	All aetiologies
Contrast-enhanced ultrasound	Ultrasound after the intravenous injection of specific contrast material	All aetiologies
US SAPI	Splenic artery pulsatile index measured with ultrasound	All aetiologies
Algorithms of non-invasive fibrosis assessment		
Bordeaux	Synchronous Fibrotest and Fibroscan	HCV
Fibropaca	Synchronous Fibrotest, APRI and Forns index	HCV
Leroy	Synchronous Fibrotest and APRI	HCV
SAFE	APRI and Fibrotest sequentially	HCV

ALP, alkaline phosphatase; APGA, AST, platelet count, gamma-glutamyl transpeptidase (GGT), α-fetoprotein; BARD, BMI, AST–ALT ratio, diabetes; CDS, Cirrhosis Discriminant Score; CTFG, connective tissue growth factor; DW-MRI, diffusion-weighted magnetic resonance imaging; GGT, gamma-glutamyl transpeptidase; GUCI, Göteborg University Cirrhosis Index; HOMA-IR, homeostatic model assessment-insulin resistance; IFG, impaired fasting glucose; INR, International Normalized Ratio; MMP1, matrix metalloproteinase-1; NIHCED, non-invasive hepatitis C-related early detection; PAPAS, Age, ALP, α-fetoprotein, AST; PGAA, prothrombin time, GGT, apolipoprotein A1, α2-macroglobulin; PIIINP, amino-terminal propeptide of type III procollagen; SAFE, sequential algorithm for fibrosis evaluation; SAPI, splenic artery pulsatile index; YKL-40, a direct marker of liver fibrosis; US, ultrasound.

TABLE 5 Diagnostic accuracy of non-invasive tests for detection of fibrosis stage ≥ F2 in patients with chronic HCV

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Indirect non-invasive serum tests					
APRI (low cut-off)	47	0.4–0.7	0.82 (0.77 to 0.86)	0.57 (0.49 to 0.65)	Bivariate random-effects model with correlation between sensitivity and specificity
APRI (high cut-off)	36	1.5	0.39 (0.32 to 0.47)	0.92 (0.89 to 0.95)	Bivariate random-effects model with correlation between sensitivity and specificity
Age–Platelet Index	1	3	0.58 (0.46 to 0.70)	0.70 (0.64 to 0.84)	Single study
AST–ALT ratio	7	0.6–1	0.44 (0.27 to 0.63)	0.71 (0.62 to 0.78)	Bivariate random-effects model with correlation between sensitivity and specificity
CDS	1	6	0.66 (0.59 to 0.73)	0.49 (0.34 to 0.64)	Single study
FIB-4 (low cut-off)	11	0.6–1.45	0.89 (0.79 to 0.95)	0.42 (0.25 to 0.61)	Random-effects model for sensitivity and specificity without correlation
FIB-4 (high cut-off)	9	1–3.25	0.59 (0.43 to 0.73)	0.74 (0.56 to 0.87)	Bivariate random-effects model with correlation between sensitivity and specificity
Forns index (low cut-off)	18	4.2–4.5	0.88 (0.83 to 0.91)	0.40 (0.33 to 0.48)	Bivariate random-effects model with correlation between sensitivity and specificity
Forns index (high cut-off)	15	6.9–8.7	0.35 (0.29 to 0.41)	0.96 (0.92 to 0.98)	Bivariate random-effects model with correlation between sensitivity and specificity
FibroQ	1	1.6	0.78 (0.71 to 0.83)	0.66 (0.51 to 0.78)	Single study
Fibrosis probability index (low cut-off)	2	0.2	0.91 (0.83 to 0.96)	0.45 (0.34 to 0.57)	Fixed-effects model for sensitivity and specificity without correlation
Fibrosis probability index (high cut-off)	2	0.8	0.42 (0.32 to 0.54)	0.95 (0.87 to 0.98)	Fixed-effects model for sensitivity and specificity without correlation
GUCI	3	0.33–1.1	0.65 (0.1 to 1.00)	0.79 (0.03 to 1.00)	Bivariate random-effects model with correlation between sensitivity and specificity
King's	1	9.87	0.84 (0.75 to 0.9)	0.70 (0.61 to 0.79)	Single study
King's (low cut-off)	1	4.46	0.62 (0.55 to 0.69)	0.81 (0.76 to 0.86)	Single study
King's (high cut-off)	1	12.3	0.58 (0.51 to 0.65)	0.79 (0.73 to 0.83)	Single study

continued

TABLE 5 Diagnostic accuracy of non-invasive tests for detection of fibrosis stage ≥ F2 in patients with chronic HCV (*continued*)

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Lok's index	4	0.2–1.67	0.67 (0.55 to 0.77)	0.55 (0.29 to 0.78)	Bivariate random-effects model with correlation between sensitivity and specificity
Platelets	10	48–182	0.50 (0.41 to 0.59)	0.89 (0.83 to 0.93)	Bivariate random-effects model with correlation between sensitivity and specificity
Pohl index	2	Positive	0.06 (0.04 to 0.1)	0.99 (0.93 to 1.00)	Fixed-effects model for sensitivity and specificity without correlation
Direct serum non-invasive serum tests					
Aminopyrine breath test	1	8.1	0.73 (0.57 to 0.85)	0.74 (0.58 to 0.85)	Single study
Hyaluronic acid	8	34–110 ng/ml	0.75 (0.64 to 0.83)	0.75 (0.68 to 0.82)	Bivariate random-effects model with correlation between sensitivity and specificity
Hepascore	10	0.31–0.5	0.73 (0.66 to 0.79)	0.73 (0.65 to 0.79)	Bivariate random-effects model with correlation between sensitivity and specificity
Hepascore (high cut-off)	1	0.84	0.33 (0.24 to 0.43)	0.92 (0.85 to 0.96)	Single study
MP3	1	0.3	0.82 (0.73 to 0.89)	0.73 (0.63 to 0.81)	Single study
PIIINP	2	8.3–9.1	0.78 (0.63 to 0.87)	0.76 (0.54 to 0.90)	Fixed-effects model for sensitivity and specificity without correlation
PIIINP/MMP1 index	1	0.3	0.65 (0.55 to 0.75)	0.85 (0.77 to 0.90)	Single study
Type IV collagen	5	110–298	0.88 (0.71 to 0.96)	0.73 (0.63 to 0.82)	Random-effects model for sensitivity and specificity without correlation
YKL-40 (low cut-off)	1	290	0.80 (0.66 to 0.89)	0.33 (0.26 to 0.41)	Single study
YKL-40 (high cut-off)	1	540	0.33 (0.21 to 0.48)	0.80 (0.73 to 0.86)	Single study
Commercial non-invasive serum tests					
ELF	1	8.75	0.84 (0.69 to 0.92)	0.70 (0.52 to 0.83)	Single study
ELF (low cut-off)	1	9.55	0.90 (0.85 to 0.93)	0.52 (0.43 to 0.61)	Single study
ELF (high cut-off)	1	11.07	0.47 (0.41 to 0.54)	0.90 (0.83 to 0.94)	Single study
Fibroindex (low cut-off)	4	1.25	0.83 (0.15 to 0.99)	0.57 (0.22 to 0.86)	Random-effects model for sensitivity and specificity without correlation

TABLE 5 Diagnostic accuracy of non-invasive tests for detection of fibrosis stage ≥ F2 in patients with chronic HCV (*continued*)

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Fibroindex (high cut-off)	4	2.25	0.24 (0.11 to 0.43)	0.98 (0.93 to 1.00)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
Fibrometer	4	0.42–0.57	0.79 (0.69 to 0.86)	0.73 (0.63 to 0.81)	Bivariate random-effects model with correlation between sensitivity and specificity
FibrospectII	5	42–72	0.78 (0.49 to 0.93)	0.71 (0.59 to 0.80)	Random-effects model for sensitivity and specificity without correlation
Fibrotest	17	0.32–0.53	0.68 (0.58 to 0.77)	0.72 (0.70 to 0.77)	Bivariate random-effects model with correlation between sensitivity and specificity
Fibrotest (low cut-off)	7	0.1–0.3	0.91 (0.86 to 0.94)	0.41 (0.37 to 0.46)	Random-effects model for sensitivity and specificity without correlation
Fibrotest (high cut-off)	10	0.6–0.7	0.57 (0.46 to 0.67)	0.85 (0.74 to 0.92)	Bivariate random-effects model with correlation between sensitivity and specificity
ARFI	3	1.21–1.34	0.79 (0.75 to 0.83)	0.89 (0.84 to 0.93)	Fixed-effects model for sensitivity and specificity without correlation
PLT–Spleen ratio	3	1750–2200	0.88 (0.62 to 0.99)	0.73 (0.41 to 0.99)	Bivariate random-effects model with correlation between sensitivity and specificity
Real-time elastography	1	2.73	0.83 (0.73 to 0.90)	0.92 (0.65 to 0.99)	Single study
Fibroscan	37	5.2–10.1	0.79 (0.74 to 0.84)	0.83 (0.77 to 0.88)	Bivariate random-effects model with correlation between sensitivity and specificity
Combination of fibrosis non-invasive tests algorithms					
Bordeaux	1	–	0.88 (0.85 to 0.91)	0.89 (0.85 to 0.92)	Single study
Fibropaca	1	–	0.85 (0.81 to 0.89)	0.90 (0.86 to 0.93)	Single study
Leroy	1	–	0.90 (0.79 to 0.96)	0.98 (0.95 to 0.99)	Single study
SAFE	4	–	1.00 (1.00 to 1.00)	0.81 (0.80 to 0.83)	Fixed-effects model for sensitivity and specificity without correlation

CDS, Cirrhosis Discriminant Score; MMP1, matrix metalloproteinase-1; MP3, metalloproteinase-3; PIIINP, amino-terminal propeptide of type III procollagen; PLT, platelet; SAFE, sequential algoritm for fibrosis evaluation; YKL-40, a direct marker of liver fibrosis.

TABLE 6 Diagnostic accuracy of non-invasive tests for detection of fibrosis stage ≥ F4 in patients with chronic HCV

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Indirect non-invasive serum tests					
APRI (low cut-off)	24	0.75–1	0.77 (0.73 to 0.81)	0.78 (0.74 to 0.81)	Bivariate random-effects model with correlation between sensitivity and specificity
APRI (high cut-off)	19	2	0.48 (0.41 to 0.56)	0.94 (0.91 to 0.95)	Bivariate random-effects model with correlation between sensitivity and specificity
AST–ALT ratio	13	1	0.49 (0.39 to 0.59)	0.87 (0.75 to 0.94)	Bivariate random-effects model with correlation between sensitivity and specificity
CDS	1	8	0.88 (0.66 to 0.97)	0.67 (0.57 to 0.77)	Single study
FIB-4 (low cut-off)	2	1.45	0.87 (0.74 to 0.94)	0.61 (0.53 to 0.69)	Fixed-effects model for sensitivity and specificity without correlation
FIB-4 (high cut-off)	3	3.25–4.44	0.51 (0.39 to 0.63)	0.86 (0.81 to 0.90)	Fixed-effects model for sensitivity and specificity without correlation
Forns index (low cut-off)	2	3.9–4.2	0.88 (0.60 to 1.00)	0.43 (0.1 to 1.00)	Fixed-effect model for sensitivity and random-effects model for specificity without correlation
Forns index (high cut-off)	1	6.9	0.67 (0.53 to 0.78)	0.91 (0.84 to 0.95)	Single study
GUCI	3	Positive	0.76 (0.07 to 0.99)	0.85 (0.78 to 0.90)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
Lok's index (low cut-off)	2	0.2–0.26	0.84 (0.88 to 1.00)	0.66 (0.01 to 100)	Fixed-effect model for sensitivity and random-effects model for specificity without correlation
Lok's index (high cut-off)	1	0.5	0.40 (0.29 to 0.52)	0.95 (0.91 to 0.97)	Single study
Platelets	10	130–196	0.68 (0.59 to 0.76)	0.86 (0.72 to 0.94)	Bivariate random-effects model with correlation between sensitivity and specificity
Direct serum non-invasive serum tests					
^{13}C-caffeine breath test	2	0.01–1.7	0.88 (0.22 to 0.99)	0.73 (0.18 to 0.97)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
Fontana	1	0.3	0.79 (0.72 to 0.84)	0.66 (0.61 to 0.71)	Single study
Hyaluronic acid	7	78–237 ng/ml	0.80 (0.61 to 0.91)	0.88 (0.78 to 0.94)	Bivariate random-effects model with correlation between sensitivity and specificity

TABLE 6 Diagnostic accuracy of non-invasive tests for detection of fibrosis stage ≥ F4 in patients with chronic HCV (*continued*)

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Hepascore	7	0.84	0.80 (0.68 to 0.88)	0.83 (0.76 to 0.89)	Bivariate random-effects model with correlation between sensitivity and specificity
Hepascore (low cut-off)	1	0.58	0.80 (0.72 to 0.86)	0.83 (0.80 to 0.85)	Single study
Hepascore (high cut-off)	1	1.159	0.39 (0.31 to 0.48)	0.99 (0.98 to 0.99)	Single study
PIIINP	3	0.8–1	0.70 (0.42 to 0.89)	0.84 (0.74 to 0.90)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
Type IV collagen	1	190	0.78 (0.65 to 0.86)	0.72 (0.61 to 0.81)	Single study
Commercial non-invasive serum tests					
ELF	1	9.4	0.93 (0.69 to 0.99)	0.79 (0.67 to 0.88)	Single study
ELF (low cut-off)	1	10.06	0.90 (0.84 to 0.94)	0.53 (0.46 to 0.59)	Single study
ELF (high cut-off)	1	11.73	0.52 (0.43 to 0.60)	0.90 (0.85 to 0.93)	Single study
Fibroindex	1	1.82	0.70 (0.52 to 0.84)	0.91 (0.82 to 0.96)	Single study
Fibrometer	2	0.88	0.72 (0.36 to 0.92)	0.88 (0.60 to 0.97)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
Fibrometer (low cut-off)	1	0.63	0.96 (0.90 to 0.98)	0.71 (0.68 to 0.74)	Single study
Fibrometer (high cut-off)	1	0.98	0.36 (0.28 to 0.45)	0.98 (0.97 to 0.99)	Single study
Fibrotest	8	0.56–0.74	0.60 (0.43 to 0.76)	0.86 (0.81 to 0.91)	Bivariate random-effects model with correlation between sensitivity and specificity
Fibrotest (low cut-off)	1	0.66	0.82 (0.74 to 0.88)	0.77 (0.74 to 0.80)	Single study
Fibrotest (high cut-off)	1	0.86	0.42 (0.34 to 0.51)	0.96 (0.94 to 0.97)	Single study
Imaging modalities					
ARFI	4	1.6–2.3	0.84 (0.72 to 0.91)	0.77 (0.50 to 0.92)	Random-effects model for sensitivity and specificity without correlation
PLT–Spleen ratio	1	Spleen > 120, PLT < 140	0.85 (0.76 to 0.91)	0.82 (0.80 to 0.84)	Single study
Real-time elastography	1	3.93	0.91 (0.73 to 0.98)	0.91 (0.80 to 0.97)	Single study

continued

TABLE 6 Diagnostic accuracy of non-invasive tests for detection of fibrosis stage ≥ F4 in patients with chronic HCV (*continued*)

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Fibroscan	36	9.2–17.3	0.89 (0.84 to 0.92)	0.91 (0.89 to 0.93)	Bivariate random-effects model with correlation between sensitivity and specificity
Combination of fibrosis non-invasive tests algorithms					
Bordeaux	1	–	0.87 (0.80 to 0.92)	0.95 (0.93 to 0.96)	Single study
Fibropaca	1	–	0.73 (0.62 to 0.81)	0.97 (0.95 to 0.98)	Single study
SAFE	4	–	0.74 (0.42 to 0.92)	0.93 (0.91 to 0.94)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation

CDS, Cirrhosis Discriminant Score; GUCI, Göteborg University Cirrhosis Index; PIIINP, amino-terminal propeptide of type III procollagen; PLT, platelet; SAFE, sequential algoritm for fibrosis evaluation.

For the diagnosis of cirrhosis, there were 37 different evaluated tests; however, only nine converged with the bivariate random-effects model (APRI low and high cut-offs, AST–ALT ratio, platelet count, hyaluronic acid, Hepascore, Fibrotest and Fibroscan).

For the diagnosis of fibrosis stage ≥ F1, there were only five tests that reported diagnostic accuracy; however, none converged with the bivariate random-effects model.

For the diagnosis of fibrosis stage ≥ F3, there were 37 different evaluated tests, of which six converged with the bivariate random-effects model (APRI high cut-off, FIB-4 low and high cut-offs, Hepascore, Fibrotest and Fibroscan).

Uninterpretable NILT results were very rare in serum markers (< 1%) and were more frequently encountered in patients who were undergoing Fibroscan examination. The rate of uninterpretable results with Fibroscan (due to < 10 valid measurements, success rate < 60% and interquartile range > 30%) was 8.5%; however, this could be underestimated due to under-reporting.

Cut-offs of non-invasive tests for specific disease stages varied among studies and were predetermined in only 51 studies (31.4%).[25,31,72–74,83,85–87,90–92,94,95,97,99,100,106,109,117,119,127,132–134,143,144,146,150,154,156,157,167,169,171,172,175,186,187,192–194,203,213–216,220,222,223,366] We did not include data on APRI for cirrhosis from some studies in the meta-analysis because some cut-offs differed significantly from what is used in the literature.[82,137,168,194] Liver biopsy was of acceptable quality (≥ 15 cm in length with ≥ 6 portal tracks) in only 20 studies (12.3%),[75,86,110,112,115–117,119,124,127,129,137,141,143,153,186,188,194,222,223] while minimum sample requirements were not reported in 84 (51.8%) studies.[26,27,72–74,77–80,84,85,88,91–93,96,100,101,103,104,106,107,111,114,118,121–123,125,126,130,131,133,135,138,140,147,148,151,152,159–163,165,167–169,172–175,177,178,180,181,183–185,190–192,196–204,208–215,218,221] Overall, only three studies[86,143,222] had a low risk of bias in all of the domains of the QUADAS-2 tool; therefore, all our estimates may be biased. Quality assessment of included studies based on QUADAS-2[62] is shown in *Table 7*. Studies that were judged as low risk of bias or unknown in the three most important QUADAS domains, namely patient selection, index test and reference standard, were still a modest fraction of the total number of studies (29 out of 152; 19%).[72,73,79,80,86,91,92,106,110,133,135,143,148,165,167,173,197–199,201,203,206–208,211,214,215,221,222]

We explored potential sources of heterogeneity as outlined in the methods section. As all but five studies were of low methodological quality,[86,127,143,186,222] this potential source of heterogeneity could not be assessed. Significant heterogeneity was found mainly in relation to the transaminases level, without,

TABLE 7 Quality assessment of studies included in the meta-analysis for chronic HCV. Quality assessment was done using the QUADAS-2 tool

	Domain 1: patient sampling		Domain 2: index test		Domain 3: reference standard		Domain 4: flow and timing
Study ID	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias
Adams 2005[71]	✓	✓	✗	✓	✗	✓	✓
Ahmad 2011[72]	?	✓	✓	✓	?	✓	✓
Al Mohri[73]	?	?	✓	✓	?	✓	✓
Anaparthy 2009[74]	✗	✗	✓	✓	?	✓	?
Arena 2008[75]	✓	✗	✗	✓	✓	✓	✓
Beckebaum 2010[76]	✗	✗	✗	✓	✗	✓	✓
Bejarano 2009[77]	✓	✓	✗	✓	?	✓	✓
Berg 2004[79]	?	?	?	?	?	✓	?
Borroni 2002[80]	?	?	?	?	?	✓	?
Bourliere 2006[81]	✓	✗	✗	✓	✗	✓	✗
Boursier 2009[82]	✓	✓	✗	✓	✗	✓	✗
Boursier 2012[83]	✓	✓	✓	✗	✗	✓	✓
Burton 2010[84]	✗	✗	✓	✓	?	✓	?
Cales 2010[85]	✗	✗	✓	✓	?	✓	✗
Cales 2010[26]	?	?	✗	✓	?	✓	✓
Calvaruso 2010[86]	✓	✓	✓	✓	✓	✓	✓
Cardoso 2012[87]	✓	✓	✓	✓	✗	✓	✗
Carrion 2006[88]	✓	✓	✗	✗	?	✓	✓
Carvalho 2008[89]	✗	✗	✗	✓	✗	✓	✓
Castera 2005[28]	✓	✓	✗	✓	✗	✓	✓
Castera 2007[91]	?	?	✓	✓	?	✓	?
Castera 2009[90]	✓	✓	✓	✓	✗	✓	✓
Ceriani 2001[92]	?	?	✓	?	?	✓	?
Chen 2008[93]	?	?	✗	✓	?	✓	?
Cheung 2008[94]	✗	✗	✓	✗	✗	✓	✓
Cho 2011[95]	✓	✓	✓	✓	✗	✓	✓
Christensen 2006[96]	?	✗	✗	✓	?	✓	✗
Chrysanthos 2006[97]	✓	✓	✓	✓	✗	✓	✓
Cobbold 2010[98]	✓	✓	✗	✓	✗	✓	✗
Colletta 2005[99]	✗	✗	✓	?	✗	✓	✗
Corradi 2009[100]	✗	✗	✓	✓	?	✓	✓
Crespo 2010[101]	?	?	✗	✓	?	✓	✓
Cross 2010[102]	✓	✓	✗	✗	✗	✓	✗
da Silva 2008[103]	?	✗	✗	✓	?	✓	✓

continued

TABLE 7 Quality assessment of studies included in the meta-analysis for chronic HCV. Quality assessment was done using the QUADAS-2 tool (*continued*)

Study ID	Domain 1: patient sampling		Domain 2: index test		Domain 3: reference standard		Domain 4: flow and timing
	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias
Danila 2011[104]	?	?	✗	?	?	✓	✓
De Ledinghen 2006[105]	✓	✓	✗	✗	✗	✓	✓
Degos 2010[106]	✓	✓	✓	✓	?	✓	✓
Dinesen 2008[107]	✗	✗	✗	✓	?	✓	?
Esmat 2007[108]	✓	✓	✗	✓	✗	✓	✗
Fabris 2006[109]	✗	✗	✓	✓	✗	✓	✓
Fahmy 2011[110]	✓	✓	?	✓	✓	✓	✓
Fontaine 2009[111]	✗	✗	✗	✓	?	✓	?
Fontana 2008[112]	✗	✗	✗	✗	✓	✓	?
Fontanges 2008[113]	✗	✗	✗	✗	✗	✓	✓
Forestier 2010[114]	✗	✗	✗	✗	?	?	✗
Forns 2002[115]	✗	✗	✗	✗	✓	✓	✓
Fraquelli 2011[116]	✗	✗	✗	✓	✓	✓	✗
Fujii 2009[117]	✗	✗	✓	✓	✓	✓	✓
Fujimoto 2011[118]	✗	✓	✗	✓	?	✓	✓
Gaia 2011[119]	✗	✓	✓	✓	✓	✓	✗
Ganne-Carrie 2006[120]	✓	✓	✗	✗	✗	✓	✓
Gara 2011[121]	✗	✗	✗	✓	?	✓	✗
Giannini 2006[122]	✓	✓	✗	✓	?	✓	✓
Gobel 2006[123]	✗	✗	✗	✓	?	✓	✓
Guechot 2010[124]	✓	✓	✗	✓	✓	✓	✓
Guechot 1996[125]	✓	✓	✗	✗	?	✓	✓
Guzelbulut 2011[126]	✓	✓	✗	✓	?	✓	✓
Halfon 2006[128]	✓	✓	✗	✓	✗	✓	✓
Halfon 2005[129]	✓	✓	✗	✓	✓	✓	✗
Halfon 2007[127]	✓	✓	✓	✓	✓	✓	✓
Harada 2008[130]	✗	✗	✗	✓	?	✓	✗
Hsieh 2012[131]	✗	✗	✗	✓	?	✓	✓
Iacobellis 2005[132]	✓	✓	✓	✓	✗	✓	✓
Imbert-Bismut 2001[23]	✗	✓	✗	✓	✗	✓	✗
Imperiale 2000[133]	✓	✓	✓	✓	?	✓	✗
Islam 2005[134]	✗	✗	✓	✓	✗	✓	✗
Iushchuk 2005[135]	?	?	?	?	?	?	?

TABLE 7 Quality assessment of studies included in the meta-analysis for chronic HCV. Quality assessment was done using the QUADAS-2 tool (*continued*)

	Domain 1: patient sampling		Domain 2: index test		Domain 3: reference standard		Domain 4: flow and timing
Study ID	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias
Jazia 2009[78]	✗	✗	✗	✓	?	✓	✓
Kalantari 2011[136]	✓	✓	✗	✓	✗	✓	✓
Kamphues 2010[137]	✗	✗	✗	✓	✓	✓	✗
Kandemir 2009[138]	✗	✗	✗	✓	?	✓	✓
Kelleher 2005[139]	✗	✓	✗	✓	✗	✓	✓
Khan 2008[140]	?	✓	✗	✓	?	✓	✓
Kim 2011[141]	✓	✓	✗	✓	✓	✓	✓
Koda 2007[24]	✓	✓	✗	✓	✗	✓	✓
Koizumi 2011[142]	✓	✓	✗	✓	✗	✓	✓
Lackner 2005[143]	✓	✓	✓	✓	✓	✓	✓
Ladero 2010[144]	✓	✓	✓	✓	✗	✓	✓
Lee 2011[145]	✓	✗	✗	✓	✗	✓	✓
Leroy 2004[147]	✓	✓	✗	✓	?	✓	✓
Leroy 2007[146]	✓	✓	✓	✓	✗	✓	✓
Leroy 2011[148]	?	?	?	✓	?	✓	✓
Lewin 2007[149]	✓	✓	✗	✓	✗	✓	✓
Lieber 2006[150]	✗	✗	✓	✓	✗	✓	✗
Liu 2006[151]	✓	✓	✗	✓	?	✓	✓
Liu 2011[153]	✓	✓	✗	✓	✓	✓	✓
Liu 2007[152]	✓	✓	✗	✓	?	✓	✓
Loko 2008[154]	✓	✓	✓	✓	✗	✓	✓
Lupsor 2008[155]	✓	✓	✗	✓	✗	✓	✓
Lupsor 2009[29]	✓	✓	✗	✓	✗	✓	✓
Macias 2006[156]	✓	✓	✓	✓	✗	✓	✓
Macias 2011[157]	?	?	✓	✓	✗	✓	?
Martinez 2011[158]	✓	✓	✗	✓	✗	✓	?
Morikawa 2011[159]	?	✓	✗	✓	?	✓	✓
Murawaki 2001[160]	✗	✗	✗	✓	?	✓	?
Myers 2002[366]	?	✓	✓	✓	✗	✓	✓
Nitta 2009[161]	✓	✓	✗	✓	?	✓	✓
Nojiri 2010[162]	?	✓	✗	✓	?	✓	?
Nunes 2005[163]	✓	✓	✗	✓	?	✓	✗
Obara 2008[164]	✗	✓	✗	✓	✗	✓	✗
Oliveira 2005[165]	?	?	?	✓	?	✓	?

continued

TABLE 7 Quality assessment of studies included in the meta-analysis for chronic HCV. Quality assessment was done using the QUADAS-2 tool (*continued*)

Study ID	Domain 1: patient sampling		Domain 2: index test		Domain 3: reference standard		Domain 4: flow and timing
	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias
Orrlachio 2011[166]	✗	?	✗	✓	✗	✓	✗
Paggi 2008[167]	✓	✓	✓	✓	?	✓	✗
Parise 2006[168]	✓	✓	✗	✓	?	✓	✓
Park 2000[169]	✗	✓	✓	✓	?	✓	✗
Parkes 2011[27]	✗	✓	✗	✓	?	✓	✗
Patel 2009[25]	?	✓	✓	✓	✗	✓	?
Patel 2011[170]	?	✓	✗	✓	✗	✓	?
Pohl 2001[171]	✗	✓	✓	✓	✗	✓	?
Poynard 2012[172]	✗	✓	✓	✓	?	✓	✗
Prati 2011[173]	?	✓	?	✓	?	✓	✗
Qiu 2004[174]	✓	✓	✗	✓	?	✓	✓
Reedy 1998[175]	✗	✓	✓	✓	?		✓
Ronot 2010[176]	✓	✓	✗	✓	✗		✓
Rossi 2003[177]	✓	✓	✗	✓	?		?
Rossini 2010[178]	?	?	✗	✓	✓		?
Said 2010[179]	✓	✓	✗	✓	✓		?
Saitou 2005[180]	?	✗	✗	✓	✓		?
Sanvisens 2009[181]	✓	✓	✗	✓	✓		✗
Schiavon 2007[182]	✗	✓	✗	✓	✓	✓	✗
Schiavon 2008[183]	✗	✓	✗	✓	✓	✓	✗
Schneider 2006[185]	?	?	✗	✓	✓	✓	?
Scneider 2005[184]	✓	✓	✗	✓	?	✓	✓
Sebastiani 2012[31]	✓	✓	✓	✓	✗	✓	✓
Sebastiani 2009[187]	✗	✓	✓	✓	✗	✓	✓
Sebastiani 2006[188]	✓	✓	✗	✓	✓	✓	✓
Sebastiani 2008[186]	✓	✓	✓	✓	✓	✓	✓
Sene 2006[189]	✗	✗	✗	✓	✗	✓	?
Sharabash 2009[190]	✗	✓	?	✓	?	✓	?
Shastry 2007[191]	?	?	✗	✓	?	✓	✓
Sheth 1997[192]	✗	✓	✓	✓	?	✓	✗
Singal 2011[193]	✓	✓	✓	✓	✗	✓	✗
Sirli 2010[194]	✗	✓	✓	✓	✓	✓	✓
Snyder 2007[196]	✓	✓	✗	✓	?	✓	✓
Snyder 2006[195]	✗	✓	✗	✓	✗	✓	✗
Sohn 2010[197]	?	✓	?	✓	?	✓	?

TABLE 7 Quality assessment of studies included in the meta-analysis for chronic HCV. Quality assessment was done using the QUADAS-2 tool (*continued*)

	Domain 1: patient sampling		Domain 2: index test		Domain 3: reference standard		Domain 4: flow and timing
Study ID	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias
Sporea 2008[199]	✓	✓	?	✓	?	✓	✗
Sporea 2010[200]	✓	✓	✗	✓	?	✓	✓
Sporea 2011[201]	?	✓	?	✓	?	✓	✓
Sporea 2011[198]	?	✓	?	✓	?	✓	?
Sterling 2006[202]	✓	✗	✗	✓	?	✓	✗
Stibbe 2011[203]	?	✓	✓	✓	?	✓	✗
Sud 2009[204]	✓	✓	✗	✓	?	✓	✓
Testa 2006[205]	?	✓	✗	✓	✗	✓	✗
Thompson 2009[208]	?	✓	?	✓	?	✓	?
Thompson 2009[206]	?	✓	?	✓	?	✓	?
Thompson 2010[207]	?	✓	?	✓	?	✓	?
Toniutto 2007[209]	✗	✗	✗	✓	?	✓	✓
Trang 2008[210]	?	✓	✗	✓	?	✓	✗
Trifan 2009[211]	✓	✓	?	✓	?	✓	✓
Trocme 2006[212]	?	✓	✗	✓	?	✓	✗
Vallet-Pichard 2007[214]	✓	✓	✓	✓	?	✓	✓
Tural 2007[213]	✓	✗	✓	✓	?	✓	✓
Valva 2011[215]	?	✓	✓	✓	?	✓	✓
Varaut 2005[216]	✗	✗	✓	✓	✗	✓	✗
Wai 2003[218]	✓	✗	✗	✓	?	✓	✗
Westin 2008[219]	✓	✓	?	✓	✗	✓	?
Wilson 2006[220]	✗	✗	✓	✓	✗	✓	?
Wong 1998[221]	✓	✓	?	✓	?	✓	?
Zaman 2004[222]	✓	✓	✓	✓	✓	✓	✓
Zarski 2012[223]	✓	✗	✓	✓	✓	✓	✗
Ziol 2005[224]	✗	✓	✗	✓	✗	✓	✗

✗, high risk of bias; ✓, low risk of bias; ?, unclear risk of bias.

however, a particular pattern in relation to the transaminase elevation. More specifically and according to source of heterogeneity, details of significant associations are:

- Full text versus abstract publication: Fibrotest for ≥ F2 (significantly higher specificity if published in full text) and borderline for ≥ F3 (likelihood ratio test, $p = 0.059$) and platelet count for F4 (significantly lower specificity if published in full text).
- Transaminases levels (comparator is normal transaminases): for ≥ F2 Fibrotest, hyaluronic acid, Hepascore, Lok's index, Fibroscan; for ≥ F3 FIB-4 low and high cut-off, TE; and for F4 AST–ALT ratio. There was no specific pattern of influence; therefore, the test could have improved or worse diagnostic accuracy if the transaminases levels were high.
- Histological score used (comparator is the METAVIR system): for Hepascore in ≥ F3, use of Ludwig scoring system in one study[76] resulted in significantly lower sensitivity. This might reflect the particular study rather than the histological score used; for AST–ALT ratio in F4, use of Ludwig scoring system resulted in significantly higher sensitivity; and for platelet count in F4, use of Scheuer or Knodell resulted in significantly higher sensitivity and specificity, respectively.

Results: hepatitis B virus

Data on patients with HBV were extracted from 52 studies.[116,119,120,200,225–272] Meta-analyses were performed separately for each non-invasive test assessed at each METAVIR stage (F1–F4). Summary sensitivity and specificity for fibrosis stages F2 and F4 are shown in *Tables 8* and *9*, whereas summary sensitivity and specificity for fibrosis stages F1 and F3 are reported in *Appendix 3*. The median prevalence (minimum–maximum) of fibrosis stages F1–F4 in included studies was for F1 0.617 (0.416–0.884), F2 0.528 (0.269–0.915), F3 0.370 (0.171–0.780) and F4 0.209 (0–0.604). Individual study characteristics, and forest plots and SROC plots of the different NILTs across fibrosis stages, are presented in *Appendices 4*, *5* and *6*, respectively.

Overall, there were 18 different non-invasive tests reported for the diagnosis of ≥ F2, of which 13 (72%) were reported in single studies. Of 18 different evaluated tests, only five converged with the bivariate random-effects model (APRI low and high cut-off, FIB-4 low cut-off, Fibrotest and Fibroscan). The most commonly evaluated non-invasive tests were Fibroscan (13 studies),[116,200,225,230,241,242,246,247,250,251,263,264,272] APRI (low cut-off, eight studies),[225,233,256–259,269,272] APRI (high cut-off, six studies)[225,256–259,270] and Fibrotest (six studies).[225,247,249,255–257]

For the diagnosis of cirrhosis, there were 14 different evaluated tests; however, only Fibroscan converged with the bivariate random-effects model.

For the diagnosis of fibrosis stage ≥ F1, there were eight tests that reported on diagnostic accuracy; however, none converged with the bivariate random-effects model.

For the diagnosis of fibrosis stage ≥ F3, there were 10 different evaluated tests, of which only Fibroscan converged with the bivariate random-effects model.

We explored potential sources of heterogeneity as outlined in the methods section. As all but one study[200] were of low methodological quality, this potential source of heterogeneity could not be assessed. Significant heterogeneity was only found in relation to the transaminases level (comparator is normal transaminases): lower specificity for FIB-4 low cut-off in F2, and borderline lower specificity for Fibrotest in F2 ($p = 0.055$).

Cut-offs for specific histological stages were predetermined in 11 studies (21%).[119,200,225,230,232,252,253,255,258,264,271] Liver biopsy was of acceptable quality in 12 studies (23%).[116,119,200,225,227,230,234,251,263,266,271,272] We did not include data from some studies on APRI for F2 and F3,[242] on Forns index for F2[269] and on AST–ALT ratio for cirrhosis[254] in the meta-analysis because of cut-offs that differed significantly from what is used in the literature. Only one study[257] had low risk of bias in all of the domains of the QUADAS-2 tool; therefore, all our estimates may be biased and should be assessed with caution. Studies that were judged as low risk of bias or unknown in the

TABLE 8 Diagnostic accuracy of non-invasive tests for detection of fibrosis states ≥ F2 in patients with chronic HBV

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Indirect non-invasive serum tests					
APGA	1	6.7	0.17 (0.10 to 0.27)	0.98 (0.95 to 0.99)	Single study
APRI (low cut-off)	8	0.4–0.6	0.80 (0.68 to 0.88)	0.65 (0.52 to 0.77)	Bivariate random-effects model with correlation between sensitivity and specificity
APRI (high cut-off)	6	1.5	0.37 (0.22 to 0.55)	0.93 (0.85 to 0.97)	Bivariate random-effects model with correlation between sensitivity and specificity
Age–Platelet Index	1	3	0.68 (0.61 to 0.74)	0.62 (0.57 to 0.67)	Single study
AST–ALT ratio	1	0.67	0.57 (0.51 to 0.64)	0.59 (0.54 to 0.63)	Single study
FIB-4 (low cut-off)	4	1.1–1.7	0.68 (0.60 to 0.75)	0.73 (0.67 to 0.79)	Bivariate random-effects model with correlation between sensitivity and specificity
FIB-4 (high cut-off)	1	3.25	0.58 (0.04 to 0.17)	0.99 (0.96 to 1.00)	Single study
Forns index (low cut-off)	1	4.2	0.58 (0.47 to 0.68)	0.77 (0.61 to 0.88)	Single study
Forns index (high cut-off)	1	6.9	0.15 (0.08 to 0.24)	1.00 (0.90 to 1.00)	Single study
GUCI	1	0.2	0.67 (0.55 to 0.76)	0.97 (0.85 to 0.99)	Single study
Hui index	1	0.15	0.50 (0.39 to 0.61)	0.91 (0.78 to 0.97)	Single study
PAPAS	1	1.67	0.73 (0.62 to 0.81)	0.78 (0.71 to 0.84)	Single study
Direct serum non-invasive serum tests					
Hyaluronic acid	1	185.3	0.84 (0.73 to 0.91)	0.83 (0.66 to 0.93)	Single study
Hepascore	1	0.5	0.79 (0.68 to 0.86)	0.74 (0.65 to 0.81)	Single study
Commercial non-invasive serum tests					
Fibrotest	6	0.40–0.48	0.66 (0.57 to 0.75)	0.80 (0.72 to 0.86)	Bivariate random-effects model with correlation between sensitivity and specificity
Imaging modalities					
ARFI	1	1.33	0.71 (0.59 to 0.80)	0.67 (0.30 to 0.90)	Single study
Real-time elastography	1	55.3	0.82 (0.67 to 0.91)	0.65 (0.49 to 0.78)	Single study
Fibroscan	13	6.3–8.9	0.71 (0.62 to 0.78)	0.84 (0.74 to 0.91)	Bivariate random-effects model with correlation between sensitivity and specificity

APGA, AST, platelet count, gamma-glutamyl transpeptidase (GGT), α-fetoprotein; PAPAS, Age, ALP, α-fetoprotein, AST.

TABLE 9 Diagnostic accuracy of non-invasive tests for detection of ≥ F4 in patients with chronic HBV

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Indirect non-invasive serum tests					
APRI (low cut-off)	4	1	0.58 (0.49 to 0.66)	0.76 (0.70 to 0.81)	Fixed-effect model for sensitivity and random-effects model for specificity without correlation
APRI (high cut-off)	3	2	0.24 (0.08 to 0.52)	0.91 (0.83 to 0.96)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
Age–Platelet Index	2	4–4.5	0.83 (0.72 to 0.90)	0.74 (0.66 to 0.80)	Fixed-effects model for sensitivity and specificity without correlation
AST–ALT ratio	3	1	0.33 (0.04 to 0.83)	0.77 (0.69 to 0.84)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
FIB-4 (low cut-off)	2	1.6–1.9	0.86 (0.79 to 0.91)	0.82 (0.77 to 0.86)	Fixed-effects model for sensitivity and specificity without correlation
FIB-4 (high cut-off)	1	3.6	0.30 (0.24 to 0.36)	0.98 (0.97 to 0.99)	Single study
GUCI	1	1	0.23 (0.10 to 0.43)	0.91 (0.83 to 0.95)	Single study
Direct serum non-invasive serum tests					
Hyaluronic acid	1	77	0.82 (0.52 to 0.95)	0.88 (0.79 to 0.93)	Single study
Hepascore	1	0.87	0.87 (0.62 to 0.96)	0.85 (0.78 to 0.89)	Single study
Type IV collagen	1	6.3	0.64 (0.35 to 0.85)	0.89 (0.81 to 0.94)	Single study
Commercial non-invasive serum tests					
Fibrotest	4	0.58–0.74	0.74 (0.25 to 0.96)	0.90 (0.83 to 0.94)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
Imaging modalities					
Platelet–spleen ratio	2	6	0.83 (0.74 to 0.90)	0.77 (0.70 to 0.83)	Random-effects model for sensitivity and specificity without correlation
Real-time elastography	1	90.3	0.71 (0.50 to 0.86)	0.80 (0.69 to 0.88)	Single study
Fibroscan	19	9.4–16.0	0.86 (0.79 to 0.91)	0.85 (0.78 to 0.89)	Bivariate random-effects model with correlation between sensitivity and specificity

three most important QUADAS domains, namely patient selection, index test and reference standard, were still a fraction of the total number of studies (11 out of 52; 23%).[225,235,236,241,242,244,246,257,263,265,271] Quality assessment of included studies based on QUADAS-2[62] is shown in *Table 10*.

Results: non-alcoholic fatty liver disease

Data on patients with NAFLD were extracted from 48 studies.[117,119,165,229,284–327] Meta-analysis was performed separately for each non-invasive test assessed at each Kleiner stage (F1–F4). Summary sensitivity and specificity for F3 and F4 are shown in *Tables 11* and *12*, while summary sensitivity and specificity for F1 and F2 are reported in *Appendix 3*. The median prevalence (minimum–maximum) of fibrosis stages F1–F4 in included studies was for F1 0.588 (0.367–0.814), F2 0.319 (0.119–0.526), F3 0.186 (0.050–0.440) and F4 0.128 (0.039–0.907). Individual study characteristics are presented in *Appendix 4*. The prevalence of F1–F4 in NAFLD is lower than the prevalence of such stages in the other evaluated aetiologies of liver disease; this is probably due to the relatively low prevalence of the progressive steatohepatitis among patients with NAFLD.[47] Forest plots and SROC plots of different NILTs across fibrosis stages are presented in *Appendices 5* and *6*, respectively.

Overall, there were 24 different non-invasive tests reported for the diagnosis of ≥ F3, of which 11 (46%) were reported in single studies.[117,229,284,286,287,289,290,299,304,306,327] Of 24 different evaluated tests, 10 converged with the bivariate random-effects model [BARD (BMI, AST-ALT ratio, diabetes), AST–ALT ratio low cut-off, NAFLD fibrosis score low and high cut-offs, FIB-4 low and high cut-off, hyaluronic acid, type IV collagen, Fibrotest and Fibroscan]. The most commonly evaluated non-invasive tests were NAFLD fibrosis score (low cut-off, 10 studies), [290,300,301,309,311,315,320,322] NAFLD fibrosis score (high cut-off, nine studies),[285,290,300,301,309,311,315,320,322] Fibroscan (eight studies)[119,236,288,296,298,308,323,324] and BARD (seven studies).[284,291,300,301,312,315,319]

For the diagnosis of cirrhosis, there were 14 different evaluated tests; however, only platelet count and Fibroscan converged with the bivariate random-effects model.

For the diagnosis of fibrosis stage ≥ F1, there were 12 tests that reported on diagnostic accuracy, and three of them converged with the bivariate random-effects model (NAFLD fibrosis score low and high cut-offs, TE).

For the diagnosis of fibrosis stage ≥ F2, there were 20 different evaluated tests, of which three converged with the bivariate random-effects model (NAFLD fibrosis score low and high cut-offs, Fibrotest).

Uninterpretable NILT results were very rare in serum markers (< 1%) and were more frequently encountered in patients who were undergoing Fibroscan examination. The rate of uninterpretable results with Fibroscan (due to < 10 valid measurements, success rate < 60% and interquartile range> 30%) was 9.6%; however, this could be underestimated due to under-reporting.

Cut-offs for specific histological stages were predetermined in 10 studies (21%).[117,119,284,291,309,311,312,315,321,322] Liver biopsy was of acceptable quality in 10 studies (21%).[117,119,229,284,286,293,308,319,321,326] We did not include data on APRI and NAFLD fibrosis scores from one study[327] in the meta-analysis because of cut-offs that differed significantly from what is used in the literature. Only one study[119] had low risk of bias in all of the domains of the QUADAS-2 tool; therefore, all our estimates may be biased. Studies that were judged as low risk of bias or unknown in the three most important QUADAS domains, namely patient selection, index test and reference standard, were 21% of the total number of studies (10 out of 48).[119,165,284,294,298,301,303,319,323] Quality assessment of included studies based on QUADAS-2[62] is shown in *Table 13*.

TABLE 10 Quality assessment of studies included in the meta-analysis for chronic HBV. Quality assessment was done using the QUADAS-2 tool

	Domain 1: patient sampling		Domain 2: index test		Domain 3: reference standard		Domain 4: flow and timing
Study ID	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias
Castera 2011[225]	?	✓	✓	✓	✓	✓	✗
Chan 2009[226]	✗	✗	✗	✓	?	✓	✓
Chen 2008[227]	✗	✗	✗	✓	✓	✓	?
Chen 2012[228]	✗	✗	?	✓	?	✓	✗
Fraquelli 2011[116]	✓	✓	✗	✓	✓	✓	?
Fung 2011[230]	✗	✗	✓	✓	✓	✓	✓
Gaia 2011[119]	✓	✓	✓	✓	✓	✓	✗
Ganne-Carrie 2006[120]	✓	✗	?	✓	?	✓	✓
Gui 2010[231]	✗	✗	✗	✓	?	✓	✓
Guo-Qiu 2010[232]	✗	✗	✓	✓	?	✓	?
Hongbo 2007[233]	✓	✓	✗	✓	?	✓	✓
Hu 2010[234]	?	✓	✗	✓	✓	✓	?
Hui 2005[235]	?	?	?	?	?	?	?
Kim 2009[238]	✓	✓	✗	✓	?	✓	✓
Kim 2010[236]	?	?	✗	?	?	✓	✗
Kim 2009[239]	✓	✓	✗	✓	✗	✓	✓
Kim 2007[237]	✓	✓	✗	✓	?	✓	?
Kim 2010[236]	?	?	?	?	?	?	?
Kwok 2009[240]	✓	✓	✗	✓	?	✓	✗
Lee 2011[241]	✓	✓	?	✓	?	✓	✓
Lesmana[242]	?	?	?	?	?	?	?
Li 2012[243]	✗	✓	✗	✓	✗	✓	✗
Liu 2011[244]	?	?	?	✓	?	✓	?
Mallet 2009[245]	✗	✓	✗	✓	✗	✓	?
Marcellin 2009[246]	?	✓	?	?	?	✓	✗
Miailhes 2011[247]	✗	✓	✗	✓	?	✓	✗
Mohamadnejad 2006[248]	?	?	✗	?	✗	✓	?
Myers 2003[249]	✓	✓	✗	✓	✗	✓	✓
Ogawa 2011[250]	?	?	✗	✓	?	✓	?
Osakabe 2011[251]	✓	✓	✗	✓	✓	✓	✓
Park 2003[253]	✗	✓	✓	✓	✗	✓	✓
Park 2004[254]	✓	✓	✗	✓	?	✓	✓
Park 2005[252]	✗	✓	✓	✓	✗	✓	✓
Poynard 2009[255]	✗	✓	✓	✓	?	✓	?

TABLE 10 Quality assessment of studies included in the meta-analysis for chronic HBV. Quality assessment was done using the QUADAS-2 tool (*continued*)

	Domain 1: patient sampling		Domain 2: index test		Domain 3: reference standard		Domain 4: flow and timing
Study ID	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias
Raftopoulos 2012[256]	?	✓	✗	✓	✗	✓	?
Sebastiani 2007[257]	✓	✓	✗	✓	?	✓	✓
Seto 2011[258]	✗	✓	✓	✓	?	✓	✓
Shin 2008[259]	?	✓	✗	✓	?	✓	?
Sinakos 2011[260]	?	✓	✗	✓	?	✓	?
Sohn 2011[261]	?	✓	✗	✓	?	✓	?
Sokucu 2010[262]	✓	✓	✗	✓	?	✓	?
Sporea 2010[263]	?	✓	?	✓	✓	✓	?
Sporea 2010[200]	✓	✓	✓	✓	✓	✓	✓
Vigano 2011[264]	?	✓	✓	✓	✗	✓	✓
Wang 2012[265]	✓	✓	?	✓	?	✓	✓
Wong 2008[267]	✓	✓	✗	✓	?	✓	✗
Wong 2010[266]	?	✓	✗	✓	✓	✓	✗
Wong 2011[268]	?	✓	✗	✓	?	✓	?
Wu 2012[269]	?	✓	✗	✓	✗	✓	✓
Zhang 2008[271]	?	✓	✓	✓	✓	✓	✗
Zhang 2011[270]	✗	✗	✗	✓	?	✓	✓
Zhu 2011[272]	✗	✗	✗	✓	✓	✓	?

✗, high risk of bias; ✓, low risk of bias; ?, unclear risk of bias.

TABLE 11 Diagnostic accuracy of non-invasive tests for detection of fibrosis stage of ≥ F3 in patients with non-alcoholic steatohepatitis

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Indirect non-invasive serum tests					
APRI	4	0.5–1.0	0.40 (0.07 to 0.86)	0.82 (0.78 to 0.6)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
Age–Platelet Index	1	6	0.66 (0.53 to 0.76)	0.78 (0.74 to 0.81)	Single study
AST–ALT ratio (low cut-off)	4	0.8	0.79 (0.51 to 0.91)	0.70 (0.55 to 0.82)	Bivariate random-effects model with correlation between sensitivity and specificity
AST–ALT ratio (high cut-off)	3	1.0	0.46 (0.29 to 0.65)	0.91 (0.85 to 0.95)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
BARD	7	2	0.84 (0.69 to 0.93)	0.61 (0.47 to 0.73)	Bivariate random-effects model with correlation between sensitivity and specificity
FIB-4 (low cut-off)	4	1.3–1.92	0.84 (0.75 to 0.90)	0.74 (0.64 to 0.83)	Bivariate random-effects model with correlation between sensitivity and specificity
FIB-4 (high cut-off)	2	3.25	0.38 (0.22 to 0.57)	0.97 (0.92 to 0.99)	Bivariate random-effects model with correlation between sensitivity and specificity
NAFLD fibrosis score (low cut-off)	10	–1.455	0.80 (0.67 to 0.89)	0.66 (0.57 to 0.74)	Bivariate random-effects model with correlation between sensitivity and specificity
NAFLD fibrosis score (high cut-off)	9	0.676	0.40 (0.20 to 0.64)	0.97 (0.94 to 0.98)	Bivariate random-effects model with correlation between sensitivity and specificity
Platelets	1		0.63 (0.57 to 0.69)	0.76 (0.74 to 0.78)	Single study
Direct serum non-invasive serum tests					
Hyaluronic acid	4	46–50	0.88 (0.58 to 0.97)	0.82 (0.75 to 0.87)	Bivariate random-effects model with correlation between sensitivity and specificity
Hepascore	1	0.37	0.75 (0.62 to 0.85)	0.84 (0.78 to 0.89)	Single study
NAFIC (low cut-off)	1	1	0.96 (0.88 to 0.98)	0.67 (0.63 to 0.71)	Single study
NAFIC (high cut-off)	1	3	0.84 (0.73 to 0.91)	0.82 (0.79 to 0.85)	Single study

TABLE 11 Diagnostic accuracy of non-invasive tests for detection of fibrosis stage of ≥ F3 in patients with non-alcoholic steatohepatitis (*continued*)

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
NDP: advanced fibrosis	1	0.24	0.88 (0.64 to 0.96)	0.70 (0.58 to 0.80)	Single study
Type IV collagen	2	5	0.79 (0.69 to 0.87)	0.80 (0.66 to 0.89)	Bivariate random-effects model with correlation between sensitivity and specificity
Commercial non-invasive serum tests					
ELF	1	10.35	0.80 (0.65 to 0.89)	0.90 (0.84 to 0.94)	Bivariate random-effects model with correlation between sensitivity and specificity
Fibrotest (low cut-off)	3	0.3	0.88 (0.68 to 0.99)	0.73 (0.56 to 0.85)	Random-effects model for sensitivity and specificity without correlation
Fibrotest (high cut-off)	4	0.57–0.70	0.40 (0.24 to 0.58)	0.96 (0.91 to 0.98)	Bivariate random-effects model with correlation between sensitivity and specificity
Imaging modalities					
ARFI	1	4.2	0.90 (0.77 to 0.96)	0.90 (0.82 to 0.94)	Single study
Fibroscan	8	7.5–10.4	0.82 (0.74 to 0.88)	0.84 (0.78 to 0.89)	Bivariate random-effects model with correlation between sensitivity and specificity
Combination of non-invasive test algorithms					
NAFLD fibrosis score and ELF (low cut-off)	1		0.91 (0.79 to 0.96)	0.96 (0.91 to 0.98)	Single study
NAFLD fibrosis score and ELF (high cut-off)	1		0.86 (0.73 to 0.94)	0.99 (0.96 to 1.00)	Single study
Fibroscan and Fibrotest	1		0.39 (0.27 to 0.53)	0.96 (0.92 to 0.98)	Single study

BARD, BMI, AST–ALT ratio, diabetes; NAFIC, ferritin, fasting insulin, type IV collagen; NDP, NAFLD diagnostic panel.

TABLE 12 Diagnostic accuracy of non-invasive tests for detection of fibrosis stage of ≥ F4 in patients with non-alcoholic steatohepatitis

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Indirect non-invasive serum tests					
APRI	2	0.54 and NA	0.78 (0.71 to 0.99)	0.71 (0.30 to 0.93)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
Age–Platelet Index	1	NA	0.89 (0.56 to 0.98)	0.83 (0.69 to 0.91)	Single study
AST–ALT ratio	1	1	0.89 (0.56 to 0.98)	0.73 (0.58 to 0.84)	Single study
BARD	1	2	0.52 (0.33 to 0.71)	0.84 (0.79 to 0.88)	Single study
CDS (low cut-off)	1	3	0.89 (0.56 to 0.98)	0.90 (0.87 to 0.96)	Single study
CDS (high cut-off)	1	5	0.33 (0.12 to 0.65)	1.00 (0.91 to 1.00)	Single study
FIB-4 (low cut-off)	1	1.92	0.74 (0.54 to 0.87)	0.71 (0.64 to 0.76)	Single study
Lok's index (low cut-off)	1	0.6	0.89 (0.56 to 0.98)	0.68 (0.53 to 0.80)	Single study
Lok's index (high cut-off)	1	0.97	0.22 (0.06 to 0.55)	1.00 (0.91 to 1.00)	Single study
Platelets	2	160,000	0.96 (0.30 to 0.99)	0.92 (0.85 to 0.96)	Bivariate random-effects model with correlation between sensitivity and specificity
Direct serum non-invasive serum tests					
^{13}C-caffeine breath test	1	1.27	0.90 (0.60 to 0.98)	0.76 (0.61 to 0.87)	Single study
Hepascore	1	0.7	0.87 (0.68 to 0.95)	0.89 (0.84 to 0.93)	Single study
Commercial non-invasive serum tests					
Fibrotest	1	0.57	0.74 (0.54 to 0.87)	0.92 (0.88 to 0.95)	Single study
Imaging modalities					
Fibroscan	4	10.3–17.5	0.96 (0.83 to 0.99)	0.89 (0.85 to 0.92)	Bivariate random-effects model with correlation between sensitivity and specificity

BARD, BMI, AST–ALT ratio, diabetes; CDS, Cirrhosis Discriminant Score; NA, not applicable.

TABLE 13 Quality assessment of studies included in the meta-analysis for NAFLD. Quality assessment was done using the QUADAS-2 tool

	Domain 1: patient sampling		Domain 2: index test		Domain 3: reference standard		Domain 4: flow and timing
Study ID	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias
Adams 2011[284]	?	✓	✓	✓	✓	✓	✗
Angulo 2007[285]	✗	✗	✗	✓	?	✓	✓
Blomme 2012[286]	✗	✗	✗	✓	✓	✓	?
Cales 2009[287]	✗	✗	?	✓	?	✓	✗
Dixon 2001[229]	✓	✓	✗	✓	✓	✓	?
Fujii 2009[117]	✗	✗	✓	✓	✓	✓	✓
Gaia 2011[119]	✓	✓	✓	✓	✓	✓	✗
Guajardo-Salinas 2010[289]	✓	✗	?	✓	?	✓	✓
Guha 2008[290]	✗	✗	✗	✓	?	✓	✓
Harrison 2008[291]	✗	✗	✓	✓	?	✓	?
Kaneda 2006[292]	✓	✓	✗	✓	?	✓	✓
Kayadibi 2009[293]	?	✓	✗	✓	✓	✓	?
Kelleher 2006[294]	?	?	?	?	?	?	?
Khosravi 2011[295]	✓	✓	✗	✓	?	✓	✓
de Ledinghen 2009[288]	?	?	✗	?	?	✓	✗
Lupsor 2010[296]	✓	✓	✗	✓	✗	✓	✓
Lydatakis 2006[297]	✓	✓	✗	✓	?	✓	?
Mahadeva 2010[298]	?	?	?	?	?	?	?
Manousou 2011[299]	✓	✓	✗	✓	?	✓	✗
McPherson 2010[301]	✓	✓	?	✓	?	✓	✓
McPherson 2011[300]	?	?	?	?	?	?	?
Obara 2008[164]	✗	✓	✗	✓	✗	✓	✗
Oliveira 2005[165]	?	?	?	✓	?	✓	?
Orlacchio 2012[302]	✗	✓	✗	✓	✗	✓	?
Pais 2011[303]	?	✓	?	?	?	✓	✗
Palmeri 2011[304]	✗	✓	✗	✓	?	✓	✗
Papalavrentios 2011[305]	?	?	✗	?	✗	✓	?
Park 2011[306]	✓	✓	✗	✓	✗	✓	✓
Pawitpok 2006[307]	?	?	✗	✓	?	✓	?
Petta 2011[308]	✓	✓	✗	✓	✓	✓	✓
Pimentel 2010[309]	✗	✓	✓	✓	✗	✓	✓

continued

TABLE 13 Quality assessment of studies included in the meta-analysis for NAFLD. Quality assessment was done using the QUADAS-2 tool (*continued*)

Study ID	Domain 1: patient sampling		Domain 2: index test		Domain 3: reference standard		Domain 4: flow and timing
	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias
Poynard 2006[310]	✓	✓	✗	✓	?	✓	✓
Qureshi 2008[311]	✗	✓	✓	✓	✗	✓	✓
Raszeja-Wyscomirska 2010[312]	✗	✓	✓	✓	?	✓	?
Ratziu 2004[313]	?	✓	✗	✓	✗	✓	?
Ratziu 2006[314]	✓	✓	✗	✓	?	✓	✓
Ruffilo 2011[315]	✗	✓	✓	✓	?	✓	✓
Sakugawa 2005[316]	?	✓	✗	✓	?	✓	?
Santos 2005[317]	?	✓	✗	✓	?	✓	?
Shimada 2007[318]	?	✓	✗	✓	?	✓	?
Sumida 2011[320]	✓	✓	✗	✓	?	✓	?
Sumida 2012[319]	?	✓	?	✓	✓	✓	?
Suzuki 2005[321]	✓	✓	✓	✓	✓	✓	✓
Wong 2008[322]	?	✓	✓	✓	✗	✓	✓
Wong 2008[323]	✓	✓	?	✓	?	✓	✓
Wong 2009[324]	✓	✓	✗	✓	?	✓	✗
Yoneda 2008[326]	?	✓	✗	✓	✓	✓	✗
Yoneda 2011[325]	?	✓	✗	✓	?	✓	?
Younossi 2011[327]	?	✓	✗	✓	✗	✓	✓

✗, high risk of bias; ✓, low risk of bias; ?, unclear risk of bias.

We explored potential sources of heterogeneity as outlined in the methods section. As all but one study were of low methodological quality, this potential source of heterogeneity could not be assessed. More specifically and according to source of heterogeneity, details of significant associations are:

- Histological score used (comparator is the Kleiner system): for NAFLD fibrosis score low cut-off in F2 and low and high cut-offs in F3, use of the Brunt scoring system resulted in significantly lower sensitivity and higher specificity.
- Full text versus abstract publication: for BARD score in F3, full-text publication was associated with significantly higher specificity.
- There was no heterogeneity identified in relation to transaminases levels.

Results: alcoholic liver disease

Data on patients with ALD were extracted from 12 studies.[114,273–283] Meta-analysis was performed separately for each non-invasive test assessed at each METAVIR stage (F1–F4). Summary sensitivity and specificity of non-invasive tests for each fibrosis stage are shown in *Table 14*. The median prevalence (minimum–maximum) of fibrosis stages F1–F4 in included studies was for F1 0.923 (single study), F2 0.633 (0.500–0.837), F3 0.509 (0.404–0.748) and F4 0.448 (0.145–0.971). Individual study characteristics, and forest plots and SROC plots of the different NILTs across fibrosis stages, are presented in *Appendices 4, 5* and *6*, respectively.

Overall, there were four different non-invasive tests reported for the diagnosis of $\geq$ F3, of which three (75%) were reported in single studies.[273,275,282] The most commonly evaluated non-invasive test was Fibroscan (four studies);[273,277,278,281] however, the results did not converge with the bivariate random-effects model. There were one, five and five NILTs evaluated for F1, F2 and F4 fibrosis stages, respectively, none of which converged with the bivariate random-effects model. APRI (high and low cut-offs) in F2 were evaluated in two studies[273,283] and Fibroscan in F4 was evaluated in six studies;[114,273,274,276,278,281] all other tests were evaluated in single studies.

There were four different non-invasive tests reported for the diagnosis of F4, of which three (75%) were reported in single studies.[273,279,280] The most commonly evaluated non-invasive test was Fibroscan (six studies),[114,273,274,276,278,281] however, with cut-offs that widely ranged from 11.4 to 25.8 kPa. The rest of the tests for F4 were PGAA [prothrombin time, gamma-glutamyl transpeptidase (GGT), apolipoprotein A1, α2-macroglobulin], Fibrotest (at a low and high cut-off) and APRI (results only reported at a high cut-off; no reason provided for not including a low cut-off).

Cut-offs for specific histological stages were predetermined in three studies (25%),[276,277,282] whereas liver biopsy was of acceptable quality in one study (8%).[273] There was no study with low risk of bias in all the domains of the QUADAS-2 tool; therefore, all of our estimates may be biased. Studies that were judged as low risk of bias or unknown in the three most important QUADAS domains, namely patient selection, index test and reference standard, were 25% of the total number of studies (3 out of 12).[276,279,282] Quality assessment of included studies based on QUADAS-2[62] is shown in *Table 15*.

Although investigation of potential sources of heterogeneity was planned, it could not be performed due to the small number of studies.

Results: liver cirrhosis

Diagnostic accuracy of NILTs for cirrhosis was also analysed irrespective of aetiology of liver disease and used for the cirrhosis economic model. Summary sensitivity and specificity for cirrhosis are shown in *Table 16*. Median prevalence of cirrhosis in the evaluated studies was 0.18 (range 0–0.97). Fibroscan was by far the most commonly evaluated non-invasive test (65 studies).[28,29,75,76,86–88,91,95,98–100,102,105,106,110,114,116,119,130,137,141,147,153,155,159,161,164,170,172,173,194,199–201,211,223,224,225,230,241,242,246,247,250,251,263,264,272–274,276,278,281,288,296,298,308,323,324,236]

We do not include a table on quality assessment of included studies to avoid repetition, as this was given separately according to disease aetiology.

TABLE 14 Diagnostic accuracy of non-invasive tests for detection of various stages of fibrosis in patients with ALD

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Non-invasive tests for diagnosis of F ≥ 1					
Fibroscan	1	5.9	0.83 (0.74 to 0.89)	0.88 (0.53 to 0.98)	Single study
Non-invasive tests for diagnosis of F ≥ 2					
Fibroscan	1	7.8	0.81 (0.7 to 0.88)	0.92 (0.76 to 0.98)	Single study
Fibrotest (high cut-off)	1	0.7	0.55 (0.47 to 0.63)	0.93 (0.85 to 0.97)	Single study
Fibrotest (low cut-off)	1	0.3	0.84 (0.77 to 0.89)	0.65 (0.55 to 0.75)	Single study
APRI (high cut-off)	2	1.5	0.54 (0.42 to 0.66)	0.78 (0.64 to 0.88)	Fixed-effects model for sensitivity and specificity without correlation
APRI (low cut-off)	2	0.5	0.72 (0.6 to 0.82)	0.46 (0.33 to 0.6)	Fixed-effects model for sensitivity and specificity without correlation
Non-invasive tests for diagnosis of F ≥ 3					
CK18	1		0.84 (0.73 to 0.92)	0.71 (0.6 to 0.79)	Single study
Forns index (high cut-off)	1	6.9	0.41 (0.23 to 0.61)	0.88 (0.66 to 0.97)	Single study
YKL-40	1	330	0.51 (0.38 to 0.63)	0.89 (0.8 to 0.94)	Single study
Fibroscan	4	11.0–12.5	0.87 (0.64 to 0.96)	0.82 (0.67 to 0.91)	Random-effects model for sensitivity and specificity without correlation
Non-invasive tests for diagnosis of F ≥ 4					
APRI (high cut-off)	1	2	0.40 (0.22 to 0.61)	0.62 (0.41 to 0.79)	Single study
Fibrotest (high cut-off)	1	0.7	0.91 (0.82 to 0.96)	0.87 (0.81 to 0.91)	Single study
Fibrotest (low cut-off)	1	0.3	1.00 (0.95 to 1.00)	0.50 (0.42 to 0.58)	Single study
PGAA	1	7	0.78 (0.64 to 0.88)	0.89 (0.85 to 0.92)	Single study
Fibroscan	6	11.4–25.8	0.86 (0.76 to 0.92)	0.83 (0.74 to 0.89)	Random-effects model for sensitivity and specificity without correlation

CK18, cytokeratin-18; PGAA, prothrombin time, gamma-glutamyl transpeptidase (GGT), apolipoprotein A1, α2-macroglobulin; YKL-40, a direct marker of liver fibrosis.

TABLE 15 Quality assessment of studies included in the meta-analysis for ALD. Quality assessment was done using the QUADAS-2 tool

	Domain 1: patient sampling		Domain 2: index test		Domain 3: reference standard		Domain 4: flow and timing
Study ID	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias	Concerns about applicability	Risk of bias
Forestier 2010[114]	✗	✗	✗	✗	?	✓	✗
Janssens 2010[273]	✗	✗	✗	✗	✓	✓	✗
Kim 2009[274]	?	?	✗	✓	✗	✓	?
Lavallard 2011[275]	✓	✓	✗	✓	✗	✓	?
Melin 2005[276]	✓	✓	✓	?	?	✓	✗
Mueller 2010[277]	✗	✗	✓	✓	✗	✓	?
Nahon 2008[278]	✓	✓	✗	✗	✗	✓	✗
Naveau 1994[279]	?	✓	?	✓	?	✓	✓
Naveau 2005[280]	✓	✓	✗	✓	?	✓	✓
Nguyen-Khac 2008[281]	✗	✓	✗	✓	✗	✓	✗
Tran 2000[282]	✓	✓	✓	✓	?	✓	?
Vanbiervliet 2005[283]	✗	✗	✗	✗	?	✓	✗

✗, high risk of bias; ✓, low risk of bias; ?, unclear risk of bias.

TABLE 16 Diagnostic accuracy of non-invasive tests for detection of cirrhosis irrespective of aetiology of liver disease

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Indirect non-invasive serum tests					
APRI (low cut-off)	27	0.75–1	0.75 (0.71 to 0.8)	0.78 (0.75 to 0.81)	Bivariate random-effects model with correlation between sensitivity and specificity
APRI (high cut-off)	23	2	0.45 (0.37 to 0.52)	0.93 (0.9 to 0.95)	Bivariate random-effects model with correlation between sensitivity and specificity
Age–Platelet Index	3		0.88 (0.08 to 1.00)	0.73 (0.43 to 0.91)	Random-effects model for sensitivity and specificity without correlation
AST–ALT ratio	13	1	0.49 (0.39 to 0.59)	0.87 (0.75 to 0.94)	Bivariate random-effects model with correlation between sensitivity and specificity
CDS	1		0.88 (0.66 to 0.97)	0.67 (0.57 to 0.77)	Single study
FIB-4 (low cut-off)	5	1.45–1.92	0.84 (0.76 to 0.89)	0.71 (0.62 to 0.79)	Bivariate random-effects model with correlation between sensitivity and specificity

continued

TABLE 16 Diagnostic accuracy of non-invasive tests for detection of cirrhosis irrespective of aetiology of liver disease (*continued*)

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
FIB-4 (high cut-off)	4	3.25–4.44	0.42 (0.2 to 0.69)	0.92 (0.58 to 0.99)	Random-effects model for sensitivity and specificity without correlation
Forns index (low cut-off)	2	4.2	0.88 (0.73 to 0.96)	0.37 (0.26 to 0.49)	Fixed-effects model for sensitivity and specificity without correlation
Forns index (high cut-off)	1	6.9	0.67 (0.53 to 0.78)	0.91 (0.84 to 0.95)	Single study
GUCI	4	0.33–1.11	0.64 (0.11 to 0.96)	0.86 (0.81 to 0.9)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
King's	1	24.3	0.74 (0.59 to 0.85)	0.90 (0.84 to 0.94)	Single study
Lok's index (low cut-off)	2	0.2–0.26	0.84 (0.09 to 1)	0.66 (0.00 to 1.00)	Fixed-effect model for sensitivity and random-effects model for specificity without correlation
Lok's index (high cut-off)	1	0.5	0.40 (0.29 to 0.52)	0.95 (0.91 to 0.97)	Single study
Platelets	12	134–196	0.72 (0.62 to 0.81)	0.88 (0.77 to 0.94)	Bivariate random-effects model with correlation between sensitivity and specificity
Direct serum non-invasive serum tests					
^{13}C-caffeine breath test	1	1.27	0.93 (0.77 to 0.98)	0.84 (0.74 to 0.91)	Single study
Fontana	1	0.2	0.79 (0.72 to 0.84)	0.66 (0.61 to 0.71)	Single study
Hyaluronic acid	8	78–237	0.81 (0.65 to 0.9)	0.88 (0.8 to 0.94)	Bivariate random-effects model with correlation between sensitivity and specificity
Hepascore	9	0.7–0.87	0.82 (0.72 to 0.88)	0.84 (0.79 to 0.88)	Bivariate random-effects model with correlation between sensitivity and specificity
Hepascore (low cut-off)	1	0.58	0.80 (0.72 to 0.86)	0.83 (0.8 to 0.85)	Single study
Hepascore (high cut-off)	1	1.16	0.39 (0.31 to 0.48)	0.99 (0.98 to 0.99)	Single study
PIIINP	3	0.8–1	0.70 (0.48 to 0.86)	0.79 (0.34 to 0.96)	Fixed-effect model for sensitivity and random-effects model for specificity without correlation
Type IV collagen	3	65–190	0.71 (0.57 to 0.82)	0.76 (0.59 to 0.87)	Bivariate random-effects model with correlation between sensitivity and specificity

TABLE 16 Diagnostic accuracy of non-invasive tests for detection of cirrhosis irrespective of aetiology of liver disease (*continued*)

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Commercial non-invasive serum tests					
ELF	1	9.4	0.93 (0.69 to 0.99)	0.79 (0.67 to 0.88)	Single study
ELF (low cut-off)	1		0.90 (0.84 to 0.94)	0.53 (0.46 to 0.59)	Single study
ELF (high cut-off)	1		0.52 (0.43 to 0.6)	0.90 (0.85 to 0.93)	Single study
Fibroindex	1	1.82	0.70 (0.52 to 0.84)	0.91 (0.82 to 0.96)	Single study
Fibrometer	2	0.88	0.72 (0.36 to 0.92)	0.88 (0.6 to 0.97)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
Fibrometer (low cut-off)	1	0.63	0.96 (0.9 to 0.98)	0.71 (0.68 to 0.74)	Single study
Fibrometer (high cut-off)	1	0.98	0.36 (0.28 to 0.45)	0.98 (0.97 to 0.99)	Single study
Fibrotest	13	0.75	0.61 (0.47 to 0.74)	0.87 (0.83 to 0.9)	Bivariate random-effects model with correlation between sensitivity and specificity
Fibrotest (low cut-off)	2	0.3	0.89 (0.29 to 0.99)	0.65 (0.01 to 1.00)	Fixed-effect model for sensitivity and random-effects model for specificity without correlation
Fibrotest (high cut-off)	3	0.86	0.73 (0.14 to 0.98)	0.94 (0.91 to 0.96)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
Imaging modalities					
ARFI	4	1.59–2	0.84 (0.72 to 0.91)	0.77 (0.5 to 0.92)	Random-effects model for sensitivity and specificity without correlation
Platelet–Spleen Index	2		0.83 (0.28 to 0.98)	0.85 (0.00 to 1.00)	Fixed-effect model for sensitivity and random-effects model for specificity without correlation
Real-time elastography	1	3.93	0.71 (0.5 to 0.86)	0.80 (0.67 to 0.88)	Single study
Fibroscan	65	9.2–26.5	0.89 (0.86 to 0.91)	0.89 (0.87 to 0.91)	Bivariate random-effects model with correlation between sensitivity and specificity
Combination of fibrosis non-invasive tests algorithms					
Bordeaux	1		0.87 (0.8 to 0.92)	0.95 (0.93 to 0.96)	Single study
Fibropaca	1		0.79 (0.72 to 0.84)	0.66 (0.61 to 0.71)	Single study
SAFE	4		0.74 (0.42 to 0.92)	0.93 (0.91 to 0.94)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation

CDS, Cirrhosis Discriminant Score; PIIINP, amino-terminal propeptide of type III procollagen; SAFE, sequential algoritm for fibrosis evaluation.

Results: imaging modalities results that were used irrespective of liver disease aetiology

Data on diagnostic accuracy of imaging modalities for the non-invasive assessment of liver fibrosis were analysed irrespective of the aetiology of liver disease, with the exception of Fibroscan and ARFI. These two tests measure liver stiffness that is associated with the amount of fibrosis and there is evidence of different cut-offs according to disease aetiology. The summary sensitivity and specificity of these imaging modalities for METAVIR stages F1–F4 are shown in *Table 17*.[30,93,98,100,101,110,118,132,145,149,152,166,167,176,184,185,227,270,305,328–363]

Certain diagnostic modalities that are not routinely used, such as MR spectroscopy,[367] double-contrast material-enhanced MRI,[368] maximal accumulative respiratory strain using ultrasound,[234] liver enhancement ratio of gadoxetic acid enhanced MRI[369] and single-photon emission CT parameters[370] and real-time elastography,[142] are not presented and were not used in the economic analysis, as they are not widely available and require further validation.

Data on ultrasound techniques using injection of contrast material (contrast-enhanced ultrasound)[344] or measuring the splenic artery pulsatility index (ultrasound SAPI)[152] were included in the analysis and in the economic modelling for HBV and HCV. These are based on ultrasound; however, they require well-trained and experienced operators and use signs that are not well validated. Moreover, all data come from specialised centres. Therefore, reported diagnostic accuracies are most probably overestimated and not reproducible in most centres. In most centres, ultrasound, CT and MRI can only diagnose cirrhosis with acceptable specificity but lower sensitivity, as early cirrhosis and lesser fibrosis stages are often missed. MR elastography is a promising tool, but is of limited use at the moment due to cost and availability.

Discussion of overall results

In total, 302 studies were selected for the meta-analysis.[23–29,31,71–327] For the fibrosis stages of interest in our models (F2 for HBV and HCV, F3 for NAFLD and F4 for ALD), there were 33 NILTs assessed in HCV,[23–29,31,71–224] 18 in HBV,[116,119,120,200,225–272] 24 in NAFLD[117,119,165,229,284–327] and four in ALD.[114,273–283] However, 19 out of 33 NILTs in HCV,[31,83,131,142,146,147,183,189,204,205,210] 13 out of 18 in HBV,[243,244,256–258,263,265] 11 out of 24 in NAFLD[117,229,284,286,287,289,290,299,304,306,327] and three out of four in ALD[273,275,282] were assessed in single studies. HCV was the disease aetiology with most studies identified, while ALD had the fewest studies assessed.

There were no data available for tests that converged using the bivariate model for ALD; therefore, the use of NILTs in such patients for treatment decisions is uncertain and requires further study. Very few tests converged using the bivariate model for HBV: only APRI (high and low cut-offs), FIB-4 (low cut-off), Fibrotest and Fibroscan converged for F2 and Fibroscan alone for F3 and F4. In patients with NAFLD there is a wider choice, as NAFLD fibrosis score, FIB-4, BARD, AST–ALT ratio, Fibrotest and Fibroscan all converged. HCV is the liver disease with the highest number of available data and, subsequently, choices on NILTs. For the diagnosis of F2, APRI low and high cut-off, AST–ALT ratio, FIB-4 low and high cut-offs, Forns index low and high cut-offs, GUCI, Lok's index, platelet count, hyaluronic acid, Hepascore, Fibrometer, Fibrotest standard, low and high cut-offs, platelet-to-spleen-diameter ratio and Fibroscan all converged.

TABLE 17 Diagnostic accuracy of imaging modalities for detection of various stages of fibrosis in patients with chronic liver disease

Test	Number of studies	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Imaging modalities for diagnosis of F ≥ 1				
MR elastography	5	0.83 (0.72 to 0.9)	0.83 (0.67 to 0.92)	Random-effects model for sensitivity and specificity without correlation
CEMRE	1	0.95 (0.75 to 0.99)	1.00 (0.57 to 1.00)	Single study
CEMRI	1	0.87 (0.78 to 0.93)	0.75 (0.51 to 0.9)	Single study
CEUS	1	0.80 (0.7 to 0.88)	0.90 (0.6 to 0.98)	Single study
CT	1	0.70 (0.52 to 0.83)	0.64 (0.43 to 0.8)	Single study
DW-MRI	1	0.79 (0.65 to 0.89)	0.83 (0.55 to 0.95)	Single study
US	1	0.77 (0.6 to 0.89)	0.89 (0.69 to 0.97)	Single study
US MARS	1	0.82 (0.64 to 0.92)	0.75 (0.41 to 0.93)	Single study
Imaging modalities for diagnosis of F ≥ 2				
CEMRI	2	0.80 (0.15 to 0.99)	0.60 (0.03 to 0.99)	Fixed-effect model for sensitivity and random-effects model for specificity without correlation
CEUS	3	0.88 (0.07 to 1.00)	0.73 (0.11 to 0.98)	Random-effects model for sensitivity and specificity without correlation
DW-MRI	5	0.78 (0.63 to 0.88)	0.78 (0.51 to 0.93)	Random-effects model for sensitivity and specificity without correlation
MR elastography	3	0.94 (0.13 to 1.00)	0.92 (0.72 to 0.98)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
US	3	0.35 (0.14 to 0.63)	0.86 (0.59 to 0.96)	Bivariate random-effects model with correlation between sensitivity and specificity
US SAPI	3	0.74 (0.69 to 0.79)	0.79 (0.72 to 0.85)	Fixed-effects model for sensitivity and specificity without correlation
US SAPI (high cut-off)	2	0.61 (0.54 to 0.68)	0.96 (0.9 to 0.98)	Fixed-effects model for sensitivity and specificity without correlation
US SAPI F2 (low cut-off)	2	0.94 (0.9 to 0.97)	0.39 (0.31 to 0.49)	Fixed-effects model for sensitivity and specificity without correlation
nDW-MRI	1	0.90 (0.74 to 0.97)	0.75 (0.3 to 0.95)	Single study
SPECT	1	0.86 (0.67 to 0.95)	0.83 (0.64 to 0.93)	Single study
US MARS	1	0.85 (0.64 to 0.95)	0.56 (0.33 to 0.77)	Single study

continued

TABLE 17 Diagnostic accuracy of imaging modalities for detection of various stages of fibrosis in patients with chronic liver disease (*continued*)

Test	Number of studies	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Imaging modalities for diagnosis of F ≥ 3				
CEUS	3	0.78 (0.14 to 0.99)	0.87 (0.24 to 0.99)	Random-effects model for sensitivity and specificity without correlation
DEMRI	1	0.93 (0.85 to 0.97)	0.86 (0.69 to 0.95)	Single study
DW-MRI	3	0.88 (0.00 to 1.00)	0.73 (0.07 to 0.99)	Random-effects model for sensitivity and specificity without correlation
MR elastography	6	0.91 (0.85 to 0.95)	0.88 (0.80 to 0.93)	Bivariate random-effects model with correlation between sensitivity and specificity
US	4	0.57 (0.31 to 0.79)	0.80 (0.67 to 0.88)	Bivariate random-effects model with correlation between sensitivity and specificity
CEMRI	1	0.75 (0.63 to 0.85)	0.50 (0.36 to 0.64)	Single study
MRI	1	0.78 (0.61 to 0.89)	0.75 (0.60 to 0.86)	Single study
nDW-MRI	1	0.96 (0.80 to 0.99)	0.67 (0.35 to 0.88)	Single study
Imaging modalities for diagnosis of F ≥ 4				
CEUS	3	0.84 (0.01 to 1.00)	0.88 (0.27 to 0.99)	Random-effects model for sensitivity and specificity without correlation
DW-MRI	2	0.88 (0.01 to 1.00)	0.73 (0.09 to 0.99)	Fixed-effect model for sensitivity and random-effects model for specificity without correlation
LSPI	4	0.91 (0.41 to 0.99)	0.88 (0.3 to 0.99)	Random-effects model for sensitivity and specificity without correlation
MR elastography	3	1.00 (0.03 to 1.00)	0.93 (0.20 to 1.00)	Random-effects model for sensitivity and specificity without correlation
MRI	2	0.75 (0.64 to 0.83)	0.80 (0.69 to 0.88)	Fixed-effects model for sensitivity and specificity without correlation
US	25	0.73 (0.66 to 0.79)	0.88 (0.83 to 0.92)	Bivariate random-effects model with correlation between sensitivity and specificity
US SAPI	2	0.73 (0.25 to 0.95)	0.67 (0.43 to 0.84)	Fixed-effects model for sensitivity and specificity without correlation
CEMRI	1	0.81 (0.65 to 0.90)	0.48 (0.37 to 0.60)	Single study
HVRI US	1	0.90 (0.74 to 0.97)	0.86 (0.78 to 0.92)	Single study
nDW-MRI	1	0.95 (0.77 to 0.99)	0.69 (0.42 to 0.87)	Single study
US MARS	1	1.00 (0.68 to 1.00)	0.50 (0.33 to 0.67)	Single study

CEMRE, contrast-enhanced magnetic resonance elastography; CEMRI, contrast-enhanced magnetic resonance imaging; CEUS, contrast-enhanced ultrasound; DEMRI, double-enhanced magnetic resonance imaging; DW-MRI, diffusion-weighted magnetic resonance imaging; HVRI, hepatic vein resistance index; LSPI, laser speckle perfusion imaging; MARS, metal artefact reduction sequence; nDW, diffusion weighted; SAPI, splenic artery pulsatility index; SPECT, single-photon emission computed tomography; US, ultrasound.

Summary sensitivity and specificity of tests that converged were as follows in HCV and HBV (≥ F2 and F4) and NAFLD (≥ F3):

- For HCV in ≥ F2: APRI low cut-off 82% and 57%, high cut-off 39% and 92%; AST–ALT ratio 44% and 71%; FIB-4 high cut-off 59% and 74%; Forns index low cut-off 88% and 40%, high cut-off 35% and 96%; GUCI 65% and 79%; Lok's index 67% and 55%; platelet count 50% and 89%; hyaluronic acid 75% and 75%; Hepascore 73% and 73%; Fibrometer 79% and 73%; Fibrotest 68% and 72%; Fibroscan 79% and 83%.
- For HCV in F4: APRI low cut-off 77% and 78%, high cut-off 48% and 94%; AST–ALT ratio 49% and 87%; platelet count 68% and 86%; hyaluronic acid 80% and 88%; Hepascore 80% and 83%; Fibrotest 60% and 86%; Fibroscan 89% and 91%.
- For HBV in ≥ F2: APRI low cut-off 80% and 65%, high cut-off 37% and 93%; FIB-4 low cut-off 68% and 73%; Fibrotest 66% and 80%; Fibroscan 71% and 84%.
- For HBV in F4: Fibroscan 86% and 85%.
- For NAFLD in ≥ F3: AST–ALT ratio 79% and 70%, BARD 84% and 61%, FIB-4 low cut-off 84% and 74%, high cut-off 38% and 97%, NAFLD fibrosis score low cut-off 80% and 66%, high cut-off 40% and 97%, hyaluronic acid 88% and 82%, Fibrotest high cut-off 40% and 96%, Fibroscan 82% and 84%.

Regarding the choice of tests, these should be based on disease aetiology, local resources and availability, cost-effectiveness and disease prevalence or pretest probability. Indirect NILTs can be applied at the point of care and can differentiate patients in low risk, high risk and indeterminate for significant fibrosis or cirrhosis. Particularly in NAFLD, they could be used to rule out patients with low risk of fibrosis, thanks to their high negative predictive value given the low prevalence of advance fibrosis in the general population of patients with steatosis/NAFLD. Direct serum tests or Fibroscan can be used either as second tier test following an indeterminate result with an indirect marker, or as a one-off test to inform further decisions on treatment or as a rule-in/rule-out test for liver biopsy. Patients who test 'negative' (true negative or false negative) with a non-invasive test should be subsequently retested in order to capture disease progression. The optimal time interval for such retesting is unknown; however, a period of 1–2 years would be safe and reasonable. There are no data on the monitoring of liver fibrosis using sequential testing with NILTs.

Cut-offs of NILTs for specific fibrosis stages were not always predetermined or sufficiently validated and this happened more often in direct serum biomarkers and Fibroscan; this represents a significant limitation in the interpretation of their results. Among all NILTs, APRI (low and high cut-offs) was the one where the established cut-offs were almost universally used in published studies. Although there are established cut-offs for Forns index, FIB-4, AST–ALT ratio and Fibrotest, these were not consistently used in all studies. NAFLD fibrosis score and BARD in NAFLD (10[285,290,300,301,309,311,315,320,322] and seven studies,[284,291,300,301,312,315,319] respectively) had consistent cut-offs used across all studies.

Fibroscan was the NILT assessed in most studies across diseases aetiologies (37 studies in HCV, 13 in HBV, eight in NAFLD and six in ALD).[28,29,75,76,86–88,91,95,98–100,102,105,106,110,116,119,130,131,141,147,153,155,159,161,164,170,172,173,194,199–201,211,223–225,230,236,241,242,246,247,250,251,263,264,272,288,296,298,308,323,324,273,277,278,281] However, there are no established and validated cut-offs for specific fibrosis stages across disease aetiologies. This represents a limitation in the use of Fibroscan; therefore, all reported sensitivities and specificities are probably overestimated. APRI was also widely assessed in HCV and HBV (47 and eight studies, respectively) but not in NAFLD or ALD.[24,31,72–74,79,81,84,85,89–91,94,97,98,100,103,107,109,121,123,126,127,130,131,134,137,140,143,144,146,150,152–154,156–158,163,164,168,182,185,187,189,194,195,209,210,218,220,223]

All non-commercial direct serum non-invasive tests assessed [hyaluronic acid, YKL-40, PIIINP (amino-terminal propeptide of type III procollagen), type IV collagen] did not have predetermined cut-offs; moreover, different enzyme-linked immunosorbent assay (ELISA) kits were used across studies. Hepascore, which is a non-commercial test that consists, among other indices, of hyaluronic acid, has predetermined cut-offs, which were not consistently used in all included studies.

Of the commercial NILTs, Fibrotest was the most widely assessed (23 studies in HCV,[23,25,31,81,85,99,111,113,127,129,146,172,177,179,186,189,206,208,211,216,220,223,366] six in HBV,[231,247,249,255–257] four in NAFLD[284,303,313,314] and one in ALD[280]); however, the predetermined cut-offs used were not always used in included studies. Fibroindex,[24,107,186] Fibrometer[82,127,223] and FibrospectII[25,96,145,170,190,196] were only evaluated in HCV for the stages of interest in our models. ELF was evaluated in three studies in HCV[27,98,158] and in a single study in NAFLD.[290]

Failures of the index test (e.g. due to high BMI for Fibroscan or haemolysis for serum tests) are not incorporated in the reported sensitivities and specificities of the NILTs. Moreover, instances where the reference standard was not adequate for analysis (insufficient sampling) is also not captured in the analysis (applicability of index test and reference standard).

Investigations of heterogeneity revealed an influence of the level of transaminases on diagnostic accuracy in some tests; however, the direction of this effect was not consistent. There was no significant heterogeneity with regard to type of publication (abstract or full text) and histological scoring systems in most tests and diseases.

Of the imaging modalities, MR elastography was assessed in three studies, with summary sensitivity and specificity of 0.94 and 0.92, respectively. Although these are very promising results, further validation of the technique and determination of disease and stage specific cut-offs is needed. Moreover, this technique is not yet widely available.

The methodological quality of included studies as assessed by the QUADAS-2 tool was poor; only 6 of the 302 studies (1.6%) were of high methodological quality.[86,127,143,186,200,222] Most common areas of high risk of bias was the conduct of the index test (cut-offs were not predetermined) and of the reference standard (liver biopsy samples were not of adequate length or did not have sufficient number of portal tracts for reliable staging). Therefore, all reported results are likely biased.

As will become apparent in the cost-effectiveness analysis, NILTs with the most robust data were not the most cost-effective. As mentioned, there is the risk of overestimating sensitivity and specificity in tests with few available data.

Chapter 5 Cost-effectiveness analysis: hepatitis B

This chapter describes the assessment of cost-effectiveness of non-invasive tests of fibrosis and cirrhosis in patients with HBV. The population of interest were HBeAg-positive and HBeAg-negative patients with suspected fibrosis or cirrhosis who would normally have a liver biopsy in order to assess eligibility for antiviral treatment, i.e. patients with increased viral load and/or elevated transaminases.

Evaluation approach for hepatitis B

Twenty-five relevant NILTs were evaluated in the first stage of the analysis, which compared the NILTs with liver biopsy alone and a 'treat all' and 'no treatment' approach. The NILTs evaluated are listed in *Table 18* and are grouped according to test categories: indirect serum makers, direct and patented serum markers and imaging modalities.

The second stage of the analysis compared a selection of tests using alternative sequential testing strategies. The criteria for selecting these tests and the assumptions used regarding combinations of tests and sequential testing strategies are detailed in *Chapter 3*.

Three of the tests evaluated in the second stage of the analysis used a combined diagnostic cut-off threshold for staging fibrosis; the test outcomes reported a number of indeterminate responses which are listed in *Table 19*. The percentage of indeterminate results was estimated using the meta-analysis data and is an aggregated value estimated from the studies for each combined test. We allowed for patients who had an indeterminate response to receive a retest with a commonly used imaging modality Fibroscan (TE). We did not choose an indirect test as the combined tests were from the indirect test category and a subsequent indirect test would not enhance the diagnostic accuracy. Of the direct tests and imaging modalities, we chose Fibroscan based on availability and current clinical practice. Overall, 56 testing strategies were compared in the second stage of the analysis.

TABLE 18 List of NILTs evaluated (HBV)

Indirect	Direct and patented	Imaging
AAR	Fibrotest	ARFI
APGA	Hyaluronic acid	CEUS
Age–Platelet Index	Hepascore	DW-MRI
APRI (high cut-off)		MR elastography
APRI (low cut-off)		CT
FIB-4 (high cut-off)		Fibroscan
FIB-4 (low cut-off)		US
Forns index (high cut-off)		US SAPI
Forns index (low cut-off)		US SAPI (high cut-off)
GUCI		US SAPI (low cut-off)
Hui index		
PAPAS		

AAR, AST–ALT ratio; APGA, AST, platelet count, GGT, α-fetoprotein; CEUS, contrast-enhanced ultrasound; DW-MRI, diffusion-weighted magnetic resonance imaging; PAPAS, age, ALP, α-fetoprotein, AST; US, ultrasound.

TABLE 19 Percentage of indeterminate results of tests with a combined cut-off applicable for HBV

Tests with a combined cut-off	% of persons with inconclusive result
APRI	41
FIB-4	31
Forns index	36

Model structure and parameters

Decision tree structure

A decision-tree model was constructed to evaluate the cost-effectiveness of the NILTs, liver biopsy, and the 'treat all' and 'no treatment' strategies. As per the schematic diagram depicting the flow of data in *Chapter 3* (see *Figure 1*), the decision tree was populated with test sensitivity, specificity and average disease prevalence from the meta-analysis (see *Chapter 4* for details), long-term costs and health outcomes from a series of Markov models, individual test costs sourced from published literature and hospital finance departments, and a measure of adverse effects associated with liver biopsy.

As discussed in *Chapter 3*, there are two stages to the decision tree analysis: the first stage where all tests are compared singly and the second stage where combinations of tests are compared using four different testing strategies. Schematic illustrations and descriptions of the sequential testing pathways are provided in *Chapter 3*. To estimate a cost and QALY for each testing strategy the long-term costs (and test costs) and QALY estimates (including disutility from liver biopsy if applicable) associated with each potential test outcome (true positive, false positive, true negative, false negative) were multiplied by the probability of each NILT returning a true positive, false positive, true negative or false negative result to give a cost and QALY estimate for each testing option.

Markov model structure

A series of Markov models were constructed to estimate the long-term costs and health outcomes associated with a correct (true positive and true negative) and incorrect (false positive and false negative) diagnosis for a hypothetical cohort of 1000 patients with HBV and suspected liver fibrosis or cirrhosis. An additional two Markov models were constructed to estimate the costs and outcomes associated with the 'treat all' and 'no treatment' approaches. Separate models were constructed for the HBeAg-positive and HBeAg-negative patient cohorts, as their natural history differs and, therefore, starting age, transition probabilities and relative risks (RRs) from treatment differed for both groups. The structural assumptions underlying the state transition models applied to both groups of patients. The models were evaluated over a lifetime period with a cycle length of 1 year. All costs were considered from the perspective of the NHS and health outcomes were measured in terms of QALYs. Costs and utilities were discounted using a rate of 3.5%. A threshold value for incremental cost-effectiveness was assumed to be £20,000–30,000 per additional QALY gained, based on UK guidelines.[66]

Figure 3 shows a schematic representation of the patient pathway, which is a modified version of previously published models of liver fibrosis and cirrhosis in patients with HBV.[371,372] Fibrosis and cirrhosis health states were defined in the Markov model according to METAVIR score: mild fibrosis (F0–1), moderate fibrosis (F2–3) and compensated cirrhosis (F4). There were five other potential health states in the Markov model: decompensated cirrhosis, hepatocellular cancer (HCC), liver transplant, post liver transplant and a health state representing 'death', which the cohort could enter from all other health states at any time. The cohort progressed through the model as per the arrows in the illustration; the circular arrows leading back into the same health state indicate that patients could remain within that health state for longer than one cycle (cycle length set at 1 year) except for the liver transplant state, where patients could only progress to a post liver transplant health state or to death. The liver transplant state comprised two events (1 month's duration for liver transplant and 11 months' duration for post-liver

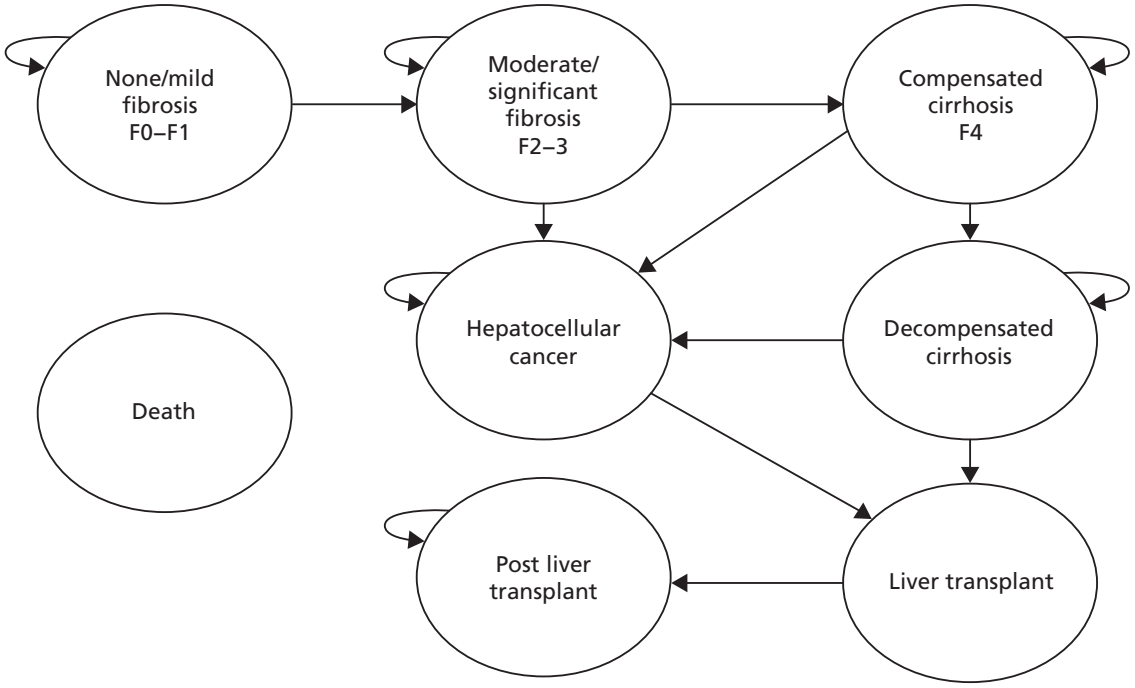

FIGURE 3 Illustration of Markov model.

transplantation care). Patients could progress to the HCC health state from the moderate, compensated cirrhosis and decompensated cirrhosis health states. Patients could not regress to an earlier health state. We note that although some recent studies show that fibrosis and even cirrhosis can regress with antiviral therapy,[373] we did not allow for this in our model.

We assumed that a METAVIR test score of $\geq$ F2 equated to a positive test outcome (true positive and false positive) and treatment with antiviral agents would commence at this stage. Conversely, a METAVIR test score of $<$ F2 indicated a negative test outcome. We incorporated treatment with peginterferon alfa-2a for 1 year for 10% of those who tested negative to reflect that a proportion of patients would receive treatment for necroinflammation.[45] The remaining 90% would undergo a policy of watchful waiting without immediate treatment.

The watchful waiting policy in the model incorporated a retest with a NILT every 2 years; the retest NILT was the same as the previous test used for the initial assessment. To allow for modelling, we assumed that the retest had perfect sensitivity and specificity and correctly diagnosed all patients. If a patient was diagnosed as positive, immediate treatment with antiviral agents would commence. We tested the robustness of this assumption in a sensitivity analysis by varying the retest sensitivity and specificity to mirror that of three commonly used NILTs: an indirect serum marker, APRI; a patented serum marker, Fibrotest; and an imaging modality, Fibroscan.

The initial starting health state for the population cohort in the models depended on the test outcome being modelled (true positive, false positive, true negative and false negative). For example, given a false positive or true negative test result, the population cohort entered the model in a F0–1 health state (no fibrosis), whereas for a true positive or false negative test result, the population cohort had an initial starting health state of F2–4, where the cohort (with fibrosis or cirrhosis) was then distributed among the health states F2–3 and F4 based on the prevalence data from the systematic review (56% and 44%, respectively).

Input parameters

The average disease prevalence used in the model (54%) was estimated as the proportion of patients with a METAVIR score ≥ F2, calculated from the meta-analysis results reported in *Chapter 4*. We used the sensitivity and specificity estimates for each NILT and the average prevalence estimate to calculate the probability of each test returning a true positive, false positive, true negative and false negative result. The probability of each NILT reporting a particular diagnostic test outcome is reported in *Appendix 7*.

Cohort data

We identified cohort characteristics (age and sex ratio) for use in the model from published NICE guidance on treatments for HBV: adefovir dipivoxil and pegylated interferon. After reviewing the sources of evidence for this guidance,[374,375] we identified a *Health Technology Assessment* report for both treatments published in 2006 by Shepherd *et al.*[371] This report included an economic analysis. Shepherd *et al.*[371] sourced cohort data from a published study undertaken by Fattovich *et al.*,[376] which reported a median age and a male-to-female ratio based on 10 published studies for HBeAg-positive chronic HBV[377–387] and four published studies for HBeAg-negative chronic HBV.[388–391] We employed the same assumptions as those used by Shepherd *et al.*:

- HBeAg-positive: starting age of 31 years and percentage of males set at 70%.
- HBeAg-negative: starting age of 40 years and percentage of males set at 90%.

Natural history: baseline transition probabilities used in model

The rate of disease progression in the models is regulated by transition probabilities. The systematic review and economic evaluation by Shepherd *et al.*[392] included a systematic review of epidemiological data for HBV.[371] We also reviewed a submission to NICE by Bristol Myers Squibb (BMS) for a single technology appraisal of entecavir treatment for HBV. Upon review, the Southampton Health Technology Assessment Centre (SHTAC) epidemiological search strategy was considered to be appropriate for our analysis. The papers used by the authors were reviewed and the search was rerun to carry out an updated search for natural history data. The submission report by BMS sourced data from the same literature sources as the SHTAC study.

Our updated epidemiological search strategy was undertaken to search from the year 2004 to 2012 (as SHTAC search was undertaken to 2004). The search was carried out in the MEDLINE database using the Ovid platform, and the search strategy is outlined in *Appendix 2* (searched on 11 May 2012). The updated search located 597 papers, whose titles were reviewed to determine if they were relevant; 24 were retrieved for abstract review. None was relevant for our study.

As none of the papers located from the updated literature search was relevant, we reviewed the studies which informed the transition probabilities in the model constructed by Shepherd *et al.*[372,376,392–401]

Some of the papers referenced by Shepherd *et al.*[371] were older published papers.[393,394] Others were published for other aetiologies of liver disease[395] and other models did not provide separate transition probabilities for the mild and moderate health states.[399] The study by Wong *et al.*[393] was identified as relevant as the authors had reported separate transition probabilities for HBeAg-positive and HBeAg-negative. As the paper by Wong *et al.*[393] was published in 1995, we conducted a search for recently published studies which had cited this paper. This search located a 2010 paper by Dakin *et al.*,[372] which assessed the cost-effectiveness of various drug treatments for HBV.

The 2010 paper by Dakin *et al.*[372] sourced transition probability data from natural history studies, economic evaluations or the placebo arms of meta-analyses or RCTs.[393,395,396,398–400,402–410] The authors used a similar set of studies to those identified by Shepherd *et al.*[371] and many of the transition probabilities elicited were either identical or similar. The progression from compensated cirrhosis to decompensated cirrhosis was the same for both studies, as Dakin *et al.*[372] used a synthesis of the same papers[396,398,399] used by Shepherd *et al.*[371]

Dakin *et al.*[372] sourced aetiology-specific data for the probability of undergoing a liver transplant while in a decompensated cirrhosis or HCC health state, using data from the UK transplant registry, specifically for chronic HBV. This was felt to be a more relevant estimate to use in our model than the value used by Shepherd *et al.*,[371] who had used a value based on data from a 1997 study by Bennett *et al.*[395] This study had sourced a value from a paper on HCV[411] and papers published in 1993[412] and 1996.[413]

We used the transition probabilities estimated by Dakin *et al.*[372] as the main source of transition probability data for the cirrhotic and post-cirrhotic health states (HCC, liver transplant and post liver transplant).

However, none of the studies reviewed elicited separate transition probability data for the precirrhotic health states (mild and moderate fibrosis). In the absence of transition probability data for these health states, we used data from a study in patients with mild chronic HCV.[414] This paper was identified in our review of data for our analysis of HCV but was also referenced as a source of cost data in the paper by Dakin *et al.*[372] We adopted this approach with regard to early transition probabilities and this was confirmed as appropriate with clinical colleagues.

Wright *et al.*[414] were unable to locate separate estimates of probabilities of progression from mild to moderate fibrosis disease stages health state from existing studies as they found that most data were based on retrospective natural history studies which do not use liver biopsy to stage fibrosis and progression. As the disease progression may not be linear, it may not be realistic to assume a constant rate of progression from mild fibrosis to cirrhosis, and so they estimated transition probabilities for mild and moderate health states using data from a trial of treatments for mild HCV undertaken in a number of London hospitals,[415] during which the liver fibrosis stage was determined by liver biopsy. Transition probabilities are listed in *Table 20*.

TABLE 20 Annual transition probabilities and sources employed within Markov model

Health state–health state	Transition probability	PSA distribution	Source
Mild–moderate	0.025	Dirichlet	Wright *et al.*[414]
Moderate–cirrhosis (HBeAg-positive)	0.037		
Moderate–cirrhosis (HBeAg-negative)	0.09		Dakin *et al.*[372]
Moderate–HCC	0.048		
Excess mortality–moderate fibrosis	0.0035		
Compensated cirrhosis–decompensated cirrhosis	0.05		
Compensated cirrhosis–HCC	0.024		
Excess mortality–compensated cirrhosis	0.051		
Decompensated cirrhosis–HCC	0.024		
Decompensated cirrhosis–liver transplant	0.016		
Decompensated cirrhosis–death	0.30		
HCC–liver transplant	0.0155		
Excess mortality–HCC	0.56		
Liver transplant–death	0.21		
Post liver transplant–death	0.057		

Mortality data

An all-cause mortality rate was calculated using the Interim Life tables for England and Wales 2008–10.[416] The risk of death increased each year according to age and the rate was weighted to allow for sex mix. The all-cause mortality rate was added to an excess mortality value (identified from study by Dakin *et al.*[372]) associated with the moderate fibrosis, compensated cirrhosis and HCC health states to provide a total risk of death per year. The all-cause mortality rate was not applied to the decompensated cirrhosis, liver transplant and post-liver-transplant health states; instead, a total mortality rate identified from the study by Dakin *et al.*[372] was applied. Mortality rates in the HBeAg-positive model ranged from 0.0007 to 0.337, and from 0.002 to 0.34 in the HBeAg-negative model. Excess mortality rates and total mortality rates employed within the models are listed in *Table 20*.

Antiviral treatment for hepatitis B: type and duration

The NICE website was reviewed to source national guidance on drug treatment for HBV.

We located guidelines for licensed drugs which included peginterferon alfa-2a (Pegasys®, Roche), entecavir (Baraclude®, BMS) and tenofovir disoproxil (Viread®, Gilead). Entecavir and tenofovir have marketing authorisation within the UK for the treatment of chronic HBV infection in adults with compensated liver disease and evidence of active viral replication, persistently elevated serum ALT levels and histological evidence of active inflammation and/or fibrosis. Peginterferon alfa-2a has UK marketing authorisation for the treatment of HBeAg-positive and HBeAg-negative chronic HBV infection in adults with compensated liver disease and evidence of viral replication, increased ALT and histologically verified liver inflammation and/or fibrosis. Interferon antiviral agents are contraindicated in chronic hepatitis patients with decompensated cirrhosis.[375,417,418]

The following treatment assumptions were employed in the model: only patients in the moderate fibrosis (F2–3), compensated cirrhosis (F4) and decompensated cirrhosis health states received treatment with antiviral agents; patients in the HCC, liver transplant and post liver transplant health states received usual standard of care.

Treatment with either entecavir or tenofovir (lifetime duration) was administered if patients tested positive (true positive or false positive); half of the eligible patients received entecavir and the other half received tenofovir.

Ten per cent of the patients who tested negative (true negative or false negative) received peginterferon alfa-2a for 1 year only, and if treatment was unsuccessful (we assumed that 30% of treated true negative patients would successfully respond to treatment and would no longer progress to any further health states except the death health state due to all-cause mortality), they would receive subsequent treatment with either entecavir or tenofovir for lifetime duration if diagnosed as positive ($\geq$ F2) at retest. The remaining 90% of true negative and false negative patients would undergo 'watchful waiting', where they would then receive treatment with either entecavir or tenofovir for lifetime duration (if diagnosed as positive during a retest).

The dosage of treatment was based on national recommendations sourced from the *British National Formulary* (BNF) 64[419] (see *Table 24*).

Treatment effectiveness

A meta-analysis by Woo *et al.*[420] which evaluated the relative efficacies of the first 12 months of treatment for chronic HBV was identified from a general search carried out using Google Scholar (http://scholar.google.com) and the search terms 'treatment effectiveness', and 'tenofovir and entecavir' (search date 12 May 2012). The paper evaluated a number of drugs – lamivudine, pegylated interferon, adefovir dipivoxil, entecavir, telbivudine and tenofovir – for use as monotherapy or in combination therapy in treatment-naive individuals.

Woo *et al.*[420] conducted a systematic literature review to locate studies of published RCTs of drugs used to treat chronic HBV as either monotherapies or combination therapies. They included studies that examined the impact of treatment in both HBeAg-positive and HBeAg-negative patients. The studies included in the review were required have examined the use of the drug using randomised, phase 3, controlled trials comparing new drug treatments with either a placebo or a licensed drug. The methodological quality of each paper was assessed independently by two reviewers using the Cochrane risk of bias tool.

The data were analysed using a Bayesian mixed-treatment comparison (MTC) analysis which allowed the authors to combine direct and indirect comparisons of the treatments. Our model used the treatment efficacy elicited from the Bayesian MTC for pegylated interferon, entecavir and tenofovir. The study did not provide separate efficacy data for peginterferon for HBeAg-negative chronic HBV, and we assumed the same RR as for HBeAg-positive. Treatment effectiveness represented by RRs and their associated CIs are displayed in *Table 21*.

We assumed that patients who tested false positive and who were in a mild state would receive the same treatment benefit from antiviral treatment as those in a moderate or cirrhotic health state. We also assumed that patients who test false negative and receive peginterferon alfa-2a for 1 year would not receive treatment benefit but would incur the costs and disutility associated with peginterferon alfa-2a treatment.

Cost data

To populate our model, we undertook a search for other relevant cost-effectiveness literature that would provide data on the costs associated with treating the different levels of fibrosis and cirrhosis in patients with HBV. To do this, we employed the search strategy devised by Shepherd *et al.*[371] and undertook an updated search using the MEDLINE database (Ovid platform, search date 10 May 2012, searched 2004–12). The search strategy is listed in *Appendix 2*. Additional search terms were added to locate papers related to other relevant treatment (entecavir, tenofovir disoproxil, and their brand names, baraclude and viread). Nine hundred and seventy-one papers were located and their titles were reviewed to determine if they were eligible. Twenty-three were retrieved for abstract review. Of these, 17 were excluded and six were retrieved for full review (*Table 22*). *Chapter 3* contains further details of the literature review inclusion criteria used when reviewing and assessing the applicability of cost data literature.

Of the six papers reviewed, four papers included cost data from a study on mild HCV by Wright *et al.*[414] This study collected resource use and cost data alongside a RCT for mild HCV.[414,415] Detailed cost data were collected from three centres based in London, Newcastle and Southampton. Resource use information collected covered inpatient and outpatient care, investigations, procedures, drug use and other services including psychiatric services.

TABLE 21 Relative risks associated with treatment

Drug	RR	95% CI
HBeAg (+ve)		
Entecavir	0.56	0.12 to 0.94
Tenofovir	0.53	0.06 to 0.95
Peginterferon alfa-2a	0.52	0.06 to 0.95
HBeAg (–ve)		
Entecavir	0.64	0.01 to 1.00
Tenofovir	0.65	0.01 to 1.00
Peginterferon alfa-2a	0.52	0.06 to 0.95

TABLE 22 Papers reviewed for cost data

Authors	Title	Journal	Source cost data used in study
Brown *et al.*[421] 2004	Hepatitis B management costs in France, Italy, Spain, and the United Kingdom	*Journal of Clinical Gastroenterology*	Questionnaire used to collect data from specialist clinicians in UK, Spain, Italy and France
Dakin *et al.*[372] 2010	Cost–utility analysis of tenofovir disoproxil fumarate in the treatment of chronic hepatitis B	*Value in Health*	Published literature: Wright *et al.* 2006[414]
Jones *et al.*[422] 2010	Tenofovir disoproxil fumarate for the treatment of chronic hepatitis B	SHTAC	Published literature: Wright *et al.* 2006[414]
Jones *et al.*[423] 2009	Adefovir dipivoxil and pegylated interferon alfa for the treatment of chronic hepatitis B: an updated systematic review and economic evaluation	*Health Technology Assessment*	Published literature: Wright *et al.* 2006,[414] Longworth 2003[424]
Takeda *et al.*[425] 2007	A systematic review and economic evaluation of adefovir dipivoxil and pegylated interferon-alfa-2a for the treatment of chronic hepatitis B	*Journal of Viral Hepatitis*	Expert opinion and published literature: Wong 1995,[393] Longworth 2003[424]
Veenstra *et al.*[426] 2007	Cost-effectiveness of peginterferon alfa-2a compared with lamivudine treatment in patients with HBe-antigen-positive chronic hepatitis B in the United Kingdom	*European Journal of Gastroenterology and Hepatology*	Expert opinion and published literature: Wright *et al.* 2006,[414] Longworth 2003,[424] Bennett *et al.* 1997[395]

One of the studies, by Veenstra *et al.*,[426] sourced additional costs from a study by Bennett *et al.*;[395] however, this study collected costs for the health states using a US population which, due to the difference in the health-care systems, would not be transferable to a UK population.

The study by Brown *et al.*[421] collected data on resource use from specialists based in four countries including the UK. The authors used the data collected to identify resource use and associated costs for the management of fibrosis and cirrhosis.

The study by Takeda *et al.*[425] also used the 1995 Wong paper[393] as a source; as this paper is an older published paper, we felt there would be more up-to-date costs available.

Three of the studies used data from the Cost-Effectiveness of Liver Transplantation (CELT) study[424] to elicit a cost for the liver transplant health states. The CELT study collected costs on adult patients listed for an isolated liver transplant (aged 16 years and over) between December 1995 and December 1996. The costs were collected and split into phases, which were determined according to when the resource use took place and according to disease aetiology (HBV, HCV, alcoholic cirrhosis). Data were split into an assessment phase, candidacy phase, transplant phase and post-transplant phases. The assessment phase started at the date of admission for assessment of suitability for liver transplantation to the date of listing for transplantation (for patients who were not listed for transplantation, the discharge date was used as the end date). The candidacy phase started at the date of listing to the date the patient was admitted for the transplant operation. The transplant phase started at the date of admission for the transplant operation to the date of discharge following the operation and the post-transplant phase started at the date of discharge following the operation onwards for a period of 2 years. Resource use data on blood products used, number of dietitian sessions, drugs used, inpatient stay, nutritional support received, outpatient visits, physiotherapy sessions, tests, length of transplant operation and key treatments and investigations were collected.

Based on the review, we concluded that the data on costs for treating HCV from the study by Wright *et al.*[414] for the mild, moderate and cirrhotic health states would be comparable with the costs for treating patients with HBV in these health states, as the resource use identified and collected (inpatient, outpatient care, procedures) should be similar. The health state costs sourced from this study did not contain the cost of antiviral treatment; therefore, they were suitable to use and we added the cost of treatment separately in the model.

From the review, we decided to source information for the decompensated cirrhosis, HCC, liver transplant and post-liver-transplant health states from the CELT study.[424] Using the raw data collected in this study, we calculated the cost for all patients with HBV who had received a transplant (sample size 24). The average length of admission for the liver transplant operation was 28 days. We approximated this to 1 month and calculated the yearly cost of a liver transplant as 11 months of post-transplant care (estimated from the average monthly cost in the first year following transplantation) plus the month the transplant operation took place. We also calculated a cost for the post-transplant health state, estimated as the average monthly cost of the second year of post-transplant care (sample size 24).

To calculate a cost for the decompensated cirrhosis and HCC health states, we used the average cost for patients with HBV awaiting liver transplant (patients in the 'assessment' and 'candidacy' disease stages were used; sample size 25).

Costs were inflated to 2012 levels using NHS inflation indices.[67] Health state costs did not include the costs associated with antiviral treatment; these were included separately. Costs are shown in *Table 23*.

As we employed two different sources of data to populate our model, there is a sizable difference between the cost for the compensated and decompensated cirrhosis health states. We conducted a sensitivity analysis where we populated our model with the cost for the decompensated cirrhosis health state (£9121) and HCC health state (£8127) from the study by Wright *et al.*[414] The main component of the costs for these two health states in the Wright *et al.*[414] model was inpatient-days. We inflated the costs to 2012 prices using NHS inflation indices.[67]

Test costs

Non-invasive liver tests using imaging modalities were sourced from published Department of Health reference costs.[427] Costs of direct and indirect serum markers were obtained from communication with finance departments based at the Royal Free Hospital. Costs for patented serum markers were sourced directly from manufacturers and via communication with finance departments based at the Royal Free Hospital (see *Appendix 5*, *Table 68*).

TABLE 23 Annual health state costs (£ 2012)

Health state	Cost per year, £	Standard error	PSA distribution	Source
Mild fibrosis	185	36.39	Gamma	Wright *et al.* 2006[414]
Moderate fibrosis	986	101.69	Gamma	
Compensated cirrhosis	1521	309.05	Gamma	
Decompensated cirrhosis	36,194	9967.19	Gamma	Longworth *et al.*[424]
HCC	36,194	9967.19	Gamma	
Liver transplant	64,122	5584.70	Gamma	
Post liver transplant	16,321	7932.51	Gamma	

The two most commonly performed liver biopsy tests are percutaneous and transjugular liver biopsy. We assumed that, for our purposes, patients received a percutaneous liver biopsy, as this tends to be associated with less severe adverse effects. We estimated that a diagnostic liver biopsy would cost £956.61.[428] Where required, costs were inflated to 2012 prices using NHS inflation indices.[67] Tests costs and sources are listed in *Appendix 9*. All NILT tests costs are based on incremental costs and exclude the capital costs of the equipment.

Medication costs

Treatment costs and recommended dosages were sourced from the BNF 64[419] and are listed in *Table 24*.

Utility data

An initial literature search for existing published utility data for HBV was undertaken using the MEDLINE database (Ovid platform, searched 11 May 2012, coverage 2004 to 2012). We supplemented this search using both the cost-effectiveness analysis (CEA) Registry (searched 11 May 2012, all dates) and EuroQol website (searched 11 May 2012, all dates).

We searched the MEDLINE database using the search strategy for quality of life devised by Shepherd *et al.*[371] We updated this search to include papers from 2004 onwards. The full search strategy is outlined in *Appendix 2*. This search returned 121 papers.

We searched the EuroQol using the general search term 'hepatitis B'. This search returned 16 papers. We also searched the CEA Registry using the search term 'hepatitis B', which returned 39 studies.

The title and abstract of each paper was reviewed and full papers were retrieved for review if they met our inclusion criteria. *Chapter 3* outlines the inclusion criteria that applied when reviewing studies for quality of life data. After excluding three duplicate papers, eight were retrieved for full review.[372,414,425,429–432]

Four of the studies retrieved[372,423,425,429] used the same study by Levy *et al.*[430] as a source for utility values.

Levy *et al.*[430] employed the standard gamble technique to collect health-related utility data for six HBV-related health states (from both infected and uninfected respondents). Hypothetical health states were developed using expert opinion and the Liver Disease Quality of Life Instrument. Data were elicited from respondents from the USA, Canada, the UK, Spain, Hong Kong and China. The study analysed 1134 respondents, of whom 100 were from the uninfected population and 93 from the infected population were from the UK.

One study used clinical judgement to elicit health-related quality of life (HRQoL) values.[431] The authors noted that the weights were arbitrary and better sources could be used.

A 2008 study by McLernon *et al.*[432] conducted a systematic review of published literature.[414,433–436] They obtained values for the compensated cirrhosis, decompensated cirrhosis and HCC health states using these data; however, the studies had mainly been carried out in HCV and the search conducted looked for quality of life data for all aetiologies, not specifically for HBV.

TABLE 24 Annual cost and dosage of medication applied in model (£ 2012)

Drug	Brand name	Dosage	Cost per year	Source
Peginterferon alfa-2a	Pegasys®	180 mg per week	£6469	BNF 64[419]
Entecavir	Baraclude®	500 mg daily	£4420	BNF 64[419]
Tenofovir	Viread®	245 mg daily	£2926	BNF 64[419]

The search also identified the 2006 study on treatment in mild HCV by Wright *et al.*[414] As none of the other studies reviewed identified separate utility values for mild and moderate health states, we adopted the same approach to sourcing data as that used for costs. We employed the same utility values used by Wright *et al.*[414] for the HCV mild, moderate and compensated cirrhosis health states. As both HBV and HCV lead to cirrhosis and related complications, it is assumed that cirrhosis would impact on the quality of life similarly for patients with HBV and HCV.

We decided to source utility values for the decompensated cirrhosis, decompensated cirrhosis, HCC, liver transplant and post liver transplant health states from the CELT study for HBV patients.[424] We used this source rather than one of the studies reviewed, as the only study reviewed which collected data directly used a standard gamble technique[430] rather than a generic preference-based measure such as the EQ-5D. As we had data available for patients with HBV (sourced from CELT study), it was felt this would be a more accurate source of data to use rather than the data elicited from the study by Levy *et al.*[430] or the systematic review by McLernon *et al.*,[432] which were applicable to all liver disease aetiologies.

Additionally, we concluded that both the Wright *et al.*[414] and the CELT study[424] would be suitable sources of data as they had both used the EQ-5D preference measure to elicit utility values within a UK population (see *Chapter 3* for inclusion criteria for quality of life data).

During the mild HCV RCT,[414] questionnaires were self-administered at baseline and treatment weeks 12, 24 and 48, and at follow-up weeks 12, 24 and 48. For the moderate and cirrhotic health states, 302 patients were sent an EQ-5D (of whom 60% of those with diagnosed as being in a moderate health state responded, and 54% of those diagnosed as having cirrhosis responded).

The CELT study[424] collected HRQoL data using the EQ-5D questionnaire. We analysed the data collected for HBV patients at the time patients were placed on the waiting list for a transplant, and at 3 months, 6 months, 12 months and 24 months post transplant.

It was assumed that utility values for the decompensated cirrhosis and HCC health states would be the same (sample size 25). This was assumed to be equivalent to the average utility value at 'listing' stage.

A utility value for the liver transplant health state was estimated using an area under the curve approach and the average utility values collected at 3, 6 and 12 months post transplant (sample size 24). A utility value for the post liver transplant disease stage was estimated from the average utility for HBV patients collected at 24 months post transplant (sample size 24). A PSA was carried out using a utility decrement approach. A beta distribution is often employed for utility values; however, this may not be appropriate for states close to death, where values of less than one are possible. For the HBV model, we performed a simple transformation of the data ($D = 1 - U$, where D is the utility decrement and U is the utility value). The decrement was constrained on the interval '0 to positive infinity' and a gamma distribution was then applied.[437] *Table 25* lists the utility data used in the model.

As we sourced utility value data from two different studies, the utility values used in the study for the decompensated cirrhosis and HCC health states were slightly higher than those used for the compensated cirrhosis health state. We undertook a sensitivity analysis where we set the utility values for the decompensated cirrhosis and HCC health states to those used in the HCV model (see *Chapter 6*) – as these were lower values than those used in the HBV model – to test if this had any effect on the robustness of the results.

Adverse effects associated with antiviral treatment

As peginterferon alfa-2a has associated side effects (such as influenza-like symptoms, depression and anxiety), we allowed for the disutility associated with antiviral treatment to be reflected in the model. We modelled a disutility decrement that was applicable during treatment, using data identified from the Wright *et al.*[414] study. This study reported HRQoL data using the EQ-5D instrument for 144 patients. They used the data from weeks 12 and 24 from the baseline date to estimate utility decrement values of

TABLE 25 Utility data

Health state	Utility value	SE	PSA distribution	Source
Mild fibrosis	0.77	0.035	Gamma	Wright *et al.*[414]
Moderate fibrosis	0.66	0.018		
Compensated cirrhosis	0.55	0.032		
Decompensated cirrhosis	0.57	0.076		Longworth *et al.*[424]
HCC	0.57	0.076		
Liver transplant	0.73	0.016		
Post liver transplant	0.78	0.064		
Mild: during treatment	0.65	0.035		
Moderate: during treatment	0.55	0.018		
Compensated cirrhosis: during treatment	0.44	0.040		
Death	0	0		Assumption

SE, systemic embolism.

0.11. The decrement was applied to all patients in the mild and moderate fibrosis and compensated cirrhosis health states as a multiplicative disutility for the duration of treatment with peginterferon alfa-2a. It was conservatively assumed that any side effects associated with treatment with entecavir or tenofovir would have no impact on quality of life.

Disutility from non-invasive liver treatment and liver biopsy

We have assumed that any adverse events associated with the NILTs would not have a significant impact on health-related utility. However, as liver biopsy is associated with morbidity and mortality risks and patient discomfort, we incorporated a utility decrement to represent expected adverse events. In the absence of identified data, this was modelled by applying a utility decrement of 0.2 based on data employed in a previous study,[428] where the authors had conducted a literature review for data on liver biopsy associated adverse events and mortality. We undertook a number of sensitivity analyses to test the robustness of the results to changes in the utility decrement.

Estimated long-term costs and quality-adjusted life-years

The estimated long-term costs and QALYs for HBeAg-positive and HBeAg-negative chronic HBV (output from Markov models) are shown in *Table 26*.

Analysis

As discussed previously in *Chapter 3*, a PSA was conducted, and an incremental analysis was carried out to estimate the most cost-effective testing option. A CEAC and CEAF were constructed to summarise the uncertainty around the results. Several one-way sensitivity analyses were also undertaken.

Treatment benefit

Our base-case analysis assumed that patients who are in a mild health state (F0–1) receive the same treatment benefit (same effectiveness) as patients in a moderate or cirrhotic health state, despite being incorrectly diagnosed and treated.[373] We tested the robustness of this assumption in a sensitivity analysis by setting the treatment benefit for patients who incorrectly receive treatment while in a mild health state to zero.

TABLE 26 Estimated long-term costs and QALYs for HBeAg-positive and HBeAg-negative chronic HBV infection

Diagnostic test outcome	Cost (£ 2012 per person)	QALY (per person)
HBeAg-positive		
TP	105,126	7.99
TN	39,201	15.61
FP	97,859	17.06
FN	99,997	7.46
Treat all persons	101,794	12.15
Treat no one	37,966	9.64
HBeAg-negative		
TP	98,145	6.65
TN	32,772	13.12
FP	90,877	15.31
FN	92,571	6.18
Treat all persons	94,815	10.62
Treat no one	37,518	8.82

FN, false negative; FP, false positive; TN, true negative; TP, true positive.

Robust test accuracy data

In the base-case analysis, all tests were included despite there being limited data available for some tests. A sensitivity analysis including only tests where the standard bivariate random-effects model used for the meta-analysis[63] converged was conducted. We conducted an analysis with the five tests where the bivariate model had converged – APRI high cut-off, Fibroscan, Fibrotest, FIB-4 low cut-off and APRI low cut-off – and compared with liver biopsy, 'treat all' and 'treat no one' strategies.

Utility value amendment

We carried out an analysis where we amended the utility values within the model for the decompensated cirrhosis, HCC, liver transplant and post liver transplant. The sensitivity analysis used the lower values from the HCV model (see *Chapter 6* for values).

Change in liver biopsy decrement

We also carried out analyses using different utility decrement values to measure the adverse effects from liver biopsy. The base-case analysis set the utility decrement value to 0.2 based on a previous analysis which had conducted a literature search for data relating to adverse events and mortality associated with liver biopsy.[428] In the sensitivity analysis we set the utility decrement at 0 and 0.3 to test the impact on the robustness of the results.

Change to health state costs

We carried out an analysis where we changed the costs for the decompensated cirrhosis and HCC health states to the costs used in the Wright *et al.*[414] study.

Average disease prevalence

Prevalence within the model is based on studies that may have been carried out largely in tertiary care centres, and the prevalence of liver fibrosis in this population may be an overestimate. To test the impact of this we undertook a sensitivity analyses using the minimum prevalence (disease prevalence modelled as ≥ F2 at 27%), maximum prevalence (92%) and the 25th and 75th quartile values (43% and 65%,

respectively), estimated from the meta-analysis of the systematic review data. In this sensitivity analysis, it is assumed that the sensitivity and specificity of the tests would be the same in populations with high and low prevalence as observed in the studies included in the review.

Sensitivity and specificity of retest

Our base-case analysis assumes that the retest (from the meta-analysis of the systematic review data in the watchful waiting strategy for patients with a negative test result) has perfect sensitivity and specificity. This is likely to overestimate the accuracy of the retest procedure. We tested the impact of this assumption by applying the sensitivity and specificity of three commonly used tests: APRI low cut-off (estimated sensitivity of 80% and specificity of 65%), Fibrotest (estimated sensitivity of 66% and specificity of 80%) and Fibroscan (estimated sensitivity of 71% and specificity of 84%).

Choice of tests for second stage of analysis

For the second stage of the analysis, the two most cost-effective tests when assessed within each specific test category singly (with and without a defined threshold) were used in the analysis of sequential testing. To test if changing the method used to choose tests for the second stage of the analysis had an effect on the overall result, we carried out an analysis where the most effective NILT from within each NILT category and the least costly NILT from each category were included in the analysis of sequential testing.

Change to hepatitis B e antigen-negative model (age and sex ratio)

We also undertook a sensitivity analysis, where we set the age-and-sex-mix data in the HBeAg-negative model equivalent to the age and sex mix in the HBeAg-positive model (base age set to 31 years and the proportion of cohort who were male set at 70%) to check if the results from both models would be similar. The only remaining differences between the HBeAg-positive and HBeAg-negative models related to the different efficacy of treatment and the increased probability of moving to a cirrhotic state from a moderate health state.

Reduction in length of time of successful response rate to peginterferon alfa-2a treatment

Our base-case analysis for true-negative patients assumes that 30% of those who receive treatment with peginterferon alfa-2a have a successful response and no longer progress to further health states in the model retaining a risk of all-cause mortality only. We tested this assumption in the model by assuming that the successful response lasts for 15 years only.

Change to non-invasive liver test costs

We carried out a sensitivity analysis where we changed the cost of the NILTs [we set the watchful waiting retest cost and all NILT costs within the model to the same cost; for comparison, we assumed an indirect serum marker test cost (we chose a commonly used test, APRI, as our comparator)]. By changing the cost of a NILT, we aimed to determine if changing the test cost (marginal cost) had an impact on the robustness of the results.

Results

Hepatitis B e antigen-positive chronic hepatitis B

At a cost-effectiveness threshold of £20,000, the cost-effective strategy is to employ a sequential testing strategy where a direct serum marker, hyaluronic acid, is combined with an imaging modality, MR elastography, with an ICER of £19,612. With this testing strategy, positive test diagnoses are confirmed with a second NILT and if the results disagree, a liver biopsy is administered to confirm the result; negative test responses undergo a process of watchful waiting.

Using a higher cost-effectiveness threshold value of £30,000, an imaging modality, MR elastography, used singly, becomes the most cost-effective option, with an ICER of £28,585. When we assessed all tests singly

in the first stage of the analysis, the most cost-effective test given a threshold value of £20,000 was GUCI, with an ICER of £19,716.

As the CEAF is more representative of the uncertainty around the optimal testing option, we present the CEAF (*Figure 4*) in the main analysis and the CEAC in *Appendix 6*.

The CEAF shows that the probability of hyaluronic acid combined with MR elastography (strategy 2) being the optimal testing option, given a cost-effectiveness threshold value of £20,000, is 4%. The CEAF also shows that MR elastography, used singly, has a 13% probability of being cost-effective for cost-effectiveness thresholds starting at £28,585, which reduces to a 3.5% probability when the cost-effectiveness threshold increases to £43,946. For cost-effectiveness threshold values greater than £43,946, the 'treat all' strategy has a 35% probability of being cost-effective.

What is noticeable about the CEAF is that two of the strategies (testing with MR elastography alone or a combination of hyaluronic acid and MR elastography) have a very low probability of being the optimal testing option, compared with a testing strategy of treating everyone.

For reasons of clarity, the CEAC presented in *Appendix 10* displays only those testing options which had a ≥ 4% probability of being cost-effective. The CEAC also displays some tests that are not picked up by the CEAF, including the Forns (high cut-off) serum marker. The reasons for this include a skew in cost data and also the fact that those tests which have the highest probability of being cost-effective may not be the optimal testing option.

Liver biopsy was a comparator for both stages of the analysis; however, this testing option is dominated by other less costly but more effective options. *Table 27* presents incremental results for the first stage of the analysis and *Table 28* incremental results for the second stage where a number of combined tests are compared using a number of different sequential testing strategies (see *Chapter 3* for details of testing strategies).

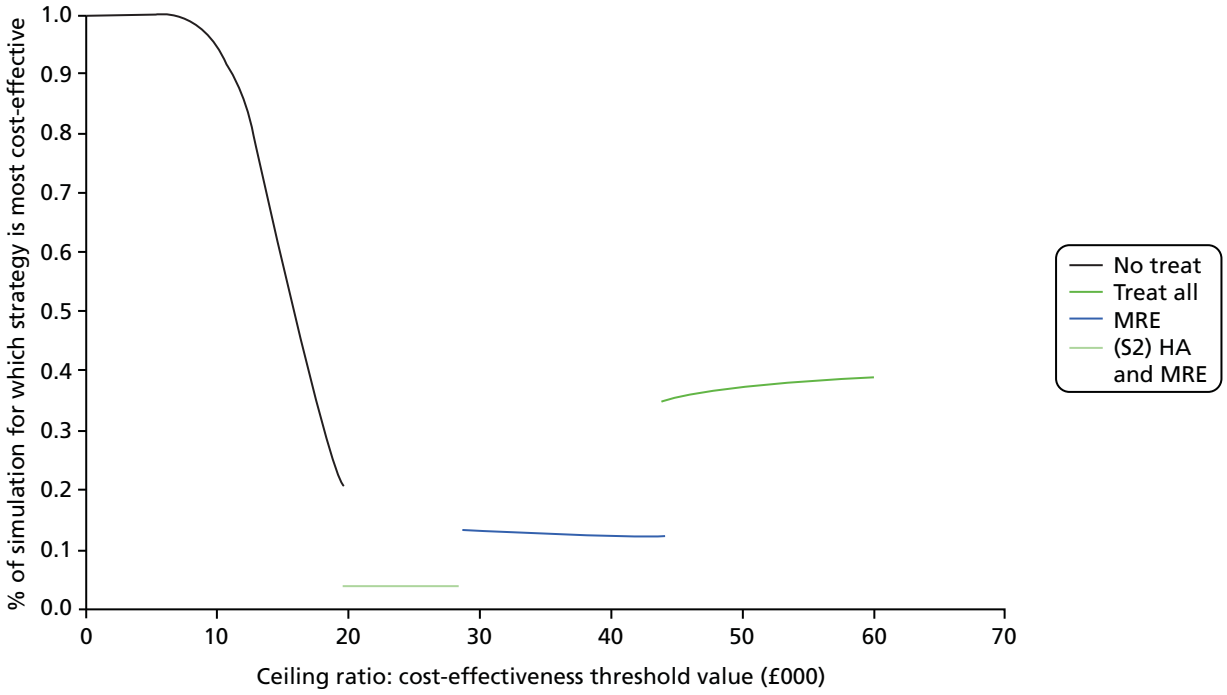

FIGURE 4 Cost-effectiveness acceptability frontier for HBeAg-positive second stage of the analysis. HA, hyaluronic acid; MRE, MR elastography.

TABLE 27 Hepatitis B e antigen-positive base-case analysis (first stage where all tests are compared singly)

Test strategy	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,831	9.64	–	–	–
FIB-4 (high cut-off)	73,028	11.33	–	–	Extendedly dominated
Forns index (high cut-off)	73,115	11.35	–	–	Extendedly dominated
APGA	73,373	11.36	–	–	Extendedly dominated
Liver biopsy	75,957	11.41	–	–	Dominated
APRI (high cut-off)	75,139	11.45	–	–	Dominated
US	77,657	11.48	–	–	Dominated
Hui index	76,047	11.50	–	–	Dominated
US SAPI (high cut-off)	75,610	11.51	–	–	Extendedly dominated
GUCI	74,921	11.52	37,090	1.88	19,716
Fibroscan	79,004	11.61	–	–	Dominated
Forns index (low cut-off)	80,008	11.61	–	–	Dominated
Fibrotest	79,519	11.62	–	–	Dominated
MR elastography	77,657	11.64	2,737	0.12	23,468
US SAPI	80,442	11.66			Extendedly dominated
PAPAS	80,223	11.65	–	–	Dominated
Hyaluronic acid	79,084	11.66	–	–	Extendedly dominated
DW-MRI	80,751	11.67	–	–	Extendedly dominated
FIB-4 (low cut-off)	81,347	11.66	–	–	Extendedly dominated
Hepascore	81,399	11.69	–	–	Extendedly dominated
ARFI	83,487	11.71	–	–	Dominated
AAR	84,951	11.72	–	–	Dominated
CEUS	82,377	11.74	–	–	Extendedly dominated
CT	84,097	11.73	–	–	Dominated
Age–Platelet Index	84,289	11.73	–	–	Dominated
APRI (low cut-off)	83,770	11.75	–	–	Extendedly dominated
US SAPI (low cut-off)	91,287	11.95	–	–	Extendedly dominated
Treat all	101,484	12.18	23,827	0.54	44,256

AAR, AST–ALT ratio; APGA, AST, platelet count, GGT, α-fetoprotein; CEUS, contrast-enhanced ultrasound; DW-MRI, diffusion-weighted magnetic resonance imaging; PAPAS, age, ALP, α-fetoprotein, AST; US, ultrasound.

TABLE 28 Hepatits B e antigen-positive base-case analysis (second stage of the analysis where tests are combined sequentially and compared with liver biopsy, 'treat all' and 'treat no one'

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,831	9.64	–	–	–
(S1) Forns index (high cut-off)	73,960	11.16	–	–	Dominated
Forns index (high cut-off)	73,084	11.35	–	–	Dominated
(S2) Forns index (high cut-off) and MR elastography	73,062	11.35	–	–	Dominated
(S2) Forns index (high cut-off) and Fibrotest	73,037	11.35	–	–	Dominated
(S2) Forns index (high cut-off) and hyaluronic acid	73,011	11.35	–	–	Extendedly dominated
(S1) US SAPI (high cut-off)	76,238	11.38	–	–	Dominated
(S1) GUCI	75,539	11.40	–	–	Dominated
Liver biopsy	75,957	11.41	–	–	Dominated
(S3) Forns index (high cut-off) and Fibrotest	75,965	11.42	–	–	Dominated
(S3) Forns index (high cut-off) and hyaluronic acid	75,939	11.43	–	–	Dominated
(S3) Forns index (high cut-off) and MR elastography	75,927	11.43	–	–	Dominated
FIB-4 (combined cut-off) and Fibroscan	74,911	11.43	–	–	Dominated
(S4) Forns index (high cut-off) and Fibrotest	74,953	11.45	–	–	Dominated
(S3) GUCI and US SAPI (high cut-off)	75,805	11.46	–	–	Dominated
(S3) Fibrotest and US SAPI (high cut-off)	76,025	11.46	–	–	Dominated
(S3) GUCI and Fibrotest	75,904	11.46	–	–	Dominated
Forns index (combined cut-off) and Fibroscan	75,253	11.47	–	–	Dominated
(S2) Fibrotest and US SAPI (high cut-off)	74,851	11.47	–	–	Dominated
(S3) Hyaluronic acid and US SAPI (high cut-off)	75,908	11.47	–	–	Dominated
(S3) GUCI and hyaluronic acid	75,803	11.48	–	–	Dominated
(S2) GUCI and US SAPI (high cut-off)	74,431	11.48	–	–	Dominated
(S4) Forns index (high cut-off) and hyaluronic acid	75,342	11.48	–	–	Dominated
(S3) GUCI and MR elastography	75,778	11.48	–	–	Dominated
(S2) GUCI and Fibrotest	74,517	11.49	–	–	Dominated
(S3) Fibrotest and MR elastography	76,198	11.49	–	–	Dominated
(S4) Forns index (high cut-off) and MR elastography	75,937	11.50	–	–	Dominated
(S2) Fibrotest and MR elastography	75,023	11.50	–	–	Dominated
(S2) GUCI and hyaluronic acid	74,416	11.50	–	–	Dominated

continued

TABLE 28 Hepatits B e antigen-positive base-case analysis (second stage of the analysis where tests are combined sequentially and compared with liver biopsy, 'treat all' and 'treat no one' (*continued*)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
(S4) Fibrotest and US SAPI (high cut-off)	75,667	11.50	–	–	Dominated
(S2) GUCI and MR elastography	74,404	11.50	–	–	Extendedly dominated
(S3) Hyaluronic acid and MR elastography	76,028	11.51	–	–	Dominated
US SAPI (high cut-off)	75,618	11.51	–	–	Dominated
(S4) GUCI and US SAPI (high cut-off)	75,450	11.51	–	–	Dominated
(S4) GUCI and Fibrotest	75,400	11.52	–	–	Dominated
(S1) Fibrotest	80,207	11.52	–	–	Dominated
(S2) Hyaluronic acid and US SAPI (high cut-off)	75,265	11.52	–	–	Dominated
GUCI	74,942	11.52	–	–	Extendedly dominated
(S4) Hyaluronic acid and US SAPI (high cut-off)	75,751	11.53	–	–	Dominated
(S4) GUCI and hyaluronic acid	75,440	11.54	–	–	Dominated
(S1) MR elastography	78,061	11.55	–	–	Dominated
(S4) Fibrotest and MR elastography	76,226	11.55	–	–	Dominated
(S4) GUCI and MR elastography	75,814	11.56	–	–	Extendedly dominated
(S2) Hyaluronic acid and MR elastography	75,386	11.56	37,555	1.9	19,612
(S4) Hyaluronic acid and MR elastography	76,069	11.57	–	–	Extendedly dominated
(S1) Hyaluronic acid	79,752	11.58	–	–	Dominated
APRI (combined cut-off) and Fibroscan	77,596	11.59	–	–	Extendedly dominated
Fibrotest	79,495	11.62	–	–	Dominated
MR elastography	77,627	11.64	2241	0.1	28,585
Hyaluronic acid	79,148	11.67	–	–	Extendedly dominated
Treat all	101,484	12.18	23,857	0.5	43,946

US, ultrasound.

A scatterplot illustrating the position of each testing strategy on the CEAC compared with the testing strategy 'treat no one' can be found in *Appendix 12*.

Hepatits B e antigen-negative chronic hepatitis B

The base-case analysis result indicates that given a cost-effectiveness threshold of £30,000, the cost-effective strategy is to adopt a strategy of 'treat all' with an ICER of £28,137. All other testing strategies are dominated by the 'treat all' strategy. As all other tests were dominated, the ICER presented represents the 'treat all' strategy versus a 'treat no one' strategy. Given a UK threshold range of £20,000, the 'treat no one' strategy was the most cost-effective option to adopt.

The CEAF (*Figure 5*) for the HBeAg-negative model shows that the probability of 'treat all' being the optimal testing option (highest expected net benefit), given a cost-effectiveness threshold value of £30,000, is 38%. The CEAF also shows that 'treat all' has a 35% probability of being cost-effective for cost-effectiveness thresholds starting at £28,137.

Appendix 10 displays the CEAC for the overall base-case analysis for HBeAg-negative model. For reasons of clarity, only those strategies that have on average a ≥ 4% probability of being the most cost-effective option have been included. The CEAC demonstrates that, given the data, there is a 38% chance that the additional cost of the 'treat all' strategy, compared with all other test strategies, is at or below £30,000 per life-year gained.

Liver biopsy was a comparator for both stages of the analysis; however, this testing option is dominated by other less costly but more effective options.

Table 29 presents incremental results for the first stage of the analysis and *Table 30* presents the incremental results for the second stage where a number of combined tests are compared using a number of different sequential testing strategies (see *Chapter 3* for details of the four testing strategies presented in the tables as S1, S2, S3 and S4).

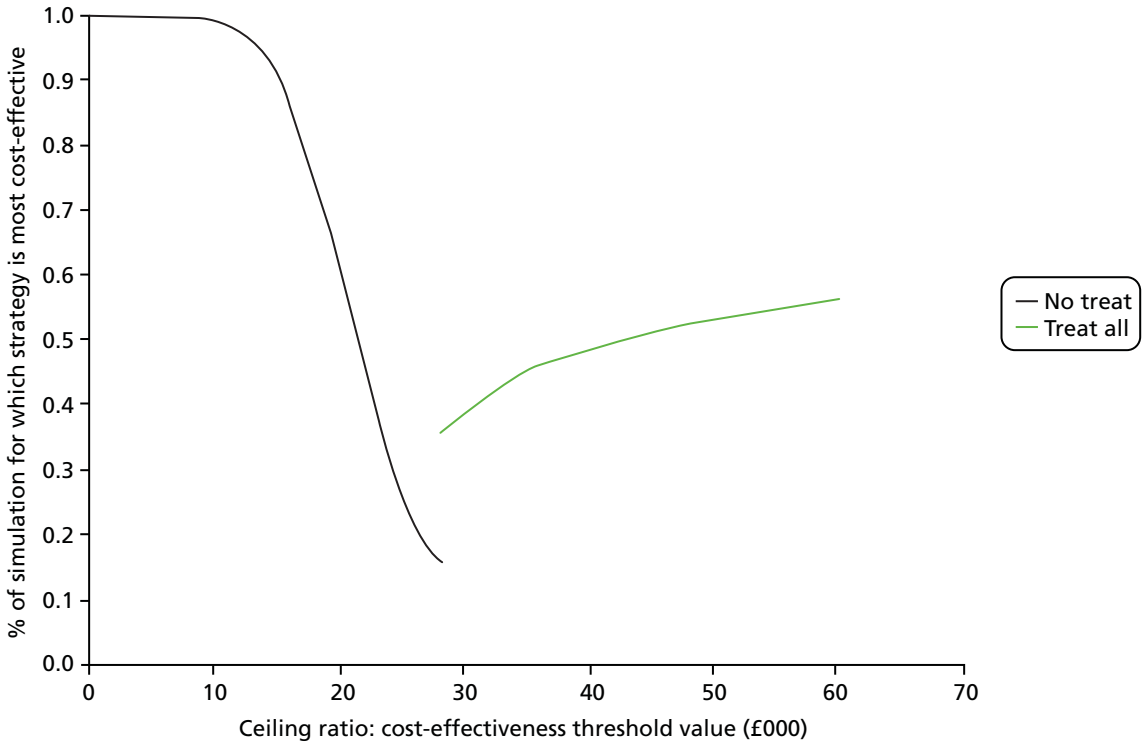

FIGURE 5 Cost-effectiveness acceptability frontier for HBeAg-negative: second stage of the analysis.

TABLE 29 Hepatitis B e antigen-negative base-case analysis (first stage where all tests are compared singly)

Test strategy	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,579	8.83	–	–	–
FIB-4 (high cut-off)	67,267	9.56	–	–	Extendedly dominated
Forns index (high cut-off)	67,348	9.58	–	–	Extendedly dominated
APGA	67,653	9.60	–	–	Extendedly dominated
Liver biopsy	70,274	9.64	–	–	Dominated
APRI (high cut-off)	69,428	9.70	–	–	Dominated
GUCI	69,196	9.76	–	–	Extendedly dominated
Hui index	70,308	9.77	–	–	Dominated
US SAPI (high cut-off)	69,635	9.78	–	–	Extendedly dominated
US	71,710	9.81	–	–	Extendedly dominated
MR elastography	71,791	9.93	–	–	Extendedly dominated
Fibroscan	73,007	9.93	–	–	Extendedly dominated
Fibrotest	73,739	9.93	–	–	Dominated
Forns index (low cut-off)	74,317	9.94	–	–	Dominated
Hyaluronic acid	73,448	9.96	–	–	Extendedly dominated
PAPAS	74,493	9.98	–	–	Extendedly dominated
US SAPI	74,508	9.99	–	–	Extendedly dominated
DW-MRI	74,774	10.01	–	–	Extendedly dominated
FIB-4 (low cut-off)	75,648	10.01	–	–	Dominated
Hepascore	75,624	10.03	–	–	Extendedly dominated
CEUS	76,577	10.09	–	–	Extendedly dominated
ARFI	77,512	10.10	–	–	Extendedly dominated
Age–Platelet Index	78,647	10.13	–	–	Dominated
CT	78,129	10.13	–	–	Dominated
AAR	79,321	10.13	–	–	Dominated
APRI (low cut-off)	78,083	10.13	–	–	Extendedly dominated
US SAPI (low cut-off)	85,587	10.45	–	–	Extendedly dominated
Treat all	96,525	10.92	58,947	2.09	28,137

AAR, AST–ALT ratio; APGA, AST, platelet count, GGT, α-fetoprotein; CEUS, contrast-enhanced ultrasound; DW-MRI, diffusion-weighted magnetic resonance imaging; PAPAS, age, ALP, α-fetoprotein, AST; US, ultrasound.

TABLE 30 Hepatitis B e antigen-negative base-case analysis (second stage of the analysis where tests are combined sequentially and compared with liver biopsy, 'treat all' and 'treat no one')

Strategies	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,579	8.83	–	–	–
(S1) Forns index (high cut-off)	68,237	9.40	–	–	Dominated
(S2) Forns index (high cut-off) and Fibrotest	66,908	9.58	–	–	Dominated
(S2) Forns index (high cut-off) and MR elastography	66,892	9.58	–	–	Dominated
(S2) Forns index (high cut-off) and hyaluronic acid	66,885	9.58	–	–	Extendedly dominated
Forns index (high cut-off)	67,361	9.58	–	–	Extendedly dominated
(S1) GUCI	69,766	9.64	–	–	Dominated
Liver biopsy	70,274	9.64	–	–	Dominated
(S3) Forns index (high cut-off) and Fibrotest	70,285	9.65	–	–	Dominated
(S3) Forns index (high cut-off) and hyaluronic acid	70,261	9.65	–	–	Dominated
(S1) US SAPI (high cut-off)	70,230	9.65	–	–	Dominated
(S3) Forns index (high cut-off) and MR elastography	70,242	9.65	–	–	Dominated
(S4) Forns index (high cut-off) and Fibrotest	69,015	9.68	–	–	Dominated
(S3) GUCI and US SAPI (high cut-off)	70,122	9.68	–	–	Dominated
FIB-4 (combined cut-off) and Fibroscan	69,189	9.69	–	–	Dominated
(S3) Fibrotest and US SAPI (high cut-off)	70,345	9.69	–	–	Dominated
(S3) GUCI and Fibrotest	70,212	9.69	–	–	Dominated
(S3) Hyaluronic acid and US SAPI (high cut-off)	70,226	9.70	–	–	Dominated
(S3) GUCI and hyaluronic acid	70,115	9.70	–	–	Dominated
(S3) GUCI and MR elastography	70,090	9.71	–	–	Dominated
(S2) GUCI and US SAPI (high cut-off)	68,562	9.71	–	–	Dominated
(S2) Fibrotest and US SAPI (high cut-off)	68,870	9.71	–	–	Dominated
(S4) Forns index (high cut-off) and hyaluronic acid	69,463	9.72	–	–	Dominated
(S2) GUCI and Fibrotest	68,637	9.72	–	–	Dominated
Forns index (combined cut-off) and Fibroscan	69,559	9.72	–	–	Dominated
(S3) Fibrotest and MR elastography	70,497	9.72	–	–	Dominated

continued

TABLE 30 Hepatitis B e antigen-negative base-case analysis (second stage of the analysis where tests are combined sequentially and compared with liver biopsy, 'treat all' and 'treat no one') (*continued*)

Strategies	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
(S2) GUCI and hyaluronic acid	68,541	9.73	–	–	Dominated
(S2) GUCI and MR elastography	68,529	9.73	–	–	Extendedly dominated
(S3) Hyaluronic acid and MR elastography	70,332	9.74	–	–	Dominated
(S2) Fibrotest and MR elastography	69,023	9.74	–	–	Extendedly dominated
(S4) GUCI and Fibrotest	69,537	9.75	–	–	Dominated
(S4) Fibrotest and US SAPI (high cut-off)	69,827	9.76	–	–	Dominated
(S4) Forns index (high cut-off) and MR elastography	70,115	9.76	–	–	Dominated
(S2) Hyaluronic acid and US SAPI (high cut-off)	69,332	9.76	–	–	Dominated
GUCI	69,167	9.76	–	–	Extendedly dominated
(S4) GUCI and US SAPI (high cut-off)	69,590	9.77	–	–	Dominated
(S4) GUCI and hyaluronic acid	69,610	9.78	–	–	Dominated
(S4) Hyaluronic acid and US SAPI (high cut-off)	69,928	9.78	–	–	Dominated
US SAPI (high cut-off)	69,609	9.78	–	–	Dominated
(S2) Hyaluronic acid and MR elastography	69,438	9.80	–	–	Extendedly dominated
(S4) Fibrotest and MR elastography	70,429	9.81	–	–	Dominated
(S4) GUCI and MR elastography	70,011	9.81	–	–	Extendedly dominated
(S4) Hyaluronic acid and MR elastography	70,277	9.83	–	–	Extendedly dominated
(S1) MR elastography	72,171	9.84	–	–	Dominated
(S1) Fibrotest	74,242	9.84	–	–	Dominated
APRI (combined cut-off) and Fibroscan	71,769	9.87	–	–	Dominated
(S1) Hyaluronic acid	73,861	9.88	–	–	Dominated
MR elastography	71,737	9.93	–	–	Extendedly dominated
Fibrotest	73,812	9.94	–	–	Dominated
Hyaluronic acid	73,426	9.96	–	–	Extendedly dominated
Treat all	96,525	10.92	58,947	2.09	28,137

US, ultrasound.

A scatterplot illustrating the position of each testing strategy on the cost-effectiveness acceptability curve compared with the testing strategy 'treat no one' can be found in *Appendix 12*.

Sensitivity analysis results

Hepatits B e antigen-positive

The base results remained robust to the majority of the sensitivity analyses, including changes to utility values, health state costs, liver biopsy disutility decrement and changes to NILT test costs; however, they were sensitive to some of the analyses. These are detailed as follows.

Removing tests from the analysis (where the studies did not converge using the bivariate model) changed the result so that the most cost-effective option, when all tests were compared singly, was to use APRI with a high cut-off; however, with an ICER of £20,673, if using a strict cost-effectiveness threshold of £20,000 this would not be adopted. If the cost-effectiveness threshold was set at £30,000, Fibroscan would be the most cost-effective test with an ICER of £23,345. The 'treat all' approach had the highest effectiveness (similar to base case) but with an ICER of £39,747 was not cost-effective at the standard UK threshold range of £20,000–30,000.

Setting treatment benefit to zero for patients who were incorrectly treated (patients in a mild health state who were diagnosed as false positive) significantly increased the ICER for 'treat all' to £550,668. The ICER for MR elastography increased to £32,220, which given a strict cost-effectiveness threshold of £20,000–30,000, would not be cost-effective. The most cost-effective option to adopt, assuming that the cost-effectiveness threshold is £20,000, is to use an indirect serum marker, GUCI, which has an ICER value of £19,934.

Changing the prevalence to the maximum prevalence estimated from the meta-analysis changed the result when all tests were compared singly; FIB-4 with a high cut-off became the most cost-effective test, with an ICER of £17,871. When the 75th quartile was used, the results changed so that treating everyone with a prior diagnostic test was the most cost-effective strategy if the cost-effectiveness threshold was at the upper bound of the recommended range (threshold £30,000, ICER £26,718).

Hepatits B e antigen-negative

The base results remained robust to the majority of sensitivity analyses; however, they were sensitive to some of the analyses which are detailed below.

Amending the sex ratio and starting age to reflect the HBeAg-negative model affected the results. Assuming a cost-effectiveness threshold of £30,000, the most cost-effective option would be to use MR elastography singly with an ICER of £25,546. However, if a strict cost-effectiveness threshold of £20,000 was employed, none of the options would be cost-effective.

Amending the average disease prevalence value used within the model changes the result; using a maximum prevalence returns a result where, given a cost-effectiveness threshold of £30,000, using MR elastography singly becomes cost-effective, with an ICER of £21,489. When the minimum disease prevalence is used, 'treat all' remains the most cost-effective given a threshold value of £30,000, but the ICER value decreases to £22,871.

Allowing for zero treatment benefit for patients who are treated incorrectly while in a mild health state (false-positive patients) changes the overall result, where MR elastography becomes the most cost-effective; however, the ICER of £32,194 is above the standard cost-effectiveness threshold value of £20,000–30,000. 'Treat all' also becomes cost-ineffective as a strategy, with an ICER of £53,660.

The base-case analysis assumes that the retest has perfect sensitivity and specificity (for modelling feasibility due to large number of tests and to ensure comparability across tests). When we amended this and set the retest sensitivity to that of three NILTs, APRI, Fibrotest and Fibroscan, the results did change for HBeAg-negative. Overall, the costs and QALYs for some of the NILTs increased, as the retest meant that some persons in mild health state (F0–1) started to receive treatment incorrectly (false-positive diagnosis), thereby incurring the associated costs and benefit (increased QALY outcome). 'Treat all' remains the most effective strategy, but is no longer the most cost-effective at standard cost-effectiveness threshold ranges. The most cost-effective test is now MR elastography, with an ICER in the upper band of the threshold range (£). MR elastography has a sensitivity of 94% of and specificity of 92%, implying that it will identify most patient correctly. This test is promising; however, as mentioned in *Chapter 1*, it is of limited use at present due to its cost and availability. Further studies to assess its diagnostic accuracy are required, as this was not one of the studies where the bivariate model converged and so the results may not be robust.

Short tables of results for the sensitivity analyses for the HBeAg-positive and HBeAg-negative models (excluded are 'dominated' and 'extendedly dominated' test strategies) are located in *Appendix 11* (full tables are available on request).

Discussion

We have estimated the cost-effectiveness of 56 testing strategies of sequential testing using 25 NILTs in patients with chronic HBV. The results differed for the HBeAg-positive and HBeAg-negative populations.

The analysis found that for HBeAg-positive patients, adopting a sequential testing strategy where patients who tested positive using hyaluronic acid were retested with a second NILT, MR elastography, to confirm results was the most cost-effective approach given a cost-effectiveness threshold of £20,000. However, though hyaluronic acid combined with MR elastography is the most cost-effective test strategy, there is considerable uncertainty around this result and the probability of it being the optimal choice (having the highest expected net benefit) is very low (4%). If the upper bound of the standard cost-effectiveness range, £30,000, was considered to be appropriate, the most cost-effective strategy would be with MR elastography singly, but again this result was associated with substantial uncertainty (13% probability of being the optimal choice). Within the £20,000–30,000 threshold range, the difference between the expected QALYs and costs was similar for many of the test combinations, and taking into account the uncertainty around the input parameters results in high levels of uncertainty.

When NILTs were assessed singly, testing with an indirect serum marker, GUCI, was the most cost-effective option at a cost-effectiveness threshold of £20,000 with a mean ICER of £19,716, but again associated with considerable uncertainty. For a threshold of £30,000, MR elastography was most cost-effective, with an ICER of £23,486.

For patients with HBeAg-negative disease, the strategy of treating all patients without testing and regardless of the degree of fibrosis offered the largest QALY gain. This is, however, at a substantial additional cost and would be cost-effective only if considered towards the upper bound of the NICE cost-effectiveness threshold range,[66] as it had an ICER of £28,317. All of the NILTs, either alone or sequentially followed by treatment, provided a QALY gain compared with a strategy of no testing and no treatment, but the gain was less than for the 'treat all' strategy, and they were dominated or extendedly dominated by the 'treat all (no prior test)' treatment strategy as their costs or ICERs were higher. There was less uncertainty in these results at the £30,000 threshold value where the probability of the 'treat all' strategy being optimum was 38%.

Similar findings for treatment in HBeAg-negative patients have been reported in assessments of the treatments conducted to inform national guidelines; the evidence review group which looked at the evidence for entecavir conducted an analysis where they analysed lifetime treatment duration for HBeAg-negative patients. This returned an ICER of £27,124 per QALY gained, similar to our base-case analysis result.[417]

This difference in results between both models highlights that treatment is more cost-effective in patients with positive disease as the HBeAg-negative population tend to be older with a higher risk of progressing from a moderate health state to a cirrhotic health state than the HBeAg-positive cohort. The impact of treatment is also more modest in this group than for the HBeAg-negative population modelled.

We conducted a sensitivity analysis to confirm that age and sex ratio drive the difference. The sensitivity analysis results showed that when all tests were compared singly, using the indirect serum marker, GUCI, had an ICER of £23,065, similar to the base-case result for the HBeAg-positive model. As the age and sex ratio had been changed, this implies that the RRs (treatment effect) drives the remaining difference between both cohorts and, indeed, is the driver behind a NILT not being cost-effective for the HBeAg-negative cohort given a strict cost-effectiveness threshold of £20,000.

Another thing to note about the results for both models is that the differences between test outcomes are very small, especially in relation to the health outcomes (QALYs), which implies that the long-term costs resulting from a particular test diagnosis (true positive, false positive, false negative or true negative) are the driving factor behind the analysis for HBeAg-positive and -negative chronic HBV.

An issue also arises regarding using the systematic literature review data to calculate the disease prevalence of liver fibrosis within a population, as most studies may have been set within a tertiary care centre rather than within a screening programme for the general population, and so disease prevalence may be overestimated. When we used the minimum prevalence for the HBeAg-positive model, this did affect the results; GUCI remained the most cost-effective test when all NILTs were compared singly, but the ICER for MR elastography increased significantly to £37,348.

Data on diagnostic accuracy of NILTs in patients with chronic HBV were limited and of low quality; therefore, our results should be interpreted with caution. Data on hyaluronic acid came from a single study, while MR elastography is not widely available and needs further validation as a diagnostic test.

Although Fibroscan is widely used in most centres, it was not the most cost-effective option, even in the sensitivity analysis when only studies that converged using the bivariate model were considered (given a strict £20,000 cost-effectiveness threshold). Our results are, therefore, indicative that sufficiently validated non-invasive testing strategies are more cost-effective in patients with HBeAg-positive chronic HBV than liver biopsy or treatment decisions based on viral load and transaminases alone, but no robust specific recommendations on the use of specific non-invasive tests can be made. Future research should further explore these possibilities.

Chapter 6 Cost-effectiveness analysis: hepatitis C

This chapter details the analysis of the cost-effectiveness of the diagnostic tests for staging fibrosis in patients with HCV. The NILTs considered for evaluation in this chapter were those considered applicable for use in people with HCV, for which data on sensitivity and specificity were available.

Evaluation approach for hepatitis C

Fifty-seven relevant NILTs were evaluated in the first stage of the analysis, which compared the NILTs with liver biopsy, 'treat all' and 'no treatment' strategies. The NILTs evaluated are listed in *Table 31*, and are grouped according to defined test categories: indirect serum makers, direct and patented serum markers and imaging modalities.

TABLE 31 List of NILTs evaluated (HCV)

Indirect	Direct and patented	Imaging
APRI	ELF	ARFI
Age-Platelet Index	ELF (high cut-off)	CT
APRI (high cut-off)	ELF (low cut-off)	EOB-MRI
APRI (low cut-off)	Fibroindex (high cut-off)	MRI
AST–ALT ratio	Fibroindex (low cut-off)	PLT–Spleen ratio
CDS	Fibrometer	Fibroscan (TE)
Fibrosis Index	Fibrospect	US
FIB-4	Fibrotest	US SAPI
FIB-4 (high cut-off)	Fibrotest (high cut-off)	US SAPI (high cut-off)
FIB-4 (low cut-off)	Fibrotest (low cut-off)	US SAPI (low cut-off)
FibroQ	Hyaluronic acid	CEUS
Forns index	Hyaluronic acid (high cut-off)	DW-MRI
Forns index (high cut-off)	Hyaluronic acid (low cut-off)	MR elastography
Forns index (low cut-off)	Hepascore	
FPI (high cut-off)	Hepascore (high cut-off)	
FPI (low cut-off)	MP3	
GUCI	PIINP/MMP-1 index	
King's	PIINP	
King's (high cut-off)	PLT	
King's (low cut-off)	Type IV collagen	
Lok's index	YKL-40 (high cut-off)	
Pohl index	YKL-40 (low cut-off)	

CDS, Cirrhosis Discriminant Score; CEUS, contrast-enhanced ultrasound; DW-MRI, diffusion-weighted magnetic resonance imaging; EOB-MRI, (gadolinium-ethoxybenzyl-diethylenetriamine-penta-acetic-acid) enhanced magnetic resonance imaging; FPI, Fibrosis Probability Index; MMP-1, matrix metalloproteinase-1; MP3, metalloproteinase-3; PLT, platelet; US, ultrasound.

The second stage of the analysis compared a selection of tests using alternative sequential testing strategies. The criteria for selecting these tests and the assumptions used regarding combinations of tests and sequential testing strategies are detailed in *Chapter 3*. The systematic review and subsequent meta-analysis of the data provided test outcome results for a number of published algorithms used for staging fibrosis [sequential algoritm for fibrosis evaluation (SAFE), Leroy, Fibropaca and Bordeaux], which were evaluated in the second stage of the analysis.

Ten tests evaluated in the second stage of the analysis used a combined diagnostic cut-off threshold for staging fibrosis; the test outcomes reported a number of indeterminate responses, which are listed in *Table 32*. We allowed for patients who had an indeterminate response to receive a retest with a commonly used imaging modality Fibroscan (TE). We did not choose an indirect test as the majority of the 10 tests with a combined diagnostic cut-off were from the indirect test category and a subsequent indirect test would not enhance the diagnostic accuracy. Of the direct tests and imaging modalities, we chose Fibroscan based on availability and current clinical practice. Overall, 56 testing strategies were compared in the second stage of the analysis.

Model structure and parameters

Decision tree structure

A decision tree model was constructed to evaluate the cost-effectiveness of the NILTs, liver biopsy, the 'treat all' and 'no treatment' strategies. As per the schematic diagram depicting the flow of data in *Chapter 3* (see *Figure 1*), the decision tree was populated with sensitivity, specificity and prevalence data from the meta-analysis of the systematic review data, cost and QALY outputs from a series of Markov models, individual test costs sourced from published literature and hospital finance departments and a measure of adverse effects associated with liver biopsy.

As discussed previously, there are two stages to the decision tree analysis: the first stage where all tests are compared singly and the second stage where combinations of tests are compared using four different strategies. Schematic illustrations and descriptions of the sequential testing pathways are provided in *Chapter 3*. To estimate costs and QALYs for each testing option, the long-term costs (and test costs) and QALY estimates (including disutility from liver biopsy if applicable) associated with each potential test outcome (true positive, false positive, true negative and false negative) were multiplied by the probability of each NILT returning a true positive, false positive, true negative or false negative result.

TABLE 32 List of tests with a combined cut-off applicable for use in HCV

Combined diagnostic threshold test	Persons with indeterminate result, %
APRI	43
ELF	41
FIB-4	29
Fibroindex	35
Fibrospect	53
Forns index	45
Fibrotest	38
Hyaluronic acid	44
Hepascore	33
YKL-40	47

Markov model structure

A series of Markov models were constructed to estimate the long-term costs and outcomes associated with a correct (true positive and true negative) and an incorrect (false positive and false negative) diagnosis for a hypothetical cohort of 1000 HCV patients with suspected liver fibrosis or cirrhosis. An additional two Markov models were constructed to estimate the costs and outcomes associated with the 'treat all and 'treat no one' approaches.

Figure 6 shows a schematic representation of the patient pathway. The structure is a modified version of previously published models of liver fibrosis in HCV.[414,438] The starting health states in the Markov models were representative of the METAVIR categorisations for staging fibrosis. There were eight potential health states in the model: mild fibrosis (F0–1), moderate/significant fibrosis (F2–3), compensated cirrhosis (F4), decompensated cirrhosis, HCC, liver transplant, post liver transplant and death. The patient cohort progressed through the model as per the arrows in the illustration; the circular arrows leading back into the same health state indicate that patients could remain within health states for longer than one cycle, except in the liver transplant state where patients could only progress to a post liver transplant health state or death. Patients could not regress back to an earlier health state. Patients could progress from all health states to the 'death' health state at any time.

On the basis of clinical advice, we assumed that a METAVIR test score of $\geq$ F2 equated to a positive test outcome (true positive and false positive) and treatment with antiviral agents would commence at this stage; conversely, a METAVIR test score of $<$ F2 indicated a negative test outcome and a policy of watchful waiting without immediate antiviral treatment would commence (see *Chapter 5* for description of the watchful waiting policy assumed in the model).

As for the HBV Markov models, the initial starting health state for the population cohort in the models depended on the test outcome being modelled [e.g. for a false positive test result, the population cohort entered the model in a mild health state (F0–1), and for a test result of true positive the population cohort entered the model in a F2–4 health state, where the patient cohort was then distributed according to the estimated disease prevalence (F2–3 62%; F4 38%)].

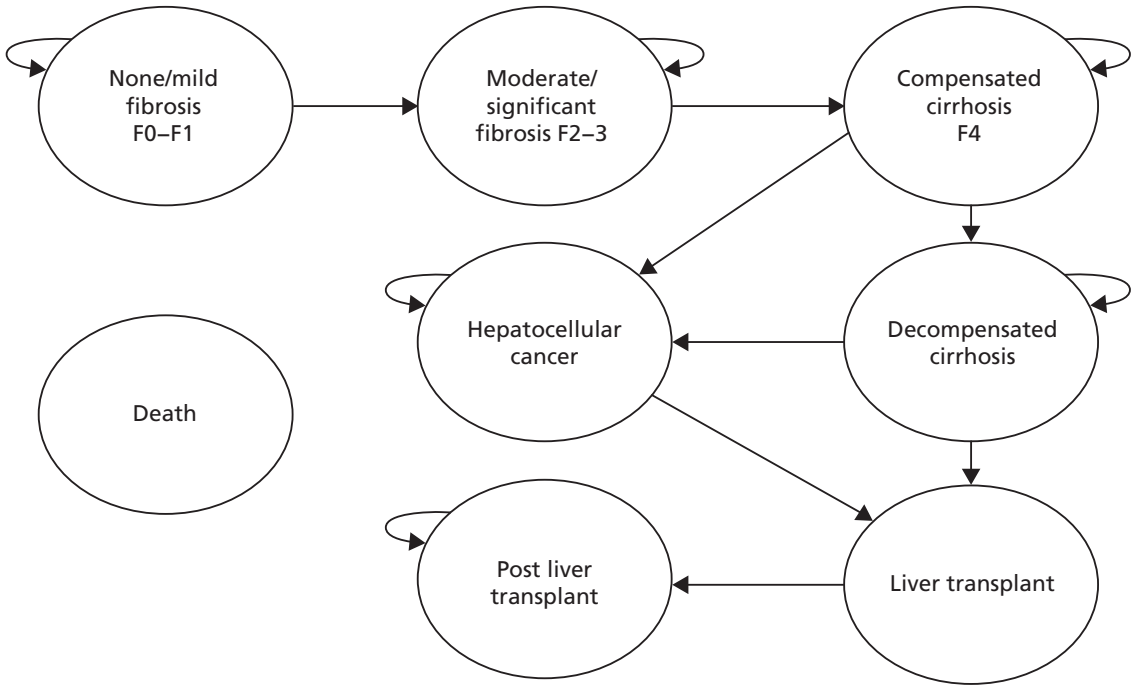

FIGURE 6 Illustration of Markov model.

Input parameters

Average disease prevalence was based on the systematic review of test accuracy (53%) using the same methods used for the HBV model (see *Chapter 5*). We used the sensitivity and specificity estimates for each NILT and the average prevalence estimate to calculate the probability of each test returning a true positive, false positive, true negative and false negative result (see *Appendix 7*).

Cohort characteristics

The characteristics of the cohort (age, sex and weight) were sourced from published literature.[414,439] Various studies have determined that treatment will have different effects according to patient genotype;[37,439] therefore, our model cohort is split to mirror the distribution of chronic HCV genotype 1, 2 and 3, and 4 infections,[439] allowing us to model the effectiveness of treatment per genotype. These are listed in *Table 33*.

Natural history: baseline transition probabilities used in model

The recent systematic review by Shepherd *et al.*[371] was reviewed and the epidemiological search strategy (see *Chapter 5* and *Appendix 2* for details) updated for HCV and rerun using the MEDLINE database (Ovid platform, searched 1 December 2012) for relevant papers related to disease progression in HCV. The search located 2343 papers, none of which was deemed relevant for our purposes. The rate of disease progression assumed in the models was based on transition probabilities sourced from a published cost-effectiveness study by Wright *et al.*,[414] which was identified in the literature search for quality of life and cost data. This study estimated transition probabilities for the mild and moderate HCV health states from a trial of treatments for mild HCV undertaken in a number of London hospitals,[414] during which the liver fibrosis stage was determined by liver biopsy. We deemed this a suitable study to use as it was a UK-based study, and noted that it does not assume that the progression between early health stages is linear.[414]

Transition probabilities for the compensated cirrhosis, decompensated cirrhosis, HCC, liver transplant and post liver transplant health states were also sourced from this study. For these health states, the authors sourced the probabilities from published literature. Data from a study by Fattovich *et al.*[409] were used for the compensated cirrhosis, decompensated cirrhosis and HCC health states, and for mortality-specific data relating to each health state. The study by Fattovich *et al.*[409] was conducted in seven European centres, including a UK tertiary centre. The study was a retrospective study looking at a cohort of 384 compensated cirrhotic patients who were assessed annually for a mean of 5 years.

Wright *et al.*[414] sourced transition probabilities relating to the liver transplant and post liver transplant health state from a study by Siebert *et al.*,[435] where the rate of liver transplants was estimated for HCV patients in the USA and revised downwards by 2% based on European transplant registry data and estimated the probability of death following a liver transplant from a survival analysis of the UK liver transplant registry data conducted by the Royal College of Surgeons.[440] Transition probabilities used in the model are listed in *Table 34*.

TABLE 33 Cohort characteristics input parameters

Cohort characteristics	Parameter value	Source
Age (years)	40	Wright *et al.*[414] Fried *et al.*[439]
Average weight	79.8 kg	
Sex		
Male, %	61	Wright *et al.*[414]
Female, %	39	
Genotype, %		
Genotype 1	66	Fried *et al.*[439]
Genotypes 2 and 3	31	
Genotype 4	3	

TABLE 34 Annual transition probabilities: HCV

Health state–health state	Transition probability	PSA distribution	Source
Mild–moderate	0.025	Dirichlet	Wright *et al.*[414]
Moderate–cirrhosis	0.037		
Cirrhosis–decompensated cirrhosis	0.04		
Cirrhosis/decompensated cirrhosis–HCC	0.14		
Decompensated cirrhosis/HCC–liver transplant	0.02		
Decompensated cirrhosis–death	0.13		
HCC–death	0.43		
Liver transplant–death (year 1)	0.15		
Post liver transplant–death	0.03		
All-cause mortality	0.0014–0.335		Interim Life Tables, 2008–10[416]

Mortality data

An all-cause mortality rate applied to all patients within the model; this was calculated using the Interim Life Tables for England and Wales, 2008–10.[416] The risk of death increased each year according to age and the rate was weighted to allow for sex mix.

Antiviral treatment: type, dosage, duration and effectiveness

In the model, patients received treatment with antiviral agents if they tested positive (true positive or false positive with a test score of ≥ F2).

If patients had a successful response to treatment (modelled using a SVR rate), they no longer progressed through the health states pathway in the model and retained a risk of all-cause mortality only. This is similar to the assumption employed by Wright *et al.*[414]

Antiviral treatment for the management of liver fibrosis and cirrhosis in patients with HCV was based on NICE guidance.[441–443] Treatment type, duration and dosage in the model varied according to HCV genotype. In accordance with NICE guidelines, our model assumed that HCV genotype 1 patients received treatment with a combination of peginterferon alfa-2a or alfa-2b, ribavirin, and either telaprevir or boceprevir.[441,442] Patients with HCV genotype 2, 3 and 4 were assumed to receive dual therapy with a combination of peginterferon alfa-2a or alfa-2b and ribavirin.[443]

Different branded products are marketed and available in the UK for ribavirin (BNF 64[419]); the summary of product characteristics for one of the brands [Rebetol®, Merck Sharpe & Dohme (MSD)] advises that it should only be used in combination with peginterferon alfa-2b (ViraferonPeg®, MSD).[444] To allow for this, we modelled two different combinations of antiviral agents in the model (combined as per their summary of product characteristics) with half of eligible patients receiving one combination and the other half of eligible patients receiving the other.

- HCV genotype 1: peginterferon alfa-2a (Pegasys®, Roche) combined with ribavirin (Copegus®, Roche) and telaprevir (Incivo®, Janssen) or peginterferon alfa-2b (ViraferonPeg®, MSD) combined with ribavirin (Rebetol®, MSD) and boceprevir (Victrelis®, MSD)
- HCV genotype 2, 3 and 4: peginterferon alfa-2a (Pegasys®, Roche) combined with ribavirin (Copegus®, Roche) or peginterferon alfa-2b (ViraferonPeg®, MSD) combined with ribavirin (Rebetol®, MSD).

Patients in the mild, moderate and compensated cirrhosis health states would receive treatment with antiviral agents if they tested positive (true positive or false positive). Patients in the decompensated cirrhosis, HCC, liver transplant and post liver transplant health states were assumed to receive no antiviral treatment, but instead received usual standard of care.[37]

Treatment dosage and treatment duration was based on those recommended in the summary of product characteristics for each drug (www.medicines.org.uk/emc); drug dosage depended on the weight of the patient; for modelling purposes we assumed that the drug dosage would be administered as per the assumed average weight in the model (79.8 kg).[439] Treatment effectiveness, represented in the model by the SVR rate, was sourced from published literature.[445] Patients within the model were assumed to be treatment naive (had not received previous treatment with antiviral agents). Different SVR rates were employed in the model and varied according to genotype and drug therapy administered. We assumed that treatment benefit would occur during the treatment year.

Different SVR rates for the mild, moderate and cirrhotic health states were not employed within the model as separate rates were not reported within the source literature,[441,445] except for HCV genotype 1 patients treated with peginterferon alfa-2b, ribavirin and boceprevir,[442] where a different SVR rate was reported for HCV patients in a cirrhotic health state. As it has been noted that the SVR rates with pegylated IFN-α and ribavirin are lower in patients with advanced fibrosis and cirrhosis than in patients with mild or moderate fibrosis,[37] we carried out a sensitivity analysis where we reduced the SVR rate for the patient cohort receiving treatment in a cirrhotic health state. The SVR rates, recommended dosage and specific duration employed in the model as part of dual or triple therapy are listed in *Table 35*.

TABLE 35 Medication used in Markov models

Drug combination	Dosage[a]	Treatment duration	SVR rate applied	Source of SVR rate
Genotype 1: treatment-naive patients				
Peginterferon alfa-2a, ribavirin, telaprevir	180 mg, 1200 mg and 2250 mg	PR and Telaprevir TW 12, PR TW 48	75%	NICE technology appraisal TA252[441]
Peginterferon alfa-2b, ribavirin, boceprevir	120 mg, 1000 mg and 2400 mg	PR for 4 weeks, PR and boceprevir TW 36, PR TW 48	66.1%	NICE technology appraisal TA253[442]
Genotype 1: cirrhotic patients (treatment naive)				
Peginterferon alfa-2b, ribavirin and boceprevir	120 mg, 1000 mg and 2400 mg	PR for 4 weeks, PR and boceprevir TW 48	41.7%	NICE technology appraisal TA253[442]
Genotype 2 and 3: treatment-naive patients				
Peginterferon alfa-2a, ribavirin	180 mg, 1200 mg	PR for 24 weeks	76%	Fried *et al.*[439]
Peginterferon alfa-2b, ribavirin	1200 mg, 1000 mg	PR for 24 weeks	82%	Manns *et al.*[445]
Genotype 4: treatment-naive patients				
Peginterferon alfa-2a, ribavirin	180 mg, 1200 mg	PR for 48 weeks	77%	Fried *et al.*[439]
Peginterferon alfa-2b, ribavirin	120 mg, 1000 mg	PR for 48 weeks	69%	Kamal *et al.*[446]

PR, peginterferon alfa-2a or alfa-2b and ribavirin; TW, till week.
a Recommended dosage listed: weekly for peginterferon alfa-2a and alfa-2b; daily for ribavirin, telaprevir and boceprevir.

Costs

Health states

We updated the search strategy for costs devised by Shepherd *et al.*[371] for HCV (see *Appendix 2* for search strategy). Using the MEDLINE database (via Ovid platform, searched 1 December 2012), the search located 21 papers, eight of which were retrieved for full review;[447–454] however, six of these papers were based outside the UK,[447–449,453,454] one contained no costs[451] and two were guidelines papers.[450,452] As none of the papers found was applicable, the costs of treating patients with mild, moderate and compensated cirrhosis were sourced from a published cost-effectiveness study by Wright *et al.*[414] identified from the literature review for quality-of-life data for HCV. This study looked at the effectiveness of antiviral agents for mild HCV, costs for which were collected during a RCT. The cost data from this study have been widely used in other recently published papers.[455–457]

The authors collected resource use and cost data alongside a mild HCV RCT. Detailed cost data were collected from three centres based in London, Newcastle and Southampton. Resource use information collected covered inpatient and outpatient care, investigations, procedures, drug use and other services including psychiatric services.

The costs associated with treating the decompensated cirrhosis, HCC, liver transplant and post liver transplant health states were sourced from a UK cost-effectiveness study of liver transplantation[424] (CELT study) which collected costs on adult patients listed for an isolated liver transplant (aged 16 years and over) between December 1995 and December 1996. Data collection was conducted between 1995 and 1999 and split into four phases: assessment, candidacy, transplant and post-transplant phases (see *Chapter 5* for description) and collected according to disease aetiology (HBV, HCV, ALD). Resource-use data on blood products used, number of dietitian sessions, drugs used, inpatient stay, nutritional support received, outpatient visits, physiotherapy sessions, tests, length of transplant operation and key treatments and investigations were collected.

We calculated the cost for HCV patients who had received a transplant (sample size 67).
The average length of time for a liver transplant procedure (from admission to discharge) was 28 days. We approximated this to 1 month and calculated the yearly cost of a liver transplant as 11 months of post-transplant care (estimated from the average monthly cost in the first year following transplantation) plus the month in which the transplant operation took place. We also calculated a cost for the post liver transplant health state, estimated from the average monthly cost for the last 12 months of post-transplant care (sample size 40).

The data were also used to estimate a cost for the decompensated cirrhosis and HCC health states. For this, the average costs of treating patients with decompensated cirrhosis were assumed to be equivalent to patients considered for a transplant. Resource-use data of patients with HCV during the 'assessment' and 'candidacy' stages of transplantation were calculated (sample size 56).

Costs were inflated to 2011–12 levels using standard NHS inflation indices.[67] Health state costs did not include the costs associated with antiviral treatment; these were included separately in the model. Health state costs are listed in *Table 36*.

Test costs

Costs of imaging modalities were sourced from published Department of Health reference costs.[427] Costs of direct and indirect serum markers were obtained from communication with finance departments based at the Royal Free Hospital. Costs for patented serum markers were sourced directly from manufacturers and via communication with finance departments based at the Royal Free Hospital. The cost of a percutaneous liver biopsy (see *Chapter 5* for further details on choice of liver biopsy) was sourced from published literature[428] (see *Chapter 5*). Test costs including sources are listed in *Appendix 9*.

TABLE 36 Health state costs used in the model (£ 2012)

Health state	Annual cost, £	Standard error	PSA distribution	Source
Mild fibrosis	185	36.39	Gamma	Wright *et al.*[414]
Moderate fibrosis	959	101.69		
Compensated cirrhosis	1521	309.05		
Decompensated cirrhosis	38,871	9410.46		Longworth *et al.*[424]
Hepatocellular cancer	38,871	9410.46		
Liver transplant	69,174	7054.86		
Post liver transplant	4356	861.57		
Death	0	0		Assumed

Medication costs

For HCV genotype 1 and 4 infection, a cost of 48 weeks of treatment was applied. For HCV genotype 2 and 3 infection, a cost approximate to 24 weeks of treatment was applied (*Table 37*). Treatment costs were sourced using the BNF 64[419] and a total cost of treatment was calculated according to the recommended dosage and duration detailed in *Table 35*.

Utility values

A search was carried out using the MEDLINE database (via Ovid platform, searched 1 December 2012) for quality of life data for use in the HCV model (see *Appendix 2* for search strategy). The search returned 459 papers; seven were retrieved for full review.[371,414,432,455,457–459] From these, the most relevant data identified were found in the published study on the Mild Hepatitis C trial[414] as it was from a UK population using the EQ-5D. As this study was also used to identify data for the mild, moderate and cirrhotic health states in the HBV model, details regarding the elicitation of utility values for the mild, moderate and compensated cirrhosis health states are reported in the chapter for HBV (see *Chapter 5*) and will not be duplicated here.

As for the HBV model, we sourced health-related utility data for the decompensated cirrhosis, HCC, liver transplant and post liver transplant health states from the CELT study.[424] The transplantation study included patients with a range of conditions that warranted liver transplantation including HCV. HRQoL data were collected using the EQ-5D.

It was assumed that utility values for the decompensated cirrhosis and HCC health states would be the same (sample size 56). This was assumed to be equivalent to the average utility value at the time patients were placed on the waiting list for a transplant.

TABLE 37 Total annual cost of combination and triple therapy used in model (£ 2012)

Treatment	Genotype 1	Genotype 2 and 3	Genotype 4
Combination therapy			
Peginterferon alfa-2a and ribavirin		4446	10,411
Peginterferon alfa-2b and ribavirin		5435	10,870
Triple therapy			
Peginterferon alfa-2a and ribavirin and telaprevir	32,809		
Peginterferon alfa-2b and ribavirin and boceprevir (treatment naive)	33,270		
Peginterferon alfa-2b and ribavirin and boceprevir (compensated cirrhosis)	41,670		

A utility value for the liver transplant health state was estimated using an area under the curve approach and the utility values collected at 3, 6 and 12 months post transplantation (sample size 67). A utility value for the post-transplant health state was estimated using the average utility for patients with HCV patients collected at 24 months post transplant (sample size 40).

Utility values during treatment and after sustained virological response

Treatment with peginterferon alfa-2a or -2b and ribavirin is associated with a number of adverse effects such as severe fatigue, depression, irritability, sleeping disorders, skin reactions, dyspnoea, neutropenia, anaemia, thrombocytopenia and ALT flares and more severe side effects such as seizures, bacterial infections, autoimmune reactions, interstitial lung disease, a neuroentinitis, bone marrow aplasia or idiopathic thrombocytopenia.[37] To reflect this in the model, we assumed a disutility value that was applicable during treatment using data identified from the Wright *et al.*[414] study. The authors collected HRQoL data using the EQ-5D (sample number: 144 patients), using the data from weeks 12 and 24 when most people were still taking treatment.

It was conservatively assumed that any adverse events associated with boceprevir or telaprevir treatment (skin rash and exacerbation of anaemia) would have no effect on HRQoL in the base-case analysis. We tested this assumption in a sensitivity analysis by applying a utility decrement value of 0.05 applicable during treatment with boceprevir and telaprevir for HCV genotype 1 patients.

This study identified a number of previously published studies[460–462] which reported that the HRQoL score for patients significantly improved after successful treatment with interferon and, therefore, assumed that therapy for mild HCV may benefit patients for non-hepatological reasons. They measured the improvement in HRQoL after a successful response to treatment by collecting data using the EQ-5D from 21 patients with mild fibrosis who had a SVR. As there were insufficient data to estimate an EQ-5D value for patients in a post-SVR moderate health state, the authors estimated a post-SVR EQ-5D value by substituting the mean estimated HRQoL for moderate disease as the baseline value into the analysis of covariance (ANCOVA) model which had been used to estimate the treatment effect for patients with mild disease. We used the same assumption as Wright *et al.*[414] and allowed for an increased EQ-5D value post SVR.

The study by Wright *et al.*[414] did not report a separate HRQoL value for patients in a cirrhotic health state receiving treatment and post SVR. For this, we used the same assumption as that used in a study by Grishchenko *et al.*,[455] who assumed that patients with cirrhosis had the same absolute gain in HRQoL as patients with mild disease.

Using the identified data, we applied a 0.11 utility decrement from baseline during treatment and a 0.05 utility increment from baseline if they had a SVR after treatment.

We tested the use of differing HRQoL values applicable during antiviral treatment and post SVR by carrying out a sensitivity analysis where we assumed that HRQoL values did not change from baseline during or after treatment. Utility values are listed in *Table 38*.

Disutility from non-invasive liver test and liver biopsy

We have assumed that no NILT has associated adverse events that would impact on HRQoL. However, as liver biopsy is associated with morbidity and mortality risks and patient discomfort, a measure of adverse effects resulting from liver biopsy was modelled by applying a utility decrement of 0.2 where applicable.[428] As the decrement value was arbitrary, we undertook a number of sensitivity analyses to test the robustness of the results to changes in the utility decrement.

Estimated long-term costs and quality-adjusted life-years

Table 39 presents the costs and QALYs from the Markov models for each potential diagnostic testing outcome.

TABLE 38 Utility values used in analysis of HCV

Health stage	Utility value	SE	PSA distribution	Source
Mild fibrosis	0.77	0.035	Beta	Wright *et al.*[414]
Moderate fibrosis	0.66	0.018		
Compensated cirrhosis	0.55	0.032		
Decompensated cirrhosis	0.49	0.056		Longworth *et al.*[421]
HCC	0.49	0.056		
Liver transplant	0.51	0.053		
Post liver transplant	0.52	0.061		
Death	0	0		Assumed
Mild fibrosis (during treatment)	0.65	0.035		Wright *et al.*[414]
Moderate fibrosis (during treatment)	0.55	0.018		
Compensated cirrhosis (during treatment)	0.44	0.04		Grishchenko *et al.*[455]
Mild fibrosis (SVR, after treatment)	0.82	0.04		Wright *et al.*[414]
Moderate fibrosis (SVR, after treatment)	0.71	0.05		
Compensated cirrhosis (SVR, after treatment)	0.60	0.04		Grishchenko *et al.*[455]

TABLE 39 Estimated long-term costs (£) and QALYs

Diagnostic test outcome	Cost (£ 2012 per person)	QALY (per person)
True positive	68,667	12.89
True negative	21,812	15.84
False positive	32,318	16.86
False negative	71,818	12.48
Treat all persons	51,374	14.77
Treat no one	55,173	12.47

Analysis

Probabilistic and one-way sensitivity analyses were undertaken. We also conducted threshold analyses around the assumptions of treatment benefit and the impact of changes to the costs and effectiveness of treatment.

Robust test accuracy data

In the base-case analysis, all tests were included despite there being limited data available for some tests. A sensitivity analysis was conducted including only tests where the bivariate model used for the meta-analysis converged (for 14 NILTs).

Changes to utility values

As mentioned previously, we carried out an analysis where we set utility values constant at baseline utility values before, during and after treatment.

We also carried out two other analyses regarding the utility values used; we set the utility values equivalent to those used in a study by Shepherd *et al.*[392] We also carried out an analysis where we increased the utility value of all health states by 0.1 to determine if this had any effect on the robustness of the results.

Average disease prevalence

Prevalence within the model is based on studies that may have been carried out largely in tertiary care centres, and the prevalence of liver fibrosis in this population may be an overestimate. To test the impact of this, we undertook sensitivity analyses using the minimum prevalence (17%) and the maximum prevalence (83%) estimated from the meta-analysis of the systematic review data.

Change to progression rates after sustained virological response while in cirrhotic health state

The base-case analysis assumes that the risk of death of patients who respond successfully to treatment (experience a SVR) is equivalent to that for the general population. We tested this assumption for patients in a compensated cirrhosis health state as it has been noted that they may retain a small risk for progressing to these health states.[37,463] We modelled this by allowing patients in the compensated cirrhosis health state to retain a small risk of decompensated cirrhosis (0.004) and HCC (0.002) after a successful response to treatment (SVR).[464]

Lower sustained virological response rate

As mentioned previously, it may be the case that patients in a cirrhotic health state have a lower SVR rate; to test this assumption, we carried out an analysis where we reduced the SVR rate by 20% (assuming the same estimate used in the study by Liu *et al.*[438]) for patients who received treatment in a cirrhotic health state.

Change in cost of non-invasive liver tests

We carried out a sensitivity analysis where we changed the cost of the NILTs (we set the watchful waiting retest cost and all NILT costs within the model to the same cost); for comparison, we assumed an indirect serum marker test cost (we chose a commonly used test, APRI, as our comparator). By changing the cost of a NILT (in some cases reducing the test cost significantly, e.g. reducing cost of ELF from £108 to £4.50) we aimed to determine if changing the test cost (marginal cost) had an impact on the robustness of the results.

Sensitivity and specificity of retest

Our base-case analysis assumes that the retest (from the meta-analysis of the systematic review data in the watchful waiting strategy for patients with a negative test result) has perfect sensitivity and specificity. We tested this assumption by applying the sensitivity and specificity of three commonly used tests: APRI (estimated sensitivity of 77% and specificity of 81%), Fibrotest (estimated sensitivity of 68% and specificity of 75%) and Fibroscan (estimated sensitivity of 79% and specificity of 83%).

Choice of tests for second stage of the analysis

For the second stage of the analysis, the two most cost-effective tests when assessed within each specific test category singly (with and without a defined threshold) were used in the analysis of sequential testing. To test if changing the method used to choose tests for the second stage of the analysis had an effect on the overall result, we carried out an analysis where we chose the most effective NILT from within each NILT category and the least costly NILT from each category.

Change in genotype distribution

We carried out an analysis where we amended the distribution of the population cohort per genotype in the model to determine if this had an impact on the results (HCV genotype 1 set at 50%, HCV genotype 2 and 3 set at 41% and HCV genotype 4 set at 9%).

Change to utility values

We also carried out analyses using different utility decrement values to represent the adverse effects from liver biopsy. The base-case analysis set the utility decrement value to 0.2; in the analyses we set the utility decrement at 0 and 0.3 to test the impact on the robustness of the results.

Adverse events

Both telaprevir and boceprevir carry a risk of adverse events. We tested the assumption that this would not have a significant impact on health-related utility by assuming a disutility decrement value of 0.05, which was applicable to all patients who received triple therapy with peginterferon alfa-2a or -2b, ribavirin and either telaprevir or boceprevir.

Threshold sensitivity analyses

Treatment benefit for patients who are incorrectly diagnosed

The base-case analysis reflects that HCV patients who are diagnosed incorrectly ('false positive') receive benefit from treatment despite only having mild disease. Treatment benefit in the model is reflected by the probability of a SVR and patients in a mild health state receive the same treatment benefit (same successful response rate measured using SVR rate) as patients who are in a moderate or cirrhotic health state. We tested the robustness of this assumption in a sensitivity analysis by undertaking a threshold sensitivity analysis where we reduced the SVR rate for patients in a mild health state (F0–1) by decrements of 10%.

Sensitivity analysis on drug costs and sustained virological response rates

We are aware that new drugs are in development for the treatment in people with HCV; for example, a new protease inhibitor has recently been investigated in phase 3 trials.[43,44] With this in mind, we conducted sensitivity analyses exploring the impact of the costs and effectiveness of treatments of fibrosis in HCV on the conclusions regarding the cost-effectiveness of the test strategies.

We increased the SVR rate for genotype 1 and 4 HCV infections to reflect the potential efficacy of new drugs suitable for treatment in HCV. Based on early results from two phase 3 studies,[43] we assumed an increased SVR rate of 90% for genotype 1 HCV infection and genotype 4 patients. No amendments were made to the SVR rates for genotypes 2 and 3 HCV infections from those in the base-case analysis based on the results of a published non-inferiority study of the same new treatment.[43] We then also increased the cost of drug treatment assuming an additional £20,000 and £40,000 cost per patient for 12 weeks of treatment with the new drug (this was added to the existing cost for peginterferon alfa-2a and alfa-2b and ribavirin used in the base-case analysis). We did not allow for different SVR rates per health state.

Results

Base case

At a standard UK threshold range, the cost-effective strategy is to adopt a 'treat all' approach with an ICER of £9204. For values below this, a patented serum maker, Fibrospect (combined cut-off), where indeterminate responses are retested with an imaging modality, Fibroscan, is the most cost-effective option.

The CEAF (*Figure 7*) shows that the probability of 'treat all' having the highest expected net benefit, given a cost-effectiveness threshold value of £20,000, is 45%. For lower threshold ranges, the CEAF also shows that there is considerable uncertainty around which test has the highest expected net benefit. For cost-effectiveness thresholds lower than £9200 [using Fibrospect (with a combined cut-off) where the inconclusive responses are retested with Fibroscan] is most likely to have the highest expected net benefit; however, there is considerable uncertainty around this result and the probability of it being optimal is $<4\%$. Liver biopsy was a comparator for both stages of the analysis; however, this testing option is dominated by other less costly but more effective options.

Appendix 10 displays the CEAC for the overall base-case analysis. For reasons of clarity, only those strategies that have a $\geq 5\%$ or greater probability of being optimal have been included. The CEAC demonstrates that, given a cost-effectiveness threshold value of £20,000, the probability that the 'treat all' strategy is cost-effective compared with the other testing strategies is 0.449. This indicates that, given the data, there is a 45% chance that the additional cost of the 'treat all' strategy, compared with all other test strategies, is at or below £20,000 per life-year gained. The CEAC also displays that liver biopsy has a high probability of being cost-effective for thresholds values below approximately £3500. This does not translate into the CEAF due to skewed data on the costs of the tests and a high level of uncertainty on differences in costs between the alternative test strategies. Often, the testing option with a high probability of being cost-effective may not be the optimal choice, which is what the CEAF represents (Fenwick *et al.*).[69]

Table 40 presents incremental results for the first stage of the analysis and *Table 41* presents incremental results for the second stage where a number of combined tests are compared using a number of different sequential testing strategies (see *Chapter 3* for details of testing strategies, S1, S2, S3 and S4). A scatterplot illustrating the position of each testing strategy on the cost-effectiveness acceptability curve compared with the testing strategy 'treat no one' can be found in *Appendix 12*.

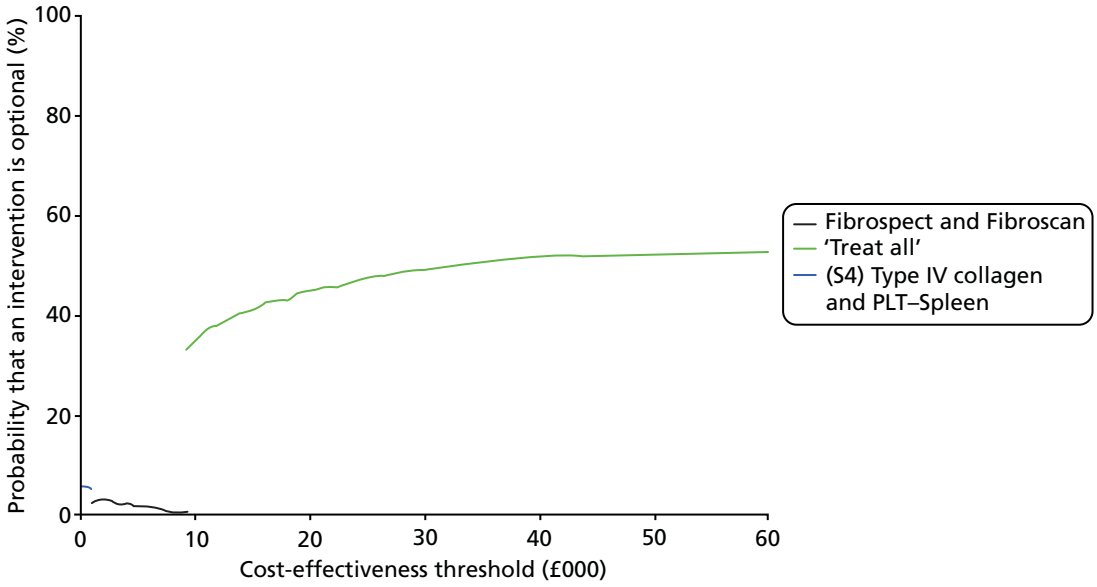

FIGURE 7 Cost-effectiveness acceptability frontier for the decision concerning most the cost-effective testing strategy to employ for HCV. PLT, platelet.

TABLE 40 Base-case analysis (first stage where all tests are compared singly)

Test strategy	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat none	54,878	12.45	–	–	Dominated
Liver biopsy	48,710	14.03	–	–	Dominated
Pohl	47,727	14.04	–	–	Dominated
Fibroindex (high cut-off)	47,769	14.08	–	–	Dominated
Forns Index (high cut-off)	47,426	14.12	–	–	Dominated
Hepascore (high cut-off)	47,897	14.13	–	–	Dominated
FPI (high cut-off)	47,335	14.14	–	–	Dominated
APRI (high cut-off)	47,525	14.14	–	–	Dominated
FIB-4	47,900	14.15	–	–	Dominated
US	48,090	14.17	–	–	Dominated
ELF (high cut-off)	47,846	14.17	–	–	Dominated
Hyaluronic acid (high cut-off)	48,969	14.18	–	–	Dominated
PLT	47,742	14.18	–	–	Dominated
US SAPI (high cut-off)	47,073	14.18	–	–	Dominated
YKL-40 (high cut-off)	48,536	14.19	–	–	Dominated
Fibrotest (high cut-off)	47,896	14.22	–	–	Dominated
PIIINP/MMP-1 index	47,724	14.24	–	–	Dominated
King's (low cut-off)	47,743	14.24	–	–	Dominated
King's (high cut-off)	47,963	14.25	–	–	Dominated
Fibrosis Index	47,423	14.25	–	–	Dominated
ARFI	47,126	14.25	–	–	Dominated
GUCI	47,791	14.25	–	–	Dominated
AST–ALT	48,629	14.26	–	–	Dominated
Age–Platelet Index	47,847	14.26	–	–	Dominated
MR	47,101	14.26	–	–	Dominated
EOB-MRI	48,054	14.26	–	–	Dominated
MR elastography	46,896	14.27	–	–	–
FIB-4 (high cut-off)	48,158	14.27	–	–	Dominated
CEUS	47,215	14.28	–	–	Extendedly dominated
APRI	47,522	14.28	–	–	Extendedly dominated
Fibroscan	47,449	14.28	–	–	Extendedly dominated
US SAPI	47,763	14.29	–	–	Extendedly dominated
DW-MRI	47,890	14.30	–	–	Dominated
Fibrotest	48,327	14.30	–	–	Dominated

TABLE 40 Base-case analysis (first stage where all tests are compared singly) (*continued*)

Test strategy	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Hyaluronic acid	48,013	14.30	–	–	Dominated
PIINP	47,921	14.30	–	–	Dominated
Hepascore	48,189	14.31	–	–	Dominated
Fibrometer	48,104	14.32	–	–	Dominated
MP3	48,008	14.33	–	–	Dominated
Fibrospect	48,210	14.33	–	–	Dominated
Type IV collagen	47,888	14.34	–	–	Dominated
Hyaluronic acid low	48,824	14.34	–	–	Dominated
King's	47,990	14.34	–	–	Dominated
ELF	48,232	14.34	–	–	Dominated
CT	48,727	14.35	–	–	Dominated
FibroQ	48,372	14.35	–	–	Dominated
PLT–Spleen	47,803	14.35	–	–	Extendedly dominated
Forns index	50,555	14.37	–	–	Dominated
Lok's index	49,077	14.38	–	–	Dominated
APRI (low cut-off)	48,713	14.40	–	–	Extendedly dominated
CDS	49,429	14.40	–	–	Dominated
Fibroindex (low cut-off)	48,872	14.40	–	–	Extendedly dominated
ELF (low cut-off)	49,041	14.44	–	–	Extendedly dominated
FPI (low cut-off)	49,232	14.47	–	–	Extendedly dominated
FIB-4 (low cut-off)	49,407	14.48	–	–	Extendedly dominated
Forns index (low cut-off)	49,571	14.49	–	–	Dominated
Fibrotest (low cut-off)	49,534	14.49	–	–	Extendedly dominated
YKL-40 (low cut-off)	50,156	14.50	–	–	Dominated
US SAPI (low cut-off)	49,561	14.51	–	–	Extendedly dominated
Treat all	51,241	14.73	4,345	0.46	9,351

CDS, Cirrhosis Discriminant Score; CEUS, contrast-enhanced ultrasound; DW-MRI, diffusion-weighted magnetic resonance imaging; EOB-MRI, (gadolinium-ethoxybenzyl-diethylenetriamine-penta-acetic-acid) enhanced magnetic resonance imaging; FPI, Fibrosis Probability Index; MMP-1, matrix metalloproteinase-1; MP3, metalloproteinase-3; PLT, platelet; US, ultrasound.

TABLE 41 Base-case analysis (second stage of analysis)

Test strategy	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat none	54,878	12.45	–	–	Dominated
Liver biopsy	48,710	14.03	–	–	Dominated
(S3) Type IV collagen and PLT–Spleen	47,099	14.16	–	–	Dominated
(S3) King's and PLT–Spleen	47,139	14.16	–	–	Dominated
(S3) King's and type IV collagen	47,113	14.16	–	–	Dominated
Hepascore (combined cut-off) and Fibroscan	47,675	14.19	–	–	Dominated
(S3) FPI (low cut-off) and type IV collagen	47,417	14.21	–	–	Dominated
(S3) FPI (low cut-off) and PLT–Spleen	47,461	14.21	–	–	Dominated
(S2) King's and type IV collagen	47,001	14.21	–	–	Dominated
(S2) King's and PLT–Spleen	46,999	14.21	–	–	Dominated
Fibroindex (combined cut-off) and Fibroscan	47,368	14.21	–	–	Dominated
(S3) Fibrotest (low cut-off) and PLT–Spleen	47,548	14.21	–	–	Dominated
(S4) King's and type IV collagen	46,978	14.21	–	–	Dominated
(S4) King's and PLT–Spleen	46,965	14.22	–	–	Dominated
(S3) King's and Fibrotest (low cut-off)	47,579	14.22	–	–	Dominated
(S3) Type IV collagen and US SAPI (low cut-off)	47,516	14.22	–	–	Dominated
(S4) Type IV collagen and PLT–Spleen	46,911	14.22	–	–	–
(S2) Type IV collagen and PLT–Spleen	46,994	14.22	–	–	Dominated
(S3) King's and US SAPI (low cut-off)	47,606	14.22	–	–	Dominated
Leroy	47,248	14.23	–	–	Dominated
(S4) FPI (low cut-off) and type IV collagen	47,328	14.24	–	–	Dominated
(S2) FPI (low cut-off) and type IV collagen	47,328	14.24	–	–	Dominated
(S2) FPI (low cut-off) and PLT–Spleen	47,359	14.24	–	–	Dominated
(S4) FPI (low cut-off) and PLT–Spleen	47,347	14.24	–	–	Dominated
(S2) Fibrotest (low cut-off) and PLT–Spleen	47,519	14.24	–	–	Dominated
(S4) Fibrotest (low cut-off) and PLT–Spleen	47,446	14.24	–	–	Dominated
(S4) King's and Fibrotest (low cut-off)	47,521	14.25	–	–	Dominated
(S4) Type IV collagen and US SAPI (low cut-off)	47,427	14.25	–	–	Dominated
Forns index (combined cut-off) and Fibroscan	47,233	14.25	–	–	Dominated
(S4) King's and US SAPI (low cut-off)	47,525	14.25	–	–	Dominated

TABLE 41 Base-case analysis (second stage of analysis) (*continued*)

Test strategy	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
(S1) Type IV collagen	48,155	14.25	–	–	Dominated
(S1) PLT–Spleen	48,233	14.26	–	–	Dominated
(S1) King's	48,375	14.26	–	–	Dominated
Fibropaca	47,545	14.26	–	–	Dominated
APRI (combined cut-off) and Fibroscan	47,355	14.26	–	–	Dominated
(S2) King's and Fibrotest (low cut-off)	47,467	14.26	–	–	Dominated
Fibrospect (combined cut-off) and Fibroscan	46,954	14.27	43.55	0.05	928
(S2) King's and US SAPI (low cut-off)	47,466	14.27	–	–	Dominated
Bordeaux	47,026	14.27	–	–	Extendedly dominated
ELF (combined cut-off) and Fibroscan	47,533	14.28	–	–	Dominated
(S2) Type IV collagen and US SAPI (low cut-off)	47,411	14.28	–	–	Extendedly dominated
Hyaluronic acid and Fibroscan	47,770	14.29	–	–	Dominated
YKL-40 (combined cut-off) and Fibroscan	48,251	14.31	–	–	Dominated
FIB-4 (combined cut-off) and Fibroscan	47,739	14.31	–	–	Dominated
Fibrotest (combined cut-off) and Fibroscan	47,748	14.31	–	–	Extendedly dominated
SAFE	47,985	14.33	–	–	Dominated
Type IV collagen	47,882	14.34	–	–	Dominated
King's	47,976	14.34	–	–	Dominated
PLT–Spleen	47,874	14.35	–	–	Extendedly dominated
(S1) FPI (low cut-off)	49,467	14.42	–	–	Dominated
(S1) Fibrotest (low cut-off)	49,699	14.44	–	–	Dominated
(S1) US SAPI (low cut-off)	49,746	14.46	–	–	Dominated
FPI (low cut-off)	49,218	14.47	–	–	Extendedly dominated
Fibrotest (low cut-off)	49,536	14.49	–	–	Dominated
US SAPI (low cut-off)	49,549	14.51	–	–	Extendedly dominated
Treat all	51,241	14.73	4,287	0.47	9204

FPI, Fibrosis Probability Index; PLT, platelet; S1, strategy 1; S2, strategy 2; S3, strategy 3; S4, strategy 4; US, ultrasound.

Sensitivity analyses

The base-case analysis result remained robust to the majority of the sensitivity analyses. One of the analyses, where we held the utility values constant at baseline values (values did not differ before, during or post treatment), did increase the ICER for the 'treat all' scenario to £16,727; however, assuming a cost-effectiveness threshold of £20,000, this would still be an acceptable ICER. All other results returned a similar ICER value to the base-case analysis.

Appendix 11 contains the short tables of incremental analysis results for all sensitivity analyses (excluding 'dominated' and 'extendedly dominated' test strategies; full tables are available on request).

Threshold sensitivity analysis

Treatment benefit varied for incorrectly diagnosed patients in mild health state.

Varying the effectiveness of treatment in patients with mild fibrosis (F0–1) affects the results of the analysis. Assuming that no patients in a mild health state (F0–1) receive benefit from treatment returns a result where the 'treat all' strategy is dominated (i.e. it is more costly and less effective than other strategies). We then reduced the effectiveness of the treatment in patients with mild fibrosis (modelled using the SVR rate) by decrements of 10%.

Reducing the SVR rates in the model by 90%, 80% and 70% returned a result where 'treat all' remained dominated by other strategies. When we reduced the SVR rate by 60%, the 'treat all' strategy was the most effective strategy but with an ICER of £723,503. The most cost-effective testing strategy to use for these stages was to test with an imaging modality (MR elastography).

When we reduced the SVR rate by 50%, 'treat all' became the most effective strategy but not the most cost-effective with an ICER of £92,995; testing with MR elastography was the most cost-effective strategy. Given a cost-effectiveness threshold of £20,000, the 'treat all' strategy is no longer cost-effective when the SVR rate is reduced by approximately 23% or more. *Figure 8* plots the decreasing ICER value for the 'treat all' strategy relative to the probability of an increased successful response rate to treatment (SVR rate) for patients in a mild health state.

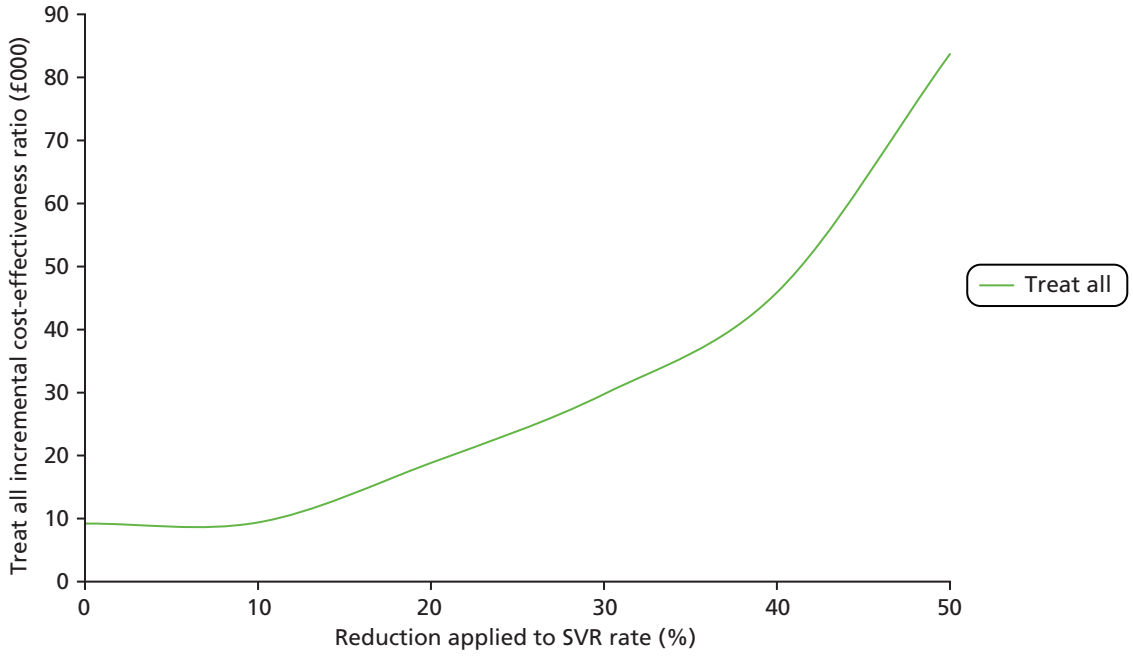

FIGURE 8 Illustration of the reduction in the 'treat all' ICER when the SVR rate for F0–1 patients is varied.

Increased effectiveness of cost and antiviral treatment

Changing both the SVR rate and increasing the cost by £20,000 (to reflect the new drug treatment) does not change the base-case results, with 'treat all' remaining the most cost-effective strategy with an ICER of £10,009.

Increasing the additional cost by £40,000 (with increased SVR rates) does change the results where the ICER for treat all is now £21,174, which would be not cost-effective given a £20,000 threshold. The most cost-effective option to adopt would be MR elastography, with an ICER of £9189.

Illustration of the reduction in the 'treat all' ICER when the SVR rate for F0–1 patients is varied.

Discussion

The analysis found that the most cost-effective option to adopt is one where all patients are treated irrespective of fibrosis stage. This result held when compared with individual tests and strategies incorporating combinations of multiple tests, and for all cost-effectiveness threshold ranges above approximately £10,000. It was also robust to most of the amendments in the sensitivity analysis. When we compared all tests singly, all other tests were either dominated or extendedly dominated, except for MR elastography which, for these very low cost-effective threshold values, was the most cost-effective strategy.

National Institute for Health and Care Excellence guidance for peginterferon alfa and ribavirin treatment in patients with HCV recommends two options for treatment in patients with mild chronic HCV: treat immediately or adopt a watchful waiting approach. The guidelines noted that initiating treatment at an earlier stage bypasses the need for an invasive liver biopsy tests. However, with the introduction of NILTs such as MR elastography and Fibroscan, less invasive testing to stage fibrosis is possible.[443]

A key driver in the cost-effectiveness results is that patients with milder degrees of fibrosis (e.g. F0–1) gain benefit from treatment, albeit at an increased cost. Therefore, the 'treat all' strategy and the tests with highest sensitivity tend to have better results. The assumption that the efficacy of antiviral therapy is similar for patients with mild HCV compared with those with histologically more advanced HCV[457,465] was tested in a threshold sensitivity analysis which found that for the 'treat all' strategy to cease to be cost-effective, the SVR rate following treatment for patients with mild fibrosis would need to be reduced by 23%. Once we reduced the SVR rate by 23%, MR elastography became the most cost-effective strategy; this was the NILT with the highest sensitivity and specificity (94% and 92%, respectively).

A 2006 study by Grieve *et al.*[457] analysed if it was cost-effective to treat patients at a mild stage compared with waiting till patients progressed to a moderate disease stage; they found that it was generally more effective to provide antiviral treatment at a mild rather than a moderate disease stage and this strategy would gain improved outcomes (QALYs) rather than treating at a later stage.

However, treating patients with mild chronic HCV (who may not necessarily require treatment) exposes this patient cohort to the risk of side effects associated with peginterferon alfa and the direct antiviral agents, boceprevir and telaprevir. Newer drugs may provide more effective treatment with fewer side effects, so waiting to treat this patient cohort may be a better option.[466] If we were to adopt this strategy, then testing all patients with MR elastography would be the most cost-effective option (when using a NILT singly).

New antiviral treatments for HCV are currently undergoing trials which indicate that these drugs may be more effective in genotype 1 and 4 HCV infections than current triple therapy.[43] However, the cost-effectiveness of such drugs will depend on the price of the new drug and robust data on effectiveness. Currently evidence is promising but limited to uncontrolled studies and a non-inferiority study demonstrating equivalence. If the SVR rates observed in the studies are borne out in practice, then the 'treat all' strategy could still be cost-effective if the overall impact on treatment costs is an increase of £20,000, but not for an increase of £40,000.

When all tests were compared singly, the tests with low threshold values for classification of fibrosis tended to have relatively higher health outcomes than tests without threshold or with high cut-off thresholds. These tests tended to have high sensitivity values and low specificity values; when we conducted the threshold sensitivity analysis around treatment benefit, NILTs with lower cut-off thresholds no longer yielded the highest health outcomes. Rather, tests with both a high sensitivity and specificity such as MR elastography had the higher QALY gain, indicating that when patients in a mild health state are treated, tests which are more likely to identify more persons as positive (i.e. treat more persons) will be more effective.

As some of the tests analysed have few studies on which their diagnostic accuracy is based, this means that some tests may have results which overestimate their effectiveness. Reducing the number of tests to those where only the bivariate model converged did not change the overall result; however, only 14 NILTs were compared at this stage, illustrating that for the majority of tests the results of the meta-analysis may not be very robust. When we removed 'treat all' from this analysis, Fibroscan was the most cost-effective testing option to use.

We have considered the SVR a good surrogate marker of treatment efficacy. This is similar to current NICE guidance for boceprevir and telaprevir; however, our results are applicable only if the SVR rate is a good surrogate marker and may change if other methods of determining treatment efficacy are developed.

The results indicate that treating all patients is cost-effective; however, this result is sensitive to changes in treatment response rates for those with mild fibrosis.

Chapter 7 Cost-effectiveness analysis: alcoholic liver disease

This chapter details the analysis approach for ALD and includes details of the model structure, inputs and results. The population considered for analysis were people with ALD who were suspected of having developed alcoholic steatohepatitis (ASH) with cirrhosis.

Background

Five NILTs were identified in the systematic review for staging liver fibrosis and cirrhosis in ALD: APRI with a (high cut-off) diagnostic threshold, Fibrotest (high cut-off), Fibrotest (low cut-off), PGAA and Fibroscan.

Current clinical management of alcoholic cirrhosis focuses on alcohol abstinence, aggressive nutritional therapy rich in calories and proteins, and primary and secondary prophylaxis of cirrhosis complications.[35] If required and available, addiction specialists, motivational therapy and anticraving drugs are also recommended.[35]

The European Association for the Study of the Liver (EASL) guidelines note that patients with ASH and cirrhosis are at risk of developing clinical decompensation, liver failure and HCC. However, prolonged abstinence can reverse previously decompensated cirrhosis to a compensated state. As the incidence of HCC among patients with alcoholic cirrhosis ranges from 7% to 16% after 5 years, the guidelines recommend that screening for HCC should be performed with ultrasound every 6 months, as recommended for any patient with cirrhosis. Screening for alcohol-induced damage in other organs (heart, kidneys) should also take place.

The guidelines do note that specific therapies have been tested in patients with alcoholic cirrhosis (including S-adenosyl-l-methionine, anabolic-androgenic steroids); however, none has any consistent beneficial effects.[35]

Liver transplantation is recommended for patients who remain abstinent for a 6-month period before being added to the waiting list for liver transplantation.[35] The 6-month period is recommended to capture those patients who may recover from their liver disease and also to identify the subset of patients who may remain abstinent after a liver transplant. NICE guidance recommends that patients who still have decompensated liver disease after 3 months of best management and abstinence from alcohol and who are suitable candidates for liver transplantation should be referred to assessment for liver transplantation and considered this to be a cost-effective treatment option.[467]

The EASL and NICE guidelines do not recommend a specific treatment applicable after diagnosis with fibrosis or cirrhosis; rather, they recommend that, regardless of the severity, abstinence and early management of alcohol abuse or dependence is warranted in all patients with ASH. Therefore, we were unable to use the same approach for a cost-effectiveness analysis as for HBV and HCV. Instead, we focus on the health economic impact of diagnosis as a result of abstinence, assuming that abstinence may increase as a result of diagnosis of fibrosis.

Literature review results

We identified a recent systematic review and economic analysis of non-invasive diagnostic assessment tools for the detection of liver fibrosis in patients with suspected alcohol-related liver disease;[428] the non-invasive tests analysed in this report were ELF, Fibrotest, Fibroscan (TE) and Fibromax.

The tests considered in the analysis by Stevenson *et al.*[428] differed from the tests found during our literature review. However, this study did not find any data on diagnostic accuracy for one of the NILTs, Fibromax. With regard to the diagnostic accuracy of the ELF test, the authors located one study that looked at the European liver fibrosis test (ELF and age)[468] which had reported findings for diagnosis of moderate to severe fibrosis; however, the diagnostic accuracy for identifying cirrhosis was not reported. The authors concluded that the data with regard to the diagnostic accuracy of the ELF test in relation to cirrhosis are not robust.

The review was considered highly relevant to our decision problem and population of interest. We carried out a literature search for cost-effectiveness studies of non-invasive tests in patients with ALD using the search strategy detailed in the Health Technology Asessment (HTA) report by Stevenson *et al.*[428] We updated the search to include the five NILTs analysed for ALD. We conducted the search using the MEDLINE database (via Ovid platform, searched 25 June 2013, all years searched). The search strategy is detailed in *Appendix 2*.

The search located 264 papers; titles were reviewed and the only relevant study found was the economic analysis detailed in the HTA report by Stevenson *et al.*[428]

The report by Stevenson *et al.*[428] estimated the incremental costs and incremental QALYs for 10 strategies which were based on using a NILT alone to confirm cirrhosis or a combination of a NILT and liver biopsy. The authors assumed that all patients in the model would receive lifestyle and abstinence advice. If the test outcome was positive, the treatment strategy would be expanded to include monitoring for HCC, hepatic encephalopathy and oesophageal varices.

Cost-effectiveness analysis

The analysis conducted by Stevenson *et al.*[428] was considered relevant, as they also considered the use of non-invasive diagnostic testing in patients with ALD. They conducted their analysis from a UK perspective and their considered population was deemed to be similar (patients with ALD suspected of having cirrhosis). The model structure employed in this study seemed appropriate given current clinical practice and the data inputs for costs and QALYs were disease specific (based on data for ALD).

We replicated the model to include the tests relevant to our decision problem and also conducted some sensitivity analyses where we assessed the impact of alternative inputs and adjustments to the model structure. To replicate the model, we constructed a decision tree model to assess the cost-effectiveness of the non-invasive tests in patients with ALD.

Decision tree model

The following section (section 1) provides a brief description of the Stevenson *et al.*[428] model, which we replicated. This includes a summary of the inputs which we also employed in our model. We then detail (see section 2) the changes we made to the model for our analysis (updated inputs, different non-invasive tests and diagnostic accuracy).

Section 1: model structure

The relevant patient population to be assessed with either a NILT or liver biopsy would be those patients suspected of having liver cirrhosis (F4).

The care pathway reflected in the model is that a positive test result (true positive or false positive) indicated that the patient would receive monitoring for HCC, varices, ascites and hepatic encephalopathy; lifestyle advice would also be given. Given a negative test result (true negative or false negative), the

patient would receive lifestyle advice only, including the recommendation to become abstinent or to reduce alcohol consumption.

Abstinence rates following diagnosis with liver biopsy

Only one study had been identified regarding abstinence rates following diagnosis, which was a small sample ($n = 96$) and had only been published in abstract form (we conducted an updated search on 26 June 2013 and were unable to find a published paper for the study). This study reported that after diagnosis with liver biopsy, 31% of those with a negative test result for cirrhosis became abstinent, whereas 62% of those with a positive test result for cirrhosis became abstinent. No specific data were found on abstinence rates following a NILT.

Probability of developing cirrhosis following diagnosis

There is a probability that patients who continue to drink after diagnosis may develop cirrhosis (false positive and true negative patients). This was set at 20% following clinical advice.

Quality-adjusted life-years

Quality-adjusted life-year estimates were sourced from a HTA report on HCC screening by Thompson Coon *et al.*[469] The QALY value reported in the Thompson Coon *et al.* study for ALD patients with cirrhosis who undergo annual serum α-fetoprotein with 6-monthly ultrasound scans was used as an estimate for patients who test true positive and stop drinking. For true-positive patients who continue to drink, QALY value was based on survival rates for cirrhotic patients who continued to drink derived from a study by Verrill *et al.*[470]

For true-negative patients who abstain from drinking, the QALY value was estimated using the average age from the study by Thompson Coon *et al.*[469] (53 years) and EQ-5D population norms (assuming that this population would live for a further 20 years). In the absence of further data, it was assumed that the QALY for false positives who abstain from drinking would be equivalent to that of true negative patients who abstain. For false-positive patients who continue to drink, the study conservatively assumed the same QALY; however, it was assumed that the QALY for the proportion of false-negative patients and true-negative patients who continue to drink and subsequently develop cirrhosis would be the same as that for true positives who continue to drink.

A QALY value for false-negative patients who abstain from drinking was sourced from the Thompson Coon *et al.*[469] study. For false-negative patients who continue to drink, the QALY value was set at 40% of the value of false-negative patients who abstain.

Estimates (where applicable) were adjusted for differing survival rates (abstainers vs. non abstainers) and for oesophageal bleeds. QALY estimates are detailed in *Table 42*.

TABLE 42 Quality-adjusted life-year estimates used in model

QALY end points	QALY
TP compensated cirrhosis (abstain drinking)	9.679
TP compensated cirrhosis (continue drinking)	4.399
FP compensated cirrhosis (abstain drinking)	11.066
FP compensated cirrhosis (continue drinking)	11.066
TN compensated cirrhosis (abstain drinking)	11.066
TN compensated cirrhosis (continue drinking)	11.066
FN compensated cirrhosis (abstain drinking)	9.359
FN compensated cirrhosis (continue drinking)	3.744

FN, false negative; FP, false positive; TN, true negative; TP, true positive.

Adverse events and mortality risk
The risk for adverse events (0.72%) and mortality (0.09%) associated with liver biopsy was estimated following a systematic review.

Liver biopsy adverse events costs and quality-adjusted life-years
A cost of £1000 (reflect hospital stay) and a QALY decrement of 0.2 associated with a serious adverse effect resulting from liver biopsy applied in the model.

Section 2: amendments to model

Test strategies assessed
We analysed 10 potential testing strategies incorporating the applicable NILTs. The strategies analysed the use of all tests either alone (biopsy or a NILT) or in combination. We also included a comparator where all patients receive treatment for cirrhosis (HCC screening) without testing (similar to 'treat all' strategies from the HBV and HCV models).

1. Biopsy all patients.
2. Test all patients with APRI high cut-off and biopsy those in whom cirrhosis is indicated.
3. Test all patients with Fibrotest high cut-off and biopsy those in whom cirrhosis is indicated.
4. Test all patients with Fibrotest low cut-off and biopsy those in whom cirrhosis is indicated.
5. Test all patients with PGAA and biopsy those in whom cirrhosis is indicated.
6. Test all patients with Fibroscan and biopsy those in whom cirrhosis is indicated.
7. Use APRI high cut-off and assume result is definite.
8. Use Fibrotest high cut-off and assume result is definite.
9. Use Fibrotest low cut-off and assume result is definite.
10. Use PGAA and assume result is definite.
11. Use Fibroscan and assume result is definite.
12. Diagnose all patients as having cirrhosis.

Abstinence rates following diagnosis with a non-invasive liver test
As abstinence rates are based on clinical practice and experience, it is plausible to assume that biopsy as an invasive procedure which may include hospitalisation and the risk of side effects such as fatal bleeding may serve as a larger warning for patients (patients may be more likely to comply after a positive biopsy result than after a positive result with a NILT). So, we assumed a lower rate of abstinence for patients who were tested with a NILT. In the base case, we assumed the rate for abstinence after diagnosis with a NILT would be 10% lower than the abstinence rates after biopsy. As the abstinence figures after a NILT are not known with certainty and these figures may potentially have a large impact on the decision tree outcomes, we undertook a number of sensitivity analyses where we decreased and increased the abstinence values applicable after diagnosis with a NILT.

The EASL position paper for ALD[55] notes that 40% of patients with ALD who have already developed fibrosis and continue to drink will develop cirrhosis. We received clinical advice on this and in the general population; the consensus is that 10–30% of patients will develop cirrhosis if they continue to drink. However, as some persons may not develop fibrosis irrespective of drinking patterns (as developing fibrosis is genetically determined), the figure may be slightly higher (40%) in those who have developed some degree of fibrosis, making them more prone to developing cirrhosis. We used the base-case value of 20% in the main analysis and conducted sensitivity analyses around this parameter.

Costs
We inflated the costs used in the study by Stevenson *et al.*[428] from 2008–9 to 2012 prices using NHS inflation indices. The costs used in the decision tree are displayed in *Table 43*.

TABLE 43 Costs of decision tree end points (includes costs of electroencephalograms, and the costs and QALY implications of screening for varices, providing prophylaxis treatment where appropriate and treating variceal bleeding)

End point decision tree costs 2012	Cost, £
True positive compensated cirrhosis (abstain drinking)	32,080
True positive compensated cirrhosis (continue drinking)	42,239
False positive compensated cirrhosis (abstain drinking)	26,916
False positive compensated cirrhosis (continue drinking)	26,916
True negative compensated cirrhosis (abstain drinking)	1070
True negative compensated cirrhosis (continue drinking)	1070
False negative compensated cirrhosis (abstain drinking)	27,928
False negative compensated cirrhosis (continue drinking)	38,628

The costs had been sourced from a HTA report which looked at the costs associated with screening persons with ALD for HCC.[469] The authors also used the costs for mild and moderate fibrosis health states for HCV based on the mild HCV trial[414] described in *Chapter 6* to modify the costs of true-positive patients progressing to a more severe state (decompensated cirrhosis) if they continued to drink and for false-positive patients who abstained from further drinking. Estimates for the true-negative end points were assumed based on clinical advice. The costs of monitoring and treatment of oesophageal bleeding and the cost of detecting a hepatic encephalopathy were included.

Estimation of probabilities of a non-invasive liver test returning a true-positive, false-positive, true-negative or false-negative result

The average prevalence of disease (defined as METAVIR score of F4) was taken from the average prevalence reported in papers included in the meta-analysis (see *Chapter 4*). We used the sensitivity and specificity estimates for each NILT (see *Chapter 4* for estimates) and the average prevalence estimate to calculate the probability of each test returning a true-positive, false-positive, true-negative and false-negative result (see *Appendix 7*).

Test costs

Costs for APRI and PGAA indirect serum markers were obtained from communication with finance departments based at the Royal Free Hospital. Costs for patented serum markers (Fibrotest) were sourced directly from manufacturers and via communication with finance departments based at the Royal Free Hospital. A test cost for Fibroscan was sourced as per clinical advice; we used the 2011–12 Department of Health reference costs for an ultrasound with duration of < 20 minutes. The cost of a percutaneous liver biopsy (see *Chapter 5* for further details on choice of liver biopsy) was sourced from published literature[428] (see *Chapter 5*). Test costs including sources are listed in *Appendix 9*.

Analysis

A PSA was undertaken and we conducted an incremental analysis of the results. A CEAC and CEAF were constructed. We also carried out univariate sensitivity analyses and conducted two analyses where we amended the structure of the ALD model. A scatterplot illustrating the position of each testing strategy on the cost-effectiveness acceptability curve compared with the least costly testing strategy can be found in *Appendix 12*.

Sensitivity analyses

Amendments to model inputs

Abstinence rates

We amended the abstinence rates assumed in the model following diagnoses with a NILT. We set the rates equivalent to those for liver biopsy and then increased the rates in increments of 10% to determine when the base-case result changed.

Probability of developing cirrhosis

We set the rate to five different values (10%, 30%, 40%, 50% and 60%) to test the impact on results. The rates chosen reflected the fact that we do not know the exact fibrosis level of patients who test true negative or false positive, and so this range allows us to test the impact of a change in this parameter on results.

Mortality rates and adverse events

We varied these in the model using a range of lower and upper estimates.

Amendments to model structure

Additional health states

We conducted an analysis where we allowed for progression to HCC and subsequent liver transplant. This is not included as an option in the Stevenson *et al.*[428] model as they argued that current evidence shows that it is of borderline cost-effectiveness and would not have an impact on the cost-effectiveness of a NILT. However, as the NILT may affect the care pathway (including decision to continue drinking which may indicate further progression to more severe health states), and as ALD is one of the main reasons for liver transplantation in the UK, we included liver transplant as an option in the care pathway. Furthermore, in the CELT study of transplantation[424] it was found that the larger incremental cost-per-QALY ratio for ALD patients was in part due to a larger proportion of ALD patients being considered unsuitable for transplantation after undergoing the assessment process. Since this study was conducted, clear guidelines have been agreed by the six liver transplant centres within the UK and endorsed by the UK Liver Advisory Group to include careful assessment of psychosocial and substance use factors for patients with a diagnosis of ALD.[471]

Cost and QALY estimates were sourced from the CELT study.[424] Upon clinical advice, we set the estimate for the percentage of patients with ALD who have a liver transplant to 4%, to reflect the number of patients with alcoholic cirrhosis who may potentially develop a tumour while abstinent.

Different starting population-advanced fibrosis (≥ F3)

An additional analysis was undertaken where we used a different section of the data as our starting population cohort: our base-case population cohort consists of patients who are suspected of having cirrhosis; the sensitivity analysis analysed the impact of substituting different data, representing patients who are suspected of having advanced fibrosis (F3), from the meta-analysis.

Results

Base-case results imply that the most cost-effective strategy is to use liver biopsy only to diagnose cirrhosis in patients with ALD, with an ICER of £822 (*Table 44*). The strategy producing the highest QALY gain is where all patients suspected of having cirrhosis receive monitoring; however, this is not the most cost-effective with an ICER of £70,861, which is above the standard UK cost-effectiveness threshold range.[66]

We constructed a CEAC (see *Appendix 10*) and CEAF (*Figure 9*). The CEAF shows that strategy 2 (liver biopsy) has the highest probability of being the optimal choice (highest expected net benefit) for threshold values of £822 and above.

TABLE 44 Alcoholic liver disease incremental analysis: all strategies compared

Strategy	Tests	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Strategy 2	APRI (high cut-off) and liver biopsy	17,415	8.79	–	–	–
Strategy 7	APRI (high cut-off)	22,463	8.87	–	–	Dominated
Strategy 10	PGAA	19,061	8.98	–	–	Dominated
Strategy 11	Fibroscan	20,009	9.02	–	–	Dominated
Strategy 8	Fibrotest (high cut-off)	19,504	9.03	–	–	Dominated
Strategy 5	PGAA and liver biopsy	17,613	9.07	198	0.28	701
Strategy 9	Fibrotest (low cut-off)	24,671	9.13	–	–	Dominated
Strategy 6	Fibroscan and liver biopsy	17,702	9.14	–	–	Extendedly dominated
Strategy 3	Fibrotest (high cut-off) and liver biopsy	17,724	9.17	–	–	Extendedly dominated
Strategy 4	Fibrotest (low cut-off) and liver biopsy	17,801	9.26	–	–	Extendedly dominated
Strategy 1	Liver biopsy	17,812	9.31	199	0.24	822
Strategy 12	All patients treated as having cirrhosis (receive HCC screening)	31,004	9.50	13,193	0.19	70,861

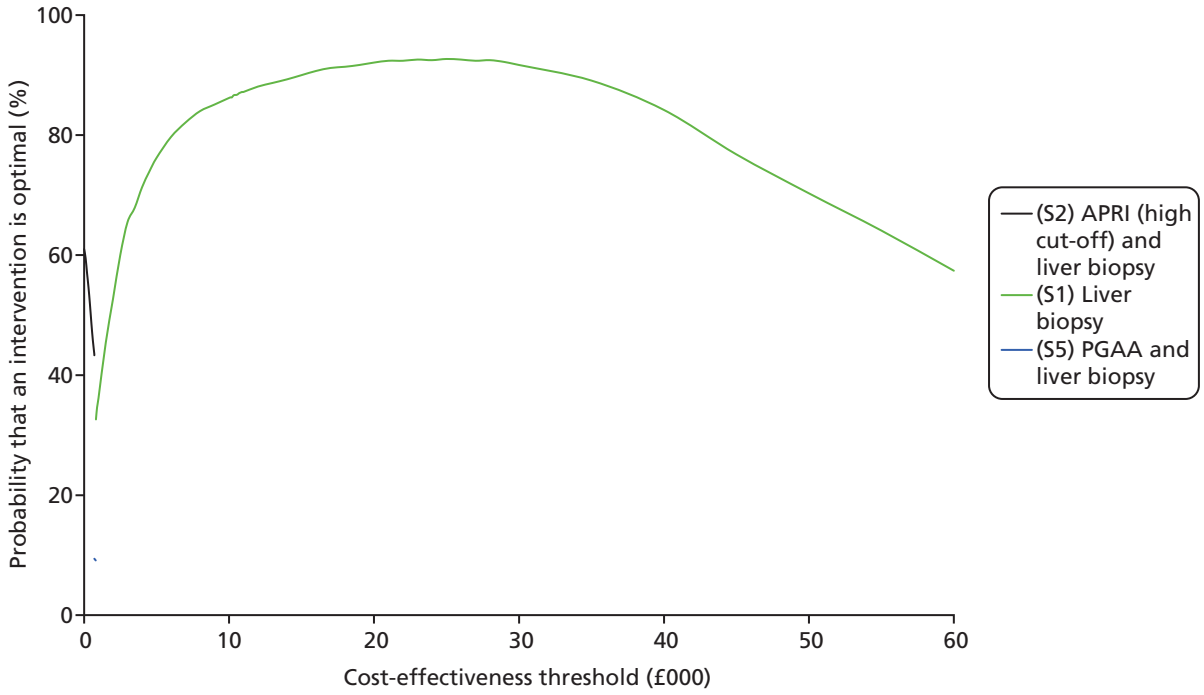

FIGURE 9 Cost-effectiveness acceptability frontier for ALD.

The CEAC presented in *Appendix 10* shows that strategy 1, liver biopsy, has the highest probability of being cost-effective (92%) given a cost-effectiveness threshold of £20,000 and this probability falls as the cost-effectiveness threshold increases (for clarity of illustration, only strategies which have a ≥ 10% probability of being the optimal strategy are shown in the CEAC).

Sensitivity analysis results

Univariate sensitivity analyses

A number of one-way sensitivity analyses were conducted around parameters in the model for which the data were assumed or advised by clinical opinion (abstinence rates, probability of developing cirrhosis if patient continues to drink). *Table 45* lists the ranges over which the values were varied and whether or not this had any impact on the base-case result.

Change in abstinence rate following diagnosis with a non-invasive liver test (diagnosis of cirrhosis and no cirrhosis)

The base-case values used in the model for abstinence rates after diagnosis with a NILT were 52% following a diagnosis of cirrhosis and 31% following a diagnosis of no cirrhosis. We conducted a number of analyses varying these rates (*Table 46*).

When the abstinence rates after diagnosis with a NILT were set equivalent to the abstinence rates after a liver biopsy (increase of 10%), the most cost-effective testing option was strategy 4, a combination of Fibrotest (low cut-off) as an initial test and liver biopsy as a second test to confirm positive diagnoses, with an ICER of £6366.

However, when we increased the abstinence rates following diagnosis with a NILT by 15% and 20%, this impacted the results where Fibrotest (high cut-off) became the most cost-effective test with an ICER of £13,115 and £5896, respectively. This test has the highest sensitivity and specificity (91% and 87%).

TABLE 45 Sensitivity analyses, parameters, ranges, result

Parameter	Base-case value	Extreme values tested	Impact on cost-effectiveness
Probability of developing cirrhosis (continue drinking FP and TN patients)	20%	10–50%	No change
Probability of developing cirrhosis (continue drinking FP and TN patients)	20%	60%	All treated as cirrhotic (receive HCC screening) becomes cost-effective option
Mortality rate following liver biopsy	0.09%	0.05–0.20%	No change
Adverse event rate following liver biopsy	0.72%	0.05–0.90%	No change

FP, false positive; TN, true negative.

TABLE 46 Increase in abstinence rates following NILT

Change in abstinence rate	Rate following diagnosis: cirrhosis	Rate following diagnosis: no cirrhosis	Impact cost-effectiveness
Increase 10%	62%	41%	Fibrotest (low cut-off) and liver biopsy becomes most cost-effective testing option
Increase 15%	67%	46%	Fibrotest (high cut-off) becomes most cost-effective test
Increase 20%	72%	51%	Fibrotest (high cut-off) becomes most cost-effective test

Increase in abstinence rate after diagnosis of no cirrhosis with a non-invasive liver test

We conducted an analysis where we only increased the probability that patients would become abstinent after a diagnosis of no cirrhosis (using a NILT), leaving all other base-case values constant. We increased the base-case value of 31% by 10% and 20%. Increasing the abstinence rate by 10% changed the result and strategy 4 [Fibrotest (low cut off) with liver biopsy] used as a second test to confirm results became the most cost-effective with an ICER of £6833. When we increased the abstinence rate by 20%, using strategy 3 [Fibrotest (high cut off) as an initial test and liver biopsy used as a second test to confirm positive results] became the most cost-effective strategy with an ICER of £4730.

Amendments to model structure

Different starting population: advanced fibrosis (≥ F3)

An analysis where we incorporated a different starting population (F3) did not change the overall result: liver biopsy remained the most cost-effective testing option. *Table 47* displays the results of the analysis.

Additional health states

We allowed for persons with cirrhosis (who remained abstinent within the model) to have a small probability of undergoing a liver transplant (4%, based on clinical advice). We incorporated cost and QALY estimates sourced from the CELT study[421] specifically for patients with ALD (cost estimate of £11,202 and QALY value of 0.58); we inflated the historic cost to 2012 prices (£17,741) using NHS inflation indices.

The results are presented in *Table 48*. Liver biopsy remains the most cost-effective testing option, with an ICER of £688. However, similar to the base-case analysis, this result is dependent on the abstinence rates used in the analysis. When we amend the abstinence rate after diagnosis with a NILT so that it became equivalent to the abstinence rate after diagnosis with a liver biopsy, the results change, so that strategy 4 [Fibrotest (low cut off with liver biopsy to confirm positive results)] becomes the most cost-effective, with an ICER of £4756.

TABLE 47 Alcoholic liver disease incremental analysis: different starting population (≥ F3)

Strategy	Test strategy	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Strategy 3	Forns index (high cut-off) and liver biopsy	21,135	8.30	–	–	–
Strategy 8	Forns index (high cut-off)	22,085	8.36	–	–	Dominated
Strategy 5	YKL-40 and liver biopsy	21,224	8.40	–	–	Extendedly dominated
Strategy 10	YKL-40	22,224	8.42	–	–	Extendedly dominated
Strategy 9	Fibroscan	23,456	8.62	–	–	Extendedly dominated
Strategy 7	CK18	24,646	8.62	–	–	Extendedly dominated
Strategy 2	CK18 and liver biopsy	21,486	8.78	351	0.49	722
Strategy 4	Fibroscan and liver biopsy	21,547	8.80	–	–	Extendedly dominated
Strategy 1	Liver biopsy	21,652	8.99	166	0.21	800
Strategy 12	All treated as cirrhotic (HCC screening)	31,963	9.14	10,311	0.15	69,522

TABLE 48 Alcoholic liver disease incremental analysis: analysis with additional health states added (HCC and liver transplant)

Strategy	Test strategy	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Strategy 2	APRI (high cut-off) and liver biopsy	17,637	8.80	–	–	–
Strategy 7	APRI (high cut-off)	22,538	8.88	–	–	Dominated
Strategy 10	PGAA	19,213	8.99	–	–	Dominated
Strategy 11	Fibroscan	20,148	9.03	–	–	Dominated
Strategy 8	Fibrotest (high cut-off)	19,638	9.04	–	–	Dominated
Strategy 5	PGAA and liver biopsy	17,928	9.09	–	–	Extendedly dominated
Strategy 9	Fibrotest (low cut-off)	24,845	9.14	–	–	Dominated
Strategy 6	Fibroscan and liver biopsy	18,027	9.16	–	–	Dominated
Strategy 3	Fibrotest (high cut-off) and liver biopsy	18,055	9.19	–	–	Dominated
Strategy 4	Fibrotest (low cut-off) and liver biopsy	18,077	9.28	–	–	Dominated
Strategy 1	Liver biopsy	18,000	9.33	363	0.53	688
Strategy 12	All treated as cirrhotic (receive HCC screening)	30,705	9.42	12,705	0.10	133,479

Cost per correct diagnosis analysis

We also constructed a simple probabilistic decision model to assess the cost per correct diagnosis of the applicable NILTs compared with liver biopsy. As for the decision tree analysis, fibrosis prevalence in the model was estimated using a section of the meta-analysis equivalent to the proportion of the population with a test score of $\geq$ F4, measured using the METAVIR score. Based on the diagnostic accuracy of each test, the patient was classified as true positive, true negative, false positive or false negative. True positives and true negatives were considered to be correct diagnoses.

Test costs

We employed the same test costs for the NILTs and liver biopsy as used in the decision tree model.

Analysis

We assumed that liver biopsy as the reference standard had perfect sensitivity and specificity, implying that biopsy accurately stages liver fibrosis. Using a hypothetical cohort of 1000 ALD patients suspected of having liver fibrosis, we estimated the incremental cost associated with each test compared with the next best alternative. Tests which had fewer numbers of true-positive or true-negative outcomes, with a higher test cost than other tests which had higher correct diagnoses, were ruled out of the analysis (dominated). We present the results separately for positive and negative diagnoses as the consequences are different for each. Results are shown in *Tables 49* and *50*.

Liver biopsy also returns the highest correct number of negative responses ($n = 634$) (see *Table 50*); however, the incremental cost per correct positive diagnosis for using liver biopsy was £13.62, compared with an incremental cost per correct diagnosis of £0.02 associated with using the next best alternative, PGAA, as the diagnostic test.

TABLE 49 Results of the cost per correct TP diagnosis

Test	Number of TPs using NILT	Test cost, £	Number of incremental correct diagnoses	Incremental cost (test only), £	Incremental cost per correct diagnosis (TP): each test compared with next best alternative (£/correct diagnosis gained)
APRI (high cut-off)	144	4.05	–	–	–
PGAA	285	7.69	141	3.64	Extendedly dominated
Fibroscan	315	51.00	–	–	Dominated
Fibrotest (high cut-off)	334	43.60	190	39.55	Extendedly dominated
Fibrotest (low cut-off)	363	43.60	219	39.55	0.18
Liver biopsy	366	956.61	4	593.92	168

TP, true positive.

TABLE 50 Results of the cost per correct TN diagnosis

Test	Number of TNs using NILT	Test cost, £	Number of incremental correct diagnoses	Incremental cost (test only), £	Incremental cost per correct diagnosis (TN): each test compared with next best alternative (£/correct diagnosis gained)
Fibrotest F4 (low cut-off)	317	43.60	–	–	Dominated
APRI F4 (high cut-off)	392	4.05	–	–	–
Fibroscan F4	526	51.00	–	–	Dominated
Fibrotest F4 (high cut-off)	551	43.60	–	–	Dominated
PGAA F4	564	7.69	172	3.64	0.02
Liver biopsy	634	956.61	70	948.92	13.62

TN, true negative; TP, true positive.

Results of cost per correct diagnosis analysis

Column 2 in *Table 49* shows the number of correct positive diagnoses using a NILT or liver biopsy. The highest number of correct positive diagnoses is given using a liver biopsy,[366] which is estimated from the disease prevalence from the meta-analysis. Using Fibrotest (low cut-off) returns the next highest number of true positives.[363] The tests are compared incrementally with tests which provide fewer true-positive results at a higher test cost ruled out (dominated) of the analysis. Liver biopsy returned the greatest number of correct diagnoses; however, the incremental cost per correct positive diagnosis for using liver biopsy was £168 compared with an incremental cost per correct diagnosis of £0.18 associated with using the next best alternative, Fibrotest (low cut-off), as the diagnostic test.

Discussion

Liver biopsy is the most cost-effective testing option to adopt given our model at the standard UK cost-effectiveness threshold range. This result is sensitive to assumptions about the abstinence rates after diagnosis, difference in abstinence between people tested with biopsy and NILT, and rates of progression to cirrhosis rates in patients who continue to drink. More accurate data on these parameters are needed in order to reduce uncertainty in these results.

Our results were particularly sensitive to changes in the abstinence rate after diagnosis with a NILT (holding the rates for liver biopsy constant). When we raise the abstinence rates after diagnosis with a NILT, the test with the highest sensitivity and specificity becomes the most cost-effective (the summary sensitivity and specificity would be comparable with liver biopsy).

When we increased the abstinence rate after diagnosis with a NILT by 20%, tests which have high summary sensitivity and specificities tend to be the most clinically effective: Fibroscan and Fibrotest (low cut-off) [Fibroscan 86% and 83% and Fibrotest (low cut-off) 100% and 50%]. This implies that when we increase the abstinence rates following diagnosis with a NILT compared with liver biopsy, tests which have a similar diagnostic accuracy to liver biopsy tend to be more effective.

Liver biopsy also appears to be the most cost-effective option, even when we amend the modelling structure to allow for a different patient cohort (F3) or additional health states (liver transplant). However, this result is also sensitive to changes in the abstinence rates applicable after a NILT; as there are few data on this, the results are uncertain.

The lack of data available regarding cessation rates after diagnosis with either a NILT or liver biopsy are a limitation in the study. One is based on a small study, the other on an assumption, and both are key drivers of the analysis which leave the results very uncertain.

The HTA report by Stevenson *et al.*[428] did not report an incremental analysis. However, they did carry out a number of threshold analyses where they decreased the abstinence rate and increased it to see at which point biopsy became cost-effective compared with a test. They found that the lower the abstinence rate after a NILT, the more likely it is that liver biopsy is cost-effective.

Liver biopsy was also a less costly testing option as we assumed that it had perfect sensitivity and specificity and could accurately identify false-positive responses from true-negative outcomes. A true-negative outcome had a lower cost than a false-positive outcome in the model. Any test that returned a high false-positive result was more costly; for example, using Fibrotest (low cut-off) was more costly and less effective than using Fibrotest to initially identify patients followed by liver biopsy to confirm diagnosis. A reduction in all-cause mortality due to abstinence is not captured in the model. This would further strengthen the robustness of using liver biopsy, if we still assumed higher abstinence rates following a liver biopsy than following a NILT.

Therefore, despite any mortality or morbidity risks associated with liver biopsy, the increased cost associated with false positives and the fact that the utility value for these patients was the same as if true negative (in other words there was no assumed treatment benefit associated with being diagnosed as false positive in the model) meant that any test which returned a high number of false-positive outcomes would be the least cost-effective.

The incremental cost of using liver biopsy compared with a NILT is high when we look at the cost per correct true negative. PGAA, which has the highest specificity, returns the highest number of true negatives which will be picked up by a NILT. However, for tests with a lower specificity, Fibrotest (low cut-off), the cost per correct diagnosis (true negative) using a liver biopsy is the cheapest option compared with using liver biopsy to diagnose true negatives for other tests.

Chapter 8 Cost-effectiveness analysis: non-alcoholic fatty liver disease

This chapter details the approach used to conduct an analysis of non-invasive liver tests for use in patients with NAFLD. The population considered for analysis were persons with NAFLD who were suspected of having developed NASH with fibrosis or cirrhosis.

Literature search

Cost-effectiveness of non-invasive liver tests in non-alcoholic fatty liver disease

We conducted a literature review using a modified version of the search strategy used for ALD to locate published cost-effectiveness studies which had analysed the use of the NILTs in a NAFLD population. We used the MEDLINE database (via Ovid platform, searched 24 June 2013, all years searched). The search strategy used is detailed in *Appendix 2* and inclusion criteria are listed in *Chapter 3*. The search returned 14 papers whose titles and abstract were reviewed. No cost-effectiveness studies were found.

Treatment in non-alcoholic fatty liver disease and effectiveness

We identified a position paper on the treatment of NAFLD and NASH published by EASL in 2010[472] which recommends that an initial treatment strategy for patients with NASH would involve treating insulin resistance and reducing body weight (particularly visceral adiposity). Potential treatment options mentioned in the guideline include weight loss and physical exercise interventions, insulin-sensitising agents, vitamin E therapy and antiobesity surgery; however, the guidelines recommend that all pharmacological medications for treatment of NASH should be considered as experimental. Additionally, treatment and monitoring of metabolic and cardiovascular comorbidities in patients with NAFLD is recommended by the EASL guidelines.

We searched the reference list of the EASL position paper and the American Association for the Study of the Liver (AASLD) practice guidelines[47] for papers providing evidence of relative treatment effects on the potential treatment strategies outlined in the guidelines and supplemented this using a general literature search conducted using Google Scholar (search terms included 'NAFLD', 'NASH' 'vitamin E therapy' and 'pioglitazone', searched 26 June 2013). We also sought clinical advice to identify up-to-date, relevant studies on potential treatments. Using this triple approach, we identified relevant papers on insulin-sensitising agents,[473–475] weight loss and exercise interventions,[476,477] behavioural interventions[478] and bariatric surgery.[479]

The EASL position paper[472] noted that no studies analysing the impact of glitazones and their effect on the cessation or decrease in progression of fibrosis/cirrhosis have shown a convincing benefit. We identified a 2010 systematic review and random-effects meta-analysis Rakoski *et al.*[473] of eight published studies[480–488] which analysed the use of insulin sensitisers [thiazolidineodiones (glitazones)] and metformin in patients with NAFLD and/or NASH. The primary outcome of interest for each paper related to histological response to treatment. When the authors looked at all nine studies together, insulin-sensitising agents showed a significant improvement in fibrosis; however, a subgroup analysis looking only at studies which investigated the use of glitazones (six studies) found that glitazones (rosiglitazone, pioglitazone) did not result in a significant improvement in fibrosis. A sensitivity analysis conducted by the authors excluding patients with diabetes from the analysis found that, when this subset of patients was removed, pioglitazone did result in a significant decrease in fibrosis. The study did not, however, provide any applicable data to use in an economic model, such as the reduced probability of patients progressing to a more severe NASH state as a result of treatment.

A 2012 cost–utility study by Mahady *et al.*[475] derived a treatment effect (RR) for histological improvement with pioglitazone from a 2011 meta-analysis of randomised trials conducted by Mahady *et al.*[489] The authors estimated a RR of 1.40 for pioglitazone versus placebo using three published studies of randomised trials where pioglitazone was used as add-on therapy to standard lifestyle advice.[481–483] However, the 2012 meta-analysis by Rakoski *et al.*[473] (reviewed above) had also analysed these three studies in their subgroup meta-analysis of glitazones and found that glitazones did not result in an improvement in fibrosis. In addition, the 2010 paper by Sanyal *et al.*,[481] which was analysed by Mahady *et al.*[489] in their meta-analysis, noted that no significant improvement was seen in fibrosis scores as a result of pioglitazone treatment. Indeed, in their paper, Mahady *et al.*[489] themselves note that the degree of improvements with thiazoidinediones (TZDs) is modest, estimated at approximately one-quarter of a grade per year for fibrosis.

Rakoski *et al.*[473] also noted that the glitazones have potential serious side effects. This point was emphasised, too, by the 2012 AASLD Practice Guidelines, which recommend that pioglitazone can be used to treat steatohepatitis in patients with NASH; however, the long-term safety and efficacy of the drug in patients with NASH is not yet established.

There are no RCTs published analysing treatment for metabolic conditions and the resulting impact on fibrosis in patients with NASH.[47]

The use of metformin was not recommended for use as specific treatment for liver disease in adults with NASH[47] and the EASL position paper[472] notes that controlled studies of metformin show no benefit resulting from this treatment.[487,488] Lavine *et al.*[490] conducted a randomised, double-blind, double-dummy, placebo-controlled clinical trial of 173 patients (aged 8–17 years) with biopsy-confirmed NALFD.[490] The study analysed whether children with NALFD would improve with use of vitamin E or metformin, and concluded that, compared with a placebo, neither therapy demonstrated significant improvements in histological features.

The 2012 AASLD guidelines[491] recommend that vitamin E (α-tocopherol) administered at daily dose of 800 IU/day for non-diabetic adults with biopsy-proven NASH may be considered as first-line pharmacotherapy. The recommendations are partially based on the results of the pioglitazone versus vitamin E versus placebo for the treatment of nondiabetic patients with non-alcoholic steatohepatitis (PIVENS) trial,[481] which found that although there was an improvement in hepatic steatosis and lobular inflammation, fibrosis scores were not significantly improved by use of vitamin E. The AASLD guidelines do not give much detail around the background to this decision; rather, they note that the primary end point of the PIVENS trial was an improvement in NAS $\geq$ 2 points, with at least 1-point improvement in hepatocellular ballooning and 1-point improvement in either the lobular inflammation or the steatosis score, and no increase in the fibrosis score. It seems that the recommendation was based on the improvement in liver histology seen during the PIVENS trial (but not necessarily an improvement in the fibrosis score).

The 2012 cost–utility study by Mahady *et al.*[475] also derived a treatment effect (RR) for vitamin E treatment. This estimate was derived from the PIVENS trial; however, as noted above, the paper for the PIVENS trial[481] notes that fibrosis scores were not significantly improved with vitamin E treatment. The RR employed in the cost–utility study (1.35) indicates that very little benefit results from treatment.

A RCT by Promrat *et al.*[492] assessed the effect of weight loss on NASH and found that weight reduction achieved through lifestyle interventions (diet and exercise programme and behavioural change) led to an improvement in steatosis, lobular inflammation and ballooning injury (measured using the NASH histological activity score); however, there was no significant change in hepatic fibrosis after 1 year of the study intervention. The study included patients with well-characterised NASH (histologically and clinically), and used a standardised, protocol-based lifestyle intervention. However, this was a small study (30 persons completed the study), a post-liver biopsy was carried out for only 90% of the participants and none of the participants had cirrhosis (16% had bridging fibrosis).[492] As the study was conducted for only 48 weeks, it may not be possible to confirm if lifestyle interventions are effective, as long-term data suggest that only

15% of participants lose > 10% of their body weight; in addition, adherence to weight-loss programmes drops after the first few months and most people regain weight.[472]

Bariatric surgery is indicated to be effective in patients with NASH and fibrosis.[479] NICE guidelines for obesity[493] recommend the use of bariatric surgery as a treatment option if patients have a BMI of ≥ 40 kg/m^2 or for patients with a BMI between 35 kg/m^2 and 40 kg/m^2 with other significant disease such as type 2 diabetes or high blood pressure. The guidelines also recommend this surgery as a first-line treatment option for adults with a BMI > 50 kg/m^2. However, the guidelines do not specifically recommend bariatric surgery for use in patients with NAFLD. In addition, the recent AASLD guidelines noted that it is premature to consider bariatric surgery as an established option to specifically treat NASH rather than obesity in general. A systematic review and meta-analysis by Mummadi *et al.*[479] concluded that fibrosis appears to improve after surgery-induced weight loss. Another 2008 study by De Freitas *et al.*[494] found that there is good evidence that bariatric surgery is associated with NAFLD regression in morbidly obese patients. However, the limitation of these studies is that they are retrospective or prospective studies. A 2010 Cochrane review[495] of bariatric surgery in obese patients with NASH did not locate any RCTs or quasi-RCTs that evaluated a bariatric procedure compared with another intervention in patients with NASH, and the authors concluded that this lack of RCT studies meant that they could not assess the benefits and harms of bariatric surgery as a potential treatment in patients with NASH.

Health-related quality of life in non-alcoholic fatty liver disease patients

The search for papers on treatment effectiveness also identified three studies that contained information on HRQoL in patients with NAFLD, two of which had reported HRQoL values.

The cost–utility study by Mahady *et al.*[475] could not locate any sources for HRQoL for patients with NASH, as no prior studies have been conducted. They used utilities from studies based on other causes of liver disease[432,435] and assumed that cirrhosis, decompensated cirrhosis and HCC represent a common pathway for liver disease and that the decrement in quality of life associated with these conditions is similar irrespective of the initial cause.

We located two papers which estimated HRQoL values in patients with NASH and NALFD. One was a study by David *et al.*[496] which assessed HRQoL of patients with diagnosed NAFLD and NASH. Using the SF-36, they found that patients with NALFD had lower reported scores and greater degrees of physical limitations than patients with HBV or HCV. They noted that the physical component summary score was similar to that of patients with HBV in a decompensated cirrhosis health state. Participants were mainly non-Hispanic white (76%) with a college education (71%), and over half (58%) had an income over $50,000. HRQoL was lower in the respondents with NASH and the authors also found that scores were worse for persons with cirrhosis than without cirrhosis. The authors reported the median SF-36 physical component score by fibrosis level, but insufficient information was provided for the other components required to enable mapping to the preference-based SF-6D in order to calculate QALYs.

The second was a published study[497] which reported HRQoL data for liver disease health states for all aetiologies including NASH. However, this was based on clinical opinion rather than empirical data reported by patients. The authors surveyed 18 general practitioners (GPs) and 12 hepatologists (Scotland and England) using a questionnaire and a Delphi approach. In addition, this report did not detail on what scale the values were estimated, and so it was not possible to interpret the reported estimates.

Summary

Currently, no pharmacological treatments or surgical interventions are explicitly recommended for use in patients with NASH by UK guidelines. The main limitation with the current published studies of potential effective interventions for NALFD/NASH is that none of the studies reviewed collected robust data on effective treatments for patients with NASH and fibrosis. Some of the recommended treatment

interventions also include physical exercise and weight-loss programmes; these are also recommended for the treatment of obesity irrespective of the degree of fibrosis progression, and so it is difficult to identify the accurate impact on fibrosis regression as any impact may be incidental and may not be explicitly captured in the programme outcomes.

The lack of published studies with relevant data on treatment options administered specifically for patients with NASH with fibrosis limited the modelling approach for NASH. As we could not identify robust cost and QALY estimates or data on treatment effectiveness, we were unable to model the long-term treatment pathway if diagnosed as true positive, false positive, true negative or false negative and were, therefore, unable to use the same modelling approach used for the HBV and HCV analyses.

Approach to analysis

We constructed a probabilistic decision model to assess the cost per correct diagnosis of the applicable NILTs compared with liver biopsy. We also conducted an exploratory analysis to assess the possible implications of using the NILTs in a primary care setting to inform the referral pathway to tertiary care for patients with NASH.

Cost per correct diagnosis

The systematic review identified data for 35 NILTs for use in NASH (using a section of the meta-analysis equivalent to the proportion of the population with a METAVIR score of $\geq$ F3). The number of true-positive, false-positive, true-negative or false-negative outcomes reported for each NILT was extracted from the meta-analysis data (see *Chapter 4*). The average disease prevalence estimated from the meta-analysis data was 19%. We calculated the probability of each test returning a true-positive, false-positive, true-negative and false-negative result using the average disease prevalence and sensitivity and specificity estimates (see *Appendix 7*).

Five of the tests evaluated in the second stage of the analysis used a combined diagnostic cut-off threshold (low and high cut-offs) for staging fibrosis; the use of a combined threshold results in a number of indeterminate responses. We assumed that patients who had an indeterminate response were retested with a commonly used imaging modality, Fibroscan, based on availability and clinical practice (choice based on clinical advice). We did not choose patented, direct or indirect tests as several of the tests with a combined diagnostic cut-off were also in these categories; our clinical advisor advised us that the same type of test would not be repeated in practice.

Test costs

Costs of imaging modalities were sourced from published Department of Health reference costs.[427] Costs of direct and indirect serum marker were obtained from communication with finance departments based at the Royal Free Hospital and costs for patented serum markers were sourced directly from manufacturers and via communication with finance departments based at the Royal Free Hospital (see *Appendix 9*, *Table 78*). The cost of a percutaneous liver biopsy (see *Chapter 5* for further details on choice of liver biopsy) was sourced from published literature.[428] Where applicable, costs were inflated to 2012 prices using NHS inflation indices.[67] All NILT test costs are based on incremental costs and exclude the capital costs of the equipment. Test costs and sources are listed in *Appendix 9*.

Analysis

We assumed that liver biopsy as the reference standard had perfect sensitivity and specificity, implying that biopsy accurately diagnoses all healthy and unhealthy patients. Using a hypothetical cohort of 1000 patients with NAFLD and suspected of having liver fibrosis, we estimated the incremental cost associated with each test compared with the next best alternative. Prevalence was based on mean prevalence data from the systematic review, which found that 19% of people tested had fibrosis level 3 disease or greater. Tests which had fewer numbers of true-positive or true-negative outcomes, with a higher test cost than other tests which had higher correct diagnoses, were ruled out of the analysis (dominated). We present the results separately for positive and negative diagnoses as the consequences are different for each. Results are shown in *Tables 51* and *52*.

TABLE 51 Results of cost per correct diagnosis (TP)

Test	Number of TPs using NILT	Test cost, £	Number of incremental correct diagnoses	Incremental cost (test only), £	Incremental cost per correct diagnosis (TP): each test compared with next best alternative (£/correct diagnosis gained)
NFS TE	15	55.95	–	–	Dominated
FIB-4 (high cut-off)	71	4.40	–	–	Dominated
Fibrotest TE	74	94.60	–	–	Dominated
NFS (high cut-off)	75	4.95	–	–	Dominated
APRI	76	4.05	–	–	Dominated
Fibrotest (high cut-off)	76	43.60	–	–	Dominated
AST–ALT (high cut-off)	88	0.90	–	–	Dominated
PLT	119	3.50	–	–	Dominated
Age–PLT Index	124	3.50	–	–	Dominated
NFS all	134	20.85	–	–	Dominated
Hepascore	143	16.24	–	–	Dominated
AST–ALT (low cut-off)	149	0.90	–	–	Dominated
FIB-4 all	149	21.09	–	–	Dominated
Type IV collagen	150	20.00	–	–	Dominated
ELF	151	108.00	–	–	Dominated
NFS (low cut-off)	151	4.95	–	–	Dominated
TE	155	51.00	–	–	Dominated
NAFIC (high cut-off)	158	28.17	–	–	Dominated
Fibrotest all	158	59.31	–	–	Dominated
FIB-4 (low cut-off)	159	4.40	–	–	Dominated
BARD	160	0.90	–	–	–
NFS ELF (high cut-off)	164	112.95	–	–	Dominated
Hyaluronic acid	165	8.00	6	7.10	1.27
NDP	166	21.18	–	–	Extendedly dominated
Fibrotest (low cut-off)	169	43.60	–	–	Dominated
ARFI	170	51.00	–	–	Dominated

continued

TABLE 51 Results of cost per correct diagnosis (TP) (*continued*)

Test	Number of TPs using NILT	Test cost, £	Number of incremental correct diagnoses	Incremental cost (test only), £	Incremental cost per correct diagnosis (TP): each test compared with next best alternative (£/correct diagnosis gained)
NFS ELF all	171	114.81	–	–	Dominated
MR elastography	172	199.00	–	–	Dominated
NFS ELF (low cut-off)	172	112.95	–	–	Dominated
NAFIC all	180	35.59	–	–	Dominated
NAFIC (low cut-off)	181	28.17	16	20.17	1.29
Liver biopsy	189	956.61	8	928.44	112.30

NAFIC, ferritin, fasting insulin, type IV collagen; NDP, NAFLD, diagnostic panel; NFS, non-alcoholic fatty liver disease fibrosis score; PLT, platelet; TP, true positive.

TABLE 52 Results of cost per correct diagnosis (TN)

Test	Number of TNs using NILT	Test cost, £	Number of incremental correct diagnoses	Incremental cost (test only), £	Incremental cost per correct diagnosis (TN): each test compared with next best alternative (£/correct diagnosis gained)
BARD	491	0.90	–	–	Dominated
NFS (low cut-off)	535	4.95	–	–	Dominated
NAFIC (low cut-off)	545	28.17	–	–	Dominated
NDP	566	21.18	–	–	Dominated
AST–ALT (low cut-off)	568	0.90	–	–	Dominated
Fibrotest (low cut-off)	593	43.60	–	–	Dominated
FIB-4 (low cut-off)	603	4.40	–	–	Dominated
PLT	617	3.50	–	–	Dominated
Age–PLT Index	632	3.50	–	–	Dominated
NAFIC all	640	35.59	–	–	Dominated
Type IV collagen	650	20.00	–	–	Dominated
NAFIC (high cut-off)	666	28.17	–	–	Dominated
Hyaluronic acid	666	8.00	–	–	Dominated
APRI	668	4.05	–	–	Dominated

TABLE 52 Results of cost per correct diagnosis (TN) (*continued*)

Test	Number of TNs using NILT	Test cost, £	Number of incremental correct diagnoses	Incremental cost (test only), £	Incremental cost per correct diagnosis (TN): each test compared with next best alternative (£/correct diagnosis gained)
TE	681	51.00	–	–	Dominated
Hepascore	682	16.24	–	–	Dominated
MR elastography	715	199.00	–	–	Dominated
ARFI	726	51.00	–	–	Dominated
ELF	730	108.00	–	–	Dominated
AST–ALT (high cut-off)	740	0.90	–	–	–
FIB-4 all	754	21.09	–	–	Dominated
NFS ELF (low cut-off)	778	112.95	–	–	Dominated
Fibrotest (high cut-off)	779	43.60	–	–	Dominated
NFS all	780	20.85	–	–	Dominated
Fibrotest TE	780	94.60	–	–	Dominated
Fibrotest all	783	59.31	–	–	Dominated
FIB-4 (high cut-off)	783	4.40	44	3.50	0.08
NFS (high cut-off)	786	4.95	3	0.55	0.21
NFS TE	795	55.95	9	51.00	5.53
NFS ELF all	805	114.81	–	–	Dominated
NFS ELF (high cut-off)	805	112.95	10	57.00	5.72
Liver biopsy	811	956.61	6	843.66	145.39

NAFIC, ferritin, fasting insulin, type IV collagen; NFS, non-alcoholic fatty liver disease fibrosis score; NPD, NAFLD diagnostic panel; PLT, platelet; TN, true negative.

Results of cost per correct diagnosis

Cost per correct diagnosis (positive): see *Table 51*

Column 2 in the table shows the number of correct positive diagnoses using a NILT or liver biopsy. The highest number of correct diagnoses is given using a liver biopsy ($n = 189$), and is assumed to be 100% accurate for all of those tested and which is estimated from the disease prevalence from the meta-analysis. An indirect serum marker, NAFIC (ferritin, fasting insulin, type IV collagen), using a combined cut-off threshold or a low cut-off threshold also returns a high number of true positives ($n = 180$ and $n = 181$, respectively). One test, an indirect serum marker, was extendedly dominated, as the incremental cost per correct diagnosis was higher for this test (£26.65) than for another indirect serum marker, NAFIC with a low diagnostic cut-off threshold, when compared with the next best alternative, hyaluronic acid. The incremental cost per correct positive diagnosis for liver biopsy was £112.30, compared with an incremental cost per correct diagnosis of £1.29 associated with using the next best alternative, NAFIC (low cut-off), as the diagnostic test.

Cost per correct diagnosis (negative): see *Table 52*

Liver biopsy returns the highest number of correct negative diagnoses (811) estimated from the disease prevalence from the meta-analysis. The majority of tests were dominated by other tests which had lower tests costs and returned higher numbers of true-negative results. A combination of NAFLD fibrosis score (NFS)–ELF test using a combined cut-off threshold returned 805 negative diagnoses at an additional cost of £5.72 per correct negative diagnosis compared with the next best alternative, NFS combined with Fibroscan. Although liver biopsy returned the highest amount of negative responses, the cost per correct negative diagnosis was £145.39 compared with the next best alternative, NFS–ELF (high cut-off). The least expensive cost per correct diagnosis compared with the next best alternative was FIB-4 (high cut-off) at £0.08 per correct true-negative response identified.

Exploratory analysis

We conducted an exploratory analysis to assess the possible costs and benefits of using the NILTs in a primary care setting to inform referral of patients to tertiary care. We constructed a simple decision model analysing the referral pathway for patients with NAFLD who have suspected fibrosis or cirrhosis.

Given the lack of data available, this analysis can only be considered as exploratory. We estimate the potential impact of the strategies on the costs of treatment based on the assumptions below. Robust data were not available to estimate the impact of the alternative strategies on health outcomes; however, we also present a sensitivity analysis exploring a range of alternative assumptions about possible impacts using a range of QALY estimates.

Scenarios assessed

Our considered population was a hypothetical cohort of 1000 patients with NAFLD with suspected fibrosis. Our analysis considered a number of scenarios, which were determined using clinical advice.

1. Immediate referral of all patients having NAFLD (irrespective of fibrosis level) to a tertiary care centre for testing, treatment and management.
2. At primary care level, test all patients with NAFLD with a NILT. If the Kleiner score is $\geq$ F3 (advanced fibrosis), refer patients to tertiary care centre for treatment and management. If the NILT score indicated low risk for advance fibrosis (Kleiner $<$ F3), treat and manage care in primary care setting.
3. Biopsy all patients, treat and manage all patients with a Kleiner score of $\geq$ F3 at a tertiary care centre. Refer those with a Kleiner score of $<$ F3 for treatment and management in primary care.

Strategy 1 was considered to most closely reflect current UK practice. We assumed that scenarios 1 and 2 would incorporate an initial diagnosis with a NILT. Following clinical advice, we modelled the use of an indirect serum marker test (FIB-4 with a combined diagnostic threshold cut-off). As this test may return a number of indeterminate results (from the meta-analysis data of the studies for FIB-4 combined, 33% of the test results would be inconclusive), we assumed that this proportion of patients would receive a second test with an imaging modality, Fibroscan, to confirm a diagnosis. Using these two tests returned an outcome where 866 patients were diagnosed as low risk and 134 patients were deemed high risk: true negative (86%), false negative (1%), true positive (4%) and false positive (9%).

Input parameters

Prevalence of population with fibrosis level of $\geq$ F3

Following clinical advice we set the average disease prevalence to 5%. We assumed that the sensitivity and specificity of the tests would be same as observed in the studies included in the systematic review in this population.

Resource use and cost

We did not identify any published studies in the literature review that provided long-term specific cost data related to treatment and management of care in patients with NASH and fibrosis. Based on the

recommended management of NASH (and potential pharmacological or surgical treatment listed in the EASL guidelines),[472] and using clinical advice, we identified resource use items for patients with NASH.

Assumptions regarding resource use

The time frame adopted in the analysis was 5 years and a discount rate of 3.5% was applied.[66]

High risk (≥ F3) treated in tertiary care

We assumed that this patient group would initially be tested with a NILT, FIB-4 (with inconclusive results retested with Fibroscan), in either a tertiary care setting (strategy 1) or a primary care setting (strategy 2). Using strategy 3, we assumed that patients would be tested with a liver biopsy in a primary care setting. Following clinical advice we assumed that patients with a METAVIR score of ≥ F3 would have two assessments per year with a consultant hepatologist. Exercise and diet programmes would be initiated and monitored by specialists (hepatologist, dietitian, physiotherapist and psychologist), with an assessment every 6 months. Treatment would involve aggressive management of metabolic syndrome components with statins, pioglitazones and antihypertensives. We modelled that a proportion of this cohort who test true positive (4%) would have a 0.04% probability of progressing to a cirrhotic health state per year.[475] Twenty per cent of patients who progressed to a cirrhotic health state would receive screening for HCC twice per year (using a combination of ultrasound and monitoring of α-fetoprotein levels). For those patients who progressed to HCC each year, we allowed for a 30-minute assessment by the hepatologist to feed back results (assuming a full assessment by a hepatologist would normally last 30 minutes, approximately). For the proportion who did not develop HCC, we allowed for 30 minutes of a hospital-based nurse's time to feed back results via letter.

We assumed that 3.5% of the patients who developed HCC would undergo liver transplantation (0.04% progression rate to HCC health state).[475]

National Institute for Health and Care Excellence guidance for obesity (CG43) recommends that persons with a BMI ≥ 40 kg/m^2 or with a BMI between 35 kg/m^2 and 40 kg/m^2 with other significant disease, who fill the following criteria, may be considered for bariatric surgery:

- individual has a BMI of ≥ 40 kg/m^2 OR individual has a BMI between 35 kg/m^2 and 40 kg/m^2 in addition to another significant disease (e.g. type 2 diabetes or high blood pressure) that could be improved if they lost weight
- all appropriate non-surgical measures have been tried but have failed to achieve or maintain adequate, clinically beneficial weight loss for at least 6 months
- the person has been receiving or will receive intensive management in a specialist obesity service
- the person is generally fit for anaesthesia and surgery
- the person commits to the need for long-term follow-up.

The guidelines also recommend that persons with a BMI > 50 kg/m^2 may be recommended for bariatric surgery as a first-line treatment.[493] In the absence of data on the percentage of persons who may be offered surgery as additional treatment if all other options have failed, we built into our model that a small proportion of the high-risk cohort will have bariatric surgery. We sourced this proportion using the costing report for the NICE clinical guidelines for obesity.[493] The costing report for the obesity guidelines assumed that of all adults who are obese nationally, 1% of these would have a BMI of 50 kg/m^2, and of those adults who meet the criteria for bariatric surgery as a first-line intervention, 80% of cases would be considered appropriate for surgical intervention. The costing report assumed that for those cases in which bariatric surgery is indicated and appropriate, 100% would choose to have the surgical procedure. We employed the same assumption as that used by the costing report in or analysis.

Low risk (< F3) treated in tertiary care

We assumed that this patient group would receive an initial test with FIB-4 (with inconclusive results retested with Fibroscan) in a tertiary care setting (strategy 1). Following clinical advice we assumed that patients with a METAVIR score of < F3 would receive an initial assessment with a consultant hepatologist

and thereafter would receive one assessment per year. Exercise and diet programmes would be initiated and monitored by a hepatologist and 20% of this population cohort would see a dietitian (initial assessment and once per year thereafter). Treatment would also involve aggressive management of metabolic syndrome components with statins, pioglitazones and antihypertensives. This patient cohort would also have annual liver function tests administered by the clinical hepatologist at the yearly assessment. We assumed that an arbitrary proportion of this group (5%) would have an additional assessment by the hepatologist to feed back results. The remaining 95% would have results fed back via letter (we identified the resource use for this as 30 minutes of a hospital-based nurse's time). This group would also receive a retest with a NILT to check progression of fibrosis at 5 years. We conservatively assumed that for the annual liver function tests and the 5-year NILT retest, a hospital-based nurse would spend 30 minutes' administration time (non-face-to-face time) for each contact episode organising tests, sending samples to laboratories for analysis and compiling results.

Low risk (< F3) treated in primary care

We assumed that this patient group either would be tested with a NILT, FIB-4 (with inconclusive results retested with Fibroscan) in a primary care setting (strategy 2), or would initially be tested with a liver biopsy in a primary care setting (strategy 3). Following clinical advice we assumed that patients who have a METAVIR score of < F3 would receive one 30-minute assessment per year with a GP who would advise on and monitor diet and exercise programmes. As in tertiary care, treatment would also involve aggressive management of metabolic syndrome components with statins, pioglitazones and antihypertensives.
This patient cohort would also have annual liver function tests (administered by a practice nurse during a 30-minute assessment) and would receive a retest with a NILT to check progression of fibrosis at 5 years. We conservatively assumed that for the annual liver function tests and the 5-year NILT retest, a nurse based at a GP practice would spend 30 minutes' administration time for each contact episode (non-face-to-face time) organising tests, sending samples to laboratories for analysis and compiling results.

Proportion of patients receiving combination of drugs or drugs alone (for treatment of metabolic conditions)

Following clinical advice, we modelled that patients could receive more than one drug: pioglitazone, vitamin E, statins or antihypertensives. The proportion or combination would remain the same if patients were treated in a tertiary care or a primary care setting. Simvastatin was chosen as the drug of choice for statins as this is the most commonly prescribed statin.[498] Lisinopril was chosen as an antihypernsive drug as per NICE guidance CG137, which advised that the first-line drug of choice for hypertension should be an angiotensin-converting enzyme (ACE) inhibitor or an angiotensin II receptor blocker.[499] We assumed that the drug would be an ACE inhibitor. We sourced the dosage for pioglitazone and vitamin E suspension (alpha tocopheryl acetate) from the BNF.[419] It was assumed in this exploratory analysis that the formulation of the drugs prescribed would be the cheapest non-proprietary drug available. Estimates of the proportions taking each type of medication are described in *Table 53*; estimates were based on clinical opinion.

TABLE 53 Proportion of patients receiving one or combination of drugs for metabolic condition

Drug(s)	Tertiary care and primary care high risk (≥ F3), %	Tertiary care and primary care low risk (< F3), %
Statins only	25	40
Statins and pioglitazone	15	10
Statins and vitamin E	20	0
Antihypertensive only	5	20
Antihypertensive and pioglitazone	15	10
Antihypertensive and vitamin E	20	0

Costs associated with resource use

National published sources of unit cost data were applied to the estimates of resource use: BNF for costs of medication;[419] unit costs of health and social care published by the Personal Social Services Research Unit (PSSRU) for NHS staff costs;[67] and Department of Health reference costs[427] for costs of surgical interventions such as bariatric surgery. The cost of a liver transplant was taken from the CELT study[421] described in *Chapter 5*, using the incremental cost of transplantation for ALD patients from date of transplant followed up over 2 years. Test costs for FIB-4, Fibroscan and liver biopsy are listed in *Appendix 9*. We assumed that test costs would be the same regardless of administration setting (primary or tertiary). *Table 54* provides a list of all identified resource use and associated costs. Where required, costs were inflated to 2012 prices using NHS inflation indices.[67]

TABLE 54 Resource use and associated costs applicable in patients with NASH (£ 2011–12)

Resource use	Unit cost	Lower cost	Upper cost	Source
Tertiary care staff: first assessment				
Consultant hepatologist	216.00	143.00	251.00	Department of Health reference costs 2011–12 (Code 306)[427]
Dietitian	91.00	15.00	148.00	Department of Health reference costs 2011–12 (Code 654)[427]
Exercise: physiotherapist	49.00	37.00	64.00	Department of Health reference costs 2011–12 (Code 650)[427]
Behavioural treatment: psychologist	89.00	66.00	66.00	Department of Health reference costs 2011–12 (Code 656)[427]
Tertiary care staff: follow-up assessment				
Consultant hepatologist	187.00	98.00	271.00	Department of Health reference costs 2011–12 (Code 306)[427]
Dietitian	93.00	16.00	119.00	Department of Health reference costs 2011–12 (Code 654)[427]
Exercise: physiotherapist	43.00	24.00	24.00	Department of Health reference costs 2011–12 (Code 650)[427]
Behavioural treatment: psychologist	313.00	51.00	204.00	Department of Health reference costs 2011–12 (Code 656)[427]
Nurse: hospital based (hour, non-face-to-face contact)	41.00	35.00		PSSRU 2012 Table 14.3[67]
Primary care staff				
GP (per minute patient contact)	3.70	3.10	–	PSSRU 2012 Table 10.8B[67]
Practice nurse (hour, face-to-face contact)	53.00	45.00	–	PSSRU 2012 Table 10.6[67]
Practice nurse (hour, non-face-to-face contact)	41.00	35.00	–	PSSRU 2012 Table 10.6[67]
Medication cost (per year)				
Statin (simvastatin)	11.86	–	–	BNF (accessed 24 June 2013)[419]
Antihypertensive (lisinopril)	14.47	–	–	BNF (accessed 24 June 2013)[419]
Diabetes medication (pioglitazone)	515.5	–	–	BNF (accessed 24 June 2013)[419]
Vitamin E	797.60	–	–	BNF (accessed 24 June 2013)[419]

continued

TABLE 54 Resource use and associated costs applicable in patients with NASH (£ 2011–12) (*continued*)

Resource use	Unit cost	Lower cost	Upper cost	Source
Test costs				
US	51.00	32.00	62.00	Department of Health reference costs 2011–12 (Code RA23Z)[427]
Monitoring of α-fetoprotein levels	1.38	–	–	Royal Free, 8 July 2013, personal communication[a]
Screening for HCC	52.38	–	–	Assumed to be cost of US and monitoring of α-fetoprotein levels
Liver function tests (AST–ALT)	0.90	–	–	AST, ALT (Royal Free, 8 July 2013, personal communication)[a]
Surgical interventions		–	–	
Bariatric surgery	6,479	–	–	NICE guideline CG43: costing report[493]
Liver transplant surgery and post care	17,741	–	–	CELT study[424]

US, ultrasound.
a See *Appendix 9, Table 78.*

Analysis exploring impact on resource use

We carried out an initial analysis where we estimated a cost per person for each scenario over for a 5-year period. We discounted the costs and QALYs using the recommended UK discount rate of 3.5%.[66] The costs were based on the resource use identified and included assessment costs, drug costs, screening costs, surgical costs and test costs. The results have been disaggregated to show the cost of treatment for those who test low risk (negative test response, true negative or false negative) or high risk (positive test response, true positive or false positive) for each group. *Table 55* presents the results of the cost analysis for the three scenarios.

Scenario 1 (refer all patients to tertiary care centre) and scenario 2 (refer only those who test ≥ F3 to tertiary care) result in similar resource use costs per person. The cost for the high-risk scenario is the same for both groups, as those diagnosed as positive in the primary care setting would be referred to a tertiary care setting. The cost for the low-risk group differs, with the cost of being treated in a primary care setting being slightly lower than that of being treated in a tertiary care setting for patients deemed low risk.

TABLE 55 Resource costs per scenario (£ 2012)

	Scenario 1			Scenario 2			Scenario 3		
Year	Cost low risk	Cost high risk	Cost total	Cost low risk	Cost high risk	Cost total	Cost low risk	Cost high risk	Cost total
Year 1	329	218	546	235	218	453	1148	129	1277
Year 2	277	227	504	211	227	438	182	85	267
Year 3	267	220	487	204	220	423	176	82	258
Year 4	258	212	471	197	212	409	170	79	250
Year 5	266	205	471	206	205	411	1000	77	1077
Total			2479			2135			3130

Allowing for a prevalence rate of 5% means that liver biopsy would accurately diagnose 95% of the cohort as true negative; in this scenario, fewer patients ($n = 50$) would be diagnosed and treated as high risk within a tertiary care centre than in scenario 1 or 2 ($n = 134$). However, liver biopsy has the most expensive resource cost per scenario; this was to be expected, as though we allowed for liver biopsy to have perfect sensitivity and specificity in the analysis, it also had the most expensive test costs (£957). We conducted a sensitivity analysis where we reduced of liver biopsy to that of Fibroscan and FIB-4 (£55.40) to determine the impact of a change in liver biopsy cost. This changed the results so that liver biopsy became the least costly scenario (£1482 per person).

The prevalence rate used in the exploratory analysis is 5% (as per clinical advice); if we amended this to 20% (as per systematic review data) the cost per arm would increase (£3170 per person for scenario 1, £2868 per person for scenario 2 and £4063 per person for scenario 3).

Analysis exploring impact on health outcomes

No data were available to accurately assess the impact of the referral pathway on the health outcomes of patients. We conducted a further tentative exploratory analysis to assess possible impacts on health outcomes. We estimated hypothetical QALYs for each test outcome (true negative, false negative, true positive and false positive) based on assumptions about possible health outcomes in patients with NASH.

We did not identify data on the expected lifetime QALY estimates of people with NAFLD. In the absence of available data, we assumed illustrative QALY values for patients with NAFLD based on the average lifetime QALYs of patients with HBV. The expected mean QALYs for each of the four type of diagnosis (true positive, false positive, true negative, false negative) from the HBeAg-negative model was assumed to represent the average QALYs for patients with NAFLD and corresponding diagnoses. It is recognised that the HRQoL and prognoses of patients with NAFLD may be lower in practice; however, data were not available to include in the analysis.

We assumed that the expected lifetime QALYs of patients diagnosed as true or false positive would be similar in a primary care setting as for patients diagnosed in tertiary care. A positive diagnosis in primary care would involve immediate referral to a tertiary centre for further investigation and so would be unlikely to lead to substantially different health outcomes. A difference in the health outcomes may be more likely for patients whose test results are negative. We assumed that the estimates of lifetime QALYs for people with true negative diagnoses treated in a tertiary care centre would be slightly higher than for those treated in a primary care setting. It is hypothesised that patients treated in tertiary centres may benefit from additional treatment such as access to behavioural therapists, physiotherapists and advice from dieticians, even if their disease has not reached a more advanced stage of progression. We hypothesise that the impact of management in primary care may be greater for patients with false-negative diagnoses, and that the health outcomes of these patients may be less in primary care than in tertiary care due to less intensive monitoring. We conducted an illustrative analysis where we assumed that the QALYs gained for the group of patients treated in primary care would either be 90% or 75% of the QALYs gained of comparable patients treated in the tertiary care setting (*Table 56*).

TABLE 56 Range of values for lifetime QALY for TN and FN patients

	Tertiary care		Primary care
Diagnostic test outcome	Lifetime QALY	Varied at 90%	Varied at 75%
FN	6.18	5.56	4.64
TN	13.12	11.81	9.84

FN, false negative; TN, true negative.

Using the assumed range of QALY values, we can estimate the lifetime QALY that may apply for scenarios 1 and 2, assuming that the NILTs used are FIB-4 combined, followed by Fibroscan as a retest for indeterminate results. The QALY gain for scenario 1 will remain constant at 13.00 as all patients will be treated within tertiary care. The results are shown in *Table 57*.

The largest decrease in QALYs for both scenarios 2 and 3 is when we assume that the quality of life is 25% lower in false negative patients and 10% lower in true-negative patients when treated in primary care as opposed to tertiary care. However, if we assume that true-negative patients have the same QALY gain regardless of setting but false-negative patients retain a QALY which is 25% less than in tertiary care, the difference in QALY gain between scenarios 1 and 2 is marginal. This implies that the true-negative response has a higher impact on the overall QALY gain. This is due to the fact that the testing strategy used (FIB-4, followed by Fibroscan retest for indeterminate patients) only has a 1% combined probability of predicting a false-negative result. Tests which have a higher probability of false-negative outcomes and a lower probability of true-negative outcomes may have different result when the values are varied.

Discussion

The lack of treatments specifically for the treatment of fibrosis progression in people with NAFLD meant that a different approach to modelling was required. We conducted an incremental cost per case detected to assess the relative cost-effectiveness of the tests. The results were analysed according to positive and negative diagnoses as the potential consequences of each are likely to be very different. The analysis of the incremental cost per correct positive diagnosis 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 (low cut-off) had an ICER of £1.29. The ICER for liver biopsy was £112.30, which means that it costs an additional £112.30 to obtain an additional correctly diagnosed positive result compared with the next best alternative. The analysis of the incremental cost per correct negative diagnosis found that FIB-4 (high cut-off) and NFS (high cut-off) had ICERs of below £1, the ICERs for NFS ELF 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. In addition, they do not take into account the potential negative impacts of biopsy on morbidity and mortality.

We conducted an exploratory analysis to tentatively assess the potential use of the tests to determine referral to tertiary care. A lack of data meant that as some of our inputs are arbitrary (frequency of appointments, exact resource use that would be incurred in primary care including access to specialists such as dieticians or behavioural therapists), we would expect the results to be sensitive to changes in assumptions or costs. The fact that they are confirms the need for well-designed long-term studies of patients with NASH, part of which would include collecting good-quality resource use data for use in economic models and further analysis.

TABLE 57 Results of exploratory analysis of referral pathways on expected QALYs

	Expected mean QALYs	
Diagnostic test outcome	Scenario 2	Scenario 3
High impact on FN (75%) and TN (90%)	11.86	11.55
High impact on FN (75%) and low impact on TN (100%)	12.98	12.80
Low impact on FN (90%) and low impact on TN (100%)	12.99	12.80
Low impact on FN (90%) and high impact on TN (90%)	11.87	11.55

FN, false negative; TN, true negative.

The QALY analysis is based on assumptions as to date there have been no studies that have collected HRQoL using a comparable measure such as an EQ-5D in patients with NAFLD. The lack of robust QALY data is also a limitation. In the absence of viable data, we have based the analysis on results for people with HBV. This assumption may have underestimated the decrement in HRQoL for patients with NASH.

One of the main limitations with regards to modelling NASH is the lack of effectiveness data relating to the impact of treatment on fibrosis in patients with NASH. Well-designed prospective studies of patients with NAFLD who are followed for long periods are lacking.[472] This lack of data does not allow us to model the long-term care pathway for NASH patients.

Treatment is often indicated for other conditions (diabetes, obesity, cardiovascular diseases) and, indeed, most people die earlier from cardiovascular disease-related comorbidities than liver-related disease.[500] It is also hard to separate out why treatment has sometimes been given (e.g. bariatric surgery, liver biopsies are performed during surgery but surgery is not necessarily carried out for a liver disease-related cause).

Although we did identify a recently published cost–utility analysis for treatment in patients with NASH, we did not feel that the existing published data of potential treatments and their effect on liver fibrosis in patients with NASH were of sufficient quality or quantity for use in a detailed economic model. Indeed, the data derived by Mahady *et al.*[489] indicate that both treatments analysed (pioglitazone and vitamin E) had only a modest effect on liver fibrosis progression.

Well-designed long-term prospective RCTs are required in patients with NASH to capture the impact of treatment and its progression on disease, to obtain valid estimates of quality of life and to obtain good-quality long-term costing estimates for treatment and management of patients with NASH.

Chapter 9 Cost-effectiveness analysis: cirrhosis

This chapter reports the cost-effectiveness analysis for cirrhosis and contains details of the literature review, modelling approach and results. The population of interest are those patients suspected of having cirrhosis, irrespective of aetiology.

Literature review

Background

Cirrhosis is the end stage of every chronic liver disease. The development of HCC may accelerate the course of the disease at any stage. Cirrhosis comprises two health states: compensated and decompensated. Decompensated cirrhosis is defined by the presence of ascites, variceal bleeding, encephalopathy and/or jaundice. HCC develops at a rate of approximately 3% per year in people with cirrhosis.[501]

We searched the NICE guidelines website for national guidance on the management and treatment of cirrhosis (applicable for all aetiologies). NICE guidelines for cirrhosis are currently in development. We also searched for recent EASL guidelines for best-practice recommendations for the treatment of cirrhosis. EASL guidance and position papers for HBV, HCV and ALD[37,45] recommend long-term monitoring for HCC in patients with cirrhosis.

Literature search

A search for relevant literature pertaining to costs and quality-of-life data for cirrhosis was conducted using the MEDLINE database (via Ovid platform, search date 26 July 2013). To search for literature related to cost data for cirrhosis we used the search terms 'cirrhosis' and 'costs' or 'Costs and Cost Analysis', and for literature relating to quality-of-life data we used the search terms 'cirrhosis' and 'Quality of Life'. The search for cost data returned 413 papers and the search for quality-of-life data returned 739 papers. We identified four papers that contained relevant data.[456,502–504]

The study by Thompson Coon *et al.*[456] reported the clinical effectiveness, cost-effectiveness and cost–utility of surveillance of HCC in patients with cirrhosis, using serum α-fetoprotein testing and/or ultrasound examination. They allowed for treatment with liver transplantation or resection. The authors[456] conducted an analysis which analysed three aetiologies (ALD, HBV and HCV) separately but also produced a mixed cohort weighted according to proportions of persons with ALD (57.6%), HBV (7.3%) and HCV (35.1%), and, using these models, estimated lifetime costs for various HCC surveillance strategies. Thompson Coon *et al.*[456] concluded that the most cost-effective strategy to adopt for screening for HCC in patients with cirrhosis (mixed aetiology) was to conduct screening using ultrasound and serum α-fetoprotein testing on a 6-monthly basis.

The 2008 paper identified by Andersson *et al.*[502] looked at a similar screening programme to Thompson Coon *et al.*[456] within a US setting. They found that semiannual ultrasound surveillance was a cost-effective strategy that improved outcomes at a reasonable cost. Bolondi *et al.*[503] reported details of an Italian cost-effectiveness analysis of a programme of monitoring HCC using ultrasound and α-fetoprotein testing. The analysis was conducted alongside a study in which a cohort of patients with liver cirrhosis ($n = 313$) received monitoring and a cohort of similar patients who acted as a control ($n = 104$). The authors concluded that the surveillance programme was not a good use of health-care resources.[503] Saab *et al.*[504] reported a decision-analytic model to compare the cost-effectiveness of ultrasound, α-fetoprotein with ultrasound and CT from a US (Medicare) perspective. The authors concluded that ultrasound is the most cost-effective strategy.

Of all the studies identified, the analysis by Thompson Coon *et al.*[456] was most relevant to NHS practice; however, none examined the use of the NILTs to determine which patients enter programmes of monitoring or screening for HCC. We therefore conducted our own analysis of the cost-effective use of the NILTs in this setting.

Cost-effectiveness analysis approach

We constructed a decision tree model to estimate the cost-effectiveness of NILTs in diagnosing cirrhosis. The population were persons who were suspected of having cirrhosis. The analysis adopted a time horizon of 4 years. The systematic review for NILTs for use in patients with cirrhosis identified 59 applicable tests (see *Chapter 4*).

Model structure

The decision tree was based on the recommended management for patients with cirrhosis, which includes screening for oesophageal varices (and ascites) and HCC. Patients received an initial NILT to diagnose the presence of cirrhosis. Screening would occur only if the test returned a positive outcome (true positive or false positive).

Model inputs

Average disease prevalence and test outcome

We estimated the average disease prevalence based on the data from the meta-analysis of the NILTs (20%). We calculated the probability of each NILT returning a true positive, true negative, false negative or false positive outcome using the sensitivity and specificity estimates from the meta-analysis data (see *Chapter 4*) and the estimated average disease prevalence. The probability of each NILT returning a specific diagnostic test outcome is listed in *Appendix 7*.

Screening tests and frequency

Following clinical advice, a large study of a surveillance programme for HCC in cirrhosis patients was identified.[505] This was a RCT analysing screening for HCC (using α-fetoprotein testing and ultrasound examination every 6 months). Patients were allocated either to screening with an α-fetoprotein test and ultrasound examination every 6 months or to no screening. The sample size was large (9373 in screening group and 9443 in control group). The authors found that biannual screening with ultrasound examination and α-fetoprotein testing was associated with a reduction in mortality (rate ratio 0.63; 95% CI 0.41 to 0.98), compared with no screening. We used this study as the source for effectiveness data related to screening (as a reduction in HCC mortality of 37%). As our clinical effectiveness data were based on the combination of the two tests used 6-monthly, we used the same screening strategy in our model. We sourced costs for the screening tests from national sources,[427] and via personal communication with finance departments at the Royal Free Hospital (see *Appendix 9*, *Table 78*). We sourced additional costs of screening from the study by Thompson Coon *et al.*[456]

Mortality from hepatocellular cancer: adjusted for screening and no screening

The study by Zhang *et al.*[505] found that, for patients who were screened for HCC, there was a 37% reduction in mortality from HCC compared with the non-screened group. We factored this into our model; we estimated that a diagnostic test outcome of false negative (no screening programme) would have a QALY value which was 37% lower than a diagnostic test outcome of true positive (screening programme).

Test costs

The cost of the NILTs were sourced from Department of Health reference costs 2011–12,[427] personal communication with finance departments based at the Royal Free Hospital and communication with manufactures of patented serum markers (see *Appendix 9*, *Table 78*). The cost for liver biopsy was sourced from published literature.[428] Where required, costs were inflated to 2012 prices using NHS inflation indices.[67]

Costs and quality-adjusted life-year end points of decision tree

We used the average costs and QALY values over a 4-year period sourced from the model developed for HBeAg-positive disease. We used this cohort as we sourced the reduction in mortality rate with screening from the study by Zhang *et al.*[505] who had conducted the RCT in a population with HBV. The time period was also defined by the Zhang *et al.*[505] study, which recruited patients over a 2-year period and with a follow-up period of up to 5 years (average time period of follow-up was 3–5 years).

We used the true-positive Markov model for HBeAg-positive as we assumed that patients would receive usual standard of care (including antiviral agents if applicable) alongside screening or no-screening programmes. Costs and QALYs values were estimated from the model for a 4-year period. These estimates also captured the progression through health states (including progression from moderate or cirrhotic health states to more advanced liver disease states including decompensated cirrhosis). The costs and QALYs also captured the probability of a patient moving from a compensated cirrhosis health state to a decompensated cirrhosis health state. The costs and screening for oesophageal varices are included within the costs and transition probabilities for the compensated cirrhosis health state. This treatment pathway also captures treatment for HCC with liver transplant, including the post liver transplant state.

We assumed that patients who tested true negative or false positive would incur the costs and QALYs of patients in a moderate to advanced fibrosis health state (F2–3). As patients who test false positive also receive screening, we included the additional cost of screening every 6 months with ultrasound examination and α-fetoprotein testing over a 4-year period in this cost.

We assumed that patients who tested true positive or false negative would incur the costs and QALYs of patients in a cirrhotic health state (F4). Patients who tested false negative were assumed to not receive screening, and, therefore, the cost of 6-monthly screening with ultrasound examination and α-fetoprotein testing was excluded from their 4-year costs.

The cost and QALY end points used in the model are detailed in *Table 58*.

Analysis

We conducted a probabilistic analysis and compared the results of each non-invasive test and liver biopsy incrementally to determine the cost-effective testing approach. We also constructed a CEAC and CEAF of the results.

We conducted a sensitivity analysis where we allowed for an increase of 10% (arbitrary value) to apply to the QALY value for false positive patients, to reflect the treatment benefit that may apply from receiving screening (although not necessarily indicated in this patient cohort, there may be some benefit due to earlier diagnosis of HCC). We also conducted a second sensitivity analysis where we removed the 6-monthly cost of testing with α-fetoprotein.

TABLE 58 Cost and QALY end point decision tree (£ 2012)

Diagnostic test outcome	Cost	QALY
TP	29,913	1.91
FN	29,703	1.19
TN	23,665	2.25
FP	23,875	2.25

FN, false negative; FP, false positive; TN, true negative; TP, true positive.

Results

The results of the incremental analysis are listed in *Table 59*. The results show that the most cost-effective test to use would be Forns index (serum marker with a sensitivity of 100% and specificity of 74% in patients with cirrhosis), with an ICER of £1926.

The CEAF (*Figure 10*) shows that given a cost-effectiveness threshold value of £20,000, Forns index used alone has a 50% probability of being the optimal test (highest expected net benefit).The CEAC (see *Appendix 10*) shows that Forns index has the highest probability of being the most cost-effective test (for illustrative clarity, only tests which have a ≥ 5% probability of being cost-effective are shown in the CEAC).

A scatterplot illustrating the position of each testing strategy on the cost-effectiveness acceptability curve compared with the least costly testing strategy can be found in *Appendix 12*.

TABLE 59 Incremental analysis of NILTs for cirrhosis

Diagnostic test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
CDS (high cut-off)	24,908	2.082	–	–	Dominated
Fibrometer (high cut-off)	24,947	2.085	–	–	Dominated
Fibrosis Index	24,907	2.089	–	–	–
Hepascore (high cut-off)	24,920	2.090	–	–	Dominated
Lok's index (high cut-off)	24,918	2.091	–	–	Extendedly dominated
FIB-4 (high cut-off)	24,921	2.096	–	–	Dominated
APRI (high cut-off)	24,920	2.098	–	–	Extendedly dominated
AST–ALT ratio	24,929	2.105	–	–	Extendedly dominated
ELF (high cut-off)	25,024	2.108	–	–	Dominated
BARD	24,935	2.109	–	–	Extendedly dominated
Fibrotest	24,972	2.122	–	–	Dominated
GUCI	24,942	2.126	–	–	Dominated
Forns index (high cut-off)	24,933	2.130	–	–	Extendedly dominated
APRI (combined cut-off) and Fibroscan	24,942	2.133	–	–	Dominated
Hepascore (combined cut-off) and Fibroscan	24,945	2.135	–	–	Dominated
PIIINP	24,976	2.135	–	–	Dominated
Fibroindex	24,974	2.136	–	–	Dominated
Type IV collagen	24,973	2.136	–	–	Dominated
US SAPI	25,019	2.138	–	–	Dominated
Fibrotest (high cut-off)	24,967	2.138	–	–	Dominated
PLT	24,940	2.138	–	–	Dominated
Fibropaca	25,419	2.139	–	–	Dominated
Fibrometer	24,976	2.139	–	–	Dominated
US	24,982	2.139	–	–	Extendedly dominated
SAFE	25,628	2.140	–	–	Dominated

TABLE 59 Incremental analysis of NILTs for cirrhosis (*continued*)

Diagnostic test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Fibrotest (combined cut-off) and Fibroscan	24,979	2.141	–	–	Dominated
King's	24,940	2.141	–	–	Extendedly dominated
MR	25,119	2.141	–	–	Dominated
MRI	25,118	2.142	–	–	Dominated
APRI (low cut-off)	24,958	2.143	–	–	Dominated
PGAA	24,945	2.147	–	–	Extendedly dominated
Fontana	24,999	2.147	–	–	Dominated
FIB-4 (combined cut-off) and Fibroscan	24,951	2.148	–	–	Dominated
APRI	24,973	2.148	–	–	Dominated
Hepascore (low cut-off)	24,962	2.149	–	–	Dominated
FIB-4	24,960	2.149	–	–	Dominated
Hyaluronic acid	24,946	2.150	–	–	Dominated
Hepascore	24,961	2.151	–	–	Dominated
FIB-4 (low cut-off)	24,972	2.154	–	–	Dominated
Lok's index (low cut-off)	24,982	2.155	–	–	Dominated
ARFI	25,001	2.155	–	–	Dominated
PLT–Spleen ratio	24,998	2.157	–	–	Dominated
ELF (combined cut-off) and Fibroscan	25,059	2.158	–	–	Dominated
Bordeaux	25,021	2.159	–	–	Dominated
CEUS	25,046	2.161	–	–	Dominated
Fibrotest (low cut-off)	25,012	2.161	–	–	Dominated
Forns index (low cut-off)	25,031	2.161	–	–	Dominated
Age–Platelet Index	24,971	2.161	–	–	Dominated
CDS	24,983	2.162	–	–	Dominated
Fibroscan	24,988	2.162	–	–	Dominated
Fibrometer (combined cut-off) and Fibroscan	24,988	2.162	–	–	Dominated
DW-MRI	25,125	2.162	–	–	Dominated
CDS (low cut-off)	24,946	2.162	39	0.07	527
ELF (low cut-off)	25,070	2.164	–	–	Dominated
Forns index (combined cut-off) and Fibroscan	24,982	2.167	–	–	Dominated
ELF	25,050	2.168	–	–	Dominated
Fibrometer (low cut-off)	25,009	2.172	–	–	Dominated
MR elastography	25,127	2.177	–	–	Dominated
Forns index	24,975	2.177	29	0.01	1926

CDS, Cirrhosis Discriminant Score; DW-MRI, diffusion-weighted magnetic resonance imaging; PLT, platelet; US, ultrasound.

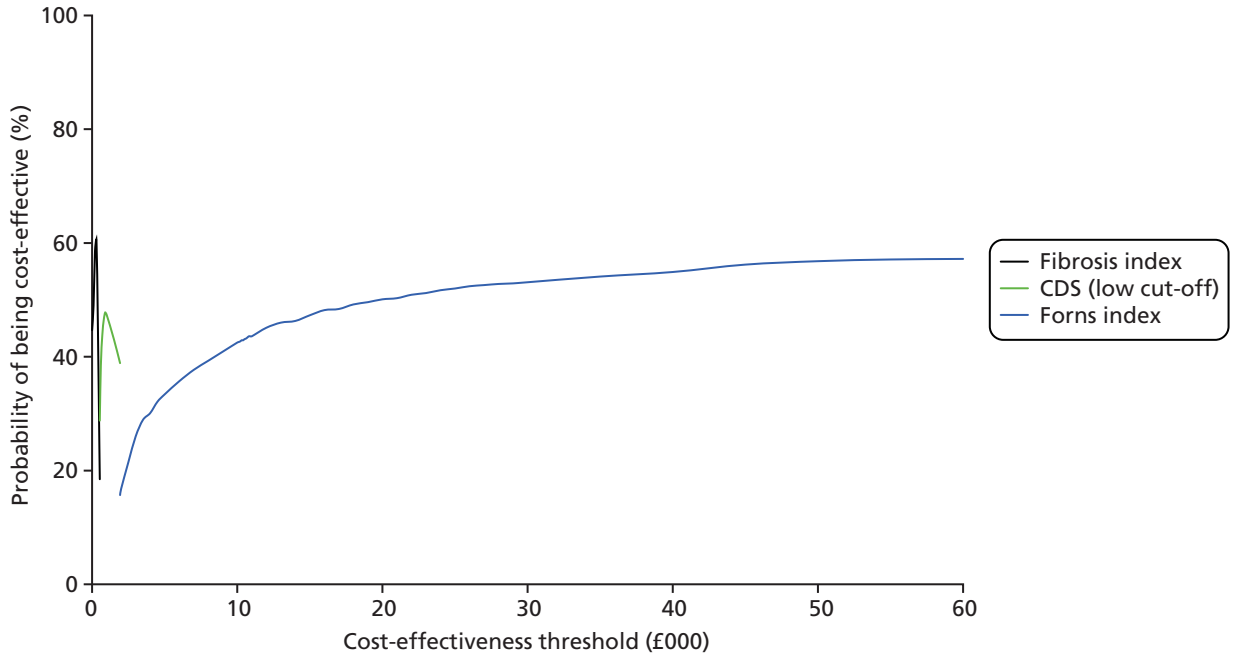

FIGURE 10 Cost-effectiveness acceptability frontier for cirrhosis. CDS, Cirrhosis Discriminant Score.

Sensitivity analysis

Allowing for a 10% increase to the QALY value for patients who are false positive does change the result so that using Forns index (low cut-off) becomes the most cost-effective with an ICER of £1106.

A test with a low diagnostic cut-off is more likely to pick up more positive results, as we are increasing the QALY outcome for patients who are false positive; it is more likely that tests with a low cut-off will have a higher QALY gain, as evidenced by results (see *Appendix 11*).

Removing the cost of α-fetoprotein from the analysis did not change the base-case results.

Discussion

The results imply that using an indirect serum marker, Forns index, which has a high sensitivity (100%) and a specificity of 74%, is the most cost-effective test, although this testing option is sensitive to changes in assumptions around benefit for those who receive treatment when diagnosed as false positive. It has been necessary to include several assumptions in order to conduct the analysis. These include that, in the absence of data on the impact of surveillance on HRQoL, the reduction in mortality from screening would translate directly into an equivalent reduction in QALYs.

The current analysis only takes a 4-year time horizon. Unfortunately, it was not possible to extrapolate the data reported in the RCT by Zhang *et al.* over a longer period.[505] If the health outcomes associated with positive diagnoses were estimated over a longer period, we would expect these to have a greater impact on the analysis, although this would be tempered by increased costs of surveillance. There is very little difference in the mid-term outcomes from the tests (cost and QALY values), implying that the difference between each test given a 4-year horizon is slight; extrapolation to lifetime would give a clearer picture as it would include the long-term impact of each test (long-term impact based on test diagnoses).

The analysis evaluates screening for HCC using α-fetoprotein testing and ultrasound examination as the clinical effectiveness data were based on a study which used this screening strategy; however, current practice may be to use ultrasound examination only. We examined the impact of removing the cost of α-fetoprotein testing in the sensitivity analysis; however, it may be the case that effectiveness is reduced with the exclusion of α-fetoprotein testing (likewise, less frequent screening would also reduce the clinical effectiveness).

Chapter 10 Discussion

We have comprehensively reviewed the accuracy and cost-effectiveness of NILTs of fibrosis and cirrhosis. We identified a substantial amount of evidence for the NILTs as a whole (114,071 potential papers identified, 302 relevant papers), and this is the first study to systematically evaluate and synthesise this evidence base. The NILTs vary substantially in terms of how they diagnose fibrosis severity, and in their cost. Some NILTs, such as the AST–ALT ratio, have been available in the NHS for a long time; however, recently there has been a rapid expansion in the introduction and use of new NILTs, which could put pressure on NHS resources.

The systematic review identified 41 NILTs for use in HCV, 21 NILTs for use in HBV, seven NILTs for use in ALD and 28 NILTs for use in NAFLD. Thirty-six NILTs were found for cirrhosis (irrespective of aetiology) and 14 imaging modalities were found that were applicable for all aetiologies. The highest number of studies identified was for HCV ($n = 162$). ALD had the lowest number of studies identified ($n = 12$). Fibroscan was the NILT assessed in most studies across disease aetiologies: 37 in HCV, 13 in HBV, eight in NAFLD and six studies in ALD.

Given limited health-care resources, the importance of considering the costs and benefits of competing interventions is now widely recognised in order to make best use of available resources. The consideration of cost-effectiveness is just as important for diagnostic tests as it is for treatments; in doing so it is necessary to consider the health and cost consequences as a result of the test. In relation to this decision problem, considering a cost per case detected would not give a complete picture as it would not reflect important differences in the consequences of an incorrect positive test compared with the consequences of an incorrect negative test. This does create a significant challenge for the cost-effectiveness analysis as the clinical studies typically focus on the accuracy of the tests rather than on effectiveness and health outcomes. We have attempted to overcome this challenge through the use of decision-analytic models and the synthesis of a range of data.

Our analyses have reflected the different causes of fibrosis and cirrhosis. Disease progression, care pathways and patient characteristics vary substantially between aetiologies; these have been analysed separately, and different modelling approaches have been required to account for these differences. For aetiologies HBV and HCV, active treatments are available which could (potentially) be instigated depending on the presence of a specific level of fibrosis, and there is a reasonable evidence base for these treatments. In these cases it is possible to reflect the potential consequences of the outcome of the fibrosis test within the analysis (i.e. the start of treatment or not). The situation for ALD and NAFLD is somewhat different as most interventions are aimed at behaviour change and would not depend on the outcome of the test (e.g. a patient presenting with a BMI of 30 kg/m^2 would likely be given dietary and exercise advice regardless of whether the NILT indicates level 1 or 3 fibrosis). There has been investigation of some pharmacological treatments of fibrosis in these patients, but generally the evidence is considered to be weak and they have not so far been included in the standard guidelines for treatment (e.g. EASL position paper for the treatment of patients with NAFLD). In practice, the pathway of care for these patients diagnosed with significant fibrosis or cirrhosis may include increased monitoring, screening for HCC and treatment of complications; however, evidence on the effectiveness of these overall packages of care is not available. The cost-effectiveness analyses of the NILTs in ALD is based on their potential impact on abstinence rates and draws heavily on a previously published HTA.[428] The analysis for NAFLD is limited to an incremental analysis of the cost per correct positive/negative diagnoses and exploratory analyses around the cost-effectiveness including longer-term outcomes. The cost-effective testing strategy for each aetiology is presented in *Table 60*.

TABLE 60 Cost-effective testing strategy for each aetiology

Aetiology	Cost-effective test for a threshold of £20,000 per QALY gained	Cost-effective test for a threshold of £30,000 per QALY gained
Incremental analysis *HBeAg-positive*		
Single test	GUCI	MR elastography
Sequential tests	Strategy 2: first NILT, hyaluronic acid; second NILT, MR elastography	MR elastography used singly
HBeAg-negative		
Single and sequential testing	No test or treatment	Treat all (no prior test)
Hepatitis C		
Single test	Treat all (no prior test)	Treat all (no prior test)
Sequential tests	Treat all (no prior test)	Treat all (no prior test)
ALD		
	Liver biopsy	Liver biopsy
Cirrhosis		
	Forns index	Forns index

	Cost-effective test for a threshold of £2 per correct diagnosis	Cost-effective test for a threshold of £10 per correct diagnosis	Cost-effective test for a threshold of £150 per correct diagnosis
Cost per correct diagnosis *NAFLD*			
True positive	NAFIC (low cut-off)	NAFIC (low cut-off)	Liver biopsy
True negative	NFS (high cut-off)	NFS ELF (high cut-off)	Liver biopsy

The results for HCV show that a strategy of treating all those suspected of fibrosis without testing is cost-effective. For conclusions regarding the population with HBV, the results were less clear. For patients with HBeAg-negative disease, the conclusion regarding cost-effectiveness depends on the specific cost-effectiveness threshold employed. In the UK, NICE specifies a cost-effectiveness threshold range of £20,000–30,000 per additional QALY, below which technologies are usually considered cost-effective and within which specific additional factors must be considered important. It is unclear which threshold is appropriate in this circumstance: we note that NICE has previously approved treatment for people with HBV with an ICER above the £20,000 lower bound;[417] however, we also acknowledge that recent research has suggested that the threshold for the NHS should be lower than that specified by NICE.[506]

Current NICE guidance for HCV recommends that all patients with mild HCV should receive treatment with antiviral agents.[443] Our analysis reinforces this recommendation; our findings indicate that that treating all patients irrespective of fibrosis stage without prior testing is the most cost-effective option. However, recent clinical practice guidelines published by EASL state that all treatment-naive patients with compensated chronic liver disease related to HCV should receive treatment (if willing).[507] The guidelines also argue that the timing of treatment in patients with minimal or no fibrosis is debatable and could be deferred pending the development and availability of new treatment (a strategy which should include regular assessment). One reason to defer would be due to the potential side effects of current triple therapy (boceprevir and telaprevir). We conducted a sensitivity analysis which incorporated a HRQoL decrement to represent potential side effects of current treatment; however, this had no effect on the

overall conclusions (see *Chapter 6*). We also conducted an exploratory analysis to evaluate potential new antiviral treatment (see *Chapter 6*), the results of which depended on the increase in treatment price. With this analysis, if treatment cost was increased by more than approximately £37,500 (with a corresponding increase in SVR rate: see *Chapter 6* for details), the strategy 'treat all' was not cost-effective; however, this analysis was exploratory and not based on actual data. The issue of whether or not treatment should be deferred until the arrival of new antiviral drugs cannot be answered conclusively given current data on efficacy and treatment cost.

Current NICE guidelines for HBV[374] recommend that antiviral treatment is considered for people with evidence of fibrosis following a liver biopsy, or following either a biopsy or diagnosis using Fibroscan for adults aged < 30 years. Our results for HBeAg-positive patients (though highly uncertain: see *Chapter 5*) found that the use of a NILT with treatment initiation if the patient tested positive was the most cost-effective option without the need for biopsy. The analysis focusing on only those tests where the bivariate model converged found that testing with Fibroscan to assess fibrosis level prior to treatment was the cost-effective option. This finding applied to all patients in the analysis and not just to young adults as in the NICE guidance. For HBeAg-negative patients, our findings were somewhat different from the current NICE guidelines. These found that treatment without prior testing was cost-effective if the upper bound of the NICE threshold was accepted, and no treatment if the lower cost-effectiveness threshold was considered to apply. Current EASL clinical practice guidelines for HBV recommend that all patients should receive treatment if they have HBV DNA levels > 2000 IU/ml and/or serum ALT levels above the ULN for the laboratory, and results from a liver biopsy or a non-invasive marker showing moderate to severe necroinflammation and/or fibrosis using a standardised scoring system (e.g. at least grade A2 or stage F2 by METAVIR scoring).[45] However, as noted above, our analysis shows that for HBeAg-negative patients, strategies without prior testing were the most cost-effective options.

The assessment of cost-effectiveness of the NILTs in HCV found that results were driven by the estimates of treatment effectiveness in this particular group. Treatments for fibrosis in these populations, as for people with HBV, have marketing authorisations for treatment of fibrosis regardless of METAVIR score, and patients with only mild fibrosis may benefit from early treatment.[443] The absolute benefit in terms of health outcomes may, however, not be as great as for patients with more severe levels of fibrosis, and the cost-effectiveness was uncertain. We assessed whether or not using the NILTs to target treatment to those with more severe fibrosis would be a cost-effective use of resources compared with a strategy of treating all those people with HBV or HCV and suspected fibrosis. The base-case analysis found that a scenario where everyone received early treatment was the most effective and cost-effective option, compared with using NILTs to target treatment for those with more severe fibrosis. However, when we conducted a sensitivity analysis around the assumption of treatment benefit in patients with mild fibrosis, treating everyone without a prior diagnostic test stopped being the most cost-effective option when we reduced treatment benefit by approximately 23%, after which it became cost-effective to use a NILT, MR elastography, which had high summary sensitivity and specificity in this population (94% and 92%). If the absolute benefit from treatment is not as high in patients with mild fibrosis as it is for patients with more advanced fibrosis, then it would be more advisable to target treatment using a NILT to identify those with advanced disease.

Given that increasingly sedentary lifestyles and changing dietary patterns mean that NAFLD poses a significant health problem, we strongly recommend that these evidence gaps are addressed. Currently, fatty liver accounts for 0.1% of all deaths in England (648 deaths annually). Fatty liver is also an underlying contributory cause for 1801 deaths per year, which is higher than for any other cause of liver disease. The prevalence of NALFD is also increasing in children, with 10–77% among those who are obese,[508] and its presence is associated with progressive liver disease including cirrhosis, which could lead to the need for liver transplantation. The implications of the increasing extent of NAFLD within the population, and in particular within the paediatric population, suggest that this will place an ever-increasing burden on the NHS.

Long-term prospective studies of target-based interventions are required in this population (both adult and paediatric) to determine effective treatments to halt the progression of fibrosis and subsequently limit the encumbrance of this disease on the health-care system.

Additionally, in NAFLD, non-invasive tests differentiate significant fibrosis but not steatohepatitis from simple steatosis. Whereas simple steatosis is usually benign, steatohepatitis could potentially progress to end-stage liver disease. Steatohepatitis is not necessarily characterised by fibrosis, meaning that it could potentially be missed and, therefore, patients and doctors could be falsely reassured. Non-invasive assays that differentiate steatohepatitis from steatosis are based in apoptosis rather than fibrosis[509] and have not been adequately validated. Subsequent development and validation of such markers is warranted given the increasing prevalence of NAFLD in the general population. Their assessment was beyond the scope of this review.

Strengths of analysis

Our meta-analysis of NILTs has been the most detailed and extensive to date, including all described serum tests and imaging modalities with no language restrictions and using state-of-the-art statistical and reporting methods. A similar recent study by Chou *et al.*[510] conducted a systematic review evaluating the diagnostic accuracy of blood tests to identify fibrosis or cirrhosis in patients with HCV infection. However, this study restricted its analysis to blood tests for fibrosis, searched fewer databases and excluded studies not in the English language. The study included serum NILTs but did not include imaging modalities such as Fibroscan in HCV. In addition, the study did not attempt to estimate the cost-effectiveness of these tests for use in HCV. One of the main strengths of our analysis is that we have analysed the cost-effectiveness of these tests in both HBV and HCV and we have based this cost-effectiveness on the long-term impact resulting from a diagnostic test outcome rather than just basing it on the cost of the test itself.

Limitations of analysis

Although the totality of the evidence was substantial for the NILTs, there was considerable variation in the amount and quality of the evidence for individual tests. A great number of NILTs for specific fibrosis stages either were assessed in single studies or had results that did not converge using the random-effects model with correlation between sensitivity and specificity. We assessed the impact of this on the conclusions from the cost-effectiveness analyses through sensitivity analyses including only those where the most robust models of test accuracy were possible. This reduced the number of tests analysed for HBV by 20 and for HCV by 43. The number of tests excluded emphasises how many NILTs had diagnostic accuracy data that were not robust. When we removed the tests, the results changed for HBeAg-positive model only.

Furthermore, only five of the included studies or 1.6% were of high methodological quality; therefore, all results are likely to be biased. This implies that further studies with better design are warranted to increase the robustness of the results.

Finally, cut-offs of NILTs for specific fibrosis stages varied in published studies and were not always predetermined or sufficiently validated. This is similar to measuring renal function with serum creatinine but not knowing the exact ULN. Apart from the practical issue of applying the NILTs in clinical practice with uncertain cut-offs, this resulted in overestimation of diagnostic accuracy of NILTs in the meta-analysis in all cases where there is a range of cut-offs.

The number of data available varied considerably between aetiologies. For example, for NAFLD we identified a position paper by EASL[472] and a practice guideline by AASLD.[47] Current treatment recommended in both papers included weight-loss programmes and treatment to ameliorate the metabolic conditions associated with NAFLD. However, we found insufficient evidence around either lifestyle

interventions or effective pharmacological treatments directed at the liver to prevent fibrosis progression specifically in patients with NASH.

Uninterpretable and indeterminate results were reported in < 50% of studies; however, they were infrequent in serum non-invasive tests (<1%) and could be considered negligible. In the case of Fibroscan, uninterpretable results were prevalent in 8–10% of examinations; this is probably an underestimated failure rate, as not all studies reported on failures while others included failures in the exclusion criteria. In a prospective study of over 10,000 Fibroscan examinations, unreliable results were reported in 15.8% of cases and were associated with obesity, age > 52 years, operator experience and presence of diabetes.[511] This significant failure rate is an important caveat in the use of Fibroscan. This NILT was the most cost-effective option when tests which did not converge (using bivariate model) were excluded from the analysis for HCV and HBeAg (positive), where the cost-effectiveness threshold was £20,000 and £30,000, respectively. However, as we did not account for indeterminate Fibroscan results in our economic analysis, its cost-effectiveness is likely to have been overestimated.

The searches for the systematic review were conducted in April 2012; as the research took place over a 2-year time period, it was not possible to conduct an updated search. This may mean that we have missed some diagnostic accuracy studies published since April 2012. However, given the robust findings for some of the aetiologies, for example that the 'treat all' strategy remained robust to the majority of sensitivity analyses for patients with HCV, any additional studies may not have a substantial impact on our analysis.

Currently, most studies report the results for diagnostic test accuracy using a 2 × 2 classification matrix, which restricts test results to either positive or negative.[512] Using these data, we can predict the summary sensitivity and specificity data required to summarise the diagnostic accuracy for each NILT. We initially attempted to construct the models for HBV and HCV using a 3 × 3 classification matrix to allow for the multiple categories of the METAVIR staging system. The restriction of studies reporting in a 2 × 2 format did not allow us to estimate with precision the proportion of people who tested incorrectly who may have had a higher or lower degree of disease. For example, if a non-invasive test returns a certain number of false-positive outcomes, and we used a section of the data which represented a by METAVIR score of F3, it is not possible from the data provided in the studies to determine the number of persons who actually have mild (F1) or moderate (F2) fibrosis. Likewise, if we use a F2-by-METAVIR section of the data, although we can estimate the probability of a test reporting a false-negative result, we do not know the split within this false-negative result that is applicable to either advanced disease (F3) or cirrhosis (F4).

A 2012 paper by Schuetz *et al.*[513] examined whether a 3 × 2 classification matrix is better than the 2 × 2 classification matrix when assessing diagnostic accuracy. They found that the parameters for diagnostic accuracy (summary and sensitivity estimates) decrease significantly if a 3 × 2 table is used. As there was a lack of studies reporting data using a 3 × 2 classification matrices, we used a F2 by METAVIR section of the data for our HBV and HCV models. Using this meant that we could not model the population cohorts who are diagnosed as F2 or F3 separately. Although we could identify the prevalence of patients with a METAVIR score of F4 from the meta-analysis data, we could not with accuracy identify the same for F3 as the reported data for F3 also included F4 scores (i.e. this section of the data represented diagnostic accuracy from the study for METAVIR scores which were ≥ F3).

Our findings, though UK based, should be applicable to health systems where the treatment pathway for patients with liver disease is similar. However, our results may not be generalisable to other settings, particularly resource-poor countries especially where the finding is to treat all patients irrespective of disease level.

A transferability issue may also arise over the estimated prevalence used in the analysis. The prevalence was estimated from the studies found during the systematic review; however, these were conducted in tertiary care settings, mainly in countries where the populations might be very different from those in countries with a lower level of health care.

Liver biopsy, although the reference standard for fibrosis assessment, is not 100% sensitive and specific, due to sample and intra- and interobserver variability. Moreover, it assesses fibrosis semiquantitatively and histological scores include both description of fibrosis and architecture. The misclassification rate of liver biopsy is the source of the myth that non-invasive fibrosis tests cannot achieve a high concordance with histological stages. However, serum non-invasive fibrosis markers have been developed and calibrated with direct reference to a set of liver biopsies. Therefore, the perfect serum marker in this case would replicate the 'golden' histological standard and could theoretically reach an AUROC of 1, replicating even the misclassifications of a liver biopsy. Imaging modalities, such as Fibroscan, that have been developed independently of liver biopsy, could potentially be affected by the misclassification rate of a liver biopsy. A potential solution would be to validate NILTs against clinical outcomes; however, this would take time and would require large cohorts of patients. Another solution would be for non-invasive fibrosis markers, which assess fibrosis quantitatively, to be ideally developed and validated with reference to a pure histological quantitative assessment of liver fibrosis. Such histological methods that quantify fibrosis by measuring liver collagen using digital image analysis have indeed been developed and could be used in future studies.

Chapter 11 Conclusion

The evidence suggests that, for HCV, treating all patients without prior diagnostic testing is the most cost-effective option. For HBV, the results differed for patients with HBeAg-positive disease and HBeAg-negative disease. The results suggests that if the upper bound of the standard UK cost-effectiveness threshold range is accepted for patients with HBeAg-negative disease, a strategy where all patients are treated regardless of fibrosis level is cost-effective.

For 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. Based on our results, using two NILTs sequentially (hyaluronic acid combined with MR elastography using the second sequential testing strategy outlined in *Chapter 2*) is most likely to be the optimal strategy at a threshold of £20,000; however, there is only a 4% probability of this being optimal.

Liver fibrosis and cirrhosis from HBV and HCV are significant health problems worldwide. The findings from the models may not be transferable to a resource setting where funds are limited and the ability to treat all patients is not a realistic option.

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 and if it is likely to be higher following an invasive test, then biopsy is very likely to be cost-effective. If there is no differential impact of the invasiveness of the test on abstinence, then Fibrotest is likely to be cost-effective (with either a high or low test threshold depending on the overall impact of fibrosis diagnosis on abstinence rates).

For NAFLD, most interventions are aimed at behavioural change and are not necessarily recommended specifically to reduce or halt fibrosis progression (e.g. weight-loss programmes for obesity). We located some studies for pharmacological interventions which had looked at the impact on fibrosis in NASH, but found that they had not conclusively demonstrated significant impact; this implies that the current potential pharmacological treatments such as pioglitazones are not effective in patients with NASH and fibrosis.

Implications for research

Further research could examine if the model is applicable to other settings, particularly resource-poor settings. Hepatitis is a global health problem and the pathways of care and expected treatment outcomes are likely to differ between settings. As such, the consequences of a false-negative test and a false-positive test may have different levels of importance according to locality. The cost-effectiveness of the non-invasive tests could be evaluated in these specific local settings, taking into account availability of treatments, local cost data and HRQoL values.

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

We were limited in our modelling approach as diagnostic studies do not report data using 3 × 2 tables, which would have allowed us to model the diagnostic test outcomes with precision. Future studies need to report all outcomes from tests rather than dichotomising into 2 × 2 tables in order to make the results more applicable for cost-effectiveness studies of diagnostic tests.

With alcoholic steatohepatitis, as abstinence is recommended at any stage of liver disease, we have assumed that the diagnosis of fibrosis impacts on abstinence rates and, furthermore, that abstinence may be further increased through more invasive tests. This was based on weak evidence, and further research could be conducted to assess the impact of testing on abstinence. It may be the case that interventions that include monitoring and support may be more effective in patients with lower degrees of fibrosis than non-invasive tests or biopsy. Further research needs to be conducted in this area.

With NAFLD, the lack of data was a significant limitation to our modelling. Considering the growing burden of the related complications with this disease on the NHS, long-term prospective studies are required that collect data on the impact of treatments on patients with NASH and fibrosis, long-term resource use and associated HRQoL using a comparable measure such as the EQ-5D. Data are also required on the relative effectiveness of management and treatment in primary care rather than secondary referrals.

Additionally, NILTs cannot differentiate simple steatosis from steatohepatitis in patients with NAFLD. 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.

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.

Acknowledgements

Contributions of authors

Catriona Crossan (Research Fellow, Health Economics) was responsible for the economic modelling and review of data inputs for the economic modelling. She contributed to the analysis and interpretation of results and led on drafting and revision of the final report.

Dr Emmanuel A Tsochatzis (Senior Clinical Lecturer and Honorary Consultant in Hepatology) was responsible for the systematic review and provided oversight to whole project (March 2013 to July 2013). He also contributed to the analysis and interpretation of results, and drafting and revision of the final report.

Dr Louise Longworth (Reader, Health Economics) oversaw the economic analysis. She also contributed to the analysis and interpretation of results, and drafting and revision of the final report.

Dr Kurinchi Gurusamy (Lecturer in Surgery) conducted the meta-analysis of results from the systematic review. He also contributed to the analysis and interpretation of results, and drafting and revision of the final report.

Professor Brian Davidson (Professor of Surgery) contributed to the analysis and interpretation of results, and drafting and revision of the final report.

Dr Manuel Rodríguez-Perálvarez (MD and PhD in Hepatology), **Dr Konstantinos Mantzoukis** (Research Fellow), **Dr Julia O'Brien** (Consultant Gastroenterologist & Hepatologist), **Dr Evangelos Thalassinos** (Clinical Research Fellow) and **Dr Vassilios Papastergiou** (Research Fellow in Hepatology) reviewed and extracted data identified in the systematic review.

Professor Andrew Burroughs (Professor of Hepatology) provided oversight to the whole project (November 2010 to March 2013), contributed to the analysis and interpretation of results, and drafting and revision of the final report.

Andrew Burroughs, **Emmanuel A Tsochatzis**, **Kurinchi Gurusamy** and **Brian Davidson** developed the project proposal and secured funding for the project.

We would like to thank Eszter Nagy for providing administrative support in the preparation and formatting of the report.

We would like to thank Ailish Higgins for reviewing the economic models.

We would like to thank Anna Noel-Storr for her contribution in developing the search strategy for the non-invasive tests in the PubMed and EMBASE databases.

References

1. Bataller R, Brenner DA. Liver fibrosis. *J Clin Invest* 2005;**115**:209–18. http://dx.doi.org/10.1172/JCI24282
2. Department of Health. *Annual Report of the Chief Medical Officer 2001*. 2001. URL: www.dh.gov.uk/en/Publicationsandstatistics/Publications/AnnualReports/DH_4005607 (accessed 8 March 2009).
3. Office for National Statistics. *Population Estimates. UK Population Grows to 60,975,000*. Office for National Statistics; 2008. URL: www.statistics.gov.uk/cci/nuggetasp?id=6 (accessed 8 March 2009).
4. NHS Blood and Transplant. *Transplant Activity in the UK 2007–2008*. URL: www.uktransplant.org.uk/ukt/statistics/transplant_activity_report/current_activity_reports/ukt/transplant_activity_uk_2007–2008.pdf (accessed 8 March 2009).
5. The Scottish Government publications. *Health in Scotland 2007: Annual Report of the Chief Medical Officer. The Chief Medical Officer's Report to First Minister on the Health of the Nation*. 2008. URL: www.scotland.gov.uk/Publications/2008/11/26155748/8 (accessed 8 March 2009).
6. Denzer U, Arnoldy A, Kanzler S, Galle PR, Dienes HP, Lohse AW. Prospective randomized comparison of minilaparoscopy and percutaneous liver biopsy: diagnosis of cirrhosis and complications. *J Clin Gastroenterol* 2007;**41**:103–10. http://dx.doi.org/10.1097/01.mcg.0000225612.86846.82
7. Terjung B, Lemnitzer I, Dumoulin FL, Effenberger W, Brackmann HH, Sauerbruch T, *et al.* Bleeding complications after percutaneous liver biopsy. An analysis of risk factors. *Digestion* 2003;**67**:138–45. http://dx.doi.org/10.1159/000071293
8. Cholongitas E, Senzolo M, Standish R, Marelli L, Quaglia A, Patch D, *et al.* A systematic review of the quality of liver biopsy specimens. *Am J Clin Pathol* 2006;**125**:710–21. http://dx.doi.org/10.1309/W3XCNT4HKFBN2G0B
9. The French METAVIR Cooperative Study Group. Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. *Hepatology* 1994;**20**:15–20.
10. Germani G, Hytiroglou P, Fotiadu A, Burroughs AK, Dhillon AP. Assessment of fibrosis and cirrhosis in liver biopsies: an update. *Semin Liver Dis* 2011;**31**:82–90. http://dx.doi.org/10.1055/s-0031-1272836
11. Standish RA, Cholongitas E, Dhillon A, Burroughs AK, Dhillon AP. An appraisal of the histopathological assessment of liver fibrosis. *Gut* 2006;**55**:569–78. http://dx.doi.org/10.1136/gut.2005.084475
12. Germani G, Burroughs AK, Dhillon AP. The relationship between liver disease stage and liver fibrosis: a tangled web. *Histopathology* 2010;**57**:773–84. http://dx.doi.org/10.1111/j.1365-2559.2010.03609.x
13. Tsochatzis EA, Manousou P, Fede G, Dhillon AP, Burroughs AK. Validating non-invasive markers of fibrosis: the need for a new histological reference standard. *Gut* 2011;**60**:1442–3; author reply 3–4. http://dx.doi.org/10.1136/gut.2010.229484
14. Isgro G, Calvaruso V, Andreana L, Luong TV, Garcovich M, Manousou P, *et al.* The relationship between transient elastography and histological collagen proportionate area for assessing fibrosis in chronic viral hepatitis. *J Gastroenterol* 2012;**48**:921–9. http://dx.doi.org/10.1007/s00535-012-0694-9

15. Manousou P, Dhillon AP, Isgro G, Calvaruso V, Luong TV, Tsochatzis E, *et al.* Digital image analysis of liver collagen predicts clinical outcome of recurrent hepatitis C virus 1 year after liver transplantation. *Liver Transpl* 2011;**17**:178–88. http://dx.doi.org/10.1002/lt.22209

16. Calvaruso V, Burroughs AK, Standish R, Manousou P, Grillo F, Leandro G, *et al.* Computer-assisted image analysis of liver collagen: relationship to Ishak scoring and hepatic venous pressure gradient. *Hepatology* 2009;**49**:1236–44. http://dx.doi.org/10.1002/hep.22745

17. Calvaruso V, Dhillon AP, Tsochatzis E, Manousou P, Grillo F, Germani G, *et al.* Liver collagen proportionate area predicts decompensation in patients with recurrent hepatitis C virus cirrhosis after liver transplantation. *J Gastroenterol Hepatol* 2012;**27**:1227–32. http://dx.doi.org/10.1111/j.1440-1746.2012.07136.x

18. Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 2003;**38**:1449–57. http://dx.doi.org/10.1016/j.hep.2003.09.022

19. Tsochatzis EA, Germani G, Hall A, Anousou PM, Dhillon AP, Burroughs AK. Noninvasive assessment of liver fibrosis: the need for better validation. *Hepatology* 2011;**53**:1781–2; author reply 2–3. http://dx.doi.org/10.1002/hep.24271

20. Martinez SM, Crespo G, Navasa M, Forns X. Noninvasive assessment of liver fibrosis. *Hepatology* 2011;**53**:325–35. http://dx.doi.org/10.1002/hep.24013

21. Castera L. Noninvasive methods to assess liver disease in patients with hepatitis B or C. *Gastroenterology* 2012;**142**:1293–302 e4.

22. Lichtinghagen R, Pietsch D, Bantel H, Manns MP, Brand K, Bahr MJ. The Enhanced Liver Fibrosis (ELF) score: normal values, influence factors and proposed cut-off values. *J Hepatol* 2013;**59**:236–42. http://dx.doi.org/10.1016/j.jhep.2013.03.016

23. Imbert-Bismut F, Ratziu V, Pieroni L, Charlotte F, Benhamou Y, Poynard T, *et al.* Biochemical markers of liver fibrosis in patients with hepatitis C virus infection: a prospective study. *Lancet* 2001;**357**:1069–75. http://dx.doi.org/10.1016/S0140-6736(00)04258-6

24. Koda M, Matunaga Y, Kawakami M, Kishimoto Y, Suou T, Murawaki Y. FibrolIndex, a practical index for predicting significant fibrosis in patients with chronic hepatitis C. *Hepatology* 2007;**45**:297–306. http://dx.doi.org/10.1002/hep.21520

25. Patel K, Benhamou Y, Yoshida EM, Kaita KD, Zeuzem S, Torbenson M, *et al.* An independent and prospective comparison of two commercial fibrosis marker panels (HCV FibroSURE and FIBROSpect II) during albinterferon alfa-2b combination therapy for chronic hepatitis C. *J Viral Hepatol* 2009;**16**:178–86. http://dx.doi.org/10.1111/j.1365-2893.2008.01062.x

26. Cales P, Halfon P, Batisse D, Carrat F, Perre P, Penaranda G, *et al.* Comparison of liver fibrosis blood tests developed for HCV with new specific tests in HIV/HCV co-infection. *J Hepatol* 2010;**52**:S405. http://dx.doi.org/10.1016/S0168-8278(10)61048-3

27. Parkes J, Guha IN, Roderick P, Harris S, Cross R, Manos MM, *et al.* Enhanced Liver Fibrosis (ELF) test accurately identifies liver fibrosis in patients with chronic hepatitis C. *J Viral Hepatol* 2011;**18**:23–31. http://dx.doi.org/10.1111/j.1365-2893.2009.01263.x

28. Castera L, Vergniol J, Foucher J, Le Bail B, Chanteloup E, Haaser M, *et al.* Prospective comparison of transient elastography, Fibrotest, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C. *Gastroenterology* 2005;**128**:343–50. http://dx.doi.org/10.1053/j.gastro.2004.11.018

29. Lupsor M, Badea R, Stefanescu H, Sparchez Z, Branda H, Serban A, *et al.* Performance of a new elastographic method (ARFI technology) compared to unidimensional transient elastography in the noninvasive assessment of chronic hepatitis C. Preliminary results. *J Gastrointestin Liv Dis* 2009;**18**:303–10.

30. Huwart L, Sempoux C, Vicaut E, Salameh N, Annet L, Danse E, *et al.* Magnetic resonance elastography for the noninvasive staging of liver fibrosis. *Gastroenterology* 2008;**135**:32–40. http://dx.doi.org/10.1053/j.gastro.2008.03.076

31. Sebastiani G, Halfon P, Castera L, Mangia A, Di Marco V, Pirisi M, *et al.* Comparison of three algorithms of non-invasive markers of fibrosis in chronic hepatitis C. *Aliment Pharmacol Ther* 2012;**35**:92–104. http://dx.doi.org/10.1111/j.1365-2036.2011.04897.x

32. Tsochatzis EA, Gurusamy KS, Ntaoula S, Cholongitas E, Davidson BR, Burroughs AK. Elastography for the diagnosis of severity of fibrosis in chronic liver disease: a meta-analysis of diagnostic accuracy. *J Hepatol* 2011;**54**:650–9. http://dx.doi.org/10.1016/j.jhep.2010.07.033

33. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. *J Hepatol* 2009;**51**:237–67. http://dx.doi.org/10.1016/j.jhep.2009.04.009

34. European Association for the Study of the Liver. EASL clinical practice guidelines for HFE hemochromatosis. *J Hepatol* 2010;**53**:3–22. http://dx.doi.org/10.1016/j.jhep.2010.03.001

35. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: Wilson's disease. *J Hepatol* 2012;**56**:671–85. http://dx.doi.org/10.1016/j.jhep.2011.11.007

36. Tsochatzis EA, Bosch J, Burroughs AK. New therapeutic paradigm for patients with cirrhosis. *Hepatology* 2012;**56**:1983–92. http://dx.doi.org/10.1002/hep.25915

37. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of hepatitis C virus infection. *J Hepatol* 2011;**55**:245–64. http://dx.doi.org/10.1016/j.jhep.2011.02.023

38. Poynard T, Bedossa P, Opolon P. Natural history of liver fibrosis progression in patients with chronic hepatitis C. The OBSVIRC, METAVIR, CLINIVIR, and DOSVIRC groups. *Lancet* 1997;**349**:825–32. http://dx.doi.org/10.1016/S0140-6736(96)07642-8

39. Jacobson IM, McHutchison JG, Dusheiko G, Di Bisceglie AM, Reddy KR, Bzowej NH, *et al.* Telaprevir for previously untreated chronic hepatitis C virus infection. *N Engl J Med* 2011;**364**:2405–16. http://dx.doi.org/10.1056/NEJMoa1012912

40. Poordad F, McCone J Jr., Bacon BR, Bruno S, Manns MP, Sulkowski MS, *et al.* Boceprevir for untreated chronic HCV genotype 1 infection. *N Engl J Med* 2011;**364**:1195–206. http://dx.doi.org/10.1056/NEJMoa1010494

41. Wilck MB, Hamel MB, Baden LR. Clinical decisions. Management of incidental hepatitis C virus infection–polling results. *N Engl J Med* 2009;**360**:e30. http://dx.doi.org/10.1056/NEJMclde0902765

42. Afdhal NH, Lok AS, Di Bisceglie AM. Clinical decisions. Management of incidental hepatitis C virus infection. *N Engl J Med* 2009;**360**:1902–6. http://dx.doi.org/10.1056/NEJMclde0900131

43. Lawitz E, Mangia A, Wyles D, Rodriguez-Torres M, Hassanein T, Gordon SC, *et al.* Sofosbuvir for previously untreated chronic hepatitis C infection. *N Engl J Med* 2013;**368**:1878–87. http://dx.doi.org/10.1056/NEJMoa1214853

44. Jacobson IM, Gordon SC, Kowdley KV, Yoshida EM, Rodriguez-Torres M, Sulkowski MS, *et al.* Sofosbuvir for hepatitis C genotype 2 or 3 in patients without treatment options. *N Engl J Med* 2013;**368**:1867–77. http://dx.doi.org/10.1056/NEJMoa1214854

45. European Association for the Study of the Liver. EASL clinical practice guidelines: management of chronic hepatitis B virus infection. *J Hepatol* 2012;**57**:167–85. http://dx.doi.org/10.1016/j.jhep.2012.02.010

46. Hadziyannis SJ, Papatheodoridis GV. Hepatitis B e antigen-negative chronic hepatitis B: natural history and treatment. *Sem Liv Dis* 2006;**26**:130–41. http://dx.doi.org/10.1055/s-2006-939751

47. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, *et al.* The diagnosis and management of non-alcoholic fatty liver disease: practice guideline by the American Gastroenterological Association, American Association for the Study of Liver Diseases, and American College of Gastroenterology. *Gastroenterology* 2012;**142**:1592–609. http://dx.doi.org/10.1053/j.gastro.2012.04.001

48. Tsochatzis E, Papatheodoridis GV, Manesis EK, Kafiri G, Tiniakos DG, Archimandritis AJ. Metabolic syndrome is associated with severe fibrosis in chronic viral hepatitis and non-alcoholic steatohepatitis. *Aliment Pharmacol Ther* 2008;**27**:80–9. http://dx.doi.org/10.1111/j.1365-2036.2007.03538.x

49. Qian MY, Yuwei JR, Angus P, Schelleman T, Johnson L, Gow P. Efficacy and cost of a hepatocellular carcinoma screening program at an Australian teaching hospital. *J Gastroenterol Hepatol* 2010;**25**:951–6. http://dx.doi.org/10.1111/j.1440-1746.2009.06203.x

50. Argo CK, Northup PG, Al-Osaimi AM, Caldwell SH. Systematic review of risk factors for fibrosis progression in non-alcoholic steatohepatitis. *J Hepatol* 2009;**51**:371–9. http://dx.doi.org/10.1016/j.jhep.2009.03.019

51. Ekstedt M, Franzen LE, Mathiesen UL, Thorelius L, Holmqvist M, Bodemar G, *et al.* Long-term follow-up of patients with NAFLD and elevated liver enzymes. *Hepatology* 2006;**44**:865–73. http://dx.doi.org/10.1002/hep.21327

52. Sanyal AJ, Banas C, Sargeant C, Luketic VA, Sterling RK, Stravitz RT, *et al.* Similarities and differences in outcomes of cirrhosis due to nonalcoholic steatohepatitis and hepatitis C. *Hepatology* 2006;**43**:682–9. http://dx.doi.org/10.1002/hep.21103

53. O'Shea RS, Dasarathy S, McCullough AJ. Alcoholic liver disease. *Hepatology* 2010;**51**:307–28. http://dx.doi.org/10.1002/hep.23258

54. Colombo M. Screening and diagnosis of hepatocellular carcinoma. *Liver Int* 2009;**29**(Suppl. 1):143–7. http://dx.doi.org/10.1111/j.1478-3231.2008.01938.x

55. European Association for the Study of the Liver. EASL clinical practical guidelines: management of alcoholic liver disease. *J Hepatol* 2012;**57**:399–420. http://dx.doi.org/10.1016/j.jhep.2012.04.004

56. Tome S, Lucey MR. Review article: current management of alcoholic liver disease. *Aliment Pharmacol Ther* 2004;**19**:707–14. http://dx.doi.org/10.1111/j.1365-2036.2004.01881.x

57. Goodman ZD. Grading and staging systems for inflammation and fibrosis in chronic liver diseases. *J Hepatol* 2007;**47**:598–607. http://dx.doi.org/10.1016/j.jhep.2007.07.006

58. Buntinx F, Aertgeerts B, Macaskill P. Guidelines for conducting systematic review of studies evaluating the accuracy of diagnostic tests. In Knottnerus J, Buntinx F, editors. *The Evidence Base of Clinical Diagnosis*. 2nd edn. Oxford and Hoboken, NJ: BMJ Books; 2009.

59. Whiting P, Westwood M, Burke M, Sterne J, Glanville J. Systematic reviews of test accuracy should search a range of databases to identify primary studies. *J Clin Epidemiol* 2008;**61**:357–64. http://dx.doi.org/10.1016/j.jclinepi.2007.05.013

60. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;**3**:25. http://dx.doi.org/10.1186/1471-2288-3-25

61. Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PN, Kleijnen J. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Med Res Methodol* 2006;**6**:9. http://dx.doi.org/10.1186/1471-2288-6-9

62. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;**155**:529–36. http://dx.doi.org/10.7326/0003-4819-155-8-201110180-00009

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

64. Takwoingi Y, Deeks J. *MetaDAS: a SAS Macro for Meta-Analysis of Diagnostic Accuracy Studies – User Guide Version 1.3*. 2010. URL: http://srdta.cochrane.org/sites/srdta.cochrane.org/files/uploads/MetaDAS%20Readme%20v13%20May%202012.pdf (accessed 25 October 2012).

65. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. Analysing and presenting results. In Deeks JJ, Bossuyt PM, Gatsonis C, editors. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy: Version 10*. The Cochrane Collaboration; 2010. URL: http://srdta.cochrane.org/files/uploads/Chapter10-Version1.0.pdf (accessed 1 January 2011).

66. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*. London: NICE; 2013.

67. Curtis L. *Unit Costs of Health and Social Care 2012*. Canterbury: PSSRU, University of Kent; 2012.

68. Drummond M, O'Brien BJ, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press; 1997.

69. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Econ* 2001;**10**:779–87. http://dx.doi.org/10.1002/hec.635

70. Barton GR, Briggs AH, Fenwick EAL. Optimal cost-effectiveness decisions: the role of the cost-effectiveness acceptability curve (CEAC), the cost-effectiveness acceptability frontier (CEAF), and the expected value of perfection information (EVPI). *Value Health* 2008;**11**:886–97. http://dx.doi.org/10.1111/j.1524-4733.2008.00358.x

71. Adams LA, Bulsara M, Rossi E, DeBoer B, Speers D, George J, *et al.* Hepascore: an accurate validated predictor of liver fibrosis in chronic hepatitis C infection. *Clin Chem* 2005;**51**:1867–73. http://dx.doi.org/10.1373/clinchem.2005.048389

72. Ahmad W, Ijaz B, Javed FT, Gull S, Kausar H, Sarwar MT, *et al.* A comparison of four fibrosis indexes in chronic HCV: development of new fibrosis-cirrhosis index (FCI). *BMC Gastroenterol* 2011;**11**:44. http://dx.doi.org/10.1186/1471-230X-11-44

73. Al-Mohri H, Cooper C, Murphy T, Klein MB. Validation of a simple model for predicting liver fibrosis in HIV/hepatitis C virus-coinfected patients. *HIV Med* 2005;**6**:375–8. http://dx.doi.org/10.1111/j.1468-1293.2005.00330.x

74. Anaparthy R, Guturu P, Snyder N, Sood G. Diagnostic utility of APRI for staging of liver disease in hepatitis C patients with ESRD on hemodialysis. *Gastroenterology* 2009;**136**(Suppl. 1):A833–4. http://dx.doi.org/10.1016/S0016-5085(09)63842-7

75. Arena U, Vizzutti F, Abraldes JG, Corti G, Stasi C, Moscarella S, *et al.* Reliability of transient elastography for the diagnosis of advanced fibrosis in chronic hepatitis C. *Gut* 2008;**57**:1288–93. http://dx.doi.org/10.1136/gut.2008.149708

76. Beckebaum S, Iacob S, Klein CG, Dechene A, Varghese J, Baba HA, *et al.* Assessment of allograft fibrosis by transient elastography and noninvasive biomarker scoring systems in liver transplant patients. *Transplantation* 2010;**89**:983–93. http://dx.doi.org/10.1097/TP.0b013e3181cc66ca

77. Bejarano G, Vergara M, Dalmau B, Puig J, Bella MR, Suarez D, *et al.* Prospective evaluation of liver fibrosis in chronic viral hepatitis C infection using the Sabadell NIHCED (Non-Invasive Hepatitis-C-Related Cirrhosis Early Detection) index. *Revista Espanola de Enfermedades Digestivas* 2009;**101**:325–35. http://dx.doi.org/10.4321/S1130-01082009000500004

78. Ben Jazia E, Kaabia N, Benabdelkader A, Khalifa M, Harrabi I, Braham A, *et al.* Noninvasive fibrosis markers for the prediction of significant fibrosis in patients with chronic hepatitis C virus infection in Tunisia. *Infect Dis Clin Pract* 2009;**17**:385–7. http://dx.doi.org/10.1097/IPC.0b013e3181bf60d3

79. Berg T, Sarrazin C, Hinrichsen H, Buggisch P, Gerlach T, Zachoval R, *et al.* Does noninvasive staging of fibrosis challenge liver biopsy as a gold standard in chronic hepatitis C? *Hepatology* 2004;**39**:1456–7; author reply 7–8. http://dx.doi.org/10.1002/hep.20226

80. Borroni G, Ceriani R, Tommasini MA, Maltempo C, Felline C, Contu L, *et al.* Platelet count: a simple marker of liver cirrhosis in chronic hepatitis C (CHC) infection. *J Hepatol* 2002;**36**(Suppl. 1):47. http://dx.doi.org/10.1016/S0168-8278(02)80146-5

81. Bourliere M, Penaranda G, Renou C, Botta-Fridlund D, Tran A, Portal I, *et al.* Validation and comparison of indexes for fibrosis and cirrhosis prediction in chronic hepatitis C patients: proposal for a pragmatic approach classification without liver biopsies. *J Viral Hepat* 2006;**13**:659–70. http://dx.doi.org/10.1111/j.1365-2893.2006.00736.x

82. Boursier J, Bacq Y, Halfon P, Leroy V, De Ledinghen V, De Muret A, *et al.* Improved diagnostic accuracy of blood tests for severe fibrosis and cirrhosis in chronic hepatitis C. *Eur J Gastroenterol Hepatol* 2009;**21**:28–38. http://dx.doi.org/10.1097/MEG.0b013e32830cebd7

83. Boursier J, de Ledinghen V, Zarski J-P, Fouchard-Hubert I, Gallois Y, Oberti F, *et al.* Comparison of eight diagnostic algorithms for liver fibrosis in hepatitis C: new algorithms are more precise and entirely noninvasive. *Hepatology* 2012;**55**:58–67. http://dx.doi.org/10.1002/hep.24654

84. Burton MJ, Pham H, Sunesara I, McGloster N, Oliver N, McGuire BM. Performance of the APRI score in African Americans with chronic hepatitis C. *Hepatology* 2010;**52**:1236A–7A.

85. Cales P, Boursier J, Bertrais S, Oberti F, Gallois Y, Fouchard-Hubert I, *et al.* Optimization and robustness of blood tests for liver fibrosis and cirrhosis. *Clin Biochem* 2010;**43**:1315–22. http://dx.doi.org/10.1016/j.clinbiochem.2010.08.010

86. Calvaruso V, Camma C, Di Marco V, Maimone S, Bronte F, Enea M, *et al.* Fibrosis staging in chronic hepatitis C: analysis of discordance between transient elastography and liver biopsy. *J Viral Hepat* 2010;**17**:469–74. http://dx.doi.org/10.1111/j.1365-2893.2009.01199.x

87. Cardoso A-C, Carvalho-Filho RJ, Stern C, Dipumpo A, Giuily N, Ripault M-P, *et al.* Direct comparison of diagnostic performance of transient elastography in patients with chronic hepatitis B and chronic hepatitis C. *Liver Int* 2012;**32**:612–21. http://dx.doi.org/10.1111/j.1478-3231.2011.02660.x

88. Carrion JA, Navasa M, Bosch J, Bruguera M, Gilabert R, Forns X. Transient elastography for diagnosis of advanced fibrosis and portal hypertension in patients with hepatitis C recurrence after liver transplantation. *Liver Transpl* 2006;**12**:1791–8. http://dx.doi.org/10.1002/lt.20857

89. Carvalho-Filho RJ, Schiavon LL, Narciso-schiavon JL, Sampaio JP, Lanzoni VP, Ferraz MLG, *et al.* Optimized cutoffs improve performance of the aspartate aminotransferase to platelet ratio index for predicting significant liver fibrosis in human immunodeficiency virus/hepatitis C virus co-infection. *Liver Int* 2008;**28**:486–93. http://dx.doi.org/10.1111/j.1478-3231.2008.01675.x

90. Castera L, Bail BL, Roudot-Thoraval F, Bernard PH, Foucher J, Merrouche W, *et al.* Early detection in routine clinical practice of cirrhosis and oesophageal varices in chronic hepatitis C: comparison of transient elastography (FibroScan) with standard laboratory tests and non-invasive scores. *J Hepatol* 2009;**50**:59–68. http://dx.doi.org/10.1016/j.jhep.2008.08.018

91. Castera L, Loko MA, Dabis F, Le Bail B, Winnock M, Coureau G, *et al.* Validation of simple indexes (FIB-4, APRI, forns index) and platelet count for non invasive prediction of liver fibrosis in HIV-HCV coinfected French patients. *Hepatology* 2007;**46**:841A–A.

92. Ceriani R, Borroni G, Tommasini MA, Maltempo C, Felline C, Contu L, *et al.* Platelet count and AST/ALT ratio can predict liver cirrhosis in chronic HCV infection. *J Hepatol* 2001;**34**:71. http://dx.doi.org/10.1016/S0168-8278(01)81119-3

93. Chen CH, Lin ST, Yang CC, Yeh YH, Kuo CL, Nien CK. The accuracy of sonography in predicting steatosis and fibrosis in chronic hepatitis C. *Dig Dis Sci* 2008;**53**:1699–706. http://dx.doi.org/10.1007/s10620-007-0048-2

94. Cheung RC, Currie S, Shen H, Bini EJ, Ho SB, Anand BS, *et al.* Can we predict the degree of fibrosis in chronic hepatitis C patients using routine blood tests in our daily practice? *J Clin Gastroenterol* 2008;**42**:827–34. http://dx.doi.org/10.1097/MCG.0b013e318046ea9a

95. Cho HJ, Seo YS, Lee KG, Hyun JJ, An H, Keum B, *et al.* Serum aminotransferase levels instead of etiology affects the accuracy of transient elastography in chronic viral hepatitis patients. *J Gastroenterol Hepatol* 2011;**26**:492–500. http://dx.doi.org/10.1111/j.1440-1746.2010.06419.x

96. Christensen C, Bruden D, Livingston S, Deubner H, Homan C, Smith K, *et al.* Diagnostic accuracy of a fibrosis serum panel (FIBROSpect II) compared with Knodell and Ishak liver biopsy scores in chronic hepatitis C patients. *J Viral Hepatol* 2006;**13**:652–8. http://dx.doi.org/10.1111/j.1365-2893.2006.00743.x

97. Chrysanthos NV, Papatheodoridis GV, Savvas S, Kafiri G, Petraki K, Manesis EK, *et al.* Aspartate aminotransferase to platelet ratio index for fibrosis evaluation in chronic viral hepatitis. *Eur J Gastroenterol Hepatol* 2006;**18**:389–96. http://dx.doi.org/10.1097/00042737-200604000-00012

98. Cobbold JFL, Crossey MME, Colman P, Goldin RD, Murphy PS, Patel N, *et al.* Optimal combinations of ultrasound-based and serum markers of disease severity in patients with chronic hepatitis C. *J Viral Hepatol* 2010;**17**:537–45. http://dx.doi.org/10.1111/j.1365-2893.2009.01209.x

99. Colletta C, Smirne C, Fabris C, Toniutto P, Rapetti R, Minisini R, *et al.* Value of two noninvasive methods to detect progression of fibrosis among HCV carriers with normal aminotransferases. *Hepatology* 2005;**42**:838–45. http://dx.doi.org/10.1002/hep.20814

100. Corradi F, Piscaglia F, Flori S, D'Errico-Grigioni A, Vasuri F, Tame MR, *et al.* Assessment of liver fibrosis in transplant recipients with recurrent HCV infection: usefulness of transient elastography. *Dig Liver Dis* 2009;**41**:217–25. http://dx.doi.org/10.1016/j.dld.2008.06.009

101. Crespo SM, Bridges MD, Keaveny AP, Nakhleh RE, McPhail A, Pungpapong S. Magnetic resonance elastography is superior to the HULF index and APRI in evaluating fibrosis due to recurrent HCV after liver transplantation. *Hepatology* 2010;**52**:843A–4A.

102. Cross TJS, Calvaruso V, Maimone S, Carey I, Chang TP, Pleguezuelo M, *et al.* Prospective comparison of Fibroscan, King's score and liver biopsy for the assessment of cirrhosis in chronic hepatitis C infection. *J Viral Hepatol* 2010;**17**:546–54. http://dx.doi.org/10.1111/j.1365-2893.2009.01210.x

103. da Silva RG Jr, Fakhouri R, do Nascimento TVB, Santos IM, Barbosa LMM. Aspartate aminotransferase-to-platelet ratio index for fibrosis and cirrhosis prediction in chronic hepatitis C patients. *Braz J Infect Dis* 2008;**12**:15–19.

104. Danila M, Sporea I, Sirli R, Bota S, Braticevici CF, Petrisor A, *et al.* Acoustic radiation force impulse elastography (ARFI) for predicting the severity of fibrosis in patients with chronic HCV hepatitis – a bicentric Romanian study. *Gastroenterology* 2011;**140**(Suppl. 1):S968.

105. de Ledinghen V, Douvin C, Kettaneh A, Ziol M, Roulot D, Marcellin P, *et al.* Diagnosis of hepatic fibrosis and cirrhosis by transient elastography in HIV/hepatitis C virus-coinfected patients. *J Acquir Immune Defic Syndr* 2006;**41**:175–9. http://dx.doi.org/10.1097/01.qai.0000194238.15831.c7

106. Degos F, Perez P, Roche B, Mahmoudi A, Asselineau J, Voitot H, *et al.* Diagnostic accuracy of FibroScan and comparison to liver fibrosis biomarkers in chronic viral hepatitis: a multicenter prospective study (the FIBROSTIC study). *J Hepatol* 2010;**53**:1013–21. http://dx.doi.org/10.1016/j.jhep.2010.05.035

107. Dinesen L, Caspary WF, Chapman RW, Dietrich CF, Sarrazin C, Braden B. C-13-methacetin-breath test compared to also noninvasive biochemical blood tests in predicting hepatic fibrosis and cirrhosis in chronic hepatitis C. *Dig Liver Dis* 2008;**40**:743–8. http://dx.doi.org/10.1016/j.dld.2008.01.013

108. Esmat G, Metwally M, Zalata KR, Gadalla S, Abdel-Hamid M, Abouzied A, *et al.* Evaluation of serum biomarkers of fibrosis and injury in Egyptian patients with chronic hepatitis C. *J Hepatol* 2007;**46**:620–7. http://dx.doi.org/10.1016/j.jhep.2006.12.010

109. Fabris C, Smirne C, Toniutto P, Colletta C, Rapetti R, Minisini R, *et al.* Assessment of liver fibrosis progression in patients with chronic hepatitis C and normal alanine aminotransferase values: the role of AST to the platelet ratio index. *Clin Biochem* 2006;**39**:339–43. http://dx.doi.org/10.1016/j.clinbiochem.2006.01.011

110. Fahmy MI, Badran HM. Comparison of transient elastography to Doppler indices in prediction of hepatitis C induced liver fibrosis and cirrhosis. *Egypt J Radiol Nucl Med* 2011;**42**:111–17. http://dx.doi.org/10.1016/j.ejrnm.2011.05.001

111. Fontaine H, Nalpas B, Vallet-Pichard A, Mallet V, Pol S. Low diagnosis accuracy of fibrotest and Fib-4 in HCV-infected hemodialysis and kidney recipients. *Hepatology* 2009;**50**:1064A.

112. Fontana RJ, Goodman ZD, Dienstag JL, Bonkovsky HL, Naishadham D, Sterling RK, *et al.* Relationship of serum fibrosis markers with liver fibrosis stage and collagen content in patients with advanced chronic hepatitis C. *Hepatology* 2008;**47**:789–98. http://dx.doi.org/10.1002/hep.22099

113. Fontanges T, Bailly F, Trepo E, Chevallier M, Maynard-Muet M, Nalet B, *et al.* Discordance between biochemical markers of liver activity and fibrosis (Actitest-Fibrotest) and liver biopsy in patients with chronic hepatitis C. *Gastroenterol Clin Biol* 2008;**32**:858–65. http://dx.doi.org/10.1016/j.gcb.2008.05.019

114. Forestier J, Dumortier J, Guillaud O, Ecochard M, Roman S, Boillot O, *et al.* Noninvasive diagnosis and prognosis of liver cirrhosis: a comparison of biological scores, elastometry, and metabolic liver function tests. *Eur J Gastroenterol Hepatol* 2010;**22**:532–40. http://dx.doi.org/10.1097/MEG.0b013e3283343f58

115. Forns X, Ampurdanes S, Llovet JM, Aponte J, Quinto L, Martinez-Bauer E, *et al.* Identification of chronic hepatitis C patients without hepatic fibrosis by a simple predictive model. *Hepatology* 2002;**36**:986–92. http://dx.doi.org/10.1053/jhep.2002.36128

116. Fraquelli M, Rigamonti C, Casazza G, Donato MF, Ronchi G, Conte D, *et al.* Etiology-related determinants of liver stiffness values in chronic viral hepatitis B or C. *J Hepatol* 2011;**54**:621–8. http://dx.doi.org/10.1016/j.jhep.2010.07.017

117. Fujii H, Enomoto M, Fukushima W, Ohfuji S, Mori M, Kobayashi S, *et al.* Noninvasive laboratory tests proposed for predicting cirrhosis in patients with chronic hepatitis C are also useful in patients with non-alcoholic steatohepatitis. *J Gastroenterology* 2009;**44**:608–14. http://dx.doi.org/10.1007/s00535-009-0046-6

118. Fujimoto K, Tonan T, Azuma S, Kage M, Nakashima O, Johkoh T, *et al.* Evaluation of the mean and entropy of apparent diffusion coefficient values in chronic hepatitis C: correlation with pathologic fibrosis stage and inflammatory activity grade. *Radiology* 2011;**258**:739–48. http://dx.doi.org/10.1148/radiol.10100853

119. Gaia S, Carenzi S, Barilli AL, Bugianesi E, Smedile A, Brunello F, *et al.* Reliability of transient elastography for the detection of fibrosis in non-alcoholic fatty liver disease and chronic viral hepatitis. *J Hepatol* 2011;**54**:64–71. http://dx.doi.org/10.1016/j.jhep.2010.06.022

120. Ganne-Carrie N, Ziol M, De Ledinghen V, Douvin C, Marcellin P, Castera L, *et al.* Accuracy of liver stiffness measurement for the diagnosis of cirrhosis in patients with chronic liver diseases. *Hepatology* 2006;**44**:1511–17. http://dx.doi.org/10.1002/hep.21420

121. Gara N, Kleiner DE, Zhao X, Koh C, Rotman Y, Heller T, *et al.* Performance of transient elastography and APRI for prediction of advanced liver disease in a north American population with chronic hepatitis C. *Hepatology* 2011;**54**:570A.

122. Giannini EG, Zaman A, Ceppa P, Mastracci L, Risso D, Testa R. A simple approach to noninvasively identifying significant fibrosis in chronic hepatitis C patients in clinical practice. *J Clin Gastroenterol* 2006;**40**:521–7. http://dx.doi.org/10.1097/00004836-200607000-00011

123. Gobel T, Vordenwulbecke S, Hauck K, Fey H, Haussinger D, Erhardt A. New multi protein patterns differentiative liver fibrosis stages and hepatocellular carcinoma in chronic hepatitis C serum samples. *World J Gastroenterol* 2006;**12**:7604–12.

124. Guechot J, Lasnier E, Sturm N, Paris A, Zarski J-P, group AHEFs. Automation of the Hepascore and validation as a biochemical index of liver fibrosis in patients with chronic hepatitis C from the ANRS HC EP 23 Fibrostar cohort. *Clin Chim Acta* 2010;**411**:86–91. http://dx.doi.org/10.1016/j.cca.2009.10.011

125. Guechot J, Laudat A, Loria A, Serfaty L, Poupon R, Giboudeau J. Diagnostic accuracy of hyaluronan and type III procollagen amino-terminal peptide serum assays as markers of liver fibrosis in chronic viral hepatitis C evaluated by ROC curve analysis. *Clin Chem* 1996;**42**:558–63.

126. Guzelbulut F, Cetinkaya ZA, Sezikli M, Yasar B, Ozkara S, Ovunc AOK. AST–platelet ratio index, Forns index and FIB-4 in the prediction of significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Turk J Gastroenterol* 2011;**22**:279–85.

127. Halfon P, Bacq Y, De Muret A, Penaranda G, Bourliere M, Ouzan D, *et al.* Comparison of test performance profile for blood tests of liver fibrosis in chronic hepatitis C. *J Hepatol* 2007;**46**:395–402. http://dx.doi.org/10.1016/j.jhep.2006.09.020

128. Halfon P, Bourliere M, Deydier R, Botta-Fridlund D, Renou C, Tran A, *et al.* Independent prospective multicenter validation of biochemical markers (fibrotest-actitest) for the prediction of liver fibrosis and activity in patients with chronic hepatitis C: the fibropaca study. *Am J Gastroenterol* 2006;**101**:547–55. http://dx.doi.org/10.1111/j.1572-0241.2006.00411.x

129. Halfon P, Bourliere M, Penaranda G, Deydier R, Renou C, Botta-Fridlund D, *et al.* Accuracy of hyaluronic acid level for predicting liver fibrosis stages in patients with hepatitis C virus. *Comp Hepatol* 2005;**4**:6. http://dx.doi.org/10.1186/1476-5926-4-6

130. Harada N, Soejima Y, Taketomi A, Yoshizumi T, Ikegami T, Yamashita Y-i, *et al.* Assessment of graft fibrosis by transient elastography in patients with recurrent hepatitis C after living donor liver transplantation. *Transplantation* 2008;**85**:69–74. http://dx.doi.org/10.1097/01.tp.0000297248.18483.16

131. Hsieh YY, Tung SY, Lee K, Wu CS, Wei KL, Shen CH, *et al.* Routine blood tests to predict liver fibrosis in chronic hepatitis C. *World J Gastroenterol* 2012;**18**:746–53. http://dx.doi.org/10.3748/wjg.v18.i8.746

132. Iacobellis A, Fusilli S, Mangia A, Clemente R, Festa V, Giacobbe A, *et al.* Ultrasonographic and biochemical parameters in the non-invasive evaluation of liver fibrosis in hepatitis C virus chronic hepatitis. *Aliment Pharmacol Ther* 2005;**22**:769–74. http://dx.doi.org/10.1111/j.1365-2036.2005.02633.x

133. Imperiale TF, Said AT, Cummings OW, Born LJ. Need for validation of clinical decision aids: use of the AST/ALT ratio in predicting cirrhosis in chronic hepatitis C. *Am J Gastroenterol* 2000;**95**:2328–32. http://dx.doi.org/10.1111/j.1572-0241.2000.02322.x

134. Islam S, Antonsson L, Westin J, Lagging M. Cirrhosis in hepatitis C virus-infected patients can be excluded using an index of standard biochemical serum markers. *Scand J Gastroenterol* 2005;**40**:867–72. http://dx.doi.org/10.1080/00365520510015674

135. Iushchuk ND, Znoiko OO, Safiullina NK, Dudina KR, Kelli EI, Klimova EA, *et al.* [Diagnostic significance of type IV collagen and hyaluronic acid in the serum of patients with chronic hepatitis C for staging hepatic fibrosis.] *Ter Arkh* 2005;**77**:50–5.

136. Kalantari H, Hoseini H, Babak A, Yaran M. Validation of hepascore as a predictor of liver fibrosis in patients with chronic hepatitis C infection. *Hepat Res Treat* 2011;**2011**:972759. http://dx.doi.org/10.1155/2011/972759

137. Kamphues C, Lotz K, Rocken C, Berg T, Eurich D, Pratschke J, *et al.* Chances and limitations of non-invasive tests in the assessment of liver fibrosis in liver transplant patients. *Clin Transplant* 2010;**24**:652–9. http://dx.doi.org/10.1111/j.1399-0012.2009.01152.x

138. Kandemir O, Polat G, Saracoglu G, Tasdelen B. The predictive role of AST level, prothrombin time, and platelet count in the detection of liver fibrosis in patients with chronic hepatitis C. *Turk J Med Sci* 2009;**39**:857–62.

139. Kelleher TB, Mehta SH, Bhaskar R, Sulkowski M, Astemborski J, Thomas DL, *et al.* Prediction of hepatic fibrosis in HIV/HCV co-infected patients using serum fibrosis markers: the SHASTA index. *J Hepatol* 2005;**43**:78–84. http://dx.doi.org/10.1016/j.jhep.2005.02.025

140. Khan DA, Fatima Tuz Z, Khan FA, Mubarak A. Evaluation of diagnostic accuracy of APRI for prediction of fibrosis in hepatitis C patients. *J Ayub Med Coll Abbottabad* 2008;**20**:122–6.

141. Kim SU, Jang HW, Cheong JY, Kim JK, Lee MH, Kim DJ, *et al.* The usefulness of liver stiffness measurement using FibroScan in chronic hepatitis C in South Korea: a multicenter, prospective study. *J Gastroenterol Hepatol* 2011;**26**:171–8. http://dx.doi.org/10.1111/j.1440-1746.2010.06385.x

142. Koizumi Y, Hirooka M, Kisaka Y, Konishi I, Abe M, Murakami H, *et al.* Liver fibrosis in patients with chronic hepatitis C: noninvasive diagnosis by means of real-time tissue elastography – establishment of the method for measurement. *Radiology* 2011;**258**:610–17. http://dx.doi.org/10.1148/radiol.10100319

143. Lackner C, Struber G, Liegl B, Leibl S, Ofner P, Bankuti C, *et al.* Comparison and validation of simple noninvasive tests for prediction of fibrosis in chronic hepatitis C. *Hepatology* 2005;**41**:1376–82. http://dx.doi.org/10.1002/hep.20717

144. Ladero JM, Delkader J, Ortega L, Fernandez C, Devesa MJ, Lopez-Alonso G, *et al.* Non-invasive evaluation of the fibrosis stage in chronic hepatitis C: a comparative analysis of nine scoring methods. *Scand J Gastroenterol* 2010;**45**:51–9. http://dx.doi.org/10.3109/00365520903305544

145. Lee VS, Miller FH, Omary RA, Wang Y, Ganger DR, Wang E, *et al.* Magnetic resonance elastography and biomarkers to assess fibrosis from recurrent hepatitis C in liver transplant recipients. *Transplantation* 2011;**92**:581–6. http://dx.doi.org/10.1097/TP.0b013e31822805fa

146. Leroy V, Hilleret M-N, Sturm N, Trocme C, Renversez J-C, Faure P, *et al.* Prospective comparison of six non-invasive scores for the diagnosis of liver fibrosis in chronic hepatitis C. *J Hepatol* 2007;**46**:775–82. http://dx.doi.org/10.1016/j.jhep.2006.12.013

147. Leroy V, Monier F, Bottari S, Trocme C, Sturm N, Hilleret M-N, *et al.* Circulating matrix metalloproteinases 1, 2, 9 and their inhibitors TIMP-1 and TIMP-2 as serum markers of liver fibrosis in patients with chronic hepatitis C: comparison with PIIINP and hyaluronic acid. *Am J Gastroenterol* 2004;**99**:271–9. http://dx.doi.org/10.1111/j.1572-0241.2004.04055.x

148. Leroy V, Sturm N, Guechot J, Zafrani ES, Paris A, Bosson JL, *et al.* Steato-hepatitis has a strong impact on liver stiffness measured by transient elastography in chronic hepatitis C patients. *J Hepatol* 2011;**54**:S135. http://dx.doi.org/10.1016/S0168-8278(11)60340-1

149. Lewin M, Poujol-Robert A, Boelle P-Y, Wendum D, Lasnier E, Viallon M, *et al.* Diffusion-weighted magnetic resonance imaging for the assessment of fibrosis in chronic hepatitis C. *Hepatology* 2007;**46**:658–65. http://dx.doi.org/10.1002/hep.21747

150. Lieber CS, Weiss DG, Morgan TR, Paronetto F. Aspartate aminotransferase to platelet ratio index in patients with alcoholic liver fibrosis. *Am J Gastroenterol* 2006;**101**:1500–8. http://dx.doi.org/10.1111/j.1572-0241.2006.00610.x

151. Liu C-H, Lin J-W, Tsai F-C, Yang P-M, Lai M-Y, Chen J-H, *et al.* Noninvasive tests for the prediction of significant hepatic fibrosis in hepatitis C virus carriers with persistently normal alanine aminotransferases. *Liver Int* 2006;**26**:1087–94. http://dx.doi.org/10.1111/j.1478-3231.2006.01355.x

152. Liu CH, Hsu SJ, Lin JW, Hwang JJ, Liu CJ, Yang PM, *et al.* Noninvasive diagnosis of hepatic fibrosis in patients with chronic hepatitis C by splenic Doppler impedance index. *Clin Gastroenterol Hepatol* 2007;**5**:1199–206. http://dx.doi.org/10.1016/j.cgh.2007.07.017

153. Liu CH, Liang CC, Huang KW, Liu CJ, Chen SI, Lin JW, *et al.* Transient elastography to assess hepatic fibrosis in hemodialysis chronic hepatitis C patients. *Clin J Am Soc Nephrol* 2011;**6**:1057–64. http://dx.doi.org/10.2215/CJN.04320510

154. Loko MA, Castera L, Dabis F, Le Bail B, Winnock M, Coureau G, *et al.* Validation and comparison of simple noninvasive indexes for predicting liver fibrosis in HIV-HCV-coinfected patients: ANRS CO3 aquitaine cohort. *Am J Gastroenterol* 2008;**103**:1973–80. http://dx.doi.org/10.1111/j.1572-0241.2008.01954.x

155. Lupsor M, Badea R, Stefanescu H, Grigorescu M, Sparchez Z, Serban A, *et al.* Analysis of histopathological changes that influence liver stiffness in chronic hepatitis C: results from a cohort of 324 patients. *J Gastrointestin Liver Dis* 2008;**17**:155–63.

156. Macias J, Giron-Gonzalez JA, Gonzalez-Serrano M, Merino D, Cano P, Mira JA, *et al.* Prediction of liver fibrosis in human immunodeficiency virus/hepatitis C virus coinfected patients by simple non-invasive indexes. *Gut* 2006;**55**:409–14. http://dx.doi.org/10.1136/gut.2005.065904

157. Macias J, Mira J, Gilabert I, Neukam K, Roldan C, Viloria M, *et al.* Combined use of aspartate aminotransferase, platelet count and matrix metalloproteinase 2 measurements to predict liver fibrosis in HIV/hepatitis C virus-coinfected patients. *HIV Med* 2011;**12**:14–21. http://dx.doi.org/10.1111/j.1468-1293.2010.00836.x

158. Martinez SM, Fernandez-Varo G, Gonzalez P, Sampson E, Bruguera M, Navasa M, *et al.* Assessment of liver fibrosis before and after antiviral therapy by different serum marker panels in patients with chronic hepatitis C. *Aliment Pharmacol Ther* 2011;**33**:138–48. http://dx.doi.org/10.1111/j.1365-2036.2010.04500.x

159. Morikawa H, Fukuda K, Kobayashi S, Fujii H, Iwai S, Enomoto M, *et al.* Real-time tissue elastography as a tool for the noninvasive assessment of liver stiffness in patients with chronic hepatitis C. *J Gastroenterology* 2011;**46**:350–8. http://dx.doi.org/10.1007/s00535-010-0301-x

160. Murawaki Y, Koda M, Okamoto K, Mimura K, Kawasaki H. Diagnostic value of serum type IV collagen test in comparison with platelet count for predicting the fibrotic stage in patients with chronic hepatitis C. *J Gastroenterol Hepatol* 2001;**16**:777–81. http://dx.doi.org/10.1046/j.1440-1746.2001.02515.x

161. Nitta Y, Kawabe N, Hashimoto S, Harata M, Komura N, Kobayashi K, *et al.* Liver stiffness measured by transient elastography correlates with fibrosis area in liver biopsy in patients with chronic hepatitis C. *Hepatol Res* 2009;**39**:675–84. http://dx.doi.org/10.1111/j.1872-034X.2009.00500.x

162. Nojiri S. Noninvasive evaluation of hepatic fibrosis in HCV-infected patients using EOB-MR imaging. *J Hepatol* 2010;**52**:S169–70. http://dx.doi.org/10.1016/S0168-8278(10)60417-5

163. Nunes D, Fleming C, Offner G, O'Brien M, Tumilty S, Fix O, *et al.* HIV infection does not affect the performance of noninvasive markers of fibrosis for the diagnosis of hepatitis C virus-related liver disease. *J Acquir Immune Defic Syndr* 2005;**40**:538–44. http://dx.doi.org/10.1097/01.qai.0000184856.31695.bf

164. Obara N, Ueno Y, Fukushima K, Nakagome Y, Kakazu E, Kimura O, *et al.* Transient elastography for measurement of liver stiffness measurement can detect early significant hepatic fibrosis in Japanese patients with viral and nonviral liver diseases. *J Gastroenterol* 2008;**43**:720–8. http://dx.doi.org/10.1007/s00535-008-2225-2

165. Oliveira AC, Santos VN, Salgado AL, Lanzoni VP, Nogueira MD, Martins JR, *et al.* Serum hyaluronic acid and the AST/ALT ratio in the diagnosis of advanced fibrosis in patients with non alcoholic liver disease and chronic hepatitis C. *J Hepatol* 2005;**42**:250. http://dx.doi.org/10.1016/S0168-8278(05)82096-3

166. Orlacchio A, Bolacchi F, Petrella MC, Pastorelli D, Bazzocchi G, Angelico M, *et al.* Liver contrast enhanced ultrasound perfusion imaging in the evaluation of chronic hepatitis C fibrosis: preliminary results. *Ultrasound Med Biol* 2011;**37**:1–6. http://dx.doi.org/10.1016/j.ultrasmedbio.2010.10.012

167. Paggi S, Colli A, Fraquelli M, Vigano M, Del Poggio P, Facciotto C, *et al.* A non-invasive algorithm accurately predicts advanced fibrosis in hepatitis C: a comparison using histology with internal-external validation. *J Hepatol* 2008;**49**:564–71. http://dx.doi.org/10.1016/j.jhep.2008.07.007

168. Parise ER, Oliveira AC, Figueiredo-Mendes C, Lanzoni V, Martins J, Nader H, *et al.* Noninvasive serum markers in the diagnosis of structural liver damage in chronic hepatitis C virus infection. *Liver Int* 2006;**26**:1095–9. http://dx.doi.org/10.1111/j.1478-3231.2006.01356.x

169. Park GJ, Lin BP, Ngu MC, Jones DB, Katelaris PH. Aspartate aminotransferase: alanine aminotransferase ratio in chronic hepatitis C infection: is it a useful predictor of cirrhosis? *J Gastroenterol Hepatol* 2000;**15**:386–90. http://dx.doi.org/10.1046/j.1440-1746.2000.02172.x

170. Patel K, Friedrich-Rust M, Lurie Y, Grigorescu M, Stanciu C, Lee C-M, *et al.* FibroSURE and FibroScan in relation to treatment response in chronic hepatitis C virus. *World J Gastroenterol* 2011;**17**:4581–9. http://dx.doi.org/10.3748/wjg.v17.i41.4581

171. Pohl A, Behling C, Oliver D, Kilani M, Monson P, Hassanein T. Serum aminotransferase levels and platelet counts as predictors of degree of fibrosis in chronic hepatitis C virus infection. *Am J Gastroenterol* 2001;**96**:3142–6. http://dx.doi.org/10.1111/j.1572-0241.2001.05268.x

172. Poynard T, De Ledinghen V, Zarski JP, Stanciu C, Munteanu M, Vergniol J, *et al.* Relative performances of FibroTest, Fibroscan, and biopsy for the assessment of the stage of liver fibrosis in patients with chronic hepatitis C: a step toward the truth in the absence of a gold standard. *J Hepatol* 2012;**56**:541–8. http://dx.doi.org/10.1016/j.jhep.2011.08.007

173. Prati GM, D'Ambrosio R, Fraquelli M, Aghemo A, Rumi MG, Ronchi G, *et al.* The diagnostic accuracy of fibroscan for cirrhosis is influenced by liver morphometry in HCV patients with an SVR. *Hepatology* 2011;**54**:560A.

174. Qiu Y, Hoshida YJ, Kato N, Moriyama M, Otsuka M, Taniguchi H, *et al.* A simple combination of serum type IV collagen and prothrombin time to diagnose cirrhosis in patients with chronic active hepatitis C. *Hepatol Res* 2004;**30**:214–20. http://dx.doi.org/10.1016/j.hepres.2004.10.006

175. Reedy DW, Loo AT, Levine RA. AST/ALT ratio <=1 is not diagnostic of cirrhosis in patients with chronic hepatitis C. *Dig Dis Sci* 1998;**43**:2156–9. http://dx.doi.org/10.1023/A:1018888021118

176. Ronot M, Asselah T, Paradis V, Michoux N, Dorvillius M, Baron G, *et al.* Liver fibrosis in chronic hepatitis C virus infection: differentiating minimal from intermediate fibrosis with perfusion CT. *Radiology* 2010;**256**:135–42. http://dx.doi.org/10.1148/radiol.10091295

177. Rossi E, Adams L, Prins A, Bulsara M, De Boer B, Garas G, *et al.* Validation of the FibroTest biochemical markers score in assessing liver fibrosis in hepatitis C patients. *Clin Chem* 2003;**49**:450–4. http://dx.doi.org/10.1373/49.3.450

178. Rossini A, Cabassa P, Gatti E, Contessi GB, Morone M, Maroldi R. Assessment of liver fibrosis by acoustic radiation force elastography in healthy subjects and patients with chronic hepatitis C. *Hepatology* 2010;**52**:1234A.

179. Said Y, Salem M, Mouelhi L, Mekki H, Houissa F, Ben Rejeb M, *et al.* Correlation between liver biopsy and fibrotest in the evaluation of hepatic fibrosis in patients with chronic hepatitis C. *Tunis Med* 2010;**88**:573–8.

180. Saitou Y, Shiraki K, Yamanaka Y, Yamaguchi Y, Kawakita T, Yamamoto N, *et al.* Noninvasive estimation of liver fibrosis and response to interferon therapy by a serum fibrogenesis marker, YKL-40, in patients with HCV-associated liver disease. *World J Gastroenterol* 2005;**11**:476–81.

181. Sanvisens A, Serra I, Tural C, Tor J, Ojanguren I, Barluenga E, *et al.* Hyaluronic acid, transforming growth factor-beta1 and hepatic fibrosis in patients with chronic hepatitis C virus and human immunodeficiency virus co-infection. *J Viral Hepat* 2009;**16**:513–18. http://dx.doi.org/10.1111/j.1365-2893.2009.01103.x

182. Schiavon LL, Filho RJC, Narciso JL, Sampaio JP, Lanzoni VP, Ferraz MLG, *et al.* Expanding the applicability of noninvasive fibrosis markers in HIV/HCV co-infected patients. *Hepatology* 2007;**45**:257–8. http://dx.doi.org/10.1002/hep.21507

183. Schiavon LL, Narciso-Schiavon JL, Carvalho Filho RJ, Sampaio JP, Medina-Pestana JO, Lanzoni VP, *et al.* Serum levels of YKL-40 and hyaluronic acid as noninvasive markers of liver fibrosis in haemodialysis patients with chronic hepatitis C virus infection. *J Viral Hepatol* 2008;**15**:666–74. http://dx.doi.org/10.1111/j.1365-2893.2008.00992.x

184. Schneider ARJ, Teuber G, Kriener S, Caspary WF. Noninvasive assessment of liver steatosis, fibrosis and inflammation in chronic hepatitis C virus infection. *Liver Int* 2005;**25**:1150–5. http://dx.doi.org/10.1111/j.1478-3231.2005.01164.x

185. Schneider ARJ, Teuber G, Paul K, Nikodem A, Duesterhoeft M, Caspary WF, *et al.* Patient age is a strong independent predictor of 13C-aminopyrine breath test results: a comparative study with histology, duplex-Doppler and a laboratory index in patients with chronic hepatitis C virus infection. *Clin Exp Pharmacol Physiol* 2006;**33**:300–4. http://dx.doi.org/10.1111/j.1440-1681.2006.04365.x

186. Sebastian G, Vario A, Guido M, Alberti A. Performance of noninvasive markers for liver fibrosis is reduced in chronic hepatitis C with normal transaminases. *J Viral Hepatol* 2008;**15**:212–18. http://dx.doi.org/10.1111/j.1365-2893.2007.00932.x

187. Sebastiani G, Halfon P, Castera L, Pol S, Thomas DL, Mangia A, *et al.* SAFE biopsy: a validated method for large-scale staging of liver fibrosis in chronic hepatitis C. *Hepatology* 2009;**49**:1821–7. http://dx.doi.org/10.1002/hep.22859

188. Sebastiani G, Vario A, Guido M, Noventa F, Plebani M, Pistis R, *et al.* Stepwise combination algorithms of non-invasive markers to diagnose significant fibrosis in chronic hepatitis C. *J Hepatol* 2006;**44**:686–93. http://dx.doi.org/10.1016/j.jhep.2006.01.007

189. Sene D, Limal N, Messous D, Ghillani-Dalbin P, Charlotte F, Thiolliere JM, *et al.* Biological markers of liver fibrosis and activity as non-invasive alternatives to liver biopsy in patients with chronic hepatitis C and associated mixed cryoglobulinemia vasculitis. *Clin Biochem* 2006;**39**:715–21. http://dx.doi.org/10.1016/j.clinbiochem.2006.04.019

190. Sharabash NM, Jensen DM, Cao D, Reau N, Mohanty SR, Satoskar RS, *et al.* Fibrospect II (FS II) score overestimates the degree of fibrosis in chronic hepatitis C (CHC) patients with significant renal dysfunction. *Hepatology* 2009;**50**:1062A.

191. Shastry L, Wilson T, Lascher S, Nord JA. The utility of aspartate aminotransferase/platelet ratio index in HIV/hepatitis C-co-infected patients. *AIDS* 2007;**21**:2541–3. http://dx.doi.org/10.1097/QAD.0b013e3282f1fe59

192. Sheth SG, Flamm SL, Gordon FD, Chopra S. AST/ALT ratio predicts cirrhosis in patients with chronic hepatitis C virus infection. *Am J Gastroenterol* 1998;**93**:44–8. http://dx.doi.org/10.1111/j.1572-0241.1998.044_c.x

193. Singal AG, Thomassen LV, Gretch DR, Shuhart MC. Use of the AST to platelet ratio index in HCV/HIV co-infected patients. *Aliment Pharmacol Ther* 2011;**33**:566–77. http://dx.doi.org/10.1111/j.1365-2036.2010.04560.x

194. Sirli R, Sporea I, Bota S, Popescu A, Cornianu M. A comparative study of non-invasive methods for fibrosis assessment in chronic HCV infection. *Hepat Mon* 2010;**10**:88–94.

195. Snyder N, Gajula L, Xiao S-Y, Grady J, Luxon B, Lau DTY, *et al.* APRI: an easy and validated predictor of hepatic fibrosis in chronic hepatitis C. *J Clin Gastroenterol* 2006;**40**:535–42. http://dx.doi.org/10.1097/00004836-200607000-00013

196. Snyder N, Nguyen A, Gajula L, Soloway R, Xiao SY, Lau DTY, *et al.* The APRI may be enhanced by the use of the FIBROSpect II in the estimation of fibrosis in chronic hepatitis C. *Clin Chim Acta* 2007;**381**:119–23. http://dx.doi.org/10.1016/j.cca.2007.02.046

197. Sohn JH, Kim TY, Jun DW, Eun CS, Jeon YC, Han DS. Platelet count/spleen diameter ratio is more accurate for prediction of significant and extensive fibrosis than AST-to-platelet ratio index (APRI) in patients with chronic hepatitis C. *J Hepatol* 2010;**52**:S173. http://dx.doi.org/10.1016/S0168-8278(10)60426-6

198. Sporea I, Sirli R, Bota S, Tanaka H, Iijima H, Badea RI, *et al.* Is ARFI elastography useful for fibrosis evaluation in patients with chronic HCV hepatitis in our clinical practice? – A multicenter international study. *Hepatology* 2011;**54**:559A–60A. http://dx.doi.org/10.1016/j.ejrad.2012.08.018

199. Sporea I, Sirli R, Deleanu A, Tudora A, Curescu M, Cornianu M, *et al.* Comparison of the liver stiffness measurement by transient elastography with the liver biopsy. *World J Gastroenterol* 2008;**14**:6513–17. http://dx.doi.org/10.3748/wjg.14.6513

200. Sporea I, Sirli R, Deleanu A, Tudora A, Popescu A, Curescu M, *et al.* Liver stiffness measurements in patients with HBV vs HCV chronic hepatitis: a comparative study. *World J Gastroenterol* 2010;**16**:4832–7. http://dx.doi.org/10.3748/wjg.v16.i38.4832

201. Sporea I, Sirli R, Popescu A, Badea R, Lupsor M, Focsa M, *et al.* Is it better to use together transient elastography (TE) and acoustic radiation force impulse elastography (ARFI) for fibrosis evaluation in patients with chronic HCV hepatitis? *Gastroenterology* 2011;**1**):S968.

202. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, *et al.* Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 2006;**43**:1317–25. http://dx.doi.org/10.1002/hep.21178

203. Stibbe KJM, Verveer C, Francke J, Hansen BE, Zondervan PE, Kuipers EJ, *et al.* Comparison of non-invasive assessment to diagnose liver fibrosis in chronic hepatitis B and C patients. *Scand J Gastroenterol* 2011;**46**:962–72. http://dx.doi.org/10.3109/00365521.2011.574725

204. Sud A, Hui JM, Farrell GC, Bandara P, Kench JG, Fung C, *et al.* Improved prediction of fibrosis in chronic hepatitis C using measures of insulin resistance in a probability index. *Hepatology* 2004;**39**:1239–47. http://dx.doi.org/10.1002/hep.20207

205. Testa R, Testa E, Giannini E, Borro P, Milazzo S, Isola L, *et al.* Noninvasive ratio indexes to evaluate fibrosis staging in chronic hepatitis C: role of platelet count/spleen diameter ratio index. *J Intern Med* 2006;**260**:142–50. http://dx.doi.org/10.1111/j.1365-2796.2006.01673.x

206. Thompson A, Elliott L, Tillmann H, McHutchison J, Patel K. Paired biopsy comparison of two liver fibrosis marker panels (HCV fibrotest, hepascore) in following histological response to antiviral therapy for chronic hepatitis C (CHC). *Hepatology* 2009;**50**:S159. http://dx.doi.org/10.1016/S0168-8278(09)60425-6

207. Thompson AJ, Clark PJ, Tillmann HL, Torbenson M, Pulkstenis E, Subramanian M, *et al.* Independent validation of the safe biopsy algorithm: analysis of the achieve cohorts. *Gastroenterology* 2010;**1**:S780.

208. Thompson AJ, Elliott L, Tillmann HL, McHutchison J, Patel K. A large independent, single center comparison of two commercial fibrosis marker panels (HCV FibroSure and HepaScore) for moderate-to-advanced stage fibrosis in CHC patients. *Gastroenterology* 2009;**136**(Suppl. 1):A834. http://dx.doi.org/10.1016/S0016-5085(09)63844-0

209. Toniutto P, Fabris C, Bitetto D, Falleti E, Avellini C, Rossi E, *et al.* Role of AST to platelet ratio index in the detection of liver fibrosis in patients with recurrent hepatitis C after liver transplantation. *J Gastroenterol Hepatol* 2007;**22**:1904–8. http://dx.doi.org/10.1111/j.1440-1746.2006.04628.x

210. Trang T, Petersen JR, Snyder N. Non-invasive markers of hepatic fibrosis in patients co-infected with HCV and HIV: comparison of the APRI and FIB-4 index. *Clin Chim Acta* 2008;**397**:51–4. http://dx.doi.org/10.1016/j.cca.2008.07.009

211. Trifan A, Cojocariu C, Sfarti C, Stanciu C. Prospective clinical trial on the accuracy of non-invasive tests to predict liver fibrosis in chronic HCV patients. *Hepatology* 2009;**50**:1074A–5A.

212. Trocme C, Leroy V, Sturm N, Hilleret MN, Bottari S, Morel F, *et al.* Longitudinal evaluation of a fibrosis index combining MMP-1 and PIIINP compared with MMP-9, TIMP-1 and hyaluronic acid in patients with chronic hepatitis C treated by interferon-alpha and ribavirin. *J Viral Hepat* 2006;**13**:643–51. http://dx.doi.org/10.1111/j.1365-2893.2006.00730.x

213. Tural C, Tor J, Sanvisens A, Perez-Alvarez N, Martinez E, Ojanguren I, *et al.* Accuracy of simple biochemical tests in identifying liver fibrosis in patients co-infected with human immunodeficiency virus and hepatitis C virus. *Clin Gastroenterol Hepatol* 2009;**7**:339–45. http://dx.doi.org/10.1016/j.cgh.2008.11.019

214. Vallet-Pichard A, Mallet V, Nalpas B, Verkarre V, Nalpas A, Dhalluin-Venier V, *et al.* FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. Comparison with liver biopsy and FibroTest. *Hepatology* 2007;**46**:32–6. http://dx.doi.org/10.1002/hep.21669

215. Valva P, Casciato P, Diaz Carrasco JM, Gadano A, Galdame O, Galoppo MC, *et al.* The role of serum biomarkers in predicting fibrosis progression in pediatric and adult hepatitis C virus chronic infection. *PLOS One* 2011;**6**:e23218. http://dx.doi.org/10.1371/journal.pone.0023218

216. Varaut A, Fontaine H, Serpaggi J, Verkarre V, Vallet-Pichard A, Nalpas B, *et al.* Diagnostic accuracy of the fibrotest in hemodialysis and renal transplant patients with chronic hepatitis C virus. *Transplantation* 2005;**80**:1550–5. http://dx.doi.org/10.1097/01.tp.0000183399.85804.02

217. Viana MSVB, Takei K, Yamaguti DCC, Guz B, Strauss E. Use of AST platelet ratio index (APRI Score) as an alternative to liver biopsy for treatment indication in chronic hepatitis C. *Ann Hepatol* 2009;**8**:26–31.

218. Wai CT, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, Conjeevaram HS, *et al.* A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology* 2003;**38**:518–26. http://dx.doi.org/10.1053/jhep.2003.50346

219. Westin J, Ydreborg M, Islam S, Alsio A, Dhillon AP, Pawlotsky JM, *et al.* A non-invasive fibrosis score predicts treatment outcome in chronic hepatitis C virus infection. *Scand J Gastroenterol* 2008;**43**:73–80. http://dx.doi.org/10.1080/00365520701514461

220. Wilson LE, Torbenson M, Astemborski J, Faruki H, Spoler C, Rai R, *et al.* Progression of liver fibrosis among injection drug users with chronic hepatitis C. *Hepatology* 2006;**43**:788–95. http://dx.doi.org/10.1002/hep.21091

221. Wong VS, Hughes V, Trull A, Wight DG, Petrik J, Alexander GJ. Serum hyaluronic acid is a useful marker of liver fibrosis in chronic hepatitis C virus infection. *J Viral Hepatol* 1998;**5**:187–92. http://dx.doi.org/10.1046/j.1365-2893.1998.00100.x

222. Zaman A, Rosen H, Oh E, Ingram K, Smith K. Prospective assessment of FIBROSpect II hepatic to detect fibrosis in patients with chronic Hepatitis C. *Am J Gastroenterol* 2004;**99**;S79.

223. Zarski JP, Sturm N, Guechot J, Paris A, Zafrani ES, Asselah T, *et al.* Comparison of nine blood tests and transient elastography for liver fibrosis in chronic hepatitis C: the ANRS HCEP-23 study. *J Hepatol* 2012;**56**:55–62. http://dx.doi.org/10.1016/j.jhep.2011.05.024

224. Ziol M, Handra-Luca A, Kettaneh A, Christidis C, Mal F, Kazemi F, *et al.* Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with chronic hepatitis C. *Hepatology* 2005;**41**:48–54. http://dx.doi.org/10.1002/hep.20506

225. Castera L, Bernard PH, Le Bail B, Foucher J, Trimoulet P, Merrouche W, *et al.* Transient elastography and biomarkers for liver fibrosis assessment and follow-up of inactive hepatitis B carriers. *Aliment Pharmacol Ther* 2011;**33**:455–65. http://dx.doi.org/10.1111/j.1365-2036.2010.04547.x

226. Chan HLY, Wong GLH, Choi PCL, Chan AWH, Chim AML, Yiu KKL, *et al.* Alanine aminotransferase-based algorithms of liver stiffness measurement by transient elastography (Fibroscan) for liver fibrosis in chronic hepatitis B. *J Viral Hepatol* 2009;**16**:36–44. http://dx.doi.org/10.1111/j.1365-2893.2008.01037.x

227. Chen YP, Dai L, Wang JL, Zhu YF, Feng XR, Hou JL. Model consisting of ultrasonographic and simple blood indexes accurately identify compensated hepatitis B cirrhosis. *J Gastroenterol Hepatol* 2008;**23**:1228–34. http://dx.doi.org/10.1111/j.1440-1746.2008.05421.x

228. Chen YP, Liang XE, Dai L, Zhang Q, Peng J, Zhu YF, *et al.* Improving transient elastography performance for detecting hepatitis B cirrhosis. *Dig Liver Dis* 2012;**44**:61–6. http://dx.doi.org/10.1016/j.dld.2011.08.004

229. Dixon JB, Bhathal PS, O'Brien PE. Nonalcoholic fatty liver disease: predictors of nonalcoholic steatohepatitis and liver fibrosis in the severely obese. *Gastroenterology* 2001;**121**:91–100. http://dx.doi.org/10.1053/gast.2001.25540

230. Fung J, Lai CL, Cheng C, Wu R, Wong DKH, Yuen MF. Mild-to-moderate elevation of alanine aminotransferase increases liver stiffness measurement by transient elastography in patients with chronic hepatitis B. *Am J Gastroenterol* 2011;**106**:492–6. http://dx.doi.org/10.1038/ajg.2010.463

231. Gui H-l, Gao C-f, Wang H, Liu X-e, Xie Q, Dewaele S, *et al.* Altered serum N-glycomics in chronic hepatitis B patients. *Liver Int* 2010;**30**:259–67. http://dx.doi.org/10.1111/j.1478-3231.2009.02170.x

232. Guo-Qiu W, Nai-Feng L, Xiao-Bo V, Linxian L, Chen Z, Lixia G, *et al.* The level of connective tissue growth factor in sera of patients with hepatitis B virus strongly correlates with stage of hepatic fibrosis. *Viral Immunol* 2010;**23**:71–8. http://dx.doi.org/10.1089/vim.2009.0067

233. Hongbo L, Xiaohui L, Hong K, Wei W, Yong Z. Assessing routine and serum markers of liver fibrosis in CHB patients using parallel and serial interpretation. *Clin Biochem* 2007;**40**:562–6. http://dx.doi.org/10.1016/j.clinbiochem.2007.01.022

234. Hu X, Shao J, Bai J, Wang J, Qian L. New noninvasive assessment of liver fibrosis in chronic hepatitis B: maximal accumulative respiration strain. *J Ultrasound Med* 2010;**29**:1213–21.

235. Hui AY, Chan HL-Y, Wong VW-S, Liew C-T, Chim AM-L, Chan FK-L, *et al.* Identification of chronic hepatitis B patients without significant liver fibrosis by a simple noninvasive predictive model. *Am J Gastroenterol* 2005;**100**:616–23. http://dx.doi.org/10.1111/j.1572-0241.2005.41289.x

236. Kim BK, Kim DY, Park JY, Ahn SH, Chon CY, Kim JK, *et al.* Validation of FIB-4 and comparison with other simple noninvasive indices for predicting liver fibrosis and cirrhosis in hepatitis B virus-infected patients. *Liver Int* 2010;**30**:546–53. http://dx.doi.org/10.1111/j.1478-3231.2009.02192.x

237. Kim BK, Kim SA, Park YN, Cheong JY, Kim HS, Park JY, *et al.* Noninvasive models to predict liver cirrhosis in patients with chronic hepatitis B. *Liver Int* 2007;**27**:969–76. http://dx.doi.org/10.1111/j.1478-3231.2007.01519.x

238. Kim DY, Kim SU, Ahn SH, Park JY, Lee JM, Park YN, *et al.* Usefulness of FibroScan for detection of early compensated liver cirrhosis in chronic hepatitis B. *Dig Dis Sci* 2009;**54**:1758–63. http://dx.doi.org/10.1007/s10620-008-0541-2

239. Kim SU, Ahn SH, Park JY, Kang W, Kim DY, Park YN, *et al.* Liver stiffness measurement in combination with noninvasive markers for the improved diagnosis of B-viral liver cirrhosis. *J Clin Gastroenterol* 2009;**43**:267–71. http://dx.doi.org/10.1097/MCG.0b013e31816f212e

240. Kwok R, Gonzalez-Arce V, Kim A, Ngu MC, Lee AU. Evaluation of hepatic fibrosis in chronic hepatitis B using transient elastography. *J Gastroenterol Hepatol* 2009;**24**:A283.

241. Lee IC, Chan C-C, Huang Y-H, Huo T-I, Chu C-J, Lai C-R, *et al.* Comparative analysis of noninvasive models to predict early liver fibrosis in hepatitis B e antigen-negative chronic hepatitis B. *J Clin Gastroenterol* 2011;**45**:278–85. http://dx.doi.org/10.1097/MCG.0b013e3181dd5357

242. Lesmana CRA, Salim S, Hasan I, Sulaiman AS, Gani RA, Pakasi LS, *et al.* Diagnostic accuracy of transient elastography (FibroScan) versus the aspartate transaminase to platelet ratio index in assessing liver fibrosis in chronic hepatitis B: the role in primary care setting. *J Clin Pathol* 2011;**64**:916–20. http://dx.doi.org/10.1136/jclinpath-2011-200044

243. Li F, Zhu C-L, Zhang H, Huang H, Wei Q, Zhu X, *et al.* Role of hyaluronic acid and laminin as serum markers for predicting significant fibrosis in patients with chronic hepatitis B. *Braz J Infect Dis* 2012;**16**:9–14. http://dx.doi.org/10.1016/S1413-8670(12)70267-2

244. Liu H-B, Zhou J-P, Zhang Y, Lv X-H, Wang W. Prediction on liver fibrosis using different APRI thresholds when patient age is a categorical marker in patients with chronic hepatitis B. *Clin Chim Acta* 2011;**412**:33–7. http://dx.doi.org/10.1016/j.cca.2010.08.032

245. Mallet V, Dhalluin-Venier V, Roussin C, Bourliere M, Pettinelli ME, Giry C, *et al.* The accuracy of the FIB-4 index for the diagnosis of mild fibrosis in chronic hepatitis B. *Aliment Pharmacol Ther* 2009;**29**:409–15. http://dx.doi.org/10.1111/j.1365-2036.2008.03895.x

246. Marcellin P, Ziol M, Bedossa P, Douvin C, Poupon R, De Ledinghen V, *et al.* Non-invasive assessment of liver fibrosis by stiffness measurement in patients with chronic hepatitis B. *Liver Int* 2009;**29**:242–7. http://dx.doi.org/10.1111/j.1478-3231.2008.01802.x

247. Miailhes P, Pradat P, Chevallier M, Lacombe K, Bailly F, Cotte L, *et al.* Proficiency of transient elastography compared to liver biopsy for the assessment of fibrosis in HIV/HBV-coinfected patients. *J Viral Hepat* 2011;**18**:61–9. http://dx.doi.org/10.1111/j.1365-2893.2010.01275.x

248. Mohamadnejad M, Montazeri G, Fazlollahi A, Zamani F, Nasiri J, Nobakht H, *et al.* Noninvasive markers of liver fibrosis and inflammation in chronic hepatitis B-virus related liver disease. *Am J Gastroenterol* 2006;**101**:2537–45. http://dx.doi.org/10.1111/j.1572-0241.2006.00788.x

249. Myers RP, Tainturier MH, Ratziu V, Piton A, Thibault V, Imbert-Bismut F, *et al.* Prediction of liver histological lesions with biochemical markers in patients with chronic hepatitis B. *J Hepatol* 2003;**39**:222–30. http://dx.doi.org/10.1016/S0168-8278(03)00171-5

250. Ogawa E, Furusyo N, Murata M, Ohnishi H, Toyoda K, Taniai H, *et al.* Longitudinal assessment of liver stiffness by transient elastography for chronic hepatitis B patients treated with nucleoside analog. *Hepatol Res* 2011;**41**:1178–88. http://dx.doi.org/10.1111/j.1872-034X.2011.00869.x

251. Osakabe K, Ichino N, Nishikawa T, Sugiyama H, Kato M, Kitahara S, *et al.* Reduction of liver stiffness by antiviral therapy in chronic hepatitis B. *J Gastroenterol* 2011;**46**:1324–34. http://dx.doi.org/10.1007/s00535-011-0444-4

252. Park GJ, Katelaris PH, Jones DB, Seow F, Lin BP, Le Couteur DG, *et al.* The C-caffeine breath test distinguishes significant fibrosis in chronic hepatitis B and reflects response to lamivudine therapy. *Aliment Pharmacol Ther* 2005;**22**:395–403. http://dx.doi.org/10.1111/j.1365-2036.2005.02623.x

253. Park JH, Park CK, Kim ES, Park SY, Jo CM, Tak WY, *et al.* [The diagnostic value of serum hyaluronic acid, 7S domain of type IV collagen and AST/ALT ratio as markers of hepatic fibrosis in chronic hepatitis B and cirrhosis patients.] *Taehan Kan Hakhoe Chi* 2003;**9**:79–88.

254. Park SY, Kang KH, Park JH, Lee JH, Cho CM, Tak WY, *et al.* [Clinical efficacy of AST/ALT ratio and platelet counts as predictors of degree of fibrosis in HBV infected patients without clinically evident liver cirrhosis.] *Korean J Gastroenterol* 2004;**43**:246–51.

255. Poynard T, Ngo Y, Marcellin P, Hadziyannis S, Ratziu V, Benhamou Y. Impact of adefovir dipivoxil on liver fibrosis and activity assessed with biochemical markers (FibroTest-ActiTest) in patients infected by hepatitis B virus. *J Viral Hepat* 2009;**16**:203–13. http://dx.doi.org/10.1111/j.1365-2893.2008.01065.x

256. Raftopoulos SC, George J, Bourliere M, Rossi E, de Boer WB, Jeffrey GP, *et al.* Comparison of noninvasive models of fibrosis in chronic hepatitis B. *Hepatol Int* 2012;**6**:457–67. http://dx.doi.org/10.1007/s12072-011-9296-5

257. Sebastiani G, Vario A, Guido M, Alberti A. Sequential algorithms combining non-invasive markers and biopsy for the assessment of liver fibrosis in chronic hepatitis B. *World J Gastroenterol* 2007;**13**:525–31. http://dx.doi.org/10.3748/wjg.v13.i4.525

258. Seto W-K, Lee C-F, Lai C-L, Ip PPC, Fong DY-T, Fung J, *et al.* A new model using routinely available clinical parameters to predict significant liver fibrosis in chronic hepatitis B. *PLOS One* 2011;**6**:e23077. http://dx.doi.org/10.1371/journal.pone.0023077

259. Shin WG, Park SH, Jang MK, Hahn TH, Kim JB, Lee MS, *et al.* Aspartate aminotransferase to platelet ratio index (APRI) can predict liver fibrosis in chronic hepatitis B. *Dig Liver Dis* 2008;**40**:267–74. http://dx.doi.org/10.1016/j.dld.2007.10.011

260. Sinakos E, Manolakopoulos S, Papatheodoridis G, Papalavrentios L, Papageorgiou MV, Papachrysos N, *et al.* Transient elastography (FibroScan) in patients with chronic hepatitis B in everyday clinical practice. *J Hepatol* 2011;**54**:S140–S1. http://dx.doi.org/10.1016/S0168-8278(11)60352-8

261. Sohn JH, Kim TY, Jeong JY, Jun DW, Eun CS, Jeon YC, *et al.* Insulin resistance correlates with the severity of fibrosis in chronic hepatitis B. *Hepatol Int* 2011;**5**:312.

262. Sokucu S, Gokce S, Gulluoglu M, Aydogan A, Celtik C, Durmaz O. The role of the non-invasive serum marker FibroTestActiTest in the prediction of histological stage of fibrosis and activity in children with nave chronic hepatitis B infection. *Scand J Infect Dis* 2010;**42**:699–703. http://dx.doi.org/10.3109/00365541003774616

263. Sporea I, Sirli R, Popescu A, Danila M. Acoustic Radiation Force Impulse (ARFI) – a new modality for the evaluation of liver fibrosis. *Med Ultrason* 2010;**12**:26–31.

264. Vigano M, Paggi S, Lampertico P, Fraquelli M, Massironi S, Ronchi G, *et al.* Dual cut-off transient elastography to assess liver fibrosis in chronic hepatitis B: a cohort study with internal validation. *Aliment Pharmacol Ther* 2011;**34**:353–62. http://dx.doi.org/10.1111/j.1365-2036.2011.04722.x

265. Wang J, Guo L, Shi XY, Pan WQ, Bai YF, Ai H. Real-time elastography with a novel quantitative technology for assessment of liver fibrosis in chronic hepatitis B. *Eur J Radiol* 2012;**81**:E31–E6. http://dx.doi.org/10.1016/j.ejrad.2010.12.013

266. Wong GLH, Wong VWS, Choi PCL, Chan AWH, Chan HLY. Development of a non-invasive algorithm with transient elastography (Fibroscan) and serum test formula for advanced liver fibrosis in chronic hepatitis B. *Aliment Pharmacol Ther* 2010;**31**:1095–103. http://dx.doi.org/10.1111/j.1365-2036.2010.04276.x

267. Wong GLH, Wong VWS, Choi PCL, Chan AWH, Chim AML, Yiu KKL, *et al.* Evaluation of alanine transaminase and hepatitis B Virus DNA to predict liver cirrhosis in hepatitis B e antigen-negative chronic hepatitis B using transient elastography. *Am J Gastroenterol* 2008;**103**:3071–81. http://dx.doi.org/10.1111/j.1572-0241.2008.02157.x

268. Wong GLH, Wong VWS, Choi PCL, Chan AWH, Chim AML, Yiu KKL, *et al.* On-treatment monitoring of liver fibrosis with transient elastography in chronic hepatitis B patients. *Antivir Ther* 2011;**16**:165–72. http://dx.doi.org/10.3851/IMP1726

269. Wu SD, Ni YJ, Liu LL, Li H, Lu LG, Wang JY. Establishment and validation of a simple noninvasive model to predict significant liver fibrosis in patients with chronic hepatitis B. *Hepatol Int* 2012;**6**:360–8. http://dx.doi.org/10.1007/s12072-011-9328-1

270. Zhang MB, Qu EZ, Liu JB, Wang JR. Quantitative assessment of hepatic fibrosis by contrast-enhanced ultrasonography. *Chin Med Sci J* 2011;**26**:208–15. http://dx.doi.org/10.1016/S1001-9294(12)60002-9

271. Zhang Y-X, Wu W-J, Zhang Y-Z, Feng Y-L, Zhou X-X, Pan Q. Noninvasive assessment of liver fibrosis with combined serum aminotransferase/platelet ratio index and hyaluronic acid in patients with chronic hepatitis B. *World J Gastroenterol* 2008;**14**:7117–21. http://dx.doi.org/10.3748/wjg.14.7117

272. Zhu X, Wang LC, Chen EQ, Chen XB, Chen LY, Liu L, *et al.* Prospective evaluation of fibroscan for the diagnosis of hepatic fibrosis in patients with chronic hepatitis B virus infection. *Hepatol Int* 2011;**5**:306.

273. Janssens F, De Suray N, Piessevaux H, Horsmans Y, De Timary P, Starkel P. Can transient elastography replace liver histology for determination of advanced fibrosis in alcoholic patients: a real-life study. *J Clin Gastroenterol* 2010;**44**:575–82. http://dx.doi.org/10.1097/MCG.0b013e3181cb4216

274. Kim SG, Kim YS, Jung SW, Kim HK, Jang JY, Moon JH, *et al.* [The usefulness of transient elastography to diagnose cirrhosis in patients with alcoholic liver disease.] *Korean J Hepatol* 2009;**15**:42–51. http://dx.doi.org/10.3350/kjhep.2009.15.1.42

275. Lavallard VJ, Bonnafous S, Patouraux S, Saint-Paul MC, Rousseau D, Anty R, *et al.* Serum markers of hepatocyte death and apoptosis are non invasive biomarkers of severe fibrosis in patients with alcoholic liver disease. *PLOS One* 2011;**6**:e17599. http://dx.doi.org/10.1371/journal.pone.0017599

276. Melin PDA, Gauchet A, Schoeny M, Fournier C, Sandrin L, Diebold MD. Cirrhosis screening in alcoholism consultation using fibroscan. *Hepatology* 2005;**42**(Suppl. 1):492A.

277. Mueller S, Millonig G, Sarovska L, Friedrich S, Reimann FM, Pritsch M, *et al.* Increased liver stiffness in alcoholic liver disease: differentiating fibrosis from steatohepatitis. *World J Gastroenterol* 2010;**16**:966–72. http://dx.doi.org/10.3748/wjg.v16.i8.966

278. Nahon P, Kettaneh A, Tengher-Barna I, Ziol M, de Ledinghen V, Douvin C, *et al.* Assessment of liver fibrosis using transient elastography in patients with alcoholic liver disease. *J Hepatol* 2008;**49**:1062–8. http://dx.doi.org/10.1016/j.jhep.2008.08.011

279. Naveau S, Poynard T, Benattar C, Bedossa P, Chaput JC. Alpha-2-macroglobulin and hepatic fibrosis. Diagnostic interest. *Dig Dis Sci* 1994;**39**:2426–32. http://dx.doi.org/10.1007/BF02087661

280. Naveau S, Raynard B, Ratziu V, Abella A, Imbert-Bismut F, Messous D, *et al.* Biomarkers for the prediction of liver fibrosis in patients with chronic alcoholic liver disease. *Clin Gastroenterol Hepatol* 2005;**3**:167–74. http://dx.doi.org/10.1016/S1542-3565(04)00625-1

281. Nguyen-Khac E, Chatelain D, Tramier B, Decrombecque C, Robert B, Joly JP, *et al.* Assessment of asymptomatic liver fibrosis in alcoholic patients using fibroscan: prospective comparison with seven non-invasive laboratory tests. *Aliment Pharmacol Ther* 2008;**28**:1188–98. http://dx.doi.org/10.1111/j.1365-2036.2008.03831.x

282. Tran A, Benzaken S, Saint-Paul MC, Guzman-Granier E, Hastier P, Pradier C, *et al.* Chondrex (YKL-40), a potential new serum fibrosis marker in patients with alcoholic liver disease. *Eur J Gastroenterol Hepatol* 2000;**12**:989–93. http://dx.doi.org/10.1097/00042737-200012090-00004

283. Vanbiervliet G, Marine-Barjoan E, Gelsi E, Anty R, Piche T, Benzaken S, *et al.* Evaluation of liver fibrosis with APRI score in patients with chronic alcoholic liver disease. *J Hepatol* 2005;**42**:255. http://dx.doi.org/10.1016/S0168-8278(05)82109-9

284. Adams LA, George J, Bugianesi E, Rossi E, De Boer WB, van der Poorten D, *et al.* Complex non-invasive fibrosis models are more accurate than simple models in non-alcoholic fatty liver disease. *J Gastroenterol Hepatol* 2011;**26**:1536–43. http://dx.doi.org/10.1111/j.1440-1746.2011.06774.x

285. Angulo P, Hui JM, Marchesini G, Bugianesi E, George J, Farrell GC, *et al.* The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology* 2007;**45**:846–54. http://dx.doi.org/10.1002/hep.21496

286. Blomme B, Francque S, Trepo E, Libbrecht L, Vanderschaeghe D, Verrijken A, *et al.* N-glycan based biomarker distinguishing non-alcoholic steatohepatitis from steatosis independently of fibrosis. *Dig Liver Dis* 2012;**44**:315–22. http://dx.doi.org/10.1016/j.dld.2011.10.015

287. Cales P, Laine F, Boursier J, Deugnier Y, Moal V, Oberti F, *et al.* Comparison of blood tests for liver fibrosis specific or not to NAFLD. *J Hepatol* 2009;**50**:165–73. http://dx.doi.org/10.1016/j.jhep.2008.07.035

288. De Ledinghen V, Wong VW, Vergniol J, Wong G, Bail BL, Chim AML, *et al.* Prediction of fibrosis in patients with NAFLD using transient elastography: a prospective multicentre study. *Hepatology* 2009;**50**:767A–8A.

289. Guajardo-Salinas GE, Hilmy A. Prevalence of nonalcoholic fatty liver disease (NAFLD) and utility of FIBROspect II to detect liver fibrosis in morbidly obese hispano-american patients undergoing gastric bypass. *Obes Surg* 2010;**20**:1647–53. http://dx.doi.org/10.1007/s11695-009-0027-0

290. Guha IN, Parkes J, Roderick P, Chattopadhyay D, Cross R, Harris S, *et al.* Noninvasive markers of fibrosis in nonalcoholic fatty liver disease: validating the European liver fibrosis panel and exploring simple markers. *Hepatology* 2008;**47**:455–60. http://dx.doi.org/10.1002/hep.21984

291. Harrison SA, Oliver D, Arnold HL, Gogia S, Neuschwander-Tetri BA. Development and validation of a simple NAFLD clinical scoring system for identifying patients without advanced disease. *Gut* 2008;**57**:1441–7. http://dx.doi.org/10.1136/gut.2007.146019

292. Kaneda H, Hashimoto E, Yatsuji S, Tokushige K, Shiratori K. Hyaluronic acid levels can predict severe fibrosis and platelet counts can predict cirrhosis in patients with nonalcoholic fatty liver disease. *J Gastroenterol Hepatol* 2006;**21**:1459–65.

293. Kayadibi H, Gultepe M, Yasar B, Ince AT, Ozcan O, Ipcioglu OM, *et al.* Diagnostic value of serum prolidase enzyme activity to predict the liver histological lesions in non-alcoholic fatty liver disease: a surrogate marker to distinguish steatohepatitis from simple steatosis. *Dig Dis Sci* 2009;**54**:1764–71. http://dx.doi.org/10.1007/s10620-008-0535-0

294. Kelleher TB, MacFarlane C, de Ledinghen V, Beaugrand M, Foucher J, Castera L, *et al.* Risk factors and hepatic elastography (fibroscan) in the prediction of hepatic fibrosis in non alcoholic steatohepatitis. *Gastroenterology* 2006;**130**:A768.

295. Khosravi S, Alavian SM, Zare A, Daryani NE, Fereshtehnejad SM, Daryani NE, *et al.* Non-alcoholic fatty liver disease and correlation of serum alanin aminotransferase level with histopathologic findings. *Hepat Mon* 2011;**11**:452–8.

296. Lupsor M, Badea R, Stefanescu H, Grigorescu M, Serban A, Radu C, *et al.* Performance of unidimensional transient elastography in staging non-alcoholic steatohepatitis. *J Gastrointestin Liver Dis* 2010;**19**:53–60.

297. Lydatakis H, Hager IP, Kostadelou E, Mpousmpoulas S, Pappas S, Diamantis I. Non-invasive markers to predict the liver fibrosis in non-alcoholic fatty liver disease. *Liver Int* 2006;**26**:864–71. http://dx.doi.org/10.1111/j.1478-3231.2006.01312.x

298. Mahadeva S, Mahfudz A, Vijayanathan A, Goh KL, Arumugam K, Cheah PL. Accuracy of liver stiffness measurement in an Asian population with non-alcoholic fatty liver diseasea preliminary report. *J Gastroenterol Hepatol* 2010;**25**:A102.

299. Manousou P, Kalambokis G, Grillo F, Watkins J, Xirouchakis E, Pleguezuelo M, *et al.* Serum ferritin is a discriminant marker for both fibrosis and inflammation in histologically proven non-alcoholic fatty liver disease patients. *Liver Int* 2011;**31**:730–9. http://dx.doi.org/10.1111/j.1478-3231.2011.02488.x

300. McPherson S, Anstee QM, Henderson E, Burt AD, Day CP. Are simple non-invasive scoring systems for fibrosis reliable in patients with nafld and normal ALT levels? *Hepatology* 2011;**54**:1125A.

301. McPherson S, Stewart SF, Henderson E, Burt AD, Day CP. Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease. *Gut* 2010;**59**:1265–9. http://dx.doi.org/10.1136/gut.2010.216077

302. Orlacchio A, Bolacchi F, Antonicoli M, Coco I, Costanzo E, Tosti D, *et al.* Liver elasticity in NASH patients evaluated with real-time elastography (RTE). *Ultrasound Med Biol* 2012;**38**:537–44. http://dx.doi.org/10.1016/j.ultrasmedbio.2011.12.023

303. Pais R, Lebray P, Fedchuk L, Charlotte F, Poynard T, Ratziu V. Validation of a simple algorithm combining serum markers and elastometry for the diagnosis of fibrosis in NAFLD. *Hepatology* 2011;**54**:1127A–8A.

304. Palmeri ML, Wang MH, Rouze NC, Abdelmalek MF, Guy CD, Moser B, *et al.* Noninvasive evaluation of hepatic fibrosis using acoustic radiation force-based shear stiffness in patients with nonalcoholic fatty liver disease. *J Hepatol* 2011;**55**:666–72. http://dx.doi.org/10.1016/j.jhep.2010.12.019

305. Papalavrentios L, Sinakos E, Hytiroglou P, Drevelegas K, Drevelegas A, Akriviadis E. 3 tesla diffusion-weighted MRI for assessing liver fibrosis in non alcoholic fatty liver disease. *Hepatology* 2011;**54**:1120A–1A.

306. Park GJH, Wiseman E, George J, Katelaris PH, Seow F, Fung C, *et al.* Non-invasive estimation of liver fibrosis in non-alcoholic fatty liver disease using the 13C-caffeine breath test. *J Gastroenterol Hepatol* 2011;**26**:1411–16. http://dx.doi.org/10.1111/j.1440-1746.2011.06760.x

307. Pawitpok C, Kongtawelert P, Ua-Arayaporn S, Punyarit P, Chutaputti A. Efficacy of serum hyaluronic acid level in predicting liver fibrosis in patients with nonalcoholic fatty liver disease. *J Gastroenterol Hepatol* 2006;**21**:A34.

308. Petta S, Di Marco V, Camma C, Butera G, Cabibi D, Craxi A. Reliability of liver stiffness measurement in non-alcoholic fatty liver disease: the effects of body mass index. *Aliment Pharmacol Ther* 2011;**33**:1350–60. http://dx.doi.org/10.1111/j.1365-2036.2011.04668.x

309. Pimentel SK, Strobel R, Goncalves CG, Sakamoto DG, Ivano FH, Coelho JCU. Evaluation of the nonalcoholic fat liver disease fibrosis score for patients undergoing bariatric surgery. *Arg Gastroenterol* 2010;**47**:170–3. http://dx.doi.org/10.1590/S0004-28032010000200010

310. Poynard T, Ratziu V, Charlotte F, Messous D, Munteanu M, Imbert-Bismut F, *et al.* Diagnostic value of biochemical markers (NashTest) for the prediction of non alcoholo steato hepatitis in patients with non-alcoholic fatty liver disease. *BMC Gastroenterol* 2006;**6**:34. http://dx.doi.org/10.1186/1471-230X-6-34

311. Qureshi K, Clements RH, Abrams GA. The utility of the 'NAFLD fibrosis score' in morbidly obese subjects with NAFLD. *Obes Surg* 2008;**18**:264–70. http://dx.doi.org/10.1007/s11695-007-9295-8

312. Raszeja-Wyszomirska J, Szymanik B, Lawniczak M, Kajor M, Chwist A, Milkiewicz P, *et al.* Validation of the BARD scoring system in Polish patients with nonalcoholic fatty liver disease (NAFLD). *BMC Gastroenterol* 2010;**10**:67. http://dx.doi.org/10.1186/1471-230X-10-67

313. Ratziu V, Le Calvez S, Messous D, Charlotte F, Bonhay L, Munteanu M, *et al.* Diagnostic value of biochemical markers (Fibrotest) for the prediction of liver fibrosis in patients with non-alcoholic fatty liver disease (NAFLD). *J Hepatol* 2004;**40**:175. http://dx.doi.org/10.1016/S0168-8278(04)90596-X

314. Ratziu V, Massard J, Charlotte F, Messous D, Imbert-Bismut F, Bonyhay L, *et al.* Diagnostic value of biochemical markers (FibroTest-FibroSURE) for the prediction of liver fibrosis in patients with non-alcoholic fatty liver disease. *BMC Gastroenterol* 2006;**6**:6. http://dx.doi.org/10.1186/1471-230X-6-6

315. Ruffillo G, Fassio E, Alvarez E, Landeira G, Longo C, Dominguez N, *et al.* Comparison of NAFLD fibrosis score and BARD score in predicting fibrosis in nonalcoholic fatty liver disease. *J Hepatol* 2011;**54**:160–3. http://dx.doi.org/10.1016/j.jhep.2010.06.028

316. Sakugawa H, Kobashigawa K, Yamashiro T, Maeshiro T, Miyagi S, Shiroma J, *et al.* Clinical usefulness of biochemical markers of liver fibrosis in patients with nonalcoholic fatty liver disease. *World J Gastroenterol* 2005;**11**:255–9. http://dx.doi.org/10.3748/wjg.v11.i2.255

317. Santos VND, Leite-Mor MMB, Kondo M, Martins JR, Nader H, Lanzoni VP, *et al.* Serum laminin, type IV collagen and hyaluronan as fibrosis markers in non-alcoholic fatty liver disease. *Braz J Med Biol Res* 2005;**38**:747–53. http://dx.doi.org/10.1590/S0100-879X2005000500012

318. Shimada M, Kawahara H, Ozaki K, Fukura M, Yano H, Tsuchishima M, *et al.* Usefulness of a combined evaluation of the serum adiponectin level, HOMA-IR, and serum type IV collagen 7S level to predict the early stage of nonalcoholic steatohepatitis. *Am J Gastroenterol* 2007;**102**:1931–8. http://dx.doi.org/10.1111/j.1572-0241.2007.01322.x

319. Sumida Y, Yoneda M, Hyogo H, Itoh Y, Ono M, Fujii H, *et al.* Validation of the FIB4 index in a Japanese nonalcoholic fatty liver disease population. *BMC Gastroenterol* 2012;**12**:2. http://dx.doi.org/10.1186/1471-230X-12-2

320. Sumida Y, Yoneda M, Hyogo H, Yamaguchi K, Ono M, Fujii H, *et al.* A simple clinical scoring system using ferritin, fasting insulin, and type IV collagen 7S for predicting steatohepatitis in nonalcoholic fatty liver disease. *J Gastroenterol* 2011;**46**:257–68. http://dx.doi.org/10.1007/s00535-010-0305-6

321. Suzuki A, Angulo P, Lymp J, Li D, Satomura S, Lindor K. Hyaluronic acid, an accurate serum marker for severe hepatic fibrosis in patients with non-alcoholic fatty liver disease. *Liver Int* 2005;**25**:779–86. http://dx.doi.org/10.1111/j.1478-3231.2005.01064.x

322. Wong VW, Wong GL, Chim AM, Tse AM, Tsang SW, Hui AY, *et al.* Validation of the NAFLD fibrosis score in a Chinese population with low prevalence of advanced fibrosis. *Am J Gastroenterol* 2008;**103**:1682–8. http://dx.doi.org/10.1111/j.1572-0241.2008.01933.x

323. Wong VW, Wong GL, Chim AM, Yiu K, Chan AW, Choi PC, *et al.* Combination of the median and variability of liver stiffness measurements by transient elastography improves the accuracy of diagnosing advanced fibrosis in nonalcoholic fatty liver disease. *J Hepatol* 2008;**48**:S367.

324. Wong VWS, Vergniol J, Chan HLY, de Ledinghen V. Diagnostic power of Fibroscan in predicting liver fibrosis in nonalcoholic fatty liver disease reply. *Hepatology* 2009;**50**:2049–50. http://dx.doi.org/10.1002/hep.23364

325. Yoneda M, Fujii H, Sumida Y, Hyogo H, Itoh Y, Ono M, *et al.* Platelet count for predicting fibrosis in nonalcoholic fatty liver disease. *J Gastroenterol* 2011;**46**:1300–6. http://dx.doi.org/10.1007/s00535-011-0436-4

326. Yoneda M, Yoneda M, Mawatari H, Fujita K, Endo H, Iida H, *et al.* Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with nonalcoholic fatty liver disease (NAFLD). *Dig Liver Dis* 2008;**40**:371–8. http://dx.doi.org/10.1016/j.dld.2007.10.019

327. Younossi ZM, Page S, Rafiq N, Birerdinc A, Stepanova M, Hossain N, *et al.* A biomarker panel for non-alcoholic steatohepatitis (NASH) and NASH-related fibrosis. *Obes Surg* 2011;**21**:431–9. http://dx.doi.org/10.1007/s11695-010-0204-1

328. Asbach P, Klatt D, Schlosser B, Biermer M, Muche M, Rieger A, *et al.* Viscoelasticity-based staging of hepatic fibrosis with multifrequency MR elastography. *Radiology* 2010;**257**:80–6. http://dx.doi.org/10.1148/radiol.10092489

329. Aube C, Oberti F, Korali N, Namour MA, Loisel D, Tanguy JY, *et al.* Ultrasonographic diagnosis of hepatic fibrosis or cirrhosis. *J Hepatol* 1999;**30**:472–8. http://dx.doi.org/10.1016/S0168-8278(99)80107-X

330. Aube C, Winkfield B, Oberti F, Vuillemin E, Rousselet MC, Caron C, *et al.* New Doppler ultrasound signs improve the non-invasive diagnosis of cirrhosis or severe liver fibrosis. *Eur J Gastroenterol Hepatol* 2004;**16**:743–51. http://dx.doi.org/10.1097/01.meg.0000108357.41221.e5

331. Awaya H, Mitchell DG, Kamishima T, Holland G, Ito K, Matsumoto T. Cirrhosis: modified caudate-right lobe ratio. *Radiology* 2002;**224**:769–74. http://dx.doi.org/10.1148/radiol.2243011495

332. Cardi M, Muttillo IA, Amadori L, Petroni R, Mingazzini P, Barillari P, *et al.* Superiority of laparoscopy compared to ultrasonography in diagnosis of widespread liver diseases. *Dig Dis Sci* 1997;**42**:546–8. http://dx.doi.org/10.1023/A:1018895009305

333. Cioni G, Dalimonte P, Zerbinati F, Ventura P, Cristani A, Vignoli A, *et al.* Duplex-Doppler ultrasonography in the evaluation of cirrhotic patients with portal-hypertension and in the analysis of their response to drugs. *J Gastroenterol Hepatol* 1992;**7**:388–92. http://dx.doi.org/10.1111/j.1440-1746.1992.tb01005.x

334. Colli A, Cocciolo M, Riva C, Martinez E, Prisco A, Pirola M, *et al.* Abnormalities of Doppler waveform of the hepatic veins in patients with chronic liver disease: correlation with histologic findings. *AJR Am J Roentgenol* 1994;**162**:833–7. http://dx.doi.org/10.2214/ajr.162.4.8141001

335. Colli A, Fraquelli M, Andreoletti M, Marino B, Zuccoli E, Conte D. Severe liver fibrosis or cirrhosis: accuracy of US for detection–analysis of 300 cases. *Radiology* 2003;**227**:89–94. http://dx.doi.org/10.1148/radiol.2272020193

336. D'Onofrio M, Martone E, Brunelli S, Faccioli N, Zamboni G, Zagni I, *et al.* Accuracy of ultrasound in the detection of liver fibrosis in chronic viral hepatitis. *Radiol Med* 2005;**110**:341–8.

337. Do RKG, Chandanara H, Felker E, Hajdu CH, Babb JS, Kim D, *et al.* Diagnosis of liver fibrosis and cirrhosis with diffusion-weighted imaging: value of normalized apparent diffusion coefficient using the spleen as reference organ. *AJR Am J Roentgenol* 2010;**195**:671–6. http://dx.doi.org/10.2214/AJR.09.3448

338. Ferral H, Male R, Cardiel M, Munoz L, Ferrari FQY. Cirrhosis – diagnosis by liver surface-analysis with high-frequency ultrasound. *Gastroint Radiol* 1992;**17**:74–8. http://dx.doi.org/10.1007/BF01888512

339. Gaia S, Carucci P, Spandre M, Cosso L, Evangelista A, Bugianesi E, *et al.* Ultrasound evaluation, liver stiffness measurement and biopsy for staging of hepatic fibrosis in patients with chronic liver disease: a comparative study. *Hepatology* 2011;**54**:1232A.

340. Gaiani S, Gramantieri L, Venturoli N, Piscaglia F, Siringo S, D'Errico A, *et al.* What is the criterion for differentiating chronic hepatitis from compensated cirrhosis? A prospective study comparing ultrasonography and percutaneous liver biopsy. *J Hepatol* 1997;**27**:979–85. http://dx.doi.org/10.1016/S0168-8278(97)80140-7

341. Gierblinski I, Przelaskowski A, Wocial T, Kazubek M, Zych W, Walewska-Zielecka B. Measurement of liver stiffness by color-coded ultrasound elastography: a preliminary clinical feasibility study. *Gastroenterologia Polska* 2008;**15**:151–5.

342. Goyal AK, Pokharna DS, Sharma SK. Ultrasonic diagnosis of cirrhosis: reference to quantitative measurements of hepatic dimensions. *Gastrointest Radiol* 1990;**15**:32–4. http://dx.doi.org/10.1007/BF01888729

343. Ibrahim HR, El-Hamid AA, Tohamy A, Habba MR. Diagnostic value of apparent diffusion coefficient calculated with diffusion-weighted MRI for quantification of liver fibrosis. *Egypt J Radiol Nucl Med* 2011;**42**:119–31. http://dx.doi.org/10.1016/j.ejrnm.2011.05.003

344. Ishibashi H, Maruyama H, Takahashi M, Shimada T, Kamesaki H, Fujiwara K, *et al.* Demonstration of intrahepatic accumulated microbubble on ultrasound represents the grade of hepatic fibrosis. *Eur Radiol* 2012;**22**:1083–90. http://dx.doi.org/10.1007/s00330-011-2346-5

345. Joseph AE, Saverymuttu SH, al-Sam S, Cook MG, Maxwell JD. Comparison of liver histology with ultrasonography in assessing diffuse parenchymal liver disease. *Clin Radiol* 1991;**43**:26–31. http://dx.doi.org/10.1016/S0009-9260(05)80350-2

346. Kim BH, Lee JM, Lee YJ, Lee KB, Suh KS, Han JK, *et al.* MR elastography for noninvasive assessment of hepatic fibrosis: experience from a tertiary center in Asia. *J Magn Reson Imaging* 2011;**34**:1110–16. http://dx.doi.org/10.1002/jmri.22723

347. Kim BK, Han K-H, Park JY, Ahn SH, Chon CY, Kim JK, *et al.* A novel liver stiffness measurement-based prediction model for cirrhosis in hepatitis B patients. *Liver Int* 2010;**30**:1073–81. http://dx.doi.org/10.1111/j.1478-3231.2010.02269.x

348. Ladenheim JA, Luba DG, Yao F, Gregory PB, Jeffrey RB, Garcia G. Limitations of liver surface US in the diagnosis of cirrhosis. *Radiology* 1992;**185**:21–3; discussion 3–4. http://dx.doi.org/10.1148/radiology.185.1.1523310

349. Lee HS, Kim JK, Cheong JY, Han EJ, An SY, Song JH, *et al.* Prediction of compensated liver cirrhosis by ultrasonography and routine blood tests in patients with chronic viral hepatitis. *Korean J Hepatol* 2010;**16**:369–75. http://dx.doi.org/10.3350/kjhep.2010.16.4.369

350. Liu CH, Lin JW, Tsai FC, Yang PM, Lai MY, Chen JH, *et al.* Noninvasive tests for the prediction of significant hepatic fibrosis in hepatitis C virus carriers with persistently normal alanine aminotransferases. *Liver Int* 2006;**26**:1087–94. http://dx.doi.org/10.1111/j.1478-3231.2006.01355.x

351. Lutz HH, Gassler N, Tischendorf FW, Trautwein C, Tischendorf JJ. Doppler ultrasound of hepatic blood flow for noninvasive evaluation of liver fibrosis compared with liver biopsy and transient elastography. *Dig Dis Sci* 2012;**57**:2222–30. http://dx.doi.org/10.1007/s10620-012-2153-0

352. Nagata N, Miyachi H, Nakano A, Nanri K, Kobayashi H, Matsuzaki S. Sonographic evaluation of the anterior liver surface in chronic liver diseases using a 7.5-MHz annular-array transducer: correlation with laparoscopic and histopathologic findings. *J Clin Ultrasound* 2003;**31**:393–400. http://dx.doi.org/10.1002/jcu.10195

353. Nishiura T, Watanabe H, Ito M, Matsuoka Y, Yano K, Daikoku M, *et al.* Ultrasound evaluation of the fibrosis stage in chronic liver disease by the simultaneous use of low and high frequency probes. *Br J Radiol* 2005;**78**:189–97. http://dx.doi.org/10.1259/bjr/75208448

354. Numminen K, Tervahartiala P, Halavaara J, Isoniemi H, Hockerstedt K. Non-invasive diagnosis of liver cirrhosis: magnetic resonance imaging presents special features. *Scand J Gastroenterol* 2005;**40**:76–82. http://dx.doi.org/10.1080/00365520410009384

355. Ong TZ, Tan HJ. Ultrasonography is not reliable in diagnosing liver cirrhosis in clinical practice. *Singapore Med J* 2003;**44**:293–5.

356. Rustogi R, Ganger D, Horowitz J, Harmath C, Wang Y, Chen ZE, *et al.* Accuracy of MR Elastography (MRE) vs. imaging features in the diagnosis of severe fibrosis and cirrhosis. *Hepatology* 2011;**54**:908A.

357. Sandrasegaran K, Akisik FM, Lin C, Tahir B, Rajan J, Saxena R, *et al.* Value of diffusion-weighted MRI for assessing liver fibrosis and cirrhosis. *AJR Am J Roentgenol* 2009;**193**:1556–60. http://dx.doi.org/10.2214/AJR.09.2436

358. Shen L, Li JQ, Zeng MD, Lu LG, Fan ST, Bao H. Correlation between ultrasonographic and pathologic diagnosis of liver fibrosis due to chronic virus hepatitis. *World J Gastroenterol* 2006;**12**:1292–5.

359. Venkatesh S, Teo L, Ang B, Ehman R. Detection of liver fibrosis: comparison of magnetic resonance elastography and diffusion-weighted MRI. *AJR Am J Roentgenol* 2010;**194**:83.

360. Viganò M, Visentin S, Aghemo A, Rumi MG, Ronchi G. US features of liver surface nodularity as a predictor of severe fibrosis in chronic hepatitis C. *Radiology* 2005;**234**:641; author reply 641. http://dx.doi.org/10.1148/radiol.2342041267

361. Wang J-H, Changchien C-S, Hung C-H, Eng H-L, Tung W-C, Kee K-M, *et al.* FibroScan and ultrasonography in the prediction of hepatic fibrosis in patients with chronic viral hepatitis. *J Gastroenterol* 2009;**44**:439–46. http://dx.doi.org/10.1007/s00535-009-0017-y

362. Xu Y, Wang B, Cao H. An ultrasound scoring system for the diagnosis of liver fibrosis and cirrhosis. *Chin Med J* 1999;**112**:1125–8.

363. Zhu NY, Chen KM, Chai WM, Li WX, Du LJ. Feasibility of diagnosing and staging liver fibrosis with diffusion weighted imaging. *Chin Med Sci J* 2008;**23**:183–6. http://dx.doi.org/10.1016/S1001-9294(09)60036-5

364. Carlson JJ, Kowdley KV, Sullivan SD, Ramsey SD, Veenstra DL. An evaluation of the potential cost-effectiveness of non-invasive testing strategies in the diagnosis of significant liver fibrosis. *J Gastroenterol Hepatol* 2009;**24**:786–91. http://dx.doi.org/10.1111/j.1440-1746.2009.05778.x

365. Hall A, Germani G, Isgro G, Burroughs AK, Dhillon AP. Fibrosis distribution in explanted cirrhotic livers. *Histopathology* 2012;**60**:270–7. http://dx.doi.org/10.1111/j.1365-2559.2011.04094.x

366. Myers RP, Ratziu V, Imbert-Bismut F, Charlotte F, Poynard T, C MGGdEMslPLaV. Biochemical markers of liver fibrosis: a comparison with historical features in patients with chronic hepatitis C. *Am J Gastroenterol* 2002;**97**:2419–25. http://dx.doi.org/10.1111/j.1572-0241.2002.05997.x

367. Friedrich-Rust M, Muller C, Winckler A, Kriener S, Herrmann E, Holtmeier J, *et al.* Assessment of liver fibrosis and steatosis in PBC with fibroscan, MRI, MR-spectroscopy, and serum markers. *J Clin Gastroenterol* 2010;**44**:58–65. http://dx.doi.org/10.1097/MCG.0b013e3181a84b8d

368. Aguirre DA, Behling CA, Alpert E, Hassanein TI, Sirlin CB. Liver fibrosis: noninvasive diagnosis with double contrast material-enhanced MR imaging. *Radiology* 2006;**239**:425–37. http://dx.doi.org/10.1148/radiol.2392050505

369. Motosugi U, Ichikawa T, Oguri M, Sano K, Sou H, Muhi A, *et al.* Staging liver fibrosis by using liver–enhancement ratio of gadoxetic acid-enhanced MR imaging: comparison with aspartate aminotransferase-to-platelet ratio index. *Magn Reson Imaging* 2011;**29**:1047–52. http://dx.doi.org/10.1016/j.mri.2011.05.007

370. Shiramizu B, Theodore T, Bassett R, Coel M, Sherman KE, Glesby IJ, *et al.* Correlation of single photon emission computed tomography parameters as a noninvasive alternative to liver biopsies in assessing liver involvement in the setting of HIV and hepatitis C virus coinfection: a multicenter trial of the adult AIDS clinical trials group. *J Acquir Immune Defic Syndr* 2003;**33**:329–35. http://dx.doi.org/10.1097/00126334-200307010-00006

371. Shepherd J, Jones J, Takeda A, Davidson P, Price A. Adefovir dipivoxil and pegylated interferon alfa-2a for the treatment of chronic hepatitis B: a systematic review and economic evaluation. *Health Technol Assess* 2006;**10**(28). http://dx.doi.org/10.3310/hta10280

372. Dakin H, Bentley A, Dusheiko G. Cost-utility analysis of tenofovir disoproxil fumarate in the treatment of chronic hepatitis B. *Value Health* 2010;**13**:922–33. http://dx.doi.org/10.1111/j.1524-4733.2010.00782.x

373. Marcellin P, Asselah T. Long-term therapy for chronic hepatitis B: hepatitis B virus DNA suppression leading to cirrhosis reversal. *J Gastroenterol Hepatol* 2013;**28**:912–23. http://dx.doi.org/10.1111/jgh.12213

374. National Institute for Health and Care Excellence. *Diagnosis and Managment of Chronic Hepatitis B in Children, Young People and Adults*. NICE clinical guideline CG165. London: NICE; 2013.

375. National Institute for Health and Care Excellence. *Adefovir Dipivoxil and Peginteferon Alfa-2a for the Treatment of Chronic Hepatitis B*. NICE technology appraisal guidance 96. London: NICE; 2006.

376. Fattovich G. Natural history and prognosis of hepatitis B. *Sem Liv Dis* 2003;**23**:47–58. http://dx.doi.org/10.1055/s-2003-37590

377. Realdi G, Alberti A, Rugge M, Bortolotti F, Rigoli AM, Tremolada F, *et al.* Seroconversion from hepatitis B e antigen to anti-HBe in chronic hepatitis B virus infection. *Gastroenterology* 1980;**79**:195–9.

378. Hoofnagle JH, Dusheiko GM, Seeff LB, Jones EA, Waggoner JG, Bales ZB. Seroconversion from hepatitis B e antigen to antibody in chronic type B hepatitis. *Ann Intern Med* 1981;**94**:744–8. http://dx.doi.org/10.7326/0003-4819-94-6-744

379. Fattovich G, Rugge M, Brollo L, Pontisso P, Noventa F, Guido M, *et al.* Clinical, virologic and histologic outcome following seroconversion from HBeAg to anti-HBe in chronic hepatitis type B. *Hepatology* 1986;**6**:167–72. http://dx.doi.org/10.1002/hep.1840060203

380. Moreno-Otero R, Garcia-Monzon C, Garcia-Sanchez A, Buey LG, Pajares JM, Di Bisceglie AM. Development of cirrhosis after chronic type B hepatitis: a clinicopathologic and follow-up study of 46 HBeAg-positive asymptomatic patients. *Am J Gastroenterol* 1991;**86**:560–4.

381. Zarski JP, Marcellin P, Cohard M, Lutz JM, Bouche C, Rais A. Comparison of anti-HBe-positive and HBe-antigen-positive chronic hepatitis B in France. *J Hepatol* 1994;**20**:636–40. http://dx.doi.org/10.1016/S0168-8278(05)80352-6

382. Di Marco V, Lo Iacono O, Cammà C, Vaccaro A, Giunta M, Martorana G, *et al.* The long-term course of chronic hepatitis B. *Hepatology* 1999;**30**:257–64. http://dx.doi.org/10.1002/hep.510300109

383. Lin SM, Sheen IS, Chien RN, Chu CM, Liaw YF. Long-term beneficial effect of interferon therapy in patients with chronic hepatitis B virus infection. *Hepatology* 1999;**29**:971–5. http://dx.doi.org/10.1002/hep.510290312

384. Yuen MF, Hui CK, Cheng CC, Wu CH, Lai YP, Lai CL. Long-term follow-up of interferon alfa treatment in Chinese patients with chronic hepatitis B infection: the effect on hepatitis B e antigen seroconversion and the development of cirrhosis-related complications. *Hepatology* 2001;**34**:139–45. http://dx.doi.org/10.1053/jhep.2001.25273

385. Chang MH, Hsu HY, Hsu HC, Ni YH, Chen JS, Chen DS. The significance of spontaneous hepatitis B e antigen seroconversion in childhood: with special emphasis on the clearance of hepatitis B e antigen before 3 years of age. *Hepatology* 1995;**22**:1387–92.

386. Sancheztapias JM, Costa J, Mas A, Pares A, Bruguera M, Rodes J. Analysis of factors predicting early seroconversion to anti-Hbe in Hbeag-positive chronic hepatitis B. *J Hepatol* 1988;**6**:15–22. http://dx.doi.org/10.1016/S0168-8278(88)80458-6

387. Liaw YF, Tai DI, Chu CM, Chen TJ. The development of cirrhosis in patients with chronic type-B hepatitis – a prospective study. *Hepatology* 1988;**8**:493–6. http://dx.doi.org/10.1002/hep.1840080310

388. Fattovich G, Farci P, Rugge M, Brollo L, Mandas A, Pontisso P, *et al.* A randomized controlled trial of lymphoblastoid interferon-α in patients with chronic hepatitis B lacking HBeAg. *Hepatology* 1992;**15**:584–9. http://dx.doi.org/10.1002/hep.1840150405

389. Lampertico P, Del Ninno E, Manzin A, Donato MF, Rumi MG, Lunghi G, *et al.* A randomized, controlled trial of a 24-month course of interferon alfa 2b in patients with chronic hepatitis B who had hepatitis B virus DNA without hepatitis B e antigen in serum. *Hepatology* 1997;**26**:1621–5. http://dx.doi.org/10.1002/hep.510260634

390. Tassopoulos NC, Volpes R, Pasture G, Heathcote J, Buti M, Goldin RD, *et al.* Efficacy of lamivudine in patients with hepatitis B e antigen- negative/hepatitis B virus DNA-positive (precore mutant) chronic hepatitis B. *Hepatology* 1999;**29**:889–96. http://dx.doi.org/10.1002/hep.510290321

391. Brunetto MR, Oliveri F, Coco B, Leandro G, Colombatto P, Gorin JM, *et al.* Outcome of anti-HBe positive chronic hepatitis B in alpha-interferon treated and untreated patients: a long term cohort study. *J Hepatol* 2002;**36**:263–70. http://dx.doi.org/10.1016/S0168-8278(01)00266-5

392. Shepherd J, Brodin H, Cave C, Waugh N, Price A, Gabbay J. Pegylated interferon α–2a and -2b in combination with ribavirin in the treatment of chronic hepatitis C: a systematic review and economic evaluation. *Health Technol Assess* 2004;**8**(39). http://dx.doi.org/10.3310/hta8390

393. Wong JB, Koff RS, Tine F, Pauker SG. Cost-effectiveness of interferon-α2b treatment for hepatitis B e antigen-positive chronic hepatitis B. *Ann Intern Med* 1995;**122**:664–75. http://dx.doi.org/10.7326/0003-4819-122-9-199505010-00004

394. Di Bisceglie AM, Rustgi VK, Hoofnagle JH, Dusheiko GM, Lotze MT. Hepatocellular carcinoma. *Ann Intern Med* 1988;**108**:390–401. http://dx.doi.org/10.7326/0003-4819-108-3-390

395. Bennett WG, Inoue Y, Beck JR, Wong JB, Pauker SG, Davis GL. Estimates of the cost-effectiveness of a single course of interferon-α2b in patients with histologically mild chronic hepatitis C. *Ann Intern Med* 1997;**127**:855–65. http://dx.doi.org/10.7326/0003-4819-127-10-199711150-00001

396. Lavanchy D. Hepatitis B virus epidemiology, disease burden, treatment, arid current and emerging prevention and control measures. *J Viral Hepat* 2004;**11**:97–107. http://dx.doi.org/10.1046/j.1365-2893.2003.00487.x

397. Lau DTY, Everhart J, Kleiner DE, Park Y, Vergalla J, Schmid P, *et al.* Long-term follow-up of patients with chronic hepatitis B treated with interferon alfa. *Gastroenterology* 1997;**113**:1660–7. http://dx.doi.org/10.1053/gast.1997.v113.pm9352870

398. Crowley S, Tognarini D, Desmond P, Lees M, Saal G. Introduction of lamivudine for the treatment of chronic hepatitis B: expected clinical and economic outcomes based on 4-year clinical trial data. *J Gastroenterol Hepatol* 2002;**17**:153–64. http://dx.doi.org/10.1046/j.1440-1746.2002.02673.x

399. Crowley SJ, Tognarini D, Desmond PV, Lees M. Cost-effectiveness analysis of lamivudine for the treatment of chronic hepatitis B. *Pharmacoeconomics* 2000;**17**:409–27. http://dx.doi.org/10.2165/00019053-200017050-00001

400. de Franchis R, Hadengue A, Lau G, Lavanchy D, Lok A, McIntyre N. EASL International Consensus Conference on Hepatitis B. 13–14 September, 2002 Geneva, Switzerland. Consensus statement (long version). *J Hepatol* 2003;**39**(Suppl. 1):S3–25.

401. Lai CL, Dienstag J, Schiff E, Leung NW, Atkins M, Hunt C, *et al.* Prevalence and clinical correlates of YMDD variants during lamviudine therapy for patients with chronic hepatitis B. *Clin Infect Dis* 2003;**36**:10. http://dx.doi.org/10.1086/368083

402. Fattovich G, Brollo L, Giustina G, Noventa F, Pontisso P, Alberti A, *et al.* Natural history and prognostic factors for chronic hepatitis type B. *Gut* 1991;**32**:294–8. http://dx.doi.org/10.1136/gut.32.3.294

403. Fattovich G, Pantalena M, Zagni I, Realdi G, Schalm SW, Christensen E. Effect of hepatitis B and C virus infections on the natural history of compensated cirrhosis: a cohort study of 297 patients. *Am J Gastroenterol* 2002;**97**:2886–95. http://dx.doi.org/10.1111/j.1572-0241.2002.07057.x

404. Liaw YF, Lin DY, Chen TJ, Chu CM. Natural course after the development of cirrhosis in patients with chronic type B hepatitis: a prospective study. *Liver* 1989;**9**:235–41. http://dx.doi.org/10.1111/j.1600-0676.1989.tb00405.x

405. De Jongh FE, Janssen HLA, De Man RA, Hop WCJ, Schalm SW, Van Blankenstein M. Survival and prognostic indicators in hepatitis B surface antigen-positive cirrhosis of the liver. *Gastroenterology* 1992;**103**:1630–5.

406. Boring CC, Squires TS, Tong T. Cancer statistics, 1993. *CA Cancer J Clin* 1993;**43**:7–26. http://dx.doi.org/10.3322/canjclin.43.1.7

407. Pereira A. Health and economic consequences of HCV lookback. *Transfusion* 2001;**41**:832–9. http://dx.doi.org/10.1046/j.1537-2995.2001.41060832.x

408. Sonnenberg FA, Gregory P, Yomtovian R, Russell LB, Tierney W, Kosmin M, *et al.* The cost-effectiveness of autologous transfusion revisited: Implications of an increased risk of bacterial infection with allogeneic transfusion. *Transfusion* 1999;**39**:808–17. http://dx.doi.org/10.1046/j.1537-2995.1999.39080808.x

409. Fattovich G, Giustina G, Degos F, Diodati G, Tremolada F, Nevens F. Effectiveness of interferon alfa on incidence of hepatocellular carcinoma and decompensation in cirrhosis type C. *J Hepatol* 1997;**27**:5. http://dx.doi.org/10.1016/S0168-8278(97)80302-9

410. Veenstra D, Sullivan SD, Iloeje UH. Estimating future hepatitis B virus (HBV) disease burden in the United States using a disease simulation model. *Value Health* 2003;**6**:264. http://dx.doi.org/10.1016/S1098-3015(10)64012-0

411. Ascher NL, Lake JR, Emond J, Roberts J. Liver transplantation for hepatitis C virus-related cirrhosis. *Hepatology* 1994;**20**:24S–7S. http://dx.doi.org/10.1002/hep.1840200708

412. Kilpe VE, Krakauer H, Wren RE. An analysis of liver transplant experience from 37 transplant centers as reported to medicare. *Transplantation* 1993;**56**:554–61. http://dx.doi.org/10.1097/00007890-199309000-00012

413. Detre KM, Belie SH, Lombardero M. Liver transplantation for chronic viral hepatitis. *Viral Hepat Rev* 1996;**2**:10.

414. Wright M, Grieve R, Roberts J, Main J, Thomas HC, Alexander G, *et al.* Health benefits of antiviral therapy for mild chronic hepatitis C: randomized controlled trial and economic evaluation. *Health Technol Assess* 2006;**10**(21). http://dx.doi.org/10.3310/hta10210

415. Sweeting MJ, De Angelis D, Neal KR, Ramsay ME, Irving WL, Wright M, *et al.* Estimated progression rates in three United Kingdom hepatitis C cohorts differed according to method of recruitment. *J Clin Epidemiol* 2006;**59**:144–52. http://dx.doi.org/10.1016/j.jclinepi.2005.06.008

416. Office for National Statistics. *Mortality Statistics. Deaths Registered in 2010*. 2011. URL: www.ons.gov.uk/ons/taxonomy/index.html?nscl=Interim+Life+Tables (accessed 5 December 2013).

417. National Institute for Health and Care Excellence. *Entecavir for the Treatment of Chronic Hepatitis B*. NICE technology appraisal guidance 153. London: NICE; 2009.

418. National Institute for Health and Care Excellence. *Tenofovir Disoproxil for the Treatment of Chronic Hepatitis B*. London: NICE; 2009.

419. Joint Formulary Committee. *British National Formulary*. 64 ed. London: BMJ Group and Pharmaceutical Press; 2012.

420. Woo G, Tomlinson G, Nishikawa Y, Kowgier M, Sherman M, Wong DKH, *et al.* Tenofovir and entecavir are the most effective antiviral agents for chronic hepatitis B: a systematic review and Bayesian meta-analyses. *Gastroenterology* 2010;**139**:1218–29. http://dx.doi.org/10.1053/j.gastro.2010.06.042

421. Brown RE, De Cock E, Colin X, Antoñanzas F, Iloeje UH. Hepatitis B management costs in France, Italy, Spain, and the United Kingdom. *J Clin Gastroenterol* 2004;**38**(Suppl. 10):S169–74. http://dx.doi.org/10.1097/00004836-200411003-00009

422. Jones J, Colquitt J, Shephard J, Harris P, Cooker K. Tenofovir disoproxil fumarate for the treatment of chronic hepatitis B infection. *Health Technol Assess* 2010;**14**(Suppl. 1):23–9. http://dx.doi.org/10.3310/hta14Suppl1/04

423. Jones J, Shepherd J, Baxter L, Gospodarevskaya E, Hartwell D, Harris P, *et al.* Adefovir dipivoxil and pegylated interferon alpha for the treatment of chronic hepatitis B: an updated systematic review and economic evaluation. *Health Technol Assess* 2009;**13**(35). http://dx.doi.org/10.3310/hta13350

424. Longworth L, Young T, Buxton MJ, Ratcliffe J, Neuberger J, Burroughs A, *et al.* Midterm cost-effectiveness of the liver transplantation program of England and Wales for three disease groups. *Liver Transpl* 2003;**9**:1295–307. http://dx.doi.org/10.1016/j.lts.2003.09.012

425. Takeda A, Jones J, Shepherd J, Davidson P, Price A. A systematic review and economic evaluation of adefovir dipivoxil and pegylated interferon-alpha-2a for the treatment of chronic hepatitis B. *J Viral Hepat* 2007;**14**:75–88. http://dx.doi.org/10.1111/j.1365-2893.2006.00808.x

426. Veenstra DL, Sullivan SD, Dusheiko GM, Jacobs M, Aledort JE, Lewis G, *et al.* Cost-effectiveness of peginterferon α–2a compared with lamivudine treatment in patients with HBe-antigen-positive chronic hepatitis B in the United Kingdom. *Eur J Gastroenterol Hepatol* 2007;**19**:631–8. http://dx.doi.org/10.1097/MEG.0b013e3281108079

427. Department of Health. *NHS Reference Costs 2011–12*. London: Department of Health; 2012.

428. Stevenson M, Lloyd-Jones M, Morgan MY, Wong R. Non-invasive diagnostic assessment tools for the detection of liver fibrosis in patients with suspected alcohol-related liver disease: a systematic review and economic evaluation. *Health Technol Assess* 2012;**16**(4). http://dx.doi.org/10.3310/hta16040

429. Shepherd J, Gospodarevskaya E, Frampton G, Cooper K. Entecavir for the treatment of chronic hepatitis B infection. *Health Technol Assess* 2009;**13**(Suppl. 3):31–6. http://dx.doi.org/10.3310/hta13suppl3/05

430. Levy AR, Kowdley KV, Iloeje U, Tafesse E, Mukherjee J, Gish R, *et al.* The impact of chronic hepatitis B on quality of life: a multinational study of utilities from infected and uninfected persons. *Value Health* 2008;**11**:527–38. http://dx.doi.org/10.1111/j.1524-4733.2007.00297.x

431. Dusheiko GM, Roberts JA. Treatment of chronic type B and C hepatitis with interferon alfa: an economic appraisal. *Hepatology* 1995;**22**:1863–73.

432. McLernon DJ, Dillon J, Donnan PT. Health-state utilities in liver disease: a systematic review. *Med Decis Making* 2008;**28**:582–92. http://dx.doi.org/10.1177/0272989X08315240

433. Chong CAKY, Gulamhussein A, Jenny Heathcote E, Lilly L, Sherman M, Naglie G, *et al.* Health-state utilities and quality of life in hepatitis C patients. *Am J Gastroenterol* 2003;**98**:630–8. http://dx.doi.org/10.1111/j.1572-0241.2003.07332.x

434. Sherman KE, Sherman SN, Chenier T, Tsevat J. Health values of patients with chronic hepatitis C infection. *Arch Intern Med* 2004;**164**:2377–82. http://dx.doi.org/10.1001/archinte.164.21.2377

435. Siebert U, Sroczynski G, Rossol S, Wasem J, Ravens-Sieberer U, Kurth BM, *et al.* Cost effectiveness of peginterferon α–2b plus ribavirin versus interferon α–2b plus ribavirin for initial treatment of chronic hepatitis C. *Gut* 2003;**52**:425–32. http://dx.doi.org/10.1136/gut.52.3.425

436. Younossi ZM, Boparai N, McCormick M, Price LL, Guyatt G. Assessment of utilities and health-related quality of life in patients with chronic liver disease. *Am J Gastroenterol* 2001;**96**:579–83. http://dx.doi.org/10.1111/j.1572-0241.2001.03537.x

437. Briggs A, Claxton K, Sculpher M. *Decision Modelling for Health Economic Evaluation*. Oxford: Oxford University Press; 2006.

438. Liu S, Schwarzinger M, Carrat F, Goldhaber-Fiebert JD. Cost effectiveness of fibrosis assessment prior to treatment for chronic hepatitis C patients. *PLOS One* 2011;**6**:e26783. http://dx.doi.org/10.1371/journal.pone.0026783

439. Fried MW, Shiffman ML, Rajender Reddy K, Smith C, Marinos G, Gonçales FL Jr, *et al.* Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N Engl J Med* 2002;**347**:975–82. http://dx.doi.org/10.1056/NEJMoa020047

440. Royal College of Surgeons of England Clinical Effectiveness Unit. *UK & Ireland Liver Transplant Audit, 2003*. London: Royal College of Surgeons; 2003.

441. National Institute for Health and Care Excellence. *Telaprevir for the Treatment of Genotype 1 Chronic Hepatitis C*. Technology appraisal TA252. London: NICE; 2012.

442. National Institute for Health and Care Excellence. *Boceprevir for the Treatment of Genotype 1 Chronic Hepatitis C*. Technology appraisal TA253. London: NICE; 2012.

443. National Institute for Health and Care Excellence. *Peginterferon Alfa and Ribavirin for the Treatment of Mild Chronic Hepatitis C*. Technology appraisal TA106. London: NICE; 2007.

444. Datapharm. *Electronic Medicines Compendium (eMC)*. URL: www.medicines.org.uk/emc (accessxed 2 December 2012).

445. Manns MP, Wedemeyer H, Cornberg M. Treating viral hepatitis C: efficacy, side effects, and complications. *Gut* 2006;**55**:1350–9. http://dx.doi.org/10.1136/gut.2005.076646

446. Kamal SM, El Tawil AA, Nakano T, He Q, Rasenack J, Hakam SA, *et al.* Peginterferon α-2b and ribavirin therapy in chronic hepatitis C genotype 4: impact of treatment duration and viral kinetics on sustained virological response. *Gut* 2005;**54**:858–66. http://dx.doi.org/10.1136/gut.2004.057182

447. Camma C, Petta S, Enea M, Bruno R, Bronte F, Capursi V, *et al.* Cost-effectiveness of boceprevir or telaprevir for untreated patients with genotype 1 chronic hepatitis C. *Hepatology* 2012;**56**:850–60. http://dx.doi.org/10.1002/hep.25734

448. Esteban R, Buti M. Triple therapy with boceprevir or telaprevir for treatment naive HCV patients. *Best Pract Res Clin Gastroenterol* 2012;**26**:445–53. http://dx.doi.org/10.1016/j.bpg.2012.09.001

449. Jacobson IM, Marcellin P, Zeuzem S, Sulkowski MS, Esteban R, Poordad F, *et al.* Refinement of stopping rules during treatment of hepatitis C genotype 1 infection with boceprevir and peginterferon/ribavirin. *Hepatology* 2012;**56**:567–75. http://dx.doi.org/10.1002/hep.25865

450. Ramachandran P, Fraser A, Agarwal K, Austin A, Brown A, Foster GR, *et al.* UK consensus guidelines for the use of the protease inhibitors boceprevir and telaprevir in genotype 1 chronic hepatitis C infected patients. *Aliment Pharmacol Ther* 2012;**35**:647–62. http://dx.doi.org/10.1111/j.1365-2036.2012.04992.x

451. Shiffman ML, Esteban R. Triple therapy for HCV genotype 1 infection: telaprevir or boceprevir? *Liver Int* 2012;**32**(Suppl. 1):54–60. http://dx.doi.org/10.1111/j.1478-3231.2011.02718.x

452. Swiss Association for the Study of the Liver. Treatment of chronic hepatitis C genotype 1 with triple therapy comprising telaprevir or boceprevir. *Swiss Med Wkly* 2012;**142**:w13516.

453. Wilby KJ, Partovi N, Ford JA, Greanya E, Yoshida EM. Review of boceprevir and telaprevir for the treatment of chronic hepatitis C. *Can J Gastroenterol* 2012;**26**:205–10.

454. Fried MW, Hadziyannis SJ, Shiffman ML, Messinger D, Zeuzem S. Rapid virological response is the most important predictor of sustained virological response across genotypes in patients with chronic hepatitis C virus infection. *J Hepatol* 2011;**55**:69–75. http://dx.doi.org/10.1016/j.jhep.2010.10.032

455. Grishchenko M, Grieve RD, Sweeting MJ, De Angelis D, Thomson BJ, Ryder SD, *et al.* Cost-effectiveness of pegylated interferon and ribavirin for patients with chronic hepatitis C treated in routine clinical practice. *Int J Technol Assess Health Care* 2009;**25**:171–80. http://dx.doi.org/10.1017/S0266462309090229

456. Thompson Coon J, Rogers G, Hewson P, Wright D, Anderson R, Cramp M, *et al.* Surveillance of cirrhosis for hepatocellular carcinoma: systematic review and economic analysis. *Health Technol Assess* 2007;**11**(34). http://dx.doi.org/10.3310/hta11340

457. Grieve R, Roberts J, Wright M, Sweeting M, DeAngelis D, Rosenberg W, *et al.* Cost effectiveness of interferon alpha or peginterferon alpha with ribavirin for histologically mild chronic hepatitis C. *Gut* 2006;**55**:1332–8. http://dx.doi.org/10.1136/gut.2005.064774

458. John-Baptiste AA, Tomlinson G, Hsu PC, Krajden M, Heathcote EJ, Laporte A, *et al.* Sustained responders have better quality of life and productivity compared with treatment failures long after antiviral therapy for hepatitis C. *Am J Gastroenterol* 2009;**104**:2439–48. http://dx.doi.org/10.1038/ajg.2009.346

459. Marcellin P, Chousterman M, Fontanges T, Ouzan D, Rotily M, Varastet M, *et al.* Adherence to treatment and quality of life during hepatitis C therapy: a prospective, real-life, observational study. *Liver Int* 2011;**31**:516–24. http://dx.doi.org/10.1111/j.1478-3231.2011.02461.x

460. Ware JE Jr, Bayliss MS, Mannocchia M, Davis GL. Health-related quality of life in chronic hepatitis C: impact of disease and treatment response. *Hepatology* 1999;**30**:550–5. http://dx.doi.org/10.1002/hep.510300203

461. Bonkovsky HL, Wolley M. Reduction of health-related quality of life in chronic hepatitis C and improvement with interferon therapy. *Hepatology* 1999;**29**:264–70. http://dx.doi.org/10.1002/hep.510290124

462. Davis GL, Balart LA, Schiff ER, Lindsay K, Bodenheimer HC Jr, Perrillo RP, *et al.* Assessing health-related quality of life in chronic hepatitis C using the sickness impact profile. *Clin Ther* 1994;**16**:334–43.

463. Cheinquer N, Cheinquer H, Wolff FH, Coelho-Borges S. Effect of sustained virologic response on the incidence of hepatocellular carcinoma in patients with HCV cirrhosis. *Braz J Infect Dis* 2010;**14**:457–61. http://dx.doi.org/10.1016/S1413-8670(10)70093-3

464. Morgan TR, Ghany MG, Kim HY, Snow KK, Shiffman ML, De Santo JL, *et al.* Outcome of sustained virological responders with histologically advanced chronic hepatitis C. *Hepatology* 2010;**52**:833–44. http://dx.doi.org/10.1002/hep.23744

465. Manns MP, McHutchison JG, Gordon SC, Rustgi VK, Shiffman M, Reindollar R, *et al.* Peginterferon alfa-2b plus ribavirin compared with interferonalfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. *Lancet* 2001;**358**:958–65. http://dx.doi.org/10.1016/S0140-6736(01)06102-5

466. Tillmann HL. Hepatitis C infection and presence of advanced fibrosis: wait or treat? Why wait? There is no time to lose, is there? *J Hepatol* 2013;**58**:412–14. http://dx.doi.org/10.1016/j.jhep.2012.12.007

467. National Institute for Health and Care Excellence. *Alcohol Use Disorders: Physical Complications*. Clinical guideline CG100. London: NICE; 2010.

468. Rosenberg WMC, Voelker M, Thiel R, Becka M, Burt A, Schuppan D, *et al.* Serum markers detect the presence of liver fibrosis: a cohort study. *Gastroenterology* 2004;**127**:1704–13. http://dx.doi.org/10.1053/j.gastro.2004.08.052

469. Thompson Coon J, Rogers G, Hewson P, Wright D, Anderson R, Jackson S, *et al.* Surveillance of cirrhosis for hepatocellular carcinoma: a cost-utility analysis. *Br J Cancer* 2008;**98**:1166–75. http://dx.doi.org/10.1038/sj.bjc.6604301

470. Verrill C, Markham H, Templeton A, Carr NJ, Sheron N. Alcohol-related cirrhosis – early abstinence is a key factor in prognosis, even in the most severe cases. *Addiction* 2009;**104**:768–74. http://dx.doi.org/10.1111/j.1360-0443.2009.02521.x

471. National Health Service (NHS). *NHS Blood and Transplant – About Transplants*. URL: www.organdonation.nhs.uk/about_transplants/organ_allocation (cited 1 August 2013).

472. Ratziua V, Bellentanib S, Cortez-Pintoc H, Dayd C, Marchesinie G. A position statement on NAFLD/NASH based on the EASL 2009 special conference. *J Hepatol* 2010;**53**:372–84. http://dx.doi.org/10.1016/j.jhep.2010.04.008

473. Rakoski MO, Singal AG, Rogers MAM, Conjeevaram H. Meta-analysis: insulin sensitizers for the treatment of non-alcoholic steatohepatitis. *Aliment Pharmacol Ther* 2010;**32**:1211–21. http://dx.doi.org/10.1111/j.1365-2036.2010.04467.x

474. Musso G, Cassader M, Rosina F, Gambino R. Impact of current treatments on liver disease, glucose metabolism and cardiovascular risk in non-alcoholic fatty liver disease (NAFLD): a systematic review and meta-analysis of randomised trials. *Diabetologia* 2012;**55**:885–904. http://dx.doi.org/10.1007/s00125-011-2446-4

475. Mahady SE, Wong G, Craig JC, George J. Pioglitazone and vitamin E for nonalcoholic steatohepatitis: a cost utility analysis. *Hepatology* 2012;**56**:2172–9. http://dx.doi.org/10.1002/hep.25887

476. Dixon JB, Bhathal PS, Hughes NR, O'Brien PE. Nonalcoholic fatty liver disease: improvement in liver histological analysis with weight loss. *Hepatology* 2004;**39**:1647–54. http://dx.doi.org/10.1002/hep.20251

477. Huang MA, Greenson JK, Chao C, Anderson L, Peterman D, Jacobson J, *et al.* One-year intense nutritional counseling results in histological improvement in patients with nonalcoholic steatohepatitis: a pilot study. *Am J Gastroenterol* 2005;**100**:1072–81. http://dx.doi.org/10.1111/j.1572-0241.2005.41334.x

478. Bellentani S, Grave RD, Suppini A, Marchesini G, Bedogni G, Bugianesi E, *et al.* Behavior therapy for nonalcoholic fatty liver disease: the need for a multidisciplinary approach. *Hepatology* 2008;**47**:746–54. http://dx.doi.org/10.1002/hep.22009

479. Mummadi RR, Kasturi KS, Chennareddygari S, Sood GK. Effect of bariatric surgery on nonalcoholic fatty liver disease: systematic review and meta-analysis. *Clin Gastroenterol Hepatol* 2008;**6**:1396–402. http://dx.doi.org/10.1016/j.cgh.2008.08.012

480. Sanyal AJ, Mofrad PS, Contos MJ, Sargeant C, Luketic VA, Sterling RK, *et al.* A pilot study of vitamin E versus vitamin E and pioglitazone for the treatment of nonalcoholic steatohepatitis. *Clin Gastroenterol Hepatol* 2004;**2**:1107–15. http://dx.doi.org/10.1016/S1542-3565(04)00457-4

481. Sanyal AJ, Chalasani N, Kowdley KV, McCullough A, Diehl AM, Bass NM, *et al.* Pioglitazone, vitamin E, or placebo for nonalcoholic steatohepatitis. *N Engl J Med* 2010;**362**:1675–85. http://dx.doi.org/10.1056/NEJMoa0907929

482. Aithal GP, Thomas JA, Kaye PV, Lawson A, Ryder SD, Spendlove I, *et al.* Randomized, placebo-controlled trial of pioglitazone in nondiabetic subjects with nonalcoholic steatohepatitis. *Gastroenterology* 2008;**135**:1176–84. http://dx.doi.org/10.1053/j.gastro.2008.06.047

483. Belfort R, Harrison SA, Brown K, Darland C, Finch J, Hardies J, *et al.* A placebo-controlled trial of pioglitazone in subjects with nonalcoholic steatohepatitis. *N Engl J Med* 2006;**355**:2297–307. http://dx.doi.org/10.1056/NEJMoa060326

484. Ratziu V, Giral P, Jacqueminet S, Charlotte F, Hartemann-Heurtier A, Serfaty L, *et al.* Rosiglitazone for nonalcoholic steatohepatitis: one-year results of the randomized placebo-controlled Fatty Liver Improvement With Rosiglitazone Therapy (FLIRT) Trial. *Gastroenterology* 2008;**135**:100–10. http://dx.doi.org/10.1053/j.gastro.2008.03.078

485. Idilman R, Mizrak D, Corapcioglu D, Bektas M, Doganay B, Sayki M, *et al.* Clinical trial: insulin-sensitizing agents may reduce consequences of insulin resistance in individuals with non-alcoholic steatohepatitis. *Aliment Pharmacol Ther* 2008;**28**:200–8. http://dx.doi.org/10.1111/j.1365-2036.2008.03723.x

486. Shields WW, Thompson KE, Grice GA, Harrison SA, Coyle WJ. The effect of metformin and standard therapy versus standard therapy alone in nondiabetic patients with insulin resistance and nonalcoholic steatohepatitis (NASH): a pilot trial. *Ther Adv Gastroenterol* 2009;**2**:157–63. http://dx.doi.org/10.1177/1756283X09105462

487. Uygun A, Kadayifci A, Isik AT, Ozgurtas T, Deveci S, Tuzun A, *et al.* Metformin in the treatment of patients with non-alcoholic steatohepatitis. *Aliment Pharmacol Ther* 2004;**19**:537–44. http://dx.doi.org/10.1111/j.1365-2036.2004.01888.x

488. Haukeland JW, Konopski Z, Linnestad P, Azimy S, Loberg EM, Haaland T, *et al.* Abnormal glucose tolerance is a predictor of steatohepatitis and fibrosis in patients with non-alcoholic fatty liver disease. *Scand J Gastroenterol* 2005;**40**:1469–77. http://dx.doi.org/10.1080/00365520500264953

489. Mahady SE, Webster AC, Walker S, Sanyal A, George J. The role of thiazolidinediones in non-alcoholic steatohepatitis – a systematic review and meta analysis. *J Hepatol* 2011;**55**:1383–90. http://dx.doi.org/10.1016/j.jhep.2011.03.016

490. Lavine JE, Schwimmer JB, Van Natta ML, Molleston JP, Murray KF, Rosenthal P, *et al.* Effect of Vitamin E or metformin for treatment of nonalcoholic fatty liver disease in children and adolescents the tonic randomized controlled trial. *JAMA* 2011;**305**:1659–68. http://dx.doi.org/10.1001/jama.2011.520

491. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, *et al.* The diagnosis and management of non-alcoholic fatty liver disease: Practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. *Hepatology* 2012;**55**:2005–23. http://dx.doi.org/10.1002/hep.25762

492. Promrat K, Kleiner DE, Niemeier HM, Jackvony E, Kearns M, Wands JR, *et al.* Randomized controlled trial testing the effects of weight loss on nonalcoholic steatohepatitis. *Hepatology* 2010;**51**:121–9. http://dx.doi.org/10.1002/hep.23276

493. National Institute for Health and Care Excellence. *Obesity Guidance on the Prevention, Identification, Assessment and Managment of Overweight and Obesity in Adults and Children*. Clinical guideline CG43. London: NICE; 2006.

494. De Freitas ACT, Campos ACL, Coelho JCU. The impact of bariatric surgery on nonalcoholic fatty liver disease. *Curr Opin Clin Nutr Metab Care* 2008;**11**:267–74. http://dx.doi.org/10.1097/MCO.0b013e3282fbd33f

495. Chavez-Tapia NC, Tellez-Avila FI, Barrientos-Gutierrez T, Mendez-Sanchez N, Lizardi-Cervera J, Uribe M. Bariatric surgery for non-alcoholic steatohepatitis in obese patients. *Cochrane Database Syst Rev* 2010;**1**:CD007340.

496. David K, Kowdley KV, Unalp A, Kanwal F, Brunt EM, Schwimmer JB, *et al.* Quality of life in adults with nonalcoholic fatty liver disease: baseline data from the nonalcoholic steatohepatitis clinical research network. *Hepatology* 2009;**49**:1904–12. http://dx.doi.org/10.1002/hep.22868

497. Donnan PT, McLernon D, Dillon JF, Ryder S, Roderick P, Sullivan F, *et al.* Development of a decision support tool for primary care management of patients with abnormal liver function tests without clinically apparent liver disease: a record-linkage population cohort study and decision analysis (ALFIE). *Health Technol Assess* 2009;**13**(25). http://dx.doi.org/10.3310/hta13250

498. NHS Business Service Authority. *Prescribing Analysis Charts – Cardiovascular System*. URL: www.nhsbsa.nhs.uk/PrescriptionServices/2582.aspx (cited 26 June 2013).

499. National Institute for Health and Care Excellence. *Hypertension: Clinical Management of Primary Hypertension in Adults*. Clinical guideline 127. London: NICE; 2011.

500. Vernon G, Baranova A, Younossi ZM. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. *Aliment Pharmacol Ther* 2011;**34**:274–85. http://dx.doi.org/10.1111/j.1365-2036.2011.04724.x

501. D'Amico G, Garcia-Tsao G, Pagliaro L. Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. *J Hepatol* 2006;**44**:217–31. http://dx.doi.org/10.1016/j.jhep.2005.10.013

502. Andersson KL, Salomon JA, Goldie SJ, Chung RT. Cost effectiveness of alternative surveillance strategies for hepatocellular carcinoma in patients with cirrhosis. *Clin Gastroenterol Hepatol* 2008;**6**:1418–24. http://dx.doi.org/10.1016/j.cgh.2008.08.005

503. Bolondi L, Sofia S, Siringo S, Gaiani S, Casali A, Zironi G, *et al.* Surveillance programme of cirrhotic patients for early diagnosis and treatment of hepatocellular carcinoma: a cost effectiveness analysis. *Gut* 2001;**48**:251–9. http://dx.doi.org/10.1136/gut.48.2.251

504. Saab S, Ly D, Nieto J, Kanwal F, Lu D, Raman S, *et al.* Hepatocellular carcinoma screening in patients waiting for liver transplantation: a decision analytic model. *Liver Transpl* 2003;**9**:672–81. http://dx.doi.org/10.1053/jlts.2003.50120

505. Zhang BH, Yang BH, Tang ZY. Randomized controlled trial of screening for hepatocellular carcinoma. *J Cancer Res Clin Oncol* 2004;**130**:417–22. http://dx.doi.org/10.1007/s00432-004-0552-0

506. Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S, *et al. Methods for the Estimation of the NICE Cost Effectiveness Threshold*. Centre for Health Economics; 2013 (CHE Research Paper).

507. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: managment of hepatitis C virus infection. *J Hepatol* 2014;**60**:392–420. http://dx.doi.org/10.1016/j.jhep.2013.11.003

508. Barshop NJ, Francis CS, Schwimmer JB, Lavine JE. Nonalcoholic fatty liver disease as a comorbidity of childhood obesity. *Ped Health* 2009;**3**:271–81. http://dx.doi.org/10.2217/phe.09.21

509. Feldstein AE, Wieckowska A, Lopez AR, Liu YC, Zein NN, McCullough AJ. Cytokeratin-18 fragment levels as noninvasive biomarkers for nonalcoholic steatohepatitis: a multicenter validation study. *Hepatology* 2009;**50**:1072–8. http://dx.doi.org/10.1002/hep.23050

510. Chou R, Wasson N. Blood tests to diagnose fibrosis or cirrhosis in patients with chronic hepatitis C virus infection: a systematic review. *Ann Intern Med* 2013;**158**:807–20. http://dx.doi.org/10.7326/0003-4819-158-11-201306040-00005

511. Castera L, Foucher J, Bernard P-H, Carvalho F, Allaix D, Merrouche W, *et al.* Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. *Hepatology* 2010;**51**:828–35. http://dx.doi.org/10.1002/hep.23425

512. Shinkins B, Thompson M, Mallett S, Perera R. Diagnostic accuracy studies: how to report and analyse inconclusive test results. *BMJ* 2013;**346**:f2778. http://dx.doi.org/10.1136/bmj.f2778

513. Schuetz GM, Schlattmann P, Dewey M. Use of 3 × 2 tables with an intention to diagnose approach to assess clinical performance of diagnostic tests: meta-analytical evaluation of coronary CT angiography studies. *BMJ* 2012;**345**:e6717. http://dx.doi.org/10.1136/bmj.e6717

Appendix 1 Literature review: diagnostic test accuracy data

Search strategy

Date of search: April 2012.

1. CT.ti,ab.
2. tomodensitometry.ti,ab.
3. PET.ti,ab.
4. MRI.ti,ab.
5. NMRI.ti,ab.
6. zeugmatogra*.ti,ab.
7. ((computed or computeri?ed or magneti* or proton or "Acoustic Radiation Force Impulse" or ARF) adj5 (tomogra* or scan or scans or imaging)).ti,ab.
8. Tomography, X-Ray Computed/
9. Magnetic Resonance Imaging/
10. elastography.ti,ab.
11. elastographies.ti,ab.
12. sonoelastography.ti,ab.
13. sonoelastographies.ti,ab.
14. sono-elastography.ti,ab.
15. sono-elastographies.ti,ab.
16. elastogram.ti,ab.
17. elastograms.ti,ab.
18. vibroacoustography.ti,ab.
19. vibroacoustographies.ti,ab.
20. vibro-acoustography.ti,ab.
21. vibro-acoustographies.ti,ab.
22. fibroscan.ti,ab.
23. elastometry.ti,ab.
24. elasticity.ti,ab.
25. "liver stiffness".ti,ab.
26. elastogra*.ti,ab.
27. echogra*.ti,ab.
28. ultrason*.ti,ab.
29. ultrasound.ti,ab.
30. Ultrasonography/
31. Elasticity Imaging Techniques/
32. Elasticity.ti,ab.
33. ((alanine* or aspartate* or glutamic*) and (transaminase or aminotransferase*)).ti,ab.
34. SGOT.ti,ab.
35. SGPT.ti,ab.
36. AST.ti,ab.
37. ALT.ti,ab.
38. Aspartate Aminotransferases/
39. Alanine Transaminase/
40. platelet.ti,ab.
41. platelets.ti,ab.
42. thrombocyte.ti,ab.

43. thrombocytes.ti,ab.
44. APRI.ti,ab.
45. Blood Platelets/
46. 46. Biological Markers/
47. ELF.ti,ab.
48. "enhanced liver fibrosis".ti,ab.
49. Fibrotest.ti,ab.
50. Fibrosure.ti,ab.
51. Fibrometer.ti,ab.
52. FIB4.ti,ab.
53. FIB-4.ti,ab.
54. BARD.ti,ab.
55. Fibrospect.ti,ab.
56. Hepascore.ti,ab.
57. "Hyaluronic acid".ti,ab.
58. hyaluronate.ti,ab.
59. Hyaluronic Acid/
60. "Forns index".ti,ab.
61. laminin.ti,ab.
62. Laminin/
63. YKL-40.ti,ab.
64. "YKL 40".ti,ab.
65. "Type IV collagen".ti,ab.
66. Collagen Type IV/
67. "Procollagen III N-peptide".ti,ab.
68. "Lok index".ti,ab.
69. MP3.ti,ab.
70. MP-3.ti,ab.
71. "Fibrosis probability index".mp. or "sydney index".ti,ab.
72. FPI.ti,ab.
73. Fibroindex.ti,ab.
74. "Virahep-C index".ti,ab.
75. "Virahep C index".ti,ab.
76. "Göteborg University Cirrhosis Index".ti,ab.
77. GUCI.ti,ab.
78. SHASTA.ti,ab.
79. Glycocirrhotest.ti,ab.
80. Glycofibrotest.ti,ab.
81. BAAT.ti,ab.
82. "NAFLD fibrosis score".ti,ab.
83. Cytokeratin-18.ti,ab.
84. "Cytokeratin 18".ti,ab.
85. M30.ti,ab.
86. M-30.ti,ab.
87. "NASJH test".ti,ab.
88. "NAFIC score".ti,ab.
89. PGA.ti,ab.
90. "PGAA index".ti,ab.
91. "Bonancini score".ti,ab.
92. "Pohl score".ti,ab.
93. "Cirrhosis discriminant score".ti,ab.
94. "Age-platelet index".ti,ab.
95. TIMP-1.ti,ab.

96. "tissue inhibitory metalloprotease".ti,ab.
97. MBT.ti,ab.
98. "C-methacetin breath test".ti,ab.
99. "Phosphoproteomic biomarker".ti,ab.
100. "Phosphoproteomic biomarkers".ti,ab.
101. PICP.ti,ab.
102. PIIINP.ti,ab.
103. PON-I.ti,ab.
104. "paraoxonase I".ti,ab.
105. MFAP-4.ti,ab.
106. "MFAP 4".ti,ab.
107. MFAP4.ti,ab.
108. "microfibril associated glycoprotein 4".ti,ab.
109. or/1-9,27-30,33-39
110. limit 109 to yr="1988 -Current"
111. or/10-26,30-32,40-108
112. limit 111 to yr="2001 -Current"
113. 110 or 112
114. cirrhosis.ti,ab.
115. cirrhoses.ti,ab.
116. fibrosis.ti,ab.
117. fibroses.ti,ab.
118. "liver disease".ti,ab.
119. (hepatitis or hepatic).ti,ab.
120. steatohepatitis.ti,ab.
121. Liver Cirrhosis/
122. Fibrosis/
123. Liver Diseases/
124. Hepatitis/
125. or/114-124
126. 113 and 125
127. exp "sensitivity and specificity"/
128. "reproducibility of results"/
129. diagnos*.ti. or diagnostic.ab.
130. di.fs.
131. sensitivit*.ab.
132. specificit*.ab.
133. (ROC or "receiver operat*").ab.
134. Area under curve/
135. ("Area under curve" or AUC).ab.
136. (sROC or "optimal cut-off").ab.
137. (accura* or ((gold* or reference) adj2 standard)).ti,ab.
138. (likelihood adj3 (ratio* or function*)).ab.
139. ((true or false) adj3 (positive* or negative*)).ab.
140. ((positive* or negative* or false or true) adj3 (rate* or predictive)).ti,ab.
141. or/127-140
142. 126 and 141
143. *liver cirrhosis/di
144. *hepatitis/
145. *fibrosis/
146. (liver or hepatitis or hepatic or fibrosis).ab.
147. di.fs.
148. 146 and 147

149. or/143-145,148
150. 113 and 149
151. 113 and 149
152. 142 or 151

Search narrative

This search strategy has been kept deliberately very broad – utilising only two main search concepts: index test(s) (concept A) – lines 1–108 – and the disease of interest (concept B)/location of disease of interest (concept B) – lines 110–120. A methodological filter (concept C) is included but does not act as a filter to all search results [it is used in parallel: (A AND B AND C) OR (A AND B-focused)].

Potential studies for inclusion were initially identified from published non-Cochrane reviews and background literature. This generated a reference set of 70 potential (and probable) studies for inclusion to use to test the search strategy detailed above. The strategy was designed without knowledge of the 70 potential studies or of the search strategies used to identify the 70 from their original publications. All 70 studies were identified by the above strategy.

The yield from the above strategy was high. However, due the large number of tests within the scope of the review, a large yield could not be avoided.

Validation string

("19196449" or "18448567" or "16823833" or "15685546" or "19013661" or "18673426" or "18410556" or "18672413" or "16394849" or "16020491" or "17255218" or "18192914" or "17258346" or "17530363" or "18987556" or "19413672" or "17663420" or "18568136" or "18637064" or "18818788" or "18930329" or "18705692" or "19261000" or "21904476" or "17608672" or "18218676" or "19030204" or "19104699" or "18544945" or "19308312" or "18832522" or "18083083" or "12883497" or "20493576" or "20180868" or "19060630" or "19013661" or "18672413" or "19758273" or "19171202" or "18339075" or "18285716" or "18339592" or "19999223" or "18796094" or "18706734" or "18482283" or "18553008" or "18692034" or "17156890" or "17321634" or "17634962" or "17914968" or "16970597" or "16538110" or "16487951" or "16863553" or "17032409" or "16118349" or "17032410" or "16737415" or "16620291" or "16825937" or "16268817" or "16109665" or "15894397" or "15915455" or "16284529" or "15122779" or "17393509" or "18390575" or "19291784").ui.)

EMBASE

Date of search: April 2012.

1. exp *Liver Cirrhosis/ or exp *Liver Fibrosis/ or exp *Liver Disease/ or exp *Hepatitis/
2. (liver or hepatic).ti,ab.
3. exp *Liver/
4. 3 or 2
5. (cirrhosis or cirrhoses or fibrosis or fibroses or liver disease or hepatitis or steatohepatitis).ti,ab.
6. 4 and 5
7. 1 or 6
8. (CT or tomodensitometry or MRI or NMRI or zeugmatogra*).ti,ab.
9. ((computed or computerised or computerized or CT or magneti* or MR or NMR or proton) and (tomogra* or scan or scans or imaging)).ti,ab.
10. exp *computer assisted tomography/
11. exp *nuclear magnetic resonance imaging/

12. (elastography or elastographies or sonoelastography or sonoelastographies or sono-elastography or sono-elastographies or elastogram or elastograms or vibroacoustography or vibroacoustographies or vibro-acoustography or vibro-acoustographies or fibroscan or elastometry or elasticity or liver stiffness or echogra* or ultrason* or ultrasound).ti,ab.
13. exp *ultrasound/
14. exp *elastography/
15. ((alanine* or aspartate* or glutamic*) and (transaminase or aminotransferase*)).ti,ab.
16. (platelet or platelets or thrombocyte or thrombocytes or APRI or ELF or enhanced liver fibrosis or Fibrotest or Fibrosure or Fibrometer or FIB4 or FIB-4 or BARD or Fibrospect or Hepascore or Hyaluronic acid or hyaluronate or Forns index or laminin or YKL-40 or YKL 40 or Type IV collagen or Procollagen III N-peptide or Lok index or MP3 or MP-3 or Fibrosis probability index or FPI or Fibroindex or Virahep-C index or Virahep C index or Göteborg University Cirrhosis Index or GUCI or SHASTA or Glycocirrhotest or Glycofibrotest or BAAT or NAFLD fibrosis score or Cytokeratin-18 or Cytokeratin 18 or M30 or M-30 or NASJH test or NAFIC score or PGA or PGAA index or Bonancini score or Pohl score or Cirrhosis discriminant score or Age-platelet index or TIMP-1 or tissue inhibitory metalloprotease or MBT or C-methacetin breath test or Phosphoproteomic biomarker or Phosphoproteomic biomarkers or PICP or PIIINP or PON-I or paraoxonase I or MFAP-4 or MFAP 4 or MFAP4 or microfibril associated glycoprotein 4).ti,ab.
17. (SGOT or SGPT or AST or ALT).ti,ab.
18. exp *alanine aminotransferase/
19. exp *aspartate aminotransferase/
20. exp *thrombocyte/
21. exp *biological marker/
22. (2001* or 2002* or 2003* or 2004* or 2005* or 2006* or 2007* or 2008* or 2009* or 2010* or 2011* or 2012*).em.
23. 12 or 14 or 16 or 20 or 21
24. 22 and 23
25. 8 or 9 or 10 or 11 or 13 or 15 or 17 or 18 or 19
26. limit 25 to yr="1988 -Current"
27. 24 or 26
28. 7 and 27

Science Citation Index expanded

Date of search: 1988 to April 2012.

#1 TS=(cirrhosis OR cirrhoses OR fibrosis OR fibroses or liver disease or hepatitis or steatohepatitis)

#2 TS=(liver or hepatic)

#3 TS=(CT OR tomodensitometry OR PET OR MRI OR NMRI OR zeugmatogra*)

#4 TS=((Acoustic Radiation Force Impulse or ARFI OR computed OR computerised OR computerized OR CT OR magneti* OR MR OR NMR OR proton) AND (tomogra* OR scan OR scans OR imaging))

#5 TS=(elastography or elastographies or sonoelastography or sonoelastographies or sono-elastography or sono-elastographies or elastogram or elastograms or vibroacoustography or vibroacoustographies or vibro-acoustography or vibro-acoustographies or fibroscan or elastometry or elasticity or liver stiffness OR echogra* OR ultrason* OR ultrasound)

#6 TS=((alanine* OR aspartate* OR glutamic*) AND (transaminase OR aminotransferase*))

#7 TS=(SGOT OR SGPT OR AST OR ALT OR platelet OR platelets OR thrombocyte OR thrombocytes OR APRI OR ELF OR enhanced liver fibrosis OR Fibrotest OR Fibrosure OR Fibrometer OR FIB4 OR FIB-4 OR BARD OR Fibrospect OR Hepascore OR Hyaluronic acid OR hyaluronate OR Forns index OR laminin OR YKL-40 OR YKL 40 OR Type IV collagen OR Procollagen III N-peptide OR Lok index OR MP3 OR MP-3 OR Fibrosis probability index OR FPI OR Fibroindex OR Virahep-C index OR Virahep C index OR Göteborg University Cirrhosis Index OR GUCI OR SHASTA OR Glycocirrhotest OR Glycofibrotest OR BAAT OR NAFLD fibrosis score OR Cytokeratin-18 OR Cytokeratin 18 OR M30 OR M-30 OR NASJH test OR NAFIC score OR PGA OR PGAA index OR Bonancini score OR Pohl score OR Cirrhosis discriminant score OR Age-platelet index OR TIMP-1 OR tissue inhibitory metalloprotease OR MBT OR C-methacetin breath test OR Phosphoproteomic biomarker OR Phosphoproteomic biomarkers OR PICP OR PIIINP OR PON-I OR paraoxonase I OR MFAP-4 OR MFAP 4 OR MFAP4 OR microfibril associated glycoprotein 4)

#8 (#3 OR #4 OR #5 OR #6 OR #7)

#9 (#1 AND #2 AND #8)

Appendix 2 Literature review: cost-effectiveness analyses (hepatitis B, hepatitis C, alcoholic liver disease, non-alcoholic liver disease, cirrhosis)

Hepatitis B

Database, platform: MEDLINE (via Ovid)

Search strategy: Natural History
Date of search: 10 May 2012.

1. *EPIDEMIOLOGY/
2. *INCIDENCE/
3. *PREVALENCE/
4. incidence.ti.
5. prevalence.ti.
6. epidemiol$.ti.
7. (etiolog$ or aetiolog$).ti.
8. or/1-7
9. exp *Hepatitis B/
10. 8 and 9
11. limit 10 to (english language and humans)
12. limit 11 to yr="2004 -Current"

Search strategy: costs
Date of search: 11 May 2012.

1. exp Hepatitis B/ or Hepatitis B, Chronic/
2. exp Hepatitis B Virus/ or exp Hepatitis B Antibodies/
3. (hbv or hepatitis-B or hepatitis B or HBeAg negative or HBeAg positive or HBsAG).mp.
4. 1 or 2 or 3
5. (pegylat$ adj3interferon$ or peg-ifn or peginterferon$ or pegasys or pegintron or viraferonpeg).mp.
6. (interferon alpha 2a or interferon alfa 2a or interferon alpha 2b or interferon alfa 2b or alpha interferon or intron$ or viraferon or roferon).mp.
7. exp interferon-alpha/
8. 6 or 7
9. exp Polyethylene Glycols/
10. polyethylene glycol$.mp. or peg$.tw.
11. 9 or 10
12. 8 and 11
13. 5 or 12
14. 13 and 4
15. limit 14 to english language
16. (adefovir dipivoxil or adefovir$ or hepsera).mp.
17. 16 and 4
18. 17
19. limit 18 to english language
20. (tenofovir disoproxil or tenofovir$ or viread).mp.
21. 20 and 4

22. limit 21 to english language
23. (entecavir or entecavir$ or baraclude).mp.
24. 23 and 4
25. limit 24 to english language
26. exp ECONOMICS/
27. exp ECONOMICS, HOSPITAL/
28. exp ECONOMICS, PHARMACEUTICAL/
29. exp ECONOMICS, NURSING/
30. exp ECONOMICS, DENTAL/
31. exp ECONOMICS, MEDICAL/
32. exp "Costs and Cost Analysis"/
33. Cost-Benefit Analysis/
34. VALUE OF LIFE/
35. exp MODELS, ECONOMIC/
36. exp FEES/ and CHARGES/
37. exp BUDGETS/
38. (economic$ or price$ or pricing or financ$ or fee$ or pharmacoeconomic$ or pharma economic$).tw.
39. (cost$ or costly or costing$ or costed).tw.
40. (cost$ adj2 (benefits$ or utilit$ or minim$ or effective$)).tw.
41. (expenditure$ not energy).tw.
42. (value adj2 (money or monetary)).tw.
43. budget$.tw.
44. (economic adj2 burden).tw.
45. "resource use".ti,ab.
46. or/26-45
47. news.pt.
48. letter.pt.
49. editorial.pt.
50. comment.pt.
51. or/47-50
52. 46 not 51
53. 52 and 4
54. 52 and 15
55. 52 and 19
56. 52 and 22
57. 52 and 25
58. 53
59. 58 and 54 and 55 and 56 and 57
60. limit 59 to english language
61. limit 60 to yr="2004 -Current"

Search strategy: quality of life

Date of search: 11 May 2012.

1. value of life/
2. quality adjusted life year/
3. quality adjusted life.ti,ab.
4. (qaly$ or qald$ or qale$ or qtime$).ti,ab.
5. disability adjusted life.ti,ab.
6. daly$.ti,ab.
7. health status indicators/
8. (sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).ti,ab.

9. (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six or short form six).ti,ab.
10. (sf12 or sf 12 or short form 12 or shortform 12 or sf twelve or sftwelve or shortform twelve or short form twelve).ti,ab.
11. (sf16 or sf 16 or short form 16 or shortform 16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).ti,ab.
12. (sf20 or sf 20 or short form 20 or shortform or sf twenty or sf twenty or sftwenty or shortform twenty or short form twenty).ti,ab.
13. (euroqol or euro qol or eq5d or eq 5d).ti,ab.
14. (hql or hqol or h qol or hrqol or hr qol).ti,ab.
15. (hye or hyes).ti,ab.
16. health$ year$ equivalent$.ti,ab.
17. health utilit$.ab.
18. (hui or hui1 or hui2 or hui3).ti,ab.
19. disutil$.ti,ab.
20. rosser.ti,ab.
21. quality of well being.ti,ab.
22. quality of wellbeing.ti,ab.
23. qwb.ti,ab.
24. willingess to pay.ti,ab.
25. standard gamble$.ti,ab.
26. time trade off.ti,ab.
27. time tradeoff.ti,ab.
28. tto.ti,ab.
29. (index adj2 well being).mp.
30. (quality adj2 well being).mp.
31. (health adj3 utilit$ ind$).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]
32. ((multiattribute$ or multi attribute$) adj3 (health ind$ or theor$ or health state$ or utilit$ or analys$)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]
33. quality adjusted life year$.mp.
34. (15D or 15 dimension$).mp.
35. (12D or 12 dimension$).mp.
36. rating scale$.mp.
37. linear scal$.mp.
38. linear analog$.mp.
39. visual analog$.mp.
40. (categor$ adj2 scal$).mp.
41. or/1-40
42. (letter or editorial or comment).pt.
43. 41 not 42
44. exp Hepatitis B/ or Hepatitis B, Chronic/
45. exp Hepatitis B Virus/ or exp Hepatitis B Antibodies/
46. (hbv or hepatitis-B or hepatitis B or HBeAg negative or HBeAg positive or HBsAG).mp.
47. 44 or 45 or 46
48. 43 and 47
49. limit 48 to english language
50. limit 49 to yr="2004 -Current

Non-alcoholic fatty liver disease

Database, platform: MEDLINE (via Ovid)

Search strategy 1
Date of search: 24 June 2013.

1. (enhanced adj liver adj fibrosis).tw.
2. (elf adj test$).tw.
3. (elf and diagnos$).tw.
4. (elf and (fibros*s or cirrhos*s)).tw.
5. elf.tw.
6. exp liver cirrhosis/ or exp liver diseases/
7. Fatty Liver/di, dh, de, dt, ec [Diagnosis, Diet Therapy, Drug Effects, Drug Therapy, Economics]
8. 6 and 7
9. 5 and 8
10. 1 or 2 or 3 or 4 or 9
11. age-plt index.tw.
12. api index.tw.
13. apri.tw.
14. arfi.tw.
15. astalt.tw.
16. ast-alt.tw.
17. bard.tw.
18. coll4.tw.
19. typeIVcollagen.tw.
20. type IV Collagen.tw.
21. FIB 4.tw.
22. Fibrotest.tw.
23. Hyaluronic Acid.tw.
24. Hepascore.tw.
25. NAFIC.tw.
26. nafic score.tw.
27. ndp.tw.
28. nedaplatin.tw.
29. nfs.tw.
30. nafld fibrosis score.tw.
31. plt.tw.
32. magentic resonance elastography.tw.
33. mre.tw.
34. (transient adj elsatograph$).tw.
35. (elastograph$ and liver).tw.
36. or/11-35
37. exp liver cirrhosis/ or exp liver diseases/
38. Fatty Liver/di, dh, de, dt, ec [Diagnosis, Diet Therapy, Drug Effects, Drug Therapy, Economics]
39. (fibros*s or chirrhos*s).tw.
40. 37 or 38 or 39
41. Biological Markers/
42. (biomarker$ or bio-marker$).tw.
43. (marker$ and (biologic$ or biochemical or serum or direct or indirect)).tw.
44. Algorithms/
45. algorithm$.tw.
46. (composite and blood).tw.

47. or/41-46
48. 36 and 47
49. Hyaluronic Acid/
50. ((hyaluronic adj acid) or (hyalauronate or hyaluronan)).tw.
51. 49 or 50
52. (procollagen or piinp or p3np or ppcp).tw.
53. ((tissue and inhibitor and metalloproteinase$) or timps).tw.
54. 51 and 52 and 53
55. 52 or 53 or 54
56. 36 and 55
57. Alpha-Macroglobulins/
58. ((alpha and macroglobulin$) or (alpha adj 2m)).tw.
59. 57 or 58
60. ((apolipoprotein$ adj a 1) or apoa 1).tw.
61. Haptoglobins/
62. haptoglobin$.tw.
63. 61 or 62
64. (bilrubin$ or hematoidin$).tw.
65. (gamma adj glutamyl adj transpeptidase$).tw.
66. (gamma adj glutamyltransferase$).tw.
67. 64 or 65 or 66
68. 59 and 60 and 63 and 64 and 67
69. 59 or 60 or 63 or 64 or 67
70. 36 and 69
71. (alanine adj (aminotransferase$ or aminotransaminase$)).tw.
72. (serum adj glutamic adj oxaloacetic adj transaminase$).tw.
73. sgpt.tw.
74. 71 or 72 or 73
75. (asparate adj (aminotransferase$ or aminotransaminase$)).tw.
76. (serum adj glutamic adj oxaloacetic adj transaminase$).tw.
77. sgot.tw.
78. 75 or 76 or 77
79. 59 and 60 and 63 and 64 and 67 and 74 and 78
80. 59 or 60 or 63 or 64 or 67 or 74 or 78
81. 36 and 80
82. exp "Sensitivity and Specificity"/
83. sensitivity.tw.
84. specificity.tw.
85. ((pre-test or pretest) adj probability).tw.
86. post-test probability.tw.
87. predictive value$.tw.
88. likelihood ratio$.tw.
89. or/82-88
90. 48 and 89
91. 56 and 89
92. 70 and 91
93. 81 and 89
94. 90 or 91 or 92 or 93
95. iqur.tw.
96. biopredictive.tw.
97. echosens.tw.
98. 95 or 96 or 97
99. 10 or 36 or 54 or 68 or 79 or 94 or 98

100. exp "Costs and Cost Analysis"/
101. Economics/
102. exp Economics, Hospital/
103. exp Economics, Medical/
104. Economics, Nursing/
105. exp models,economic/
106. Economics, Pharamceutical/
107. exp "Fees and Charges"/
108. exp Budgets/
109. budget$.tw.
110. ec.fs.
111. cost$.ti.
112. (cost$ adj2 (effective$ or utilit$ or benefit$or minimi$)).ab.
113. (economic$ or pharmaceconomic$ or pharmaco-economic$).ti.
114. (price$ or pricing$).tw.
115. (financial or finance or finances or financed).tw.
116. (fee or fees).tw.
117. (value adj2 (money or monetary)).tw.
118. quality-adjusted life years/
119. (qaly or qalys).af.
120. (quality adjusted life year or quality adjusted life years).af.
121. 101 or 120
122. 99 and 121
123. Liver Cirrhosis/ or Middle Aged/ or Aged/ or Liver Diseases, Alcoholic/ or Hepatitis/ or Fatty Liver/ or Adult/ or Liver Diseases/ or non alcoholic liver disease.mp. or Liver/
124. 122 and 123

Search strategy 2

Date of search: 26 July 2013.

1. cost effectiveness.mp. or Cost-Benefit Analysis/
2. Hepatitis/ or Fatty Liver/ or non alcoholic steatohepatitis.mp. or Liver/
3. 1 and 2

Alcoholic liver disease

Database, platform: MEDLINE (via Ovid)

Search strategy

Date of search: 21 June 2013.

1. (enhanced adj liver adj fibrosis).tw.
2. (elf adj tests$).tw.
3. (elf and diagnos$).tw.
4. (elf and (fibros* or cirrhos*s)).tw.
5. elf.tw.
6. exp liver cirrhosis/ or exp liver diseases, alcoholic/
7. 5 and 6
8. 1 or 2 or 3 or 4 or 7
9. Cytokeratin-18.tw.
10. Forns.tw.
11. Fibroscan.tw.

12. YKL-40.tw.
13. (transient adj elastograph$).tw.
14. (elastograph$ and liver).tw.
15. or/9-14
16. exp liver cirrhosis/ or exp liver diseases, alcoholic/
17. (fibros* or cirrhos*s).tw.
18. 16 or 17
19. Biological Markers/
20. (biomarker$ or bio-markers$).tw.
21. (marker$ and (biologic$ or biochemical or serum or direct or indirect)).tw.
22. Algorithms/
23. algorithm$.tw.
24. (composite and blood).tw.
25. or/19-24
26. 18 and 25
27. Hyaluronic Acid/
28. ((hyaluronic adj acid) or (hyaluronate or hyaluronan)).tw.
29. 27 or 28
30. (procollagen or piinp or p3np or ppcp).tw.
31. ((tissue and inhibitor and metalloproteinase$) or timps).tw.
32. 29 and 30 and 31
33. 30 or 31 or 32
34. 18 or 33
35. Alpha-Macroglobulins/
36. ((apha and macroglobulin$) or (alpha adj 2m)).tw.
37. 35 or 36
38. ((apolipoprotein$ adj a 1) or apoa 1).tw.
39. Haptoglobins/
40. haptoglobin$.tw.
41. 39 or 40
42. (bilirubin$ or hematoidin$).tw.
43. (gamma adj glutamyl adj transpeptidase$).tw.
44. (gamma adj glutamyltransferase$).tw.
45. ((gamma adj gt) or ggt or ggtp).tw.
46. 43 or 44 or 45
47. 37 and 38 and 41 and 42 and 46
48. 37 or 38 or 41 or 42 or 46
49. 18 and 48
50. (alanine adj (aminotransferase$ or aminotransaminase$)).tw.
51. (serum adj glutamic adj pyruvic adj transaminase$).tw.
52. sgpt.tw.
53. 50 or 51 or 52
54. (aspartate adj (aminotransferase$ or aminotransaminase$)).tw.
55. (serum adj glutamic adj oxaloacetic adj transaminase$).tw.
56. sgot.tw.
57. 54 or 55 or 56
58. 37 and 38 and 41 and 42 and 46 and 53 and 57
59. 37 or 38 or 41 or 42 or 46 or 53 or 57
60. 18 and 59
61. exp "Sensitivity and Specificity"/
62. sensitivity.tw.
63. specificity.tw.
64. ((pre-test or pretest) adj probability).tw.

65. post-test probability.tw.
66. predictive value$.tw.
67. likelihood ratio$.tw.
68. or/61-67
69. 26 and 68
70. 34 and 68
71. 49 and 68
72. 60 and 68
73. 69 or 70 or 71 or 72
74. iqur.tw.
75. biopredictive.tw.
76. echosens.tw.
77. 74 or 75 or 76
78. 7 or 15 or 32 or 47 or 58 or 73 or 77
79. exp "Costs and Cost Analysis"/
80. Economics/
81. exp Economics, Hospital/
82. exp Economics, Medical/
83. Economics, Nursing/
84. exp models, economic/
85. Economoics, Pharmaceutical/
86. exp "Fees and Charges"/
87. exp Budgets/
88. budget$.tw.
89. ec.fs.
90. cost$.ti.
91. (cost$ adj2 (effective$ or utilit$ or benefit$ or minimi$)).ab.
92. (economic$ or pharmacoeconomic$ or pharmaco-economic$).ti.
93. (prices$ or pricing$).tw.
94. (financial or finance or finances or financed).tw.
95. (fee or fees).tw.
96. (value adj2 (money or monetary)).tw.
97. quality-adjusted life years/
98. (qaly or qalys).af.
99. (quality adjusted life year or quality adjusted life years).af.
100. or/79-99
101. 78 and 100
102. limit 101 to english language

Hepatitis C

Database, platform: MEDLINE (via Ovid)

Natural history

Search strategy

Date of search: 1 December 2012.

1. *EPIDEMIOLOGY/
2. *INCIDENCE/
3. *PREVALENCE/
4. incidence.ti.

5. prevalence.ti.
6. epidemiol$.ti.
7. (etiolog$ or aetiolog$).ti.
8. or/1-7
9. exp *Hepatitis C/
10. 8 and 9
11. limit 10 to (human and english language)
12. limit 11 to yr="2004 -Current"

Costs

Search strategy

Date of search: 1 December 2012.

1. exp Hepatitis C/ or Hepatitsi C, Chronic.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]
2. exp Hepatitis C Virus/ or exp Hepatitis C Antibodies/
3. (hcv or hepatitis-C or heptatitis C).mp.
4. 1 or 2 or 3
5. ((pegylat$ adj3 interferon$) or peg-ifn or peginterferon$ or peg-interferon$ or pegasys or pegintron or viraferonpeg).mp.
6. (interferon alpha 2a or interferon alfa 2a or interferon alpha 2b or interferon alfa 2b or alpha interferon or intron$ or viraferon or roferon).mp.
7. exp interferon-alpha/
8. 6 or 7
9. exp Polyethylene Glycols/
10. polyethylene glycol$.mp. or peg$.tw.
11. 9 or 10
12. 8 and 11
13. 5 or 12
14. 4 and 13
15. limit 14 to english language
16. (Ribavirin or ribavirin$ or copegus or rebetol).mp.
17. 4 and 16
18. 17
19. limit 18 to english language
20. (Telaprevir or telaprevir$ or incivo).mp.
21. 4 and 20
22. limit 21 to english language
23. (Boceprevir or Boceprevir$ or victrelis).mp.
24. 4 and 23
25. limit 24 to english language
26. exp ECONOMICS/
27. exp ECONOMICS, HOSPITAL/
28. exp ECONOMICS, PHARMACEUTICAL/
29. exp ECONOMICS, NURSING/
30. exp ECONOMICS, DENTAL/
31. exp ECONOMICS, MEDICAL/
32. exp "Costs and Cost Analysis"/
33. Cost-Benefit Analysis/
34. VALUE OF LIFE/
35. exp MODELS, ECONOMIC/

36. exp FEES/ and CHARGES/
37. exp BUDGETS/
38. (economics$ or price$ or pricing or fianc$ or fee$ or pharamacoenomics$ or pharma economics$).tw.
39. (cost$ or costly or costing$ or costed).tw.
40. (cost$ adj2 (benefit$ or utilit$ or minim$ or effective$)).tw.
41. (expenditure$ not energy).tw.
42. (value adj (money or monetary)).tw.
43. budget$.tw.
44. (economic adj2 burden).tw.
45. "resource use".ti,ab.
46. or/26-45
47. news.pt.
48. editorial.pt.
49. comment.pt.
50. letter.pt.
51. or/47-50
52. 46 not 51
53. 52 and 4
54. 52 and 15
55. 52 and 19
56. 52 and 23
57. 52 and 25
58. 53 and 54 and 55 and 56 and 57
59. limit 58 to english language

Quality of life

Search strategy

Date of search: 1 December 2012.

1. value of life/
2. quality adjusted life year/
3. quality adjusted life.ti,ab.
4. (qaly$ or qald$ or qale$ or qtime$).ti,ab.
5. disability adjusted life.ti,ab.
6. daly$.ti,ab.
7. health status indicators/
8. (sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).ti,ab.
9. (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six or short form six).ti,ab.
10. (sf12 or sf 12 or short form 12 or shortform 12 or sf twelve or sftwelve or shortform twelve or short form twelve).ti,ab.
11. (sf16 or sf 16 or short form 16 or shortform 16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).ti,ab.
12. (sf20 or sf 20 or short form 20 or shortform or sf twenty or sf twenty or sftwenty or shortform twenty or short form twenty).ti,ab.
13. (euroqol or euro qol or eq5d or eq 5d).ti,ab.
14. (hql or hqol or h qol or hrqol or hr qol).ti,ab.
15. (hye or hyes).ti,ab.
16. health$ year$ equivalent$.ti,ab.
17. health utilit$.ab.
18. (hui or hui1 or hui2 or hui3).ti,ab.

19. disutil$.ti,ab.
20. rosser.ti,ab.
21. quality of well being.ti,ab.
22. quality of wellbeing.ti,ab.
23. qwb.ti,ab.
24. willingess to pay.ti,ab.
25. standard gamble$.ti,ab.
26. time trade off.ti,ab.
27. time tradeoff.ti,ab.
28. tto.ti,ab.
29. (index adj2 well being).mp.
30. (quality adj2 well being).mp.
31. (health adj3 utilit$ ind$).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]
32. ((multiattribute$ or multi attribute$) adj3 (health ind$ or theor$ or health state$ or utilit$ or analys$)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]
33. quality adjusted life year$.mp.
34. (15D or 15 dimension$).mp.
35. (12D or 12 dimension$).mp.
36. rating scale$.mp.
37. linear scal$.mp.
38. linear analog$.mp.
39. visual analog$.mp.
40. (categor$ adj2 scal$).mp.
41. or/1-40
42. (letter or editorial or comment).pt.
43. 41 not 42
44. exp Hepatitis C/ or Hepatitis C, Chronic/
45. exp Hepatitis B Virus/ or exp Hepatitis B Antibodies/
46. (hbv or hepatitis-C or hepatitis C).mp.
47. 44 or 45 or 46
48. 43 and 47
49. limit 48 to english language

Cirrhosis

Platform: MEDLINE (via Ovid)

Date of search: 26 July 2013.

Search strategy

Costs

1. cirrhosis.mp.
2. costs.mp. or "Costs and Cost Analysis"/
3. 1 and 2
4. limit 3 to english language

Quality of life

1. cirrhosis.mp.
2. quality of life.mp. or "Quality of Life"/
3. 1 and 4
4. limit 5 to english language

Appendix 3 Results: meta-analysis data

TABLE 61 Diagnostic accuracy of non-invasive tests for detection of fibrosis stage of ≥ F1 in patients with chronic HCV

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Indirect serum non-invasive tests					
King's	1	24.3	0.80 (0.74 to 0.85)	0.67 (0.30 to 0.90)	Single study
Direct serum non-invasive serum tests					
Hyaluronic acid	1	16	0.36 (0.28 to 0.45)	0.92 (0.87 to 0.95)	Single study
Commercial non-invasive serum tests					
FibrospectII	1	42	0.86 (0.49 to 0.97)	0.69 (0.50 to 0.83)	Single study
Imaging modalities					
ARFI	3	1.04–1.19	0.71 (0.65 to 0.77)	0.86 (0.70 to 0.94)	Fixed-effect model for sensitivity and random-effects model for specificity without correlation
Fibroscan	8	4.5–8.8	0.85 (0.75 to 0.91)	0.87 (0.75 to 0.91)	Random-effects model for sensitivity and specificity without correlation

TABLE 62 Diagnostic accuracy of non-invasive tests for detection of fibrosis stage of ≥ F3 in patients with chronic HCV

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Indirect non-invasive serum tests					
APRI (low cut-off)	18	0.5–1.0	0.84 (0.82 to 0.86)	0.56 (0.44 to 0.68)	Fixed-effect model for sensitivity and random-effects model for specificity without correlation
APRI (high cut-off)	15	1.5–2.0	0.53 (0.43 to 0.62)	0.86 (0.79 to 0.91)	Bivariate random-effects model with correlation between sensitivity and specificity
CDS	2	8	0.54 (0.43 to 0.65)	0.74 (0.66 to 0.81)	Fixed-effects model for sensitivity and specificity without correlation
FIB-4 (low cut-off)	11	1.45	0.80 (0.72 to 0.86)	0.64 (0.56 to 0.72)	Bivariate random-effects model with correlation between sensitivity and specificity
FIB-4 (high cut-off)	11	3.25	0.37 (0.28 to 0.46)	0.94 (0.90 to 0.97)	Bivariate random-effects model with correlation between sensitivity and specificity
Forns index (low cut-off)	1	4.2	0.92 (0.81 to 0.97)	0.34 (0.26 to 0.43)	Single study
Forns index (high cut-off)	1	6.9	0.55 (0.41 to 0.68)	0.87 (0.80 to 0.92)	Single study

continued

TABLE 62 Diagnostic accuracy of non-invasive tests for detection of fibrosis stage of ≥ F3 in patients with chronic HCV (*continued*)

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
FibroQ	1	1.6	0.86 (0.78 to 0.91)	0.44 (0.35 to 0.52)	Single study
GUCI	1	0.26	0.58 (0.39 to 0.76)	0.73 (0.58 to 0.84)	Single study
King's	1	24.3	0.74 (0.59 to 0.85)	0.90 (0.84 to 0.94)	Single study
Lok's index (low cut-off)	2	0.2	0.90 (0.85 to 0.95)	0.33 (0.27 to 0.39)	Fixed-effects model for sensitivity and specificity without correlation
Lok's index (high cut-off)	2	0.5–0.58	0.50 (0.40 to 0.59)	0.84 (0.77 to 0.89)	Fixed-effects model for sensitivity and specificity without correlation
NIHCED	1	6	0.72 (0.65 to 0.78)	0.75 (0.67 to 0.81)	Single study
Platelets	2	140–150	0.53 (0.43 to 0.99)	0.88 (0.47 to 0.98)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
Pohl index	3	Positive	0.15 (0.04 to 0.42)	0.98 (0.96 to 0.99)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
Direct serum non-invasive serum tests					
^{13}C-caffeine breath test	1	0.021	0.75 (0.63 to 0.85)	0.79 (0.64 to 0.89)	Single study
Hyaluronic acid	4	20–85	0.79 (0.52 to 0.93)	0.72 (0.65 to 0.78)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
Hyaluronic acid (low cut-off)	1	48	0.77 (0.50 to 0.92)	1.00 (0.98 to 1.00)	Single study
Hyaluronic acid (high cut-off)	1	160	0.22 (0.13 to 0.37)	1.00 (0.98 to 1.00)	Single study
Hepascore	7	0.5–0.83	0.81 (0.71 to 0.87)	0.76 (0.68 to 0.83)	Bivariate random-effects model with correlation between sensitivity and specificity
Hepascore (low cut-off)	2	0.49	0.81 (0.41 to 0.96)	0.75 (0.18 to 0.98)	Fixed-effect model for sensitivity and random-effects model for specificity without correlation
Hepascore (high cut-off)	2	0.84–0.90	0.48 (0.40 to 0.56)	0.93 (0.90 to 0.95)	Fixed-effects model for sensitivity and specificity without correlation
PIIINP	2	8–9.1	0.71 (0.58 to 0.81)	0.63 (0.54 to 0.71)	Fixed-effects model for sensitivity and specificity without correlation
PIIINP/MMP1 index	1	0.3	0.86 (0.71 to 0.94)	0.74 (0.67 to 0.80)	Single study
Type IV collagen	1	130	0.67 (0.53 to 0.78)	0.76 (0.60 to 0.87)	Single study
YKL-40	1	100	0.82 (0.70 to 0.90)	0.57 (0.49 to 0.65)	Single study

TABLE 62 Diagnostic accuracy of non-invasive tests for detection of fibrosis stage of ≥ F3 in patients with chronic HCV (*continued*)

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Commercial non-invasive serum tests					
ELF (low cut-off)	1	9.59	0.85 (0.77 to 0.90)	0.63 (0.57 to 0.70)	Single study
ELF (high cut-off)	1	10.22	0.70 (0.61 to 0.78)	0.85 (0.80 to 0.89)	Single study
Fibroindex	1	1.35	0.52 (0.33 to 0.70)	0.92 (0.75 to 0.98)	Single study
Fibrometer	2	0.63–0.67	0.84 (0.77 to 0.89)	0.78 (0.75 to 0.81)	Fixed-effects model for sensitivity and specificity without correlation
FibrospectII	1	0.5	0.85 (0.72 to 0.92)	0.72 (0.62 to 0.80)	Single study
Fibrotest	9	0.32–0.67	0.73 (0.56 to 0.85)	0.69 (0.55 to 0.80)	Bivariate random-effects model with correlation between sensitivity and specificity
Fibrotest (low cut-off)	2	0.22	0.85 (0.44 to 0.98)	0.58 (0.54 to 0.99)	Fixed-effect model for sensitivity and random-effects model for specificity without correlation
Fibrotest (high cut-off)	2	0.59–0.63	0.69 (0.59 to 0.74)	0.84 (0.81 to 0.87)	Fixed-effects model for sensitivity and specificity without correlation
Imaging modalities					
ARFI	4	1.49–2.11	0.85 (0.69 to 0.94)	0.89 (0.72 to 0.97)	Random-effects model for sensitivity and specificity without correlation
Real-time elastography	1	3.25	0.86 (0.72 to 0.93)	0.96 (0.82 to 0.99)	Single study
Fibroscan	19	8.6–15.4	0.88 (0.82 to 0.92)	0.90 (0.85 to 0.93)	Bivariate random-effects model with correlation between sensitivity and specificity

CDS, Cirrhosis Discriminant Score; MMP-1, matrix metalloproteinase-1; NIHCED, non-invasive hepatitis C-related early detection.

TABLE 63 Diagnostic accuracy of non-invasive tests for detection of ≥F1 in patients with chronic HBV

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Indirect non-invasive serum tests					
APRI (low cut-off)	1	0.4	0.70 (0.59 to 0.78)	0.83 (0.68 to 0.92)	Single study
APRI (high cut-off)	1	1.5	0.37 (0.26 to 0.50)	0.82 (0.73 to 0.89)	Single study
CDS	1	4	0.28 (0.19 to 0.41)	0.90 (0.80 to 0.95)	Single study
Lok's index	1	0.87	0.48 (0.36 to 0.61)	0.90 (0.80 to 0.95)	Single study
Direct serum non-invasive tests					
CTGF	1	125.6	0.61 (0.49 to 0.71)	0.71 (0.47 to 0.87)	Single study
Commercial non-invasive serum tests					
Fibrotest	1	–	0.72 (0.57 to 0.83)	0.64 (0.49 to 0.76)	Single study
Imaging modalities					
Real-time elastography	1	–	0.87 (0.76 to 0.94)	0.85 (0.64 to 0.95)	Single study
Fibroscan	2	6.1	0.69 (0.53 to 0.82)	0.62 (0.39 to 0.80)	Fixed-effects model for sensitivity and specificity without correlation

CDS, Cirrhosis Discriminant Score.

TABLE 64 Diagnostic accuracy of non-invasive tests for detection of ≥ F3 in patients with chronic HBV

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Indirect non-invasive serum tests					
FIB-4 (low cut-off)	2	0.67–1.00	0.89 (0.84 to 0.93)	0.76 (0.69 to 0.81)	Fixed-effects model for sensitivity and specificity without correlation
FIB-4 (high cut-off)	1	2.65	0.38 (0.33 to 0.44)	0.98 (0.96 to 0.99)	Single study
Forns index (low cut-off)	2	5.2	0.99 (0.85 to 1.00)	0.20 (0.12 to 0.32)	Fixed-effects model for sensitivity and specificity without correlation
Forns index (high cut-off)	2	8.4	0.32 (0.20 to 0.46)	0.92 (0.83 to 0.97)	Fixed-effects model for sensitivity and specificity without correlation
Hui index	2	0.15	0.88 (0.76 to 0.94)	0.51 (0.37 to 0.65)	Fixed-effects model for sensitivity and specificity without correlation
Direct serum non-invasive tests					
^{13}C-caffeine breath test	1	1.49	1.00 (0.76 to 1.00)	0.72 (0.56 to 0.84)	Single study
CTGF	1	141	0.69 (0.44 to 0.86)	0.85 (0.75 to 0.92)	Single study
Commercial non-invasive serum tests					
Fibrotest	3	0.31–0.42	0.49 (0.01 to 0.99)	0.71 (0.53 to 0.84)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation
Imaging modalities					
Real-time elastography	1	80.7	0.73 (0.54 to 0.86)	0.76 (0.62 to 0.85)	Single study
Fibroscan	13	7.3–10.7	0.69 (0.58 to 0.78)	0.84 (0.79 to 0.89)	Bivariate random-effects model with correlation between sensitivity and specificity

CTGF, connective tissue growth factor.

TABLE 65 Diagnostic accuracy of non-invasive tests for detection of fibrosis stage of ≥ F1 in patients with non-alcoholic steatohepatitis

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Indirect non-invasive serum tests					
APRI	1	0.5	0.05 (0.01 to 0.17)	0.97 (0.87 to 0.99)	Single study
NAFLD fibrosis score (low cut-off)	3	–0.1657 to –1.456	0.82 (0.77 to 0.87)	0.48 (0.40 to 0.56)	Bivariate random-effects model with correlation between sensitivity and specificity
NAFLD fibrosis score (high cut-off)	2	0.676	0.29 (0.22 to 0.36)	0.92 (0.85 to 0.96)	Bivariate random-effects model with correlation between sensitivity and specificity
Direct serum non-invasive serum tests					
Hyaluronic acid	1	24.6	0.82 (0.52 to 0.95)	0.68 (0.46 to 0.85)	Single study
Laminin	1	282	0.82 (0.52 to 0.95)	0.89 (0.69 to 0.97)	Single study
NAFLD diagnostic panel	1	0.42	0.61 (0.46 to 0.75)	0.72 (0.57 to 0.84)	Single study
Type IV collagen	1	145	0.64 (0.35 to 0.85)	0.89 (0.69 to 0.97)	Single study
Commercial non-invasive serum tests					
ELF	1	9.8	0.61 (0.52 to 0.69)	0.80 (0.70 to 0.87)	Single study
Imaging modalities					
Real-time elastography	1	102	0.79 (0.65 to 0.88)	0.90 (0.60 to 0.98)	Single study
Fibroscan	3	5.3–5.9	0.87 (0.81 to 0.92)	0.76 (0.57 to 0.88)	Bivariate random-effects model with correlation between sensitivity and specificity
Combination of non-invasive test algorithms					
NAFLD fibrosis score and ELF (low cut-off)	1		0.92 (0.86 to 0.96)	0.52 (0.41 to 0.63)	Single study
NAFLD fibrosis score and ELF (high cut-off)	1		0.60 (0.51 to 0.69)	0.91 (0.83 to 0.96)	Single study

TABLE 66 Diagnostic accuracy of non-invasive tests for detection of fibrosis stage of ≥ F2 in patients with non-alcoholic steatohepatitis

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Indirect non-invasive serum tests					
APRI	2	0.43–0.5	0.69 (0.21 to 0.95)	0.82 (0.07 to 0.97)	Fixed-effect model for sensitivity and random-effects model for specificity without correlation
BARD	1	2	0.44 (0.35 to 0.54)	0.70 (0.62 to 0.77)	Single study
FIB-4	1	1.45	0.55 (0.45 to 0.64)	0.87 (0.81 to 0.92)	Single study
NAFLD fibrosis score (low cut-off)	4	–1.455	0.79 (0.56 to 0.92)	0.65 (0.46 to 0.80)	Bivariate random-effects model with correlation between sensitivity and specificity
NAFLD fibrosis score (high cut-off)	5	0.676	0.29 (0.07 to 0.68)	0.95 (0.87 to 0.98)	Bivariate random-effects model with correlation between sensitivity and specificity
Direct serum non-invasive serum tests					
Hepascore	1	0.44	0.51 (0.41 to 0.60)	0.88 (0.82 to 0.93)	Single study
Hyaluronic acid	1	218	0.78 (0.45 to 0.94)	0.89 (0.67 to 0.97)	Single study
NAFIC (low cut-off)	1	0	0.95 (0.91 to 0.98)	0.33 (0.29 to 0.37)	Single study
NAFIC (high cut-off)	1	2	0.84 (0.77 to 0.89)	0.75 (0.71 to 0.78)	Single study
Commercial non-invasive serum tests					
ELF	1	9.9	0.70 (0.59 to 0.79)	0.80 (0.72 to 0.86)	Single study
Fibrometer	1	0.490	0.78 (0.67 to 0.87)	0.96 (0.92 to 0.98)	Single study
FibrospectII	1	20	1.00 (0.95 to 1.00)	0.42 (0.32 to 0.52)	Single study
Fibrotest (low cut-off)	3	0.30–0.34	0.70 (0.56 to 0.81)	0.75 (0.70 to 0.80)	Bivariate random-effects model with correlation between sensitivity and specificity
Fibrotest (high cut-off)	2	0.7	0.15 (0.03 to 0.90)	0.98 (0.90 to 0.99)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation

continued

TABLE 66 Diagnostic accuracy of non-invasive tests for detection of fibrosis stage of ≥ F2 in patients with non-alcoholic steatohepatitis (*continued*)

Test	Number of studies	Cut-off	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Statistics
Imaging modalities					
Real-time elastography	1	94	0.84 (0.65 to 0.94)	1.00 (0.87 to 1.00)	Single study
Fibroscan	7	6.8–10.0	0.79 (0.72 to 0.85)	0.76 (0.71 to 0.80)	Random-effects model for sensitivity and fixed-effect model for specificity without correlation – the studies were clustered; fixed effects model for both sensitivity and specificity did not alter the mean but altered the CI by about 2%
Combination of non-invasive test algorithms					
NAFLD fibrosis score and ELF (low cut-off)	1		0.90 (0.81 to 0.95)	0.86 (0.78 to 0.91)	Single study
NAFLD fibrosis score and ELF (high cut-off)	1		0.79 (0.69 to 0.87)	0.91 (0.85 to 0.95)	Single study
NAFLD fibrosis score and Fibroscan	1		0.65 (0.51 to 0.76)	0.64 (0.56 to 0.71)	Single study
Fibrotest and Fibroscan	1		0.71 (0.57 to 0.81)	0.76 (0.68 to 0.82)	Single study

Appendix 4 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Adams 2005 training[71]	Hepascore_F2	Hepascore	F2	0.5	67	92	34	5	17	61
Adams 2005 training[71]	Hepascore_F3	Hepascore	F3	0.5	95	81	21	18	1	77
Adams 2005 training[71]	Hepascore_F4	Hepascore	F4	0.84	71	84	5	18	2	92
Adams 2005 validation[71]	Hepascore_F2	Hepascore	F2	0.5	63	89	37	5	22	40
Adams 2005 validation[71]	Hepascore_F3	Hepascore	F3	0.5	88	74	40	15	5	44
Adams 2005 validation[71]	Hepascore_F4	Hepascore	F4	0.84	71	89	12	10	5	77
Ahmad 2011[72]	APRI_F2_combined	APRI	F2	0.5 and 1.5			31	22	2	13
Ahmad 2011[72]	APRI_F2_high	APRI	F2	1.5	43	68	31	22	58	46
Ahmad 2011[72]	APRI_F2_low	APRI	F2	0.5			87	55	2	13
Ahmad 2011[72]	AST_ALT_ratio_F4	AST–ALT ratio	F4	1	35	68	9	44	12	92
Ahmad 2011[72]	FI_F2	Fibrosis Index	F2	2.1	100	58	52	0	37	68
Ahmad 2011[72]	FI_F4	Fibrosis Index	F4	3.3	38	100	8	0	13	136
Ahmad 2011[72]	FIB4_F3_combined	FIB-4	F3	1.45 and 3.25			32	18	8	52
Ahmad 2011[72]	FIB4_F3_high	FIB-4	F3	3.25	59	82	32	18	23	84
Ahmad 2011[72]	FIB4_F3_low	FIB-4	F3	1.45			47	50	8	52
Al Mohri 2005[73]	APRI_F2_combined	APRI	F2	0.5, 1.5			17	0	6	6
Al Mohri 2005[73]	APRI_F2_high	APRI	F2	1.5	52	100	17	0	16	13
Al Mohri 2005[73]	APRI_F2_low	APRI	F2	0.5	82	46	27	7	6	6
Al Mohri 2005[73]	FIB4_F2_combined	FIB-4	F2	1.45 and 3.25			11	0	7	11
Al Mohri 2005[73]	FIB4_F2_high	FIB-4	F2	3.25	33	100	11	0	22	13
Al Mohri 2005[73]	FIB4_F2_low	FIB-4	F2	1.45	79	85	26	2	7	11

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Anaparthy 2009[74]	APRI_F2_combined	APRI	F2	0.5 and 1.5			4	2	7	17
Anaparthy 2009[74]	APRI_F2_high	APRI	F2	1.5	20	93	4	2	16	28
Anaparthy 2009[74]	APRI_F2_low	APRI	F2	0.5	65	57	13	13	7	17
Arena 2008[75]	TE_F2	Fibroscan	F2	7.8	83	82	70	12	14	54
Arena 2008[75]	TE_F3	Fibroscan	F3	10.8	91	94	51	6	5	88
Arena 2008[75]	TE_F4	Fibroscan	F4	14.8	94	92	27	10	2	111
Beckebaum 2010[76]	APRI_F3_high	APRI	F3	1.47			13	2	12	23
Beckebaum 2010[76]	FIB4_F3_high	FIB-4	F3	3.54			8	2	17	23
Beckebaum 2010[76]	FinroIndex_F3	Fibroindex	F3	1.85			13	2	12	23
Beckebaum 2010[76]	FT_F3	Fibrotest	F3	0.67			10	1	15	24
Beckebaum 2010[76]	Hepascore_F3	Hepascore	F3	0.83			13	2	12	23
Beckebaum 2010[76]	Lok_F3_high	Lok's index	F3	0.58			10	1	15	24
Beckebaum 2010[76]	TE_F1	TE	F1	4.7	89.10	100.00	43	0	5	2
Beckebaum 2010[76]	TE_F2	Fibroscan	F2	7.1	73.00	100.00	27	0	10	13
Beckebaum 2010[76]	TE_F3	Fibroscan	F3	10.9	75.00	95.80	19	1	6	24
Beckebaum 2010[76]	TE_F4	Fibroscan	F4	17.3	100.00	97.30	4	1	0	45
Bejarano 2009[77]	NIHCED_F3	NIHCED	F3	6	72	75	137	33	53	98
Berg 2004[79]	APRI_F2_high	APRI	F2	1.5	37	93	93	16	160	215
Berg 2004[79]	APRI_F2_low	APRI	F2	0.5	82	53	207	109	46	122
Berg 2004[79]	APRI_F3_combined	APRI	F3				93	16	46	122
Berg 2004[79]	APRI_F3_combined	APRI	F3				50	25	39	286
Berg 2004[79]	APRI_F3_high	APRI	F3	2	39	93	50	25	77	332
Berg 2004[79]	APRI_F3_low	APRI	F3	1	69	80	88	71	39	286

continued

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Berg 2004[79]	APRI_F4_combined	APRI	F4				30	45	13	310
Berg 2004[79]	APRI_F4_high	APRI	F4	2	48	89	30	45	32	377
Berg 2004[79]	APRI_F4_low	APRI	F4	1	76	74	47	112	15	310
Boroni 2002[80]	PLT_F4	Platelet count	F4	134	59	100	19	0	13	188
Bourliere 2006[81]	APRI_F2_combined	APRI	F2	0.5 and 1.5			22	5	30	75
Bourliere 2006[81]	APRI_F2_high	APRI	F2	1.5	22	95	22	5	77	131
Bourliere 2006[81]	APRI_F2_Low	APRI	F2	0.5	70	55	69	61	30	75
Bourliere 2006[81]	APRI_F4_combined	APRI	F4	1 and 2			6	9	5	180
Bourliere 2006[81]	APRI_F4_high	APRI	F4	2	38	96	6	9	10	210
Bourliere 2006[81]	APRI_F4_low	APRI	F4	1	69	82	11	39	5	180
Bourliere 2006[81]	Forns_F2_combined	Forns index	F2	4.2 and 6.9			30	5	20	73
Bourliere 2006[81]	Forns_F2_high	Forns index	F2	6.9	30	96	30	5	69	131
Bourliere 2006[81]	Forns_F2_low	Forns index	F2	4.2	80	54	79	63	20	73
Bourliere 2006[81]	FT_F2_combined	Fibrotest	F2				54	14	3	27
Bourliere 2006[81]	FT_F2_high	Fibrotest	F2	0.6	55	90	54	14	45	122
Bourliere 2006[81]	FT_F2_low	Fibrotest	F2	0.1	97	20	96	109	3	27
Boursier 2009[82]	APRI_F3_low	APRI	F3	0.581	78	75	207	198	59	593
Boursier 2009[82]	Fibrometer_F3	FibroMeter	F3	0.628	84	79	223	166	43	625
Boursier 2009[82]	Fibrometer_F4_combined	FibroMeter	F4				42	19	5	667
Boursier 2009[82]	Fibrometer_F4_high	FibroMeter	F4	0.979	36	98	42	19	76	920
Boursier 2009[82]	Fibrometer_F4_low	FibroMeter	F4	0.628	96	71	113	272	5	667
Boursier 2009[82]	FT_F3_combined	Fibrotest	F3	0.44 and 0.63			178	127	43	562
Boursier 2009[82]	FT_F3_high	Fibrotest	F3	0.631	67	84	178	127	88	664

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Boursier 2009[82]	FT_F3_low	Fibrotest	F3	0.448	84	71	223	229	43	562
Boursier 2009[82]	FT_F4_combined	Fibrotest	F4				50	38	21	723
Boursier 2009[82]	FT_F4_high	Fibrotest	F4	0.862	42	96	50	38	68	901
Boursier 2009[82]	FT_F4_low	Fibrotest	F4	0.66	82	77	97	216	21	723
Boursier 2009[82]	Hepascore_F3_combined	Hepascore	F3	0.497 and 0.904			128	55	48	562
Boursier 2009[82]	Hepascore_F3_high	Hepascore	F3	0.904	48	93	128	55	138	736
Boursier 2009[82]	Hepascore_F3_low	Hepascore	F3	0.497	82	71	218	229	48	562
Boursier 2009[82]	Hepascore_F4_combined						46	9	24	779
Boursier 2009[82]	Hepascore_F4_high	Hepascore	F4	1.159	39	99	46	9	72	930
Boursier 2009[82]	Hepascore_F4_low	Hepascore	F4	0.581	80	83	94	160	24	779
Boursier 2012[83]	Bordeaux_F2	Bordeaux algorithm	F2		88	89	374	33	51	271
Boursier 2012[83]	Bordeaux_F4	Bordeaux algorithm	F4		87	95	95	31	14	589
Boursier 2012[83]	SAFE_F2	SAFE algorithm	F2		100	88	977	97	0	712
Boursier 2012[83]	SAFE_F4	SAFE algorithm	F4		61	93	138	109	89	1450
Burton 2010[84]	APRI_F2_low	APRI	F2	0.5	71	69	23	15	9	33
Cales 2010[85]	Fibrometer_F2	FibroMeter	F2	0.419	80	76	441	121	110	384
Cales 2010[26]	APRI_F2_low	APRI	F2	0.7	62	78	74	14	46	49
Cales 2010[26]	FIB4_F2_low	FIB-4	F2	1.28	72	67	74	14	46	49
Cales 2010[26]	Fibrometer_F2	FibroMeter	F2	0.48	76	72	91	18	29	45
Cales 2010[26]	FT_F2_high	Fibrotest	F2	0.65	61	83	73	11	47	52
Cales 2010[26]	Hepascore_F2		F2	0.31	90	59	108	26	12	37
Calvaruso 2010[86]	TE_F2	Fibroscan	F2	7.1			113	21	76	21
Cardoso 2012[87]	TE_F2	Fibroscan	F2	7.1	68	88	133	20	63	147

continued

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Cardoso 2012[87]	TE_F3	Fibroscan	F3	9.5	67	92	58	22	29	254
Cardoso 2012[87]	TE_F4	Fibroscan	F4	12.5	84	94	26	20	5	312
Carrion 2006[88]	TE_F2	Fibroscan	F2	8.5	90	81	66	18	7	78
Carrion 2006[88]	TE_F4	Fibroscan	F4	12.5	100	87	19	20	0	131
Carvalho 2008[89]	APRI_F2_combined	APRI	F2	0.5, 1.5			23	12	3	21
Carvalho 2008[89]	APRI_F2_high	APRI	F2	1.5	51	82	23	12	22	54
Carvalho 2008[89]	APRI_F2_low	APRI	F2	0.5	94	32	42	45	3	21
Carvalho 2008[89]	APRI_F4_combined	APRI	F4	1, 2			11	9	2	61
Carvalho 2008[89]	APRI_F4_high	APRI	F4	2	55	90	11	9	9	82
Carvalho 2008[89]	APRI_F4_low	APRI	F4	1	90	67	18	30	2	61
Carvalho 2008[89]	FIB4_F2_low	FIB-4	F2	1	91	33	41	44	4	22
Castera 2005[28]	TE_F2	Fibroscan	F2	7.1	67	89	91	5	45	42
Castera 2005[28]	TE_F3	Fibroscan	F3	9.5	73	91	61	9	22	91
Castera 2005[28]	TE_F4	Fibroscan	F4	12.5	87	91	40	12	6	125
Castera 2007[91]	APRI_F2_combined	APRI	F2	0.5, 1.5			44	1	35	28
Castera 2007[91]	APRI_F2_high	APRI	F2	1.5	28	98	44	1	113	42
Castera 2007[91]	APRI_F2_low	APRI	F2	0.5	78	65	122	15	35	28
Castera 2007[91]	APRI_F4_combined	APRI	F4	1, 2			16	14	9	114
Castera 2007[91]	APRI_F4_high	APRI	F4	2	41	91	16	14	23	147
Castera 2007[91]	APRI_F4_low	APRI	F4	1	77	71	30	47	9	114
Castera 2007[91]	FIB4_F2_combined	FIB-4	F2	1.45, 3.25			49	3	42	30
Castera 2007[91]	FIB4_F2_high	FIB-4	F2	3.25	31	93	49	3	108	40
Castera 2007[91]	FIB4_F2_low	FIB-4	F2	1.45	73	70	115	13	42	30

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Castera 2007[91]	Forns_F2_combined	Forns index	F2	4.2, 6.9			36	0	27	15
Castera 2007[91]	Forns_F2_high	Forns index	F2	6.9	35	100	36	0	121	43
Castera 2007[91]	Forns_F2_low	Forns index	F2	4.2	83	35	130	28	27	15
Castera 2007[91]	PLT_F4	Platelet count	F4	130	59	91	23	14	16	147
Castera 2009[90]	APRI_F4_combined	APRI	F4				21	13	25	186
Castera 2009[90]	APRI_F4_high	APRI	F4	2	30	94	21	13	49	215
Castera 2009[90]	APRI_F4_low	APRI	F4	1	64	81	45	42	25	186
Castera 2009[90]	AST_ALT_ratio_F4	AST/ALT ratio	F4	1	31	89	22	25	48	204
Castera 2009[90]	FT_F4	Fibrotest	F4	0.75	55	86	39	31	31	217
Castera 2009[90]	Lok_F4_combined	Lok's index	F4				28	13	10	105
Castera 2009[90]	Lok_F4_high	Lok's index	F4	0.5			28	13	42	225
Castera 2009[90]	Lok_F4_low	Lok's index	F4	0.2	86	46	50	123	10	105
Castera 2009[90]	PLT_F4	Platelet count	F4	150	41	94	29	14	41	214
Castera 2009[90]	TE_F4	Fibroscan	F4	12.6	83	95	55	10	11	212
Ceriani 2001[92]	AST_ALT_ratio_F4	AST/ALT ratio	F4	1	30	98	6	3	14	119
Cheung 2008[94]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			84	7	67	68
Cheung 2008[94]	APRI_F2_high	APRI	F2	1.5	26	96	84	7	239	160
Cheung 2008[94]	APRI_F2_low	APRI	F2	0.5	21	60	100	255	67	68
Cheung 2008[94]	APRI_F3_combined	APRI	F3				69	21	21	113
Cheung 2008[94]	APRI_F3_high	APRI	F3	1.5	37	93	69	21	118	282
Cheung 2008[94]	APRI_F3_low	APRI	F3	0.5	11	63	166	190	21	113
Cheung 2008[94]	AST_ALT_ratio_F2	AST–ALT ratio	F2	1	20	82	65	30	258	137

continued

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Cheung 2008[94]	AST_ALT_ratio_F3	AST–ALT ratio	F3	1	21	82	39	55	148	248
Cheung 2008[94]	Lok_F3_combined	Lok's index	F3				95	52	13	95
Cheung 2008[94]	Lok_F3_high	Lok's index	F3	0.5	51	83	95	52	92	251
Cheung 2008[94]	Lok_F3_low	Lok's index	F3	0.2	93	31	174	209	13	94
Cheung 2008[94]	PLT_F2	Platelet count	F2	150	28	92	90	13	233	154
Cheung 2008[94]	PLT_F3	Platelet count	F3	150	39	90	73	30	114	273
Cheung 2008[94]	Pohl_F2	Pohl score	F2	Positive	7	98	23	3	300	164
Cheung 2008[94]	Pohl_F3	Pohl score	F3	Positive	9	98	17	6	170	297
Cho 2011[95]	TE_F2	Fibroscan	F2	7.4	88	90	49	3	7	27
Cho 2011[95]	TE_F3	Fibroscan	F3	9.7	90	87	28	7	3	48
Cho 2011[95]	TE_F4	Fibroscan	F4	14.7	100	89	6	9	0	71
Christensen 2006[96]	FibrospectII_F3	Fibrospect II	F3	0.5	85.2	72.7	44	25	8	65
Chrysanthos 2006[97]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			44	16	31	64
Chrysanthos 2006[97]	APRI_F2_high	APRI	F2	1.5	30	88	44	16	102	122
Chrysanthos 2006[97]	APRI_F2_low	APRI	F2	0.5	79	46	115	74	31	64
Chrysanthos 2006[97]	APRI_F4_combined	APRI	F4				22	20	23	162
Chrysanthos 2006[97]	APRI_F4_high	APRI	F4	2	27	81	22	20	36	206
Chrysanthos 2006[97]	APRI_F4_low	APRI	F4	1	60	72	35	64	23	162
Cobbold 2010[98]	APRI_F2_low	APRI	F2	0.66	83	78	31	7	6	23
Cobbold 2010[98]	APRI_F4_low	APRI	F4	0.92	86	77	12	12	2	41
Cobbold 2010[98]	ELF_F2	ELF	F2	8.75	84	70	31	9	6	21
Cobbold 2010[98]	ELF_F4	ELF	F4	9.4	93	79	13	11	1	42
Cobbold 2010[98]	TE_F2	Fibroscan	F2	8	90	65	33	11	4	20

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Cobbold 2010 [98]	TE_F4	Fibroscan	F4	10	79	87	11	7	3	46
Colletta 2005[99]	FT_F2	Fibrotest		0.48	21	81	3	5	11	21
Colletta 2005[99]	TE_F2	Fibroscan	F2	8.74	100	100	14	0	0	26
Corradi 2009[100]	APRI_F2_high	APRI	F2	1.5	59	74	8	6	5	17
Corradi 2009[100]	Forns_F2_low	Forns index		4.2	100	8	13	21	0	2
Corradi 2009[100]	FT_F2_high	Fibrotest	F2	0.6	67	30	9	16	4	7
Corradi 2009[100]	TE_F2	Fibroscan	F2	8.7	71	61	11	15	5	23
Crespo 2010[101]	APRI_F3_high	APRI	F3	1.5	40	61	4	16	7	24
Cross 2010[102]	Kings_F1	King's Score	F1	7.6	80	69	145	2	36	4
Cross 2010[102]	Kings_F2	King's Score	F2	9.87	84	70	75	29	14	69
Cross 2010[102]	Kings_F3	King's Score	F4	24.3	74	90	29	15	10	133
Cross 2010[102]	TE_F1	Fibroscan	F1	6.75	68	91	123	1	58	5
Cross 2010[102]	TE_F2	Fibroscan	F2	8.85	74	88	66	12	23	86
Cross 2010[102]	TE_F4	Fibroscan	F4	10.05	93	88	36	18	3	130
da Silva 2008[103]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			13	0	2	11
da Silva 2008[103]	APRI_F2_high	APRI	F2	1.5	46.4	100	13	0	15	22
da Silva 2008[103]	APRI_F2_low	APRI	F2	0.5	92.9	50	26	11	2	11
da Silva 2008[103]	APRI_F4_combined	APRI	F4	1 and 2			6	1	1	27
da Silva 2008[103]	APRI_F4_high	APRI	F4	≥ 2.0	46.2	97.3	6	1	7	36
da Silva 2008[103]	APRI_F4_low	APRI	F4	≤ 1.0	92.3	73	12	10	1	27
Danila 2011[104]	ARFI_F1	ARFI	F1	1.04	89.8	93.3	153	1	17	14
Danila 2011[104]	ARFI_F2	ARFI	F2	1.21	82.5	93.9	125	2	27	31
Danila 2011[104]	ARFI_F3	ARFI	F3	1.49	81.8	90.4	73	9	16	87

continued

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Danila 2011[104]	ARFI_F4	ARFI	F4	1.82	82.8	87.1	28	19	6	132
Degos 2010[106]	TE_F2	Fibroscan	F2	5.2	90	32	506	239	56	112
Degos 2010[106]	TE_F4	Fibroscan	F4	12.9	72	89	91	87	35	700
DeLedinghen 2006[105]	TE_F2	Fibroscan	F2	4.5	93	18	41	23	3	5
DeLedinghen 2006[105]	TE_F4	Fibroscan	F4	11.8	100	93	17	4	0	51
Dinesen 2008[107]	13CBT_F3	^{13}Cmethacetin breath test	F3	0.021	75.4	79.5	43	8	14	31
Dinesen 2008[107]	13CBT_F4	^{13}Cmethacetin breath test	F4	0.0146	92.5	84.1	25	11	2	58
Dinesen 2008[107]	APRI_F3_low	APRI	F3	0.75	64.9	84.6	37	6	20	33
Dinesen 2008[107]	APRI_F4_low	APRI	F4	1	66.7	75.4	18	17	9	52
Dinesen 2008[107]	AST_ALT_ratio_F3	AST–ALT ratio	F3	0.85	70.2	48.7	40	20	17	19
Dinesen 2008[107]	AST_ALT_ratio_F4	AST–ALT ratio	F4	1	63	59.4	17	28	10	41
Dinesen 2008[107]	Fibroindex_F3	Fibroindex	F3	1.35	66.7	84.6	38	6	19	33
Dinesen 2008[107]	Fibroindex_F4	Fibroindex	F4	1.82	70.4	91.3	19	6	8	63
Esmat 2007[108]	HA_F3	Hyaluronic acid	F3	20	88	68	45	48	6	101
Esmat 2007[108]	YKL_40_F3	YKL-40	F3	100	82	57	42	64	9	85
Fabris 2006[109]	APRI_F2_low	APRI	F2	0.4	55	83	6	5	5	24
Fahmy 2011[110]	TE_F2	Fibroscan	F2	7 kPa	87	86	58	6	9	37
Fahmy 2011[110]	TE_F4	Fibroscan	F4	16.5 kPa	87	91	19	8	3	80
Fontaine 2009[111]	FIB4_F3_low	FIB-4	F3		53	65	13	30	12	55
Fontaine 2009[111]	FT_F2	Fibrotest	F2	None given	56	63	26	23	21	40
Fontaine 2009[111]	FT_F3	Fibrotest	F3		42	84	11	14	15	71

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Fontaine 2009[111]	FT_F4	Fibrotest	F4		25	94	2	6	6	96
Fontana 2008[112]	Fontana_F4	Fontana	F4	0.2–0.3	79	66	152	108	41	211
Fontanges 2008[113]	FT_F2	Fibrotest	F2	0.385	74	65	37	16	13	30
Fontanges 2008[113]	FT_F3	Fibrotest	F3	0.455	86	68	24	3	4	5
Forestier 2010[114]	TE_F4	Fibroscan	F4	12.6 kPa	85	93	25	4	4	54
Forns 2002 training[115]	Forns_F2_combined	Forns index	F2				37	11	5	120
Forns 2002 training[115]	Forns_F2_high	Forns index	F2	6.9	44	96	37	11	48	255
Forns 2002 training[115]	Forns_F2_low	Forns index	F2	4.2	94	45	80	146	5	120
Forns 2002 validation[115]	Forns_F2_combined	Forns index	F2				10	5	2	47
Forns 2002 validation[115]	Forns_F2_high	Forns index	F2	6.9	30	95	10	5	23	87
Forns 2002 validation[115]	Forns_F2_low	Forns index	F2	4.2	94	51	31	45	2	47
Fraquelli 2011[116]	TE_F2	Fibroscan	F2	8.8 kPa	81	77	197	49	47	160
Fraquelli 2011[116]	TE_F4	Fibroscan	F4	14.6 kPa	100	88	44	49	0	360
Fuji 2009[117]	AST_ALT_ratio_F4	AST–ALT ratio	F4	1	65	56	11	37	6	46
Fuji 2009[117]	APRI_F4	APRI	F4		82	70	14	25	3	58
Fuji 2009[117]	CDS_F4	CDS	F4		88	67	15	27	2	56
Gaia 2011[119]	TE_F1	Fibroscan	F1	4.5 kPa	90	33	67	2	7	1
Gaia 2011[119]	TE_F2	Fibroscan	F2	7.5 kPa	74	79	29	8	10	30
Gaia 2011[119]	TE_F3	Fibroscan	F3	10.1 kPa	77	90	13	6	4	54
Gaia 2011[119]	TE_F4	Fibroscan	F4	11.5 kPa	69	93	9	4	4	60
Ganne-Carrie 2006[120]	TE_F4	Fibroscan	F4	10.4	88	85	26	40	4	228
Gara 2011[121]	APRI_F4_low	APRI	F4	1.0	79	78	12	23	3	81
Gara 2011[121]	TE_F4	Fibroscan	F4	13.1 kPa	100	89	15	11	0	93

continued

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Giannini 2006[122]	AST_ALT_ratio_F2	AST–ALT ratio	F2	0.66	73.7	65	129	82	46	152
Giannini 2006[122]	PLT_F2	Platelet count	F2	163,000	61.7	80.8	108	45	67	189
Gobel 2006[123]	APRI_F2_high	APRI	F2	1.5	75	87	33	5	11	34
Guechot 2010[124]	Hepascore_F2	Hepascore	F2	0.5	77	70	190	80	57	186
Guechot 2010[124]	Hepascore_F3	Hepascore	F3	0.6	80	70	124	107	31	250
Guechot 2010[124]	Hepascore_F4	Hepascore	F4	0.84	84	73	64	118	12	318
Guechot 2010[124]	HA_F3	Hyaluronic acid	F3	85 μg/l	60	74	66	59	44	167
Guechot 2010[124]	HA_F4	Hyaluronic acid	F4	110 μg/l	79.2	89.4	42	29	11	244
Guechot 2010[124]	PIIINP_F3	PIIINP	F3	0.8 kU/l	70	63.4	77	83	33	143
Guechot 2010[124]	PIIINP_F4	PIIINP	F4	1.0 kU/l	64.5	91.2	34	24	19	249
Guzelbulut 2011[126]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			36	6	13	30
Guzelbulut 2011[126]	APRI_F2_high	APRI	F2	1.5	43.37	91.04	36	6	47	61
Guzelbulut 2011[126]	APRI_F2_low	APRI	F2	0.5	84.34	44.78	70	37	13	30
Guzelbulut 2011[126]	APRI_F4_combined	APRI	F4				22	5	14	80
Guzelbulut 2011[126]	APRI_F4_high	APRI	F4	2	43.14	94.95	22	5	29	94
Guzelbulut 2011[126]	APRI_F4_low	APRI	F4	1	72.55	80.81	37	19	14	80
Guzelbulut 2011[126]	FIB4_F2_combined	FIB-4	F2				76	47	0	7
Guzelbulut 2011[126]	FIB4_F4_combined	FIB-4	F4				28	8	5	57
Guzelbulut 2011[126]	Forns_F2_combined	Forns index	F2				39	4	5	23
Guzelbulut 2011[126]	Forns_F4_combined	Forns index	F4				34	9	1	27
Guzelbulut 2011[126]	FIB4_F2_high	FIB-4	F2	1	91.57	29.85	76	47	7	20
Guzelbulut 2011[126]	FIB4_F2_low	FIB-4	F2	0.6	100	10.45	83	60	0	7
Guzelbulut 2011[126]	FIB4_F4_high	FIB-4	F4	3.25	54.9	91.92	28	8	23	91

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Guzelbulut 2011[126]	FIB4_F4_low	FIB-4	F4	1.45	90.2	57.58	46	42	5	57
Guzelbulut 2011[126]	Forns_F2_high	Forns index	F2	6.9	46.99	94.03	39	4	44	63
Guzelbulut 2011[126]	Forns_F2_low	Forns index	F2	4.2	93.98	34.33	78	44	5	23
Guzelbulut 2011[126]	Forns_F4_high	Forns index	F4	6.9	66.67	90.91	34	9	17	90
Guzelbulut 2011[126]	Forns_F4_low	Forns index	F4	4.2	98.04	27.27	50	72	1	27
Halfon 2006[128]	FT_F2	Fibrotest	F2	0.36	73	72	168	77	62	197
Halfon 2005 validation[129]	HA_F1	Hyaluronic acid	F1	16	91	36	42	12	76	142
Halfon 2005 validation[129]	HA_F2_high	Hyaluronic acid	F2	121	14	99	17	1	101	135
Halfon 2005 validation[129]	HA_F2_low	Hyaluronic acid	F2	25	78	53	32	43	28	151
Halfon 2005 validation[129]	HA_F3_high	Hyaluronic acid	F3	160	22	100	13	0	47	194
Halfon 2005 validation[129]	HA_F3_low	Hyaluronic acid	F3	50	100	79	10	0	3	241
Halfon 2005 validation[129]	HA_F4	Hyaluronic acid	F4	237	31	99	4	2	9	239
Halfon 2007[127]	APRI_F2_low	APRI	F2	0.39	77	66	112	71	34	139
Halfon 2007[127]	APRI_F3_low	APRI	F3	0.58	75	76	38	73	13	232
Halfon 2007[127]	APRI_F4_low	APRI	F4	0.83	100	83	13	58	0	285
Halfon 2007[127]	Fibrometer_F2	FibroMeter	F2	0.57	64	81	93	40	53	170
Halfon 2007[127]	Fibrometer_F3	FibroMeter	F3	0.667	82	76	42	73	9	232
Halfon 2007[127]	Fibrometer_F4	FibroMeter	F4	0.88	92	87	12	45	1	298
Halfon 2007[127]	FT_F2	Fibrotest	F2	0.44	67	80	98	42	48	168
Halfon 2007[127]	FT_F3	Fibrotest	F3	0.45	84	69	43	95	8	210
Halfon 2007[127]	FT_F4	Fibrotest	F4	0.56	85	74	11	89	2	254
Halfon 2007[127]	Hepascore_F2	Hepascore	F2	0.32	77	63	112	78	34	132
Halfon 2007[127]	Hepascore_F3	Hepascore	F3	0.53	78	72	40	85	11	220

continued

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Halfon 2007[127]	Hepascore_F4	Hepascore	F4	0.61	92	72	12	96	1	247
Harada 2008[130]	APRI_F2_low	APRI	F2	0.84	73	91	15	3	6	32
Harada 2008[130]	TE_F1	Fibroscan	F1	8.8 kPa	68	100	23	0	11	22
Harada 2008[130]	TE_F2	Fibroscan	F2	9.9 kPa	90	91	19	3	2	32
Harada 2008[130]	TE_F3	Fibroscan	F3	15.4 kPa	75	95	9	2	3	42
Harada 2008[130]	TE_F4	Fibroscan	F4	26.5 kPa	100	98	5	1	0	50
Harada 2008[130]	HA_F2	Hyaluronic acid	F2	103 ng/ml	38	83	8	6	13	29
Harada 2008[130]	PLT_F2	Platelet count	F2	48 U/l	38	89	8	4	13	31
Harada 2008[130]	typeIVcoll_F2	Type IV collagen	F2	298 ng/ml	52	83	11	6	10	29
Hsieh 2012[131]	AST_ALT_ratio_F2	AST–ALT ratio	F2	0.6	67.9	70.7	133	12	63	29
Hsieh 2012[131]	AST_ALT_ratio_F3	AST–ALT ratio	F3	0.6	77.5	53.2	86	59	25	67
Hsieh 2012[131]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			111	17	6	4
Hsieh 2012[131]	APRI_F2_high	APRI	F2	1.5	56.6	58.5	111	17	85	24
Hsieh 2012[131]	APRI_F2_low	APRI	F2	0.5	96.9	9.7	190	37	6	4
Hsieh 2012[131]	APRI_F3_combined	APRI	F3				77	51	1	9
Hsieh 2012[131]	APRI_F3_high	APRI	F3	1.5	69.4	59.5	77	51	34	75
Hsieh 2012[131]	APRI_F3_low	APRI	F3	0.5	99.1	7.1	110	117	1	9
Hsieh 2012[131]	CDS_F2	CDS	F2	6	66.3	48.8	130	21	66	20
Hsieh 2012[131]	CDS_F3	CDS	F3	6	73.9	45.2	82	69	29	57
Hsieh 2012[131]	FIB4_F2_combined	FIB-4	F2				93	4	36	22
Hsieh 2012[131]	FIB4_F2_high	FIB-4	F2	3.25	47.4	90.2	93	4	103	37
Hsieh 2012[131]	FIB4_F2_low	FIB-4	F2	1.45	81.6	53.7	160	19	36	22
Hsieh 2012[131]	FIB4_F3_combined	FIB-4	F3				68	29	14	44

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Hsieh 2012[131]	FIB4_F3_high	FIB-4	F3	3.25	61.3	77	68	29	43	97
Hsieh 2012[131]	FIB4_F3_low	FIB-4	F3	1.45	87.4	34.9	97	82	14	44
Hsieh 2012[131]	FibroQ_F2	FibroQ	F2	1.6	77.6	65.9	152	14	44	27
Hsieh 2012[131]	FibroQ_F3	FibroQ	F3	1.6	85.66	43.7	95	71	16	55
Hsieh 2012[131]	Lok_F2	Lok's index	F2	0.2	81.6	58.8	160	17	36	24
Hsieh 2012[131]	Lok_F3_low	Lok's index	F3	0.2	88.3	37.3	98	79	13	47
Hsieh 2012[131]	Pohl_F2	Pohl score	F2	3.8	4.59	100	9	0	187	41
Hsieh 2012[131]	Pohl_F3		F3	3.8	7.2	99.2	8	1	103	125
Iacobellis 2005[132]	PLT_F2	Platelet count	F2	140.000/µl	51	90	330	50	318	446
Iacobellis 2005[132]	PLT_F4	Platelet count	F4		82	87	67	138	15	923
Iacobellis 2005[132]	PLTspleen_F2	Platelet count–spleen diameter ratio	F2		33	92	214	40	434	455
Iacobellis 2005[132]	PLTspleen_F4	Platelet count–spleen diameter ratio	F4		85	82	70	191	12	870
Imbert-Bismut 2001[23]	FT_F2		F2	40	78	69	47	23	13	51
Imbert-Bismut 2001[23]	FT_F2_high	Fibrotest	F2	70	62	95	37	4	23	70
Imbert-Bismut 2001[23]	FT_F2_low	Fibrotest	F2	20	92	46	55	40	5	34
Imperiale 2000[133]	AST_ALT_ratio_F4	AST–ALT ratio	F4	1	52	91	15	12	14	116
Islam 2005[134]	APRI_F4_low	APRI	F4	1	78	75	16	40	5	119
Islam 2005[134]	GUCI_F4	GUCI	F4	1	80	78	17	35	4	123
Islam 2005[134]	PLT_F4	Platelet count	F4	190	80	77	17	36	4	122
Iushchuk 2005[135]	HA_F4	Hyaluronic acid	F4	100 ng/ml	100	84.6	28	18	0	102
Jazia 2009[78]	APRI_F2	APRI	F2	0.72	93	58	25	3	2	5
Kalantari 2011[136]	Hepascore_F2	Hepascore	F2	0.34	67	56	29	16	14	21

continued

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Kalantari 2011[136]	Hepascore_F3	Hepascore	F3	0.61	82	86	19	8	4	49
Kalantari 2011[136]	Hepascore_F4	Hepascore	F4	0.84	100	97	16	2	0	62
Kamphues 2010[137]	APRI_F2_low	APRI	F2	0.4845	70	63	45	11	19	19
Kamphues 2010[137]	FIB4_F2_high	FIB-4	F2	2.8	44	87	28	4	36	26
Kamphues 2010[137]	FIB4_F4_high	FIB-4	F4	4.44	44	84	4	14	5	71
Kamphues 2010[137]	TE_F2	Fibroscan	F2	8.5 kPa	72	83	46	5	18	25
Kamphues 2010[137]	TE_F4	Fibroscan	F4	10.5 kPa	100	65	9	30	0	55
Kandemir 2009[138]	GUCI__F3	GUCI	F3	0.261	58.33	72.73	14	12	10	32
Khan 2008[140]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			26	3	11	32
Khan 2008[140]	APRI_F2_high	APRI	F2	1.5	41	95	26	3	38	53
Khan 2008[140]	APRI_F2_low	APRI	F2	0.5	83	66	53	24	11	32
Khan 2008[140]	APRI_F3_combined	APRI	F3	0.9 and 1.75			17	5	4	63
Khan 2008[140]	APRI_F3_high	APRI	F3	1.75	56	94	17	5	13	85
Khan 2008[140]	APRI_F3_low	APRI	F3	0.9	90	70	26	27	4	63
Kim 2011[141]	TE_F2	Fibroscan	F2	6.2	76	97.5	38	1	12	40
Kim 2011[141]	TE_F3	Fibroscan	F3	7.7	100	95.7	21	3	0	67
Kim 2011[141]	TE_F4	Fibroscan	F4	11	77.8	93.5	7	5	2	77
Koda 2007 training[24]	APRI_F2_low	APRI	F2	0.36	26.5	95.1	117	86	6	31
Koda 2007 training[24]	APRI_F2_combined	APRI	F2	< 0.36 and > 0.85			19	5	1	19
Koda 2007 training[24]	APRI_F2_high	APRI	F2	0.85	31.6	91.7	19	5	41	55
Koda 2007 training[24]	APRI_F2_low	APRI	F2	0.36	31.6	98.3	59	41	1	19
Koda 2007 training[24]	Fibroindex_F2_combined	Fibroindex	F2	1.25 and 2.25			18	2	2	24
Koda 2007 training[24]	Fibroindex_F2_high	Fibroindex	F2	2.25	30	96.7	18	2	42	58

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Koda 2007 training[24]	Fibroindex_F2_low	Fibroindex	F2	1.25	40	96.3	58	36	2	24
Koda 2007 training[24]	Forns_F2_combined	Forns index	F2	4.5 and 8.7			13	1	4	15
Koda 2007 training[24]	Forns_F2_high	Forns index	F2	8.7	21.7	98.3	13	1	47	59
Koda 2007 training[24]	Forns_F2_low	Forns index	F2	4.5	25.6	93.3	56	45	4	15
Koda 2007 training[24]	APRI_F2_combined	APRI	F2	0.36 and 0.85			42	5	6	31
Koda 2007 training[24]	APRI_F2_high	APRI	F2	0.85	34.1	95.7	42	5	81	112
Koda 2007 training[24]	Fibroindex_F2_combined	Fibroindex	F2	1.25 and 2.25			44	3	7	47
Koda 2007 training[24]	Fibroindex_F2_high	Fibroindex	F2	2.25	35.8	97.4	44	3	79	114
Koda 2007 training[24]	Fibroindex_F2_low	Fibroindex	F2	1.25	40.2	94.3	116	70	7	47
Koda 2007 training[24]	Forns_F2_combined	Forns index	F2	4.5 and 8.7			30	4	6	31
Koda 2007 training[24]	Forns_F2_high	Forns index	F2	8.7	24.3	96.6	30	4	93	113
Koda 2007 training[24]	Forns_F2_low	Forns index	F2	4.5	25.6	97.6	120	87	3	30
Lackner 2005[143]	APRI_F2_high	APRI	F2	1.5	44	96	45	60	6	69
Lackner 2005[143]	APRI_F2_low	APRI	F2	0.5	88	44	48	101	3	28
Lackner 2005[143]	APRI_F4_high	APRI	F4	2	55	93	38	21	13	108
Lackner 2005[143]	APRI_F4_low	APRI	F5–F6	1	93	70	44	33	7	96
Lackner 2005[143]	AST_ALT_ratio_F4	AST–ALT ratio	F5–F6	1	36	90	7	18	12	158
Lackner 2005[143]	CDS_F3	CDS	F3	8	10	100	5	0	45	144
Lackner 2005[143]	PLT_F2	Platelet count	Ishak F3	150	42	97	41	3	56	94
Lackner 2005[143]	PLT_F4	Platelet count	F4	150	77	88	25	19	7	143
Lackner 2005[143]	Pohl_F3	Pohl score	F3	Positive	18	98	9	3	41	141
Ladero 2010[144]	APRI_F2_low	APRI	F2	0.5	54.7	80.6	102	47	84	196
Ladero 2010[144]	FIB4_F2_low	FIB-4	F2	1.35	78.6	58.1	168	171	18	72

continued

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Ladero 2010[144]	Forns_F2_low	Forns index	F2	4.2	50.6	75.8	141	120	45	123
Ladero 2010[144]	GUCI_F2	GUCI	F2	0.3	29.6	90.3	168	171	18	72
Ladero 2010[144]	Kings_F2_high	King's Score	F2	4.46	22.8	93.5	108	52	78	191
Ladero 2010[144]	Kings_F2_low	King's Score	F2	12.3	81.5	62.4	116	45	70	198
Lee 2011[145]	FibrospectII_F1	FibroSpect II	F1	42	87.5	70	6	8	1	18
Leroy 2004[147]	PIIINP/MMP-1 index_F2	PIIINP/MMP-1 index	F2	0.3	65	85	55	17	29	94
Leroy 2004[147]	PIIINP/MMP-1 index_F3	PIIINP/MMP-1 index	F3	0.3	85	74	31	41	5	117
Leroy 2007[146]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			66	11	8	24
Leroy 2007[146]	APRI_F2_high	APRI	F2	1.5	72.3	87.8	66	11	25	78
Leroy 2007[146]	APRI_F2_low	APRI	F2	0.5	91.6	26.8	83	65	8	24
Leroy 2007[146]	APRI_F3_combined	APRI	F3	1 and 2			38	21	6	69
Leroy 2007[146]	APRI_F3_high	APRI	F3	2	73.9	84	38	21	13	108
Leroy 2007[146]	APRI_F3_low	APRI	F3	1	89.1	53.8	45	60	6	69
Leroy 2007[146]	Forns_F2_combined	Forns index	F2	4.2 and 6.9			38	6	11	38
Leroy 2007[146]	Forns_F2_high	Forns index	F2	6.9	41.9	92.9	38	6	53	83
Leroy 2007[146]	Forns_F2_low	Forns index	F2	4.2	88.4	42.4	80	51	11	38
Leroy 2007[146]	Forns_F3_combined	Forns index	F3	4.2 and 6.9			28	17	4	44
Leroy 2007[146]	Forns_F3_high	Forns index	F3	6.9	54.2	87	28	17	23	112
Leroy 2007[146]	Forns_F3_Low	Forns index	F3	4.2	91.7	34.1	47	85	4	44
Leroy 2007[146]	FT_F2	Fibrotest	F2	0.32	75.8	74.2	69	23	22	66
Leroy 2007[146]	FT_F2_high	Fibrotest	F2	0.59	45.1	89.9	41	9	50	80
Leroy 2007[146]	FT_F2_low	Fibrotest	F2	0.22	89	52.8	81	42	10	47
Leroy 2007[146]	FT_F3	Fibrotest	F3	0.32	90.2	64.3	46	46	5	83

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Leroy 2007[146]	FT_F3_high	Fibrotest	F3	0.59	66.7	87.6	34	16	17	113
Leroy 2007[146]	FT_F3_low	Fibrotest	F3	0.22	94.1	41.9	48	75	3	54
Leroy 2007[146]	Hepascore_F2	Hepascore	F2	0.5	53.8	83.9	49	14	42	75
Leroy 2007[146]	Hepascore_F2_combined	Hepascore	F2	0.5 and 0.84			30	7	42	75
Leroy 2007[146]	Hepascore_F2_high	Hepascore	F2	0.84	33	92	30	7	61	82
Leroy 2007[146]	Hepascore_F3_combined	Hepascore	F3	0.5 and 0.84			24	13	12	105
Leroy 2007[146]	Hepascore_F3_high	Hepascore	F3	0.84	47.1	89.8	24	13	27	116
Leroy 2007[146]	Hepascore_F3_low	Hepascore	F3	0.5	76.5	81.1	39	24	12	105
Leroy 2007[146]	MP3_F2	MP3	F2	0.3	82.4	72.7	75	24	16	65
Leroy 2007[146]	MP3_F3	MP3	F3	0.3	92.2	59.4	47	52	4	77
Leroy 2011[148]	TE_F2	Fibroscan	F2	7.6	80	80	142	45	45	184
Lieber 2006[150]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			21	9	12	22
Lieber 2006[150]	APRI_F2_high	APRI	F2	1.5			21	9	38	65
Lieber 2006[150]	APRI_F2_low		F2	0.5			47	52	12	22
Liu 2006[151]	APRI_F2_high	APRI	F2	1.5	0	100	0	0	21	58
Liu 2006[151]	APRI_F2_low	APRI	F2	0.5	28.6	94	6	3	15	55
Liu 2006[151]	AST_ALT_ratio_F2	AST–ALT ratio	F2	1	45.3	62.1	10	22	11	36
Liu 2011[153]	APRI_F2_high	APRI	F2	1.5	3	100	3	0	98	183
Liu 2011[153]	APRI_F2_low	APRI	F2	0.5	79	70	80	55	21	128
Liu 2011[153]	APRI_F3	APRI	F3	0.75	93	90	37	24	3	220
Liu 2011[153]	APRI_F4_high	APRI	F4	2	0	99	0	3	14	267
Liu 2011[153]	APRI_F4_low	APRI	F4	1	43	92	6	22	8	248
Liu 2011[153]	TE_F2	Fibroscan	F2	5.3	93	88	94	22	7	161

continued

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Liu 2011[153]	TE_F3	Fibroscan	F3	8.3	95	99	38	2	2	242
Liu 2011[153]	TE_F4	Fibroscan	F4	9.2	100	96	14	11	0	259
Loko 2008[154]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			57	2	19	21
Loko 2008[154]	APRI_F2_high	APRI	F2	1.5	36.1	95.4	57	2	100	41
Loko 2008[154]	APRI_F2_low	APRI	F2	0.5	87.9	48.8	138	22	19	21
Loko 2008[154]	APRI_F4_combined	APRI	F4	1 and 2			19	25	6	100
Loko 2008[154]	APRI_F4_high	APRI	F4	2	47.5	84.4	19	25	21	135
Loko 2008[154]	APRI_F4_low	APRI	F4	1	85	62.5	34	6	6	100
Loko 2008[154]	FIB4_F2_high	FIB-4	F2	1	83.4	53.5	131	20	26	23
Loko 2008[154]	FIB4_F2_low	FIB-4	F2	0.6	98.1	20.9	154	34	3	9
Loko 2008[154]	FIB4_F3_combined	FIB-4	F3	1.45 and 3.25			22	9	19	90
Loko 2008[154]	FIB4_F3_high	FIB-4	F3	3.25	31	93	22	9	49	120
Loko 2008[154]	FIB4_F3_low	FIB-4	F3	1.45	73.2	69.8	52	39	19	90
Loko 2008[154]	FIB4_F4_combined	FIB-4	F4	1.45 and 3.25			19	25	7	102
Loko 2008[154]	FIB4_F4_high	FIB-4	F4	3.25	40	90.6	19	25	21	135
Loko 2008[154]	FIB4_F4_low	FIB-4	F4	1.45	82.5	63.7	33	58	7	102
Loko 2008[154]	Forns_F2_combined	Forns index	F2	4.2 and 6.9			29	0	20	9
Loko 2008[154]	Forns_F2_high	Forns index	F2	6.9	23	100	29	0	97	26
Loko 2008[154]	Forns_F2_low	Forns index	F2	4.2	84.1	34.6	106	17	20	9
Loko 2008[154]	PLT_F4	Platelet count	F4	150	67.5	77.5	27	26	13	91
Lupsor 2008[155]	TE_F1	Fibroscan	F1	4.9	87.5	92.3	272	1	39	12
Lupsor 2008[155]	TE_F2	Fibroscan	F2	7.4	75.9	83.6	159	19	51	95
Lupsor 2008[155]	TE_F3	Fibroscan	F3	9.1	86.8	83.9	92	35	14	183

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Lupsor 2008[155]	TE_F4	Fibroscan	F4	11.85	86.9	90.7	60	24	9	231
Lupsor 2009[29]	ARFI_F1	ARFI	F1	1.19	62	86	61	2	37	12
Lupsor 2009[29]	ARFI_F2	ARFI	F2	1.34	68	93	46	3	22	41
Lupsor 2009[29]	ARFI_F3	ARFI	F3	1.61	79	95	40	3	11	59
Lupsor 2009[29]	ARFI_F4	ARFI	F4	2	80	95	34	4	8	67
Lupsor 2009[29]	TE_F1	Fibroscan	F1	5.2	85	93	83	1	15	13
Lupsor 2009[29]	TE_F2	Fibroscan	F2	8.1	85	95	58	2	10	42
Lupsor 2009[29]	TE_F3	Fibroscan	F3	9.6	96	87	48	8	2	54
Lupsor 2009[29]	TE_F4	Fibroscan	F4	13.1	95	89	40	8	2	62
Macias 2006[156]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			78	10	12	36
Macias 2006[156]	APRI_F2_high	APRI	F2	1.5	51	91	78	10	75	100
Macias 2006[156]	APRI_F2_low	APRI	F2	0.5	92	33	141	74	12	36
Macias 2006[156]	APRI_F4_combined	APRI	F4				21	25	9	127
Macias 2006[156]	APRI_F4_high	APRI	F4	2	53	89	21	25	19	198
Macias 2006[156]	APRI_F4_low	APRI	F4	1	78	57	31	96	9	127
Macias 2006[156]	AST_ALT_RATIO_F4	AST–ALT ratio	F4	1	38	77	15	51	25	172
Macias 2006[156]	Bonacini_F4	BONACINI MODEL	F4	7	43	83	17	38	23	185
Macias 2006[156]	Forns_F2_combined	Forns index	F2	4.2 and 6.9			66	4	34	42
Macias 2006[156]	Forns_F2_high	Forns index	F2	6.9	43	96	66	4	87	106
Macias 2006[156]	Forns_F2_low	Forns index	F2	4.2	78	38	119	68	34	42
Macias 2006[156]	PLT_F4	Platelet count	F4	150	63	37	25	140	15	83
Macias 2011[157]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			75	20	57	111
Macias 2011[157]	APRI_F2_high	APRI	F2	1.5	22	92	75	20	189	235

continued

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Macias 2011[157]	APRI_F2_low		F2	0.5	78	44	207	144	57	111
Macias 2011[157]	Forns_F2_combined	Forns index	F2				84	31	37	66
Macias 2011[157]	Forns_F2_high	Forns index	F2	6.9	32	88	84	31	180	224
Macias 2011[157]	Forns_F2_low	Forns index	F2	4.2	86	26	227	189	37	66
Martinez 2011[158]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			108	8	21	57
Martinez 2011[158]	APRI_F2_high	APRI	F2	1.5	47	93	108	8	121	103
Martinez 2011[158]	APRI_F2_low	APRI	F2	0.5	91	51	208	54	21	57
Martinez 2011[158]	APRI_F4_combined	APRI	F4	1 and 2			61	19	22	160
Martinez 2011[158]	APRI_F4_high	APRI	F4	2	49	91	61	19	63	197
Martinez 2011[158]	APRI_F4_low	APRI	F4	1	82	74	102	56	22	160
Martinez 2011[158]	ELF_F2_combined	ELF	F2	–0.45 and 1.07			108	11	23	58
Martinez 2011[158]	ELF_F2_high	ELF	F2	1.07	47	90	108	11	121	100
Martinez 2011[158]	ELF_F2_low	ELF	F2	–0.45	90	52	206	53	23	58
Martinez 2011[158]	ELF_F4_combined	ELF	F4	0.06 and 1.73			64	22	12	114
Martinez 2011[158]	ELF_F4_high	ELF	F4	1.73	52	90	64	22	60	194
Martinez 2011[158]	ELF_F4_low	ELF	F4	0.06	90	53	112	102	12	114
Martinez 2011[158]	FIB4_F3_combined	FIB-4	F3	1.45 and 3.25			84	17	12	118
Martinez 2011[158]	FIB4_F3_high	FIB-4	F3	3.25	54	91	84	17	71	168
Martinez 2011[158]	FIB4_F3_low	FIB-4	F3	1.45	92	64	143	67	12	118
Martinez 2011[158]	Forns_F2_combined	Forns index	F2				101	8	25	64
Martinez 2011[158]	Forns_F2_high	Forns index	F2	6.9	44	93	101	8	128	103
Martinez 2011[158]	Forns_F2_low	Forns index	F2	4.2	89	58	204	47	25	64
Morikawa 2011[159]	TE_F2	Fibroscan	F2	10.1	88.6	86.5	42	7	5	47

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Morikawa 2011[159]	TE_F3	Fibroscan	F3	13.3	89.7	86.6	29	9	3	60
Morikawa 2011[159]	TE_F4	Fibroscan	F4	16.3	85.7	85.4	14	12	2	73
Murawaki 2001[160]	PLT_F2	Platelet count	F2	160	68	71	60	22	28	55
Murawaki 2001[160]	PLT_F3	Platelet count	F3	140	68	74	33	10	15	27
Murawaki 2001[160]	typeIVcoll_F2	Type IV collagen	F2	110	77	73	68	21	20	56
Murawaki 2001[160]	typeIVcoll_F3	Type IV collagen	F3	130	66	75	32	9	16	28
Myers 2002[366]	FT_F2_combined	Fibrotest	F2				42	11	10	67
Myers 2002[366]	FT_F2_high	Fibrotest	F2	0.6	50	91	42	11	42	116
Myers 2002[366]	FT_F2_low	Fibrotest	F2	0.2	88	53	74	60	10	67
Nitta 2009[161]	TE_F2	Fibroscan	F2	7.1	81	80	80	13	19	53
Nitta 2009[161]	TE_F3	Fibroscan	F3	9.6	88	82	50	19	7	89
Nitta 2009[161]	TE_F4	Fibroscan	F4	11.6	92	78	22	31	2	110
Nunes 2005[163]	APRI_F4_low	APRI	F4	0.75	74	68	13	12	5	27
Nunes 2005[163]	Forns_F4_low	Forns index	F4	3.9	61	62	11	15	7	24
Nunes 2005[163]	HA_F4	Hyaluronic acid	F4	107	61	73	11	11	7	28
Nunes 2005[163]	PIIINP_F4	PIIINP	F4	0.8	79	69	14	12	4	27
Nunes 2005[163]	PLT_F4	Platelet count	F4	196	63	66	11	13	7	26
Obara 2008[164]	APRI_F2_low	APRI	F2	0.70	89	87	25	3	3	20
Obara 2008[164]	TE_F1	Fibroscan	F1	5.6	93	78	7	9	1	34
Obara 2008[164]	TE_F2	Fibroscan	F2	9.5	89	83	25	4	3	19
Obara 2008[164]	TE_F3	Fibroscan	F3	10.3	94	69	15	11	1	24
Obara 2008[164]	TE_F4	Fibroscan	F4	17.2	80	88	8	5	2	36
Obara 2008[164]	HA_F2	Hyaluronic acid	F2	96.6	82	78	23	5	5	18

continued

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Obara 2008[164]	PLT_F2	Platelet count	F2	122.0	64	96	18	1	10	22
Obara 2008[164]	typeIVcoll_F2	Type IV collagen	F2	128.0	89	74	25	6	3	17
Oliveira 2005[165]	AST_ALT_ratio_F3	AST–ALT ratio	F3	1	27	78	6	13	16	48
Oliveira 2005[165]	HA_F3	Hyaluronic acid	F3		77	77	17	14	5	47
Paggi 2008[167]	APRI_F3_combined	APRI	F3	1 and 2			58	22	34	189
Paggi 2008[167]	APRI_F3_high	APRI	F3	2	36	92	58	22	102	248
Paggi 2008[167]	APRI_F3_low	APRI	F3	1	70	79	137	81	34	189
Parise 2006[168]	APRI_F2_low	APRI	F2	0.7	85	66	73	41	13	79
Parise 2006[168]	AST_ALT_ratio_F2	AST–ALT ratio	F2	0.8	52	61	45	47	41	73
Parise 2006[168]	AST_ALT_ratio_F4	AST–ALT ratio	F4	1	36	82	16	29	28	133
Parise 2006[168]	HA_F2	Hyaluronic acid	F2	34.2	85	71	73	35	13	85
Parise 2006[168]	HA_F4	Hyaluronic acid	F4	78.6	91	81	40	31	4	131
Park 2000[169]	AST_ALT_ratio_F4	AST–ALT ratio	F4	1	47	96	14	5	16	118
Parkes 2011[27]	ELF_F3_combined	ELF	F3				78	35	17	149
Parkes 2011[27]	ELF_F3_high	ELF	F3	10.22	70	85	78	35	33	201
Parkes 2011[27]	ELF_F3_low	ELF	F3	9.59	85	63	94	87	17	149
Patel 2009[25]	FibrospectII_F2	FibroSpect II	F2		71	65	138	20	56	38
Patel 2009[25]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5	90	98	9	1	1	46
Patel 2009[25]	FIB4_F3_combined	FIB-4	F3	1.45 and 3.25	75	95	6	3	2	54
Patel 2009[25]	FibrospectII_F2	FibroSpect II	F2		95	66	21	25	1	48
Patel 2009[25]	Forns_F2_combined	Forns index	F2	4.21 and 6.9	84	88	11	5	2	38
Patel 2009[25]	FT_F2	Fibrotest	F2		100	61	18	26	0	40
Patel 2011[170]	TE_F2	Fibroscan	F2	10.1	77	88	33	21	10	150

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Patel 2011[170]	TE_F4	Fibroscan	F4	11.7	94	88	17	24	1	172
Pohl 2001[171]	AST_ALT_ratio_F3	AST–ALT ratio	F3	1	47	82	17	21	19	96
Pohl 2001[171]	Pohl_F3	Pohl score	F3		41	99	16	1	22	116
Poynard 2012[172]	FT_F2	Fibrotest	F2	0.48	66	85	520	75	268	426
Poynard 2012[172]	FT_F4	Fibrotest	F4	0.74	68	89	135	120	64	970
Poynard 2012[172]	TE_F2	Fibroscan	F2	8.8	48	93	378	35	410	466
Poynard 2012[172]	TE_F4	Fibroscan	F4	14.5	65	95	129	55	70	1036
Prati 2011[173]	TE_F4	Fibroscan	F4	12	54	94	6	1	5	16
Qui 2004[174]	typeIVcpll_F4	Type IV collagen	F4	190	77	72	45	23	13	59
Reedy 1998[175]	AST_ALT_ratio_F4	AST–ALT ratio	F4	1	44	94	10	3	13	51
Rossi 2003[177]	FT_F2_combined	Fibrotest	F2				20	5	4	22
Rossi 2003[177]	FT_F2_high	Fibrotest	F2	0.6	42	94	20	5	28	72
Rossi 2003[177]	FT_F2_low	Fibrotest	F2	0.1	92	29	44	55	4	22
Rossini 2009[178]	APRI_F3_low	APRI	F3	0.6			14	3	4	19
Rossini 2009[178]	ARFI_F3	ARFI	F3	2.11			15	2	3	20
Rossini 2009[178]	ARFI_F4	ARFI	F4	2.33			9	7	1	23
Said 2010[179]	FT_F2	Fibrotest	F2	0.5	85.1	72.2	40	5	7	13
Said 2010[179]	FT_F3	Fibrotest	F3	0.52	92.5	53.8	24	18	2	21
Said 2010[179]	FT_F4	Fibrotest	F4	0.75	85.7	70.7	6	17	1	41
Saitu 2005[180]	HA_F2	Hyaluronic acid	F2	75.7 ng/ml	75	81	58	6	19	26
Saitu 2005[180]	HA_F4	Hyaluronic acid	F4	183.5 ng/ml	80	80	24	16	6	63
Saitu 2005[180]	PIIINP_F2	PIIINP	F2	0.835 U/ml	78	75	60	8	17	24
Saitu 2005[180]	PIIINP_F4	PIIINP	F4	0.995 U/ml	77	66	23	27	7	52

continued

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Saitu 2005[180]	typeIVcoll_F2	Type IV collagen	F2	5.75 ng/ml	65	69	50	10	27	22
Saitu 2005[180]	typeIVcoll_F4	Type IV collagen	F4	6.55 ng/ml	60	61	18	31	12	48
Sanvisens 2009[181]	HA_F3	Hyaluronic acid	F3	48 µg/l	87	70	20	14	3	32
Schiavon 2007[182]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			21	11	6	84
Schiavon 2007[182]	APRI_F2_high	APRI	F2	0.95	44	93	21	11	27	144
Schiavon 2007[182]	APRI_F2_low	APRI	F2	0.4	88	54	42	71	6	84
Schiavon 2007[182]	APRI_F3_combined	APRI	F3	0.55 and 1.0			8	20	1	127
Schiavon 2007[182]	APRI_F3_high	APRI	F3	1	42	89	8	20	11	164
Schiavon 2007[182]	APRI_F3_low	APRI	F3	0.55	95	69	18	57	1	127
Schiavon 2008[183]	HA_F2_combined	Hyaluronic acid	F2				16	22	9	57
Schiavon 2008[183]	HA_F2_high	Hyaluronic acid	F2		36	84	16	22	29	118
Schiavon 2008[183]	HA_F2_low	Hyaluronic acid	F2		80	41	36	83	9	57
Schiavon 2008[183]	YKL_40_F2_combined	YKL-40	F2				15	28	9	46
Schiavon 2008[183]	YKL_40_F2_high	YKL-40	F2		33	80	15	28	30	112
Schiavon 2008[183]	YKL-40_F2_low	YKL-40	F2		80	33	36	94	9	46
Schneider 2006[185]	13CBT_F4	^{13}Cmethacetin breath test	F4	13C50 1.7	83	63	16	24	3	40
Schneider 2006[185]	APRI_F2_low	APRI	F2	0.75	81	65	38	13	9	23
Schneider 2006[185]	APRI_F4_low	APRI	F4	1	77	63	15	24	4	40
Sebastiani 2012[31]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			160	23	171	337
Sebastiani 2012[31]	APRI_F2_high	APRI	F2	1.5	29	95	160	23	392	438
Sebastiani 2012[31]	APRI_F2_low	APRI	F2	0.5	69	73	381	124	171	337
Sebastiani 2012[31]	APRI_F4_combined	APRI	F4	1 and 2			46	63	29	711

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Sebastiani 2012[31]	APRI_F4_high	APRI	F4	2	41	93	46	63	67	837
Sebastiani 2012[31]	APRI_F4_low	APRI	F4	1	74	79	84	189	29	711
Sebastiani 2012[31]	Fibropaca_F2	Fibropaca algorithm	F2		86	90	238	25	41	225
Sebastiani 2012[31]	Fibropaca_F4	Fibropaca algorithm	F4		73	97	56	21	21	672
Sebastiani 2012[31]	Forns_F2_combined	Forns index	F2	4.2 and 6.9			337	157	33	92
Sebastiani 2012[31]	Forns_F2_high	Forns index	F2	6.9	61	66	337	157	215	304
Sebastiani 2012[31]	Forns_F2_low	Forns index	F2	4.2	94	20	519	369	33	92
Sebastiani 2012[31]	FT_F2	Fibrotest	F2	0.49	62	81	342	88	210	373
Sebastiani 2012[31]	FT_F4	Fibrotest	F4	0.75	30	89	34	99	79	801
Sebastiani 2012[31]	Leroy_F2	Leroy algorithm	F2		90	98	45	5	5	245
Sebastiani 2012[31]	SAFE_F2	SAFE algorithm	F2		100	78	236	45	0	159
Sebastiani 2012[31]	SAFE_F4	SAFE algorithm	F4		82	92	75	57	16	655
Sebastiani 2009[187]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			255	40	306	810
Sebastiani 2009[187]	APRI_F2_high	APRI	F2	1.5	27	96	255	40	676	1064
Sebastiani 2009[187]	APRI_F2_low	APRI	F2	0.5	67	73	625	294	306	810
Sebastiani 2009[187]	APRI_F4_combined	APRI	F4				90	101	42	1542
Sebastiani 2009[187]	APRI_F4_high	APRI	F4	2	47	94	90	101	101	1743
Sebastiani 2009[187]	APRI_F4_low	APRI	F4	1	78	84	149	302	42	1542
Sebastiani 2009[187]	SAFE_F2	SAFE algorithm	F2		100	77	517	99	0	330
Sebastiani 2009[187]	SAFE_F4	SAFE algorithm	F4		90	93	129	100	14	1415
Sebastiani 2006[188]	FT_F4	Fibrotest	F4		50	93	15	11	15	150
Sebastiani 2008 EALT[186]	AST_ALT_ratio_F2	AST–ALT ratio	F2	1	43	58	49	21	66	28
Sebastiani 2008 EALT[186]	Fibroindex_F2_combined	Fibroindex	F2	1.25 and 2.25			22	0	37	34

continued

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Sebastiani 2008 EALT[186]	Fibroindex_F2_high	Fibroindex	F2	2.25	19	100	22	0	93	49
Sebastiani 2008 EALT[186]	Fibroindex_F2_low	Fibroindex	F2	1.25	68	69	78	15	37	34
Sebastiani 2008 EALT[186]	Forns_F2_combined	Forns index	F2	4.2 and 6.9			24	1	17	24
Sebastiani 2008 EALT[186]	Forns_F2_high	Forns index	F2	6.9	21	98	24	1	91	48
Sebastiani 2008 EALT[186]	Forns_F2_low	Forns index	F2	4.2	85	49	98	25	17	24
Sebastiani 2008 EALT[186]	FT_F2	Fibrotest	F2	0.49	82	72	94	14	21	35
Sebastiani 2008 NALT[186]	AST_ALT_ratio_F2	AST–ALT ratio	F2	1	12	88	4	6	28	42
Sebastiani 2008 NALT[186]	Fibroindex_F2_combined	Fibroindex	F2	1.25 and 2.25			3	0	19	37
Sebastiani 2008 NALT[186]	Fibroindex_F2_high	Fibroindex	F2	2.25	10	100	3	0	29	48
Sebastiani 2008 NALT[186]	Fibroindex_F2_low	Fibroindex	F2	1.25	41	77	13	11	19	37
Sebastiani 2008 NALT[186]	Forns_F2_combined	Forns index	F2	4.2 and 6.9			16	0	14	33
Sebastiani 2008 NALT[186]	Forns_F2_high	Forns index	F2	6.9	50	100	16	0	16	48
Sebastiani 2008 NALT[186]	Forns_F2_low	Forns index	F2	4.2	57	67	18	15	14	33
Sebastiani 2008 NALT[186]	FT_F2	Fibrotest	F2	0.49	67	85	21	7	11	41
Sene 2006[189]	APRI_F2_low	APRI	F2	0.38	72	75	38	18	27	55
Sene 2006[189]	FT_F2	Fibrotest	F2	0.37	85	75	55	18	10	55
Sene 2006[189]	HA_F2	Hyaluronic acid	F2	45	67	64	44	26	21	47
Sene 2006[189]	PLT_F2	Platelet count	F2	182	57	79	37	15	28	58
Sharabash 2009[190]	FibrospectII_F2	Fibrospect II	F2	72	83	86	5	2	1	12
Shastry 2007[191]	APRI_F3_low	APRI	F3	0.6			12	19	0	17
Sheth 1997[192]	AST_ALT_ratio_F4	AST–ALT ratio	F4	1	53	100	25	0	22	92
Singal 2011[193]	APRI_F3_combined	APRI	F3				12	5	1	37
Singal 2011[193]	APRI_F3_combined	APRI	F3				7	6	3	36

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Singal 2011[193]	APRI_F3_high	APRI	F3	1.5	48	94	12	5	13	76
Singal 2011[193]	APRI_F3_low	APRI	F3	0.5	96	46	24	44	1	37
Singal 2011[193]	APRI_F3_high	APRI	F3	1.5	29	93	7	6	17	75
Singal 2011[193]	APRI_F3_low	APRI	F3	0.5	88	44	21	45	3	36
Sirli 2010[194]	APRI_F2_low	APRI	F2	0.52	70	81	94	3	40	13
Sirli 2010[194]	FIB4_F2	FIB-4	F2	2.14	36	100	48	0	86	16
Sirli 2010[194]	FIB4_F4	FIB-4	F4	2.31	80	78	12	30	3	105
Sirli 2010[194]	Forns_F2_low	Forns index	F2	4.57	72	68	96	5	38	11
Sirli 2010[194]	Forns_F4	Forns index	F4	5.93	100	74	15	35	0	100
Sirli 2010[194]	Lok_F2	Lok's index	F2	0.17	58	81	78	3	56	13
Sirli 2010[194]	Lok_F4_low	Lok's index	F4	0.26	87	82	13	24	2	111
Sirli 2010[194]	PLT_F2	Platelet count	F2	176	37	100	50	0	84	16
Sirli 2010[194]	PLT_F4	Platelet count	F4	155	87	84	13	22	2	113
Sirli 2010[194]	TE_F2	Fibroscan	F2	6.8	61	73	82	4	52	12
Sirli 2010[194]	TE_F4	Fibroscan	F4	13.3	93	96	14	5	1	130
Snyder 2007[196]	FibrospectII_F2	Fibrospect II	F2	55	82	77	41	10	9	33
Snyder 2007[196]	FibrospectII_F2_combined	Fibrospect II	F2	25 and 85			26	0	0	18
Snyder 2006 prosp[195]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			34	4	10	45
Snyder 2006 prosp[195]	APRI_F2_high	APRI	F2	1.5	44	94	34	4	44	68
Snyder 2006 prosp[195]	APRI_F2_low	APRI	F2	0.5	87	62	68	27	10	45
Snyder 2006 prosp[195]	APRI_F3_combined	APRI	F3				53	137	4	117
Snyder 2006 prosp[195]	APRI_F3_high	APRI	F3	1.2	83	73	41	28	8	74
Snyder 2006 prosp[195]	APRI_F3_low	APRI	F3	0.5	96	48	47	53	2	49

continued

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Snyder 2006 prosp[195]	APRI_F4_high	APRI	F4	2	50	94	13	7	13	118
Snyder 2006 prosp[195]	AST_ALT_ratio_F4	AST–ALT ratio	F4	1	88	41	23	74	3	51
Snyder 2006 retro[195]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			54	7	28	94
Snyder 2006 retro[195]	APRI_F2_high	APRI	F2	1.5	31	96	54	7	120	169
Snyder 2006 retro[195]	APRI_F2_low	APRI	F2	0.5	83	54	148	80	28	94
Snyder 2006 retro[195]	APRI_F3_combined	APRI	F3				41	28	2	49
Snyder 2006 retro[195]	APRI_F3_high	APRI	F3	1.2	81	50	53	137	13	137
Snyder 2006 retro[195]	APRI_F3_low	APRI	F3	0.5	94	43	62	156	4	117
Sohn 2010[197]	PLTspleen_F2	Platelet count–spleen diameter ratio	F2	2200	81.5	71.4	28	4	6	10
Sporea 2008[199]	TE_F2	Fibroscan	F2	6.8	59.6	93.3	97	2	64	28
Sporea 2010[200]	TE_F2	Fibroscan	F2	6.8	60	88	167	5	111	34
Sporea 2010[200]	TE_F3	Fibroscan	F3	8.6	62	81	82	35	50	150
Sporea 2010[200]	TE_F4	Fibroscan	F4	13.3	77	93	30	19	9	259
Sporea 2011[201]	TE_F2	Fibroscan	F2	6.7	77.5	86.7	117	6	35	39
Sporea 2011[201]	TE_F4	Fibroscan	F4	12.2	96.2	89.6	96	10	4	87
Sporea 2011[198]	ARFI_F1	ARFI	F1	1.19	68.5	83.3	433	9	204	45
Sporea 2011[198]	ARFI_F2	ARFI	F2	1.29	79.7	87.5	367	28	92	204
Sporea 2011[198]	ARFI_F3	ARFI	F3	1.57	90.2	85.3	280	57	31	323
Sporea 2011[198]	ARFI_F4	ARFI	F4	1.59	83.7	80	144	104	28	415
Sterling 2006 Total cohort[202]	FIB4_F2_combined	FIB-4	F2				366	126	43	71
Sterling 2006 Total cohort[202]	FIB4_F2_high	FIB-4	F2	1	69.4	58.4	366	126	161	177
Sterling 2006 Total cohort[202]	FIB4_F2_low	FIB-4	F2	0.6	91.8	23.4	484	232	43	71

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Sterling 2006 Total cohort[202]	FIB4_F3_combined	FIB-4	F3				40	22	58	467
Sterling 2006 Total cohort[202]	FIB4_F3_high	FIB-4	F3	3.25	23	96.6	40	22	134	634
Sterling 2006 Total cohort[202]	FIB4_F3_low	FIB-4	F3	1.45	66.7	71.2	116	189	58	467
Sterling 2006 Validation[202]	FIB4_F2_high	FIB-4	F2	1	64.5	57.1	110	45	62	60
Sterling 2006 Validation[202]	FIB4_F2_low	FIB-4	F2	0.6	89.5	23.8	153	80	19	25
Sterling 2006 Validation[202]	FIB4_F3_high	FIB-4	F3	3.25	21.7	96.8	13	6	48	210
Sterling 2006 Validation[202]	FIB4_F3_low	FIB-4	F3	1.45	70	73.7	43	56	18	160
Stibbe 2011[203]	FIB4_F3_combined	FIB-4	F3				5	0	5	16
Stibbe 2011[203]	FIB4_F3_high	FIB-4	F3	3.25	28	100	5	0	13	23
Stibbe 2011[203]	FIB4_F3_low	FIB-4	F3	1.45	72	70	13	7	5	16
Sud 2009[204]	FPI_F2_high	FPI	F2	0.8	43	94	36	6	48	86
Sud 2009[204]	FPI_F2_low	FPI	F2	0.2	96	44	81	52	3	40
Sud 2009[204]	FPI_F2_high	FPI	F2	0.8	42	98	31	1	43	51
Sud 2009[204]	FPI_F2_low	FPI	F2	0.2	85	48	63	27	11	25
Testa 2006[205]	Aminobreathtest_F2	Aminobreath test	Ishak F2	8.1	73	73.7	27	10	10	28
Testa 2006[205]	PLTspleen_F2	Platelet count–spleen diameter ratio	Ishak F2	1750	78.4	78.9	29	8	8	30
Thompson 2009[208]	FT_F2	Fibrotest	F2		65	63	70	102	37	175
Thompson 2009[208]	FT_F3	Fibrotest	F3		67	38	57	185	28	114
Thompson 2009[208]	FT_F3	Fibrotest	F3		35	84	13	56	24	291
Thompson 2009[208]	Hepascore_F2	Hepascore	F2		73		78	86	29	191
Thompson 2009[208]	Hepascore_F3	Hepascore	F3				74	105	11	194
Thompson 2009[208]	Hepascore_F4	Hepascore	F4				22	56	15	291

continued

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Thompson 2009[208]	FT_F2	Fibrotest	F2		62	60	29	35	18	52
Thompson 2009[208]	Hepascore_F2	Hepascore	F2		71	62	33	33	14	54
Thompson 2010[207]	SAFE_F2	SAFE algorithm	F2		100	80	422	360	0	1439
Thompson 2010[207]	SAFE_F4	SAFE algorithm	F4		52	92	64	168	60	1929
Toniutto 2007[209]	APRI_F2_high	APRI	F2	1.4	76	77	25	16	8	53
Trang 2008[210]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			18	4	3	9
Trang 2008[210]	APRI_F2_high	APRI	F2	1.5	36	87.1	18	4	32	27
Trang 2008[210]	APRI_F2_low	APRI	F2	0.42	94	29	47	22	3	9
Trang 2008[210]	APRI_F3_combined	APRI	F3				8	5	3	24
Trang 2008[210]	APRI_F3_high	APRI	F3	1.85	28	90.3	8	5	20	48
Trang 2008[210]	APRI_F3_low	APRI	F3	0.71	89.3	45.3	25	29	3	24
Trang 2008[210]	FIB4_F2_combined	FIB-4	F2				25	6	8	18
Trang 2008[210]	FIB4_F2_high	FIB-4	F2	2.05	50	81.3	25	6	25	25
Trang 2008[210]	FIB4_F2_low	FIB-4	F2	1.39	84	58.1	42	13	8	18
Trang 2008[210]	FIB4_F3_combined	FIB-4	F3				9	5	1	25
Trang 2008[210]	FIB4_F3_high	FIB-4	F3	3.25	33.3	90.7	9	5	19	48
Trang 2008[210]	FIB4_F3_low	FIB-4	F3	1.45	96.4	47.2	27	28	1	25
Trifan 2009[211]	Lok_F2	Lok's index	F2	–1.67	80.7	32.2	121	109	57	25
Trifan 2009[211]	Lok_F2	Lok's index	F2	–0.63	55	66.8	29	85	24	174
Trifan 2009[211]	APRI_F2	APRI	F2	0.75	61	71.3	52	82	126	52
Trifan 2009[211]	APRI_F3	APRI	F3	0.76	72	59	38	106	15	153
Trifan 2009[211]	FIB4_F2	FIB-4	F2	1.66	52.5	75	45	70	134	64
Trifan 2009[211]	FIB4_F3	FIB-4	F3	1.36	83	53	44	122	9	137

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Trifan 2009[211]	FT_F2	Fibrotest	F2	0.53	49.1	81.6	32	66	146	68
Trifan 2009[211]	FT_F3	Fibrotest	F3	0.45	86	56.3	46	114	7	145
Trifan 2009[211]	Forns_F2	Forns index	F2	5.35	61.2	70.5	53	82	125	52
Trifan 2009[211]	Forns_F3	Forns index	F3	6.38	56	80.3	30	52	23	207
Trifan 2009[211]	TE_F2	Fibroscan	F2	8.65	61	74.3	46	82	132	52
Trifan 2009[211]	TE_F3	Fibroscan	F3	12.5	94	84.2	50	41	3	218
Trocme 2006[212]	FI_F2	Fibrosis Index	F2		82	70	43	22	9	50
Trocme 2006[212]	FI_F3	Fibrosis Index	F3		82	80	21	11	5	42
Tural 2009[213]	APRI_F3_combined	APRI	F3	0.5, 1.5			47	35	6	64
Tural 2009[213]	APRI_F3_high	APRI	F3	1.5			47	35	47	195
Tural 2009[213]	APRI_F3_low	APRI	F3	0.5			88	166	6	64
Tural 2009[213]	FIB4_F3_combined	FIB-4	F3	1.45, 3.25			22	8	20	156
Tural 2009[213]	FIB4_F3_high		F3	3.25			22	8	72	222
Tural 2009[213]	FIB4_F3_low			1.45			74	74	20	156
Tural 2009[213]	Forns_F3_combined	Forns index	F3	4.2, 6.9			37	22	10	70
Tural 2009[213]	Forns_F3_high	Forns index	F3	6.9			37	22	57	208
Tural 2009[213]	Forns_F3_low	Forns index	F3	4.2			84	160	10	70
Vallet-Pichard 2007[214]	FIB4_F3_combined	FIB-4	F3				55	14	38	561
Vallet-Pichard 2007[214]	FIB4_F3_high	FIB-4	F3	3.25	37.6	98.2	55	14	91	687
Vallet-Pichard 2007[214]	FIB4_F3_low	FIB-4	F3	1.45	74.3	80.1	108	140	38	561
Valva 2011[215]	HA_F2	Hyaluronic acid	F2	103.1	66.7	90	8	1	4	9
Valva 2011[215]	HA_F2	Hyaluronic acid	F2	109.7	100	82.3	5	3	0	14
Valva 2011[215]	PIIINP_F2	PIIINP	F2	9.1	75	80	9	2	3	8

continued

TABLE 67 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HCV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Valva 2011[215]	PIIINP_F3	PIIINP	F3	9.1	100	64.7	5	6	0	11
Varaut 2005[216]	FT_F2_high	Fibrotest	F2	0.6	24	92	12	5	38	55
Varaut 2005[216]	FT_F2_low	Fibrotest	F2	0.2	84	45	42	33	8	27
Wai 2003 training[218]	APRI_F2_combined	APRI	F2	< 0.5 and > 1.5			37	5	8	47
Wai 2003 training[218]	APRI_F2_high	APRI	Ishak F3	1.5	41	95	37	5	54	96
Wai 2003 training[218]	APRI_F2_low	APRI	Ishak F3	0.5	91	47	83	54	8	47
Wai 2003 training[218]	APRI_F4_combined	APRI	Ishak F5				16	11	3	123
Wai 2003 training[218]	APRI_F4_high	APRI	Ishak F5	2	57	93	16	11	12	153
Wai 2003 training[218]	APRI_F4_low	APRI	Ishak F5	1	89	75	25	41	3	123
Wai 2003 validation[218]	APRI_F2_combined	APRI	Ishak F3	< 0.5 and > 1.5			35	4	4	35
Wai 2003 validation[218]	APRI_F2_high	APRI	Ishak F3	1.5	71	86	35	4	14	25
Wai 2003 validation[218]	APRI_F2_low	APRI	Ishak F3	0.5	86	71	25	14	4	35
Wai 2003 validation[218]	APRI_F4_combined	APRI	Ishak F5				8	5	0	65
Wai 2003 validation[218]	APRI_F4_high	APRI	Ishak F5	2	73	92	8	5	3	62
Wai 2003 validation[218]	APRI_F4_low	APRI	Ishak F5	1	100	88	4	9	0	65
Westin 2008[219]	GUCI_F2	GUCI	Ishak F3	0.33	68	72	36	49	17	128
Westin 2008[219]	GUCI_F2	GUCI	Ishak F3	1.11	59	80	60	25	42	103
Westin 2008[219]	GUCI_F4	GUCI	Ishak F5	0.33	100	87	6	29	0	195
Westin 2008[219]	GUCI_F4	GUCI	Ishak F5	1.11	44	89	14	21	18	177
Wilson 2006[220]	APRI_F2_combined	APRI	Ishak F3	< 0.5 and > 1.5			2	7	3	68
Wilson 2006[220]	APRI_F2_high	APRI	Ishak F3	1.5	18	94	2	7	9	109
Wilson 2006[220]	APRI_F2_low	APRI	Ishak F3	0.5	73	59	8	48	3	68
Wilson 2006[220]	FT_F2_combined	Fibrotest	Ishak F3				6	41	1	57

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Wilson 2006[220]	FT_F2_high	Fibrotest	Ishak F3	0.48	56	65	6	41	5	75
Wilson 2006[220]	FT_F2_low	Fibrotest	Ishak F3	0.31	89	49	9	59	1	57
Wong 1998[221]	HA_F4	Hyaluronic acid	F4		86	88	18	13	3	96
Zaman 2004[222]	FibrospectII_F2	Fibrospect II	F2		71.8	73.9	28	18	11	51
Zarski 2012[223]	APRI_F2_low	APRI	F2	0.5	33.1	96.6	59	6	119	198
Zarski 2012[223]	APRI_F4_high	APRI	F4	2	77.1	99.7	93	0	28	261
Zarski 2012[223]	Fibrometer_F2	FibroMeter	F2	0.411	87.6	56.4	157	90	21	114
Zarski 2012[223]	Fibrometer_F4	FibroMeter	F4	0.88	69.6	88.7	85	29	36	232
Zarski 2012[223]	FT_F2_high	Fibrotest	F2	0.48	75.8	66.2	135	69	43	135
Zarski 2012[223]	FT_F4	Fibrotest	F4	0.74	71.4	81	86	50	35	211
Zarski 2012[223]	Hepascore_F2	Hepascore	F2	0.5	74.7	72.5	134	57	45	147
Zarski 2012[223]	Hepascore_F4	Hepascore	F4	0.84	76.8	81.3	93	50	28	211
Zarski 2012[223]	TE_F2	Fibroscan	F2	5.2	96.6	34.8	173	133	5	71
Zarski 2012[223]	TE_F4	Fibroscan	F4	12.9	76.8	89.6	93	26	28	235
Ziol 2005[224]	TE_F2	Fibroscan	F2	8.8	56	91	91	8	72	80
Ziol 2005[224]	TE_F3	Fibroscan	F3	9.6	86	85	65	26	11	149
Ziol 2005[224]	TE_F4	Fibroscan	F4	14.6	86	96	42	8	7	194

CDS, Cirrhosis Discriminant Score; EALT, elevated ALT; FN, false negative; FP, false positive; FPI, Fibrosis Probability Index; MMP-1, matrix metalloproteinase-1; MP3, metalloproteinase-3; NALT, normal ALT; NIHCED, non-invasive hepatitis C-related early detection; PLT, platelet; prosp, prospective; retro, retrospective; sens., sensitivity; spec., specificity; TP, true positive; TN, true negative.

TABLE 68 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HBV

Study ID	Test	Index test assessed	Fibrosis stage assessed	Cut-off	Sens.	Spec.	TP	FP	FN	TN
Castera 2011[225]	APRI_F2_combined	APRI	F2		62	100	6	0	17	10
Castera 2011[225]	APRI_F2_high	APRI	F2	> 1.5	14	100	6	0	38	16
Castera 2011[225]	APRI_F2_low	APRI	F2	< 0.5	62	64	27	6	17	10
Castera 2011[225]	APRI_F4_combined	APRI	F4		47	96	2	2	8	36
Castera 2011[225]	APRI_F4_high	APRI	F4	> 2	13	96	2	2	13	43
Castera 2011[225]	APRI_F4_low	APRI	F4	< 1	47	80	7	9	8	36
Castera 2011[225]	Fibrotest_F2	Fibrotest	F2	0.48	62	81	27	3	17	13
Castera 2011[225]	Fibrotest_F4	Fibrotest	F4	0.74	47	91	7	4	8	41
Castera 2011[225]	TE_F2	Fibroscan	F2	7.1	68	63	30	6	14	10
Castera 2011[225]	TE_F4	Fibroscan	F3	9.6	87	80	13	9	2	36
Chan 2009[226]	TE_F1	Fibroscan	F1		72	80	6	30	3	122
Chan 2009[226]	TE_F3	Fibroscan	F3		84	76	32	30	6	93
Chan 2009[226]	TE_F4	Fibroscan	F3		60	93	24	8	16	113
Chen 2012[227]	TE_F4	Fibroscan	F4	10.4	93	71	69	70	5	171
Fraquelli 2011[116]	TE_F2	Fibroscan	F2		77	77	42	11	12	38
Fraquelli 2011[116]	TE_F4	Fibroscan	F4		1	83	16	14	0	73
Fung 2011[230]	TE_F2	Fibroscan	F2		83	63	9	10	2	17
Fung 2011[230]	TE_F4	Fibroscan	F4		100	68	0	12	0	26
Gaia 2011[119]	TE_F3	Fibroscan	F3		64	84	22	16	4	28
Gaia 2011[119]	TE_F4	Fibroscan	F4		48	87	19	25	3	23
Gaia 2011[119]	TE_F1	Fibroscan	F1		61	72	27	13	10	20
Ganne-Carrie 2006[120]	TE_F4	Fibroscan	F4	10.4	82	85	13	16	3	90
Gui 2010[231]	Fibrotest_F1	Fibrotest	F1		70	63	31	16	12	28

Study ID	Test	Index test assessed	Fibrosis stage assessed	Cut-off	Sens.	Spec.	TP	FP	FN	TN
Gui 2010[231]	Fibrotest_F3	Fibrotest	F3		72	67	62	29	24	58
Hongbo 2007[233]	APRI_F2_low	APRI	F2	0.4	72	74.56	68	58	26	170
Hongbo 2007[233]	APRI_F1_low	APRI	F1	0.4	70	82.93	60	6	26	30
Hui 2005[235]	Hui_F3	Hui index	F3		93	49	86	28	6	27
Hui 2005 – validation[235]	Hui_F3	Hui index	F3		75	53	30	16	10	19
Kim 2009[238]	TE_F4	Fibroscan	F4		64	75	25	13	14	39
Kim 2009[239]	API_F4	Age-Platelet Index	F4	4.5	76.1	71.4	51	18	16	45
Kim 2010[236]	TE_F4	Fibroscan	F4	10.1	76.1	81	51	12	16	51
Kim 2007[237]	AAR_F4	AST–ALT ratio	F4	1	52	72	41	75	38	192
Kim 2007[237]	API_F4	Age-Platelet Index	F4	4	88.6	74.1	70	69	9	198
Kim 2007[237]	APRI_F4_combined	APRI	F4				16	20	34	210
Kim 2007[237]	APRI_F4_high	APRI	F4	2	20.3	92.5	16	20	63	247
Kim 2007[237]	APRI_F4_low	APRI	F4	1	57	78.6	45	57	34	210
Kim 2010[236]	FIB4_F3_combined	FIB-4	F3				127	7	29	260
Kim 2010[236]	FIB4_F3_high	FIB-4	F3	> 2.65	38.5	97.9	127	7	203	331
Kim 2010[236]	FIB4_F3_low	FIB-4	F3	< 1	91.2	72.8	301	78	29	260
Kim 2010[236]	FIB4_F4_combined	FIB-4	F4				69	7	27	368
Kim 2010[236]	FIB4_F4_high	FIB-4	F4	3.6	30	98.4	69	7	161	431
Kim 2010[236]	FIB4_F4_low	FIB-4	F4	1.6	88.2	84	203	70	27	368
Kwok 2009[240]	TE_F3	Fibroscan	F3		80	88.6	5	3	1	26
Lee 2011[241]	CDS_F1	CDS	F1	4	28	90.2	17	6	43	55
Lee 2011[241]	Lok's model_F1	Lok's index	F1	0.87	48.3	90.2	29	6	31	55
Lee 2011[241]	TE_F2	Fibroscan	F2		44.9	100	66	0	81	61

continued

TABLE 68 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HBV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Cut-off	Sens.	Spec.	TP	FP	FN	TN
Lee 2011[241]	TE_F4	Fibroscan	F4		100	58.7	24	76	0	108
Lesmana 2011[242]	TE_F2	Fibroscan	F2		60.3	63.6	44	16	29	28
Lesmana 2011[242]	TE_F3	Fibroscan	F3		65.5	80.7	18	17	10	72
Li 2012[243]	HA_F2	Hyaluronic acid	F2	185.3	84	83	48	5	9	25
Liu 2011[244]	AAR_F2	AST–ALT ratio	F2	0.67	57.2	58.7	123	169	92	239
Liu 2011[244]	API_F2	Age–Platelet Index	F2	3	67.9	62	146	155	69	253
Liu 2011[244]	FIB4_F2_low	FIB-4	F2	1.1	73.5	68.1	158	130	57	278
Mallet 2009[245]	FIB4_F3_low	FIB-4	F3	0.67	0.71	0.73	30	28	11	69
Marcellin 2009[246]	TE_F2	Fibroscan	F2		70	83	61	15	26	71
Marcellin 2009[246]	TE_F3	Fibroscan	F3		86	85	37	20	6	111
Marcellin 2009[246]	TE_F4	Fibroscan	F4		93	87	13	21	1	138
Miailhes 2011[247]	Fibrotest_F2	Fibrotest	F2	0.38	77	86	28	3	8	18
Miailhes 2011[247]	Fibrotest_F3	Fibrotest	F3	0.42	94	77	19	9	1	28
Miailhes 2011[247]	Fibrotest_F4	Fibrotest	F4	0.58	100	81	12	9	0	36
Miailhes 2011[247]	TE_F2	Fibroscan	F2		81	87	29	3	7	18
Miailhes 2011[247]	TE_F3	Fibroscan	F3		85	87	17	5	3	32
Miailhes 2011[247]	TE_F4	Fibroscan	F4		92	94	11	3	1	42
Myers 2003[249]	Fibrotest_F2	Fibrotest	F2	0.4	54	80	33	30	28	118
Ogawa 2011[250]	TE_F1	Fibroscan	F1		66	71	25	2	13	4
Ogawa 2011[250]	TE_F2	Fibroscan	F2		95	74	19	6	1	18
Ogawa 2011[250]	TE_F3	Fibroscan	F3		87	75	7	9	1	27
Ogawa 2011[250]	TE_F4	Fibroscan	F4		75	89	3	4	1	36
Osakabe 2011[251]	TE_F2	Fibroscan	F2		73	100	33	0	12	6

Study ID	Test	Index test assessed	Fibrosis stage assessed	Cut-off	Sens.	Spec.	TP	FP	FN	TN
Osakabe 2011[251]	TE_F3	Fibroscan	F3		70	87	19	3	8	21
Osakabe 2011[251]	TE_F4	Fibroscan	F4		79	92	11	3	3	34
Park 2003[253]	collIV_F4	Type IV collagen	F4	6.3	63.6	88.6	7	10	4	79
Park 2003[253]	HA_F4	Hyaluronic acid	F4	77	81.8	87.3	9	11	2	78
Park 2004[254]	AAR_F4	AST–ALT ratio	F4	1	39	76	32	58	50	183
Park 2005[252]	Caffeine Breath test_F3	Caffeine breath test	F3		100	72	12	10	0	26
Poynard 2009[255]	Fibrotest_F2	Fibrotest	F2	0.48	66	69	112	90	58	202
Raftopoulos 2012[256]	APRI_F2_combined	APRI	F2				21	2	16	68
Raftopoulos 2012[256]	APRI_F2_high	APRI	F2	1.5	28	98	21	2	54	102
Raftopoulos 2012[256]	APRI_F2_low	APRI	F2	0.5	79	65	59	36	16	68
Raftopoulos 2012[256]	APRI_F4_low	APRI	F4	1.0			10	31	5	133
Raftopoulos 2012[256]	Fibrotest_F2	Fibrotest	F2	0.48	54	82	41	19	35	85
Raftopoulos 2012[256]	Fibrotest_F4	Fibrotest	F4	0.73			12	18	3	146
Raftopoulos 2012[256]	Hepascore_F2	Hepascore	F2	0.5	79	94	59	27	16	77
Raftopoulos 2012[256]	Hepascore_F4	Hepascore	F4	0.87			13	25	2	139
Sebastiani 2007[257]	AAR_F4	AST–ALT ratio	F4	1	7	95	2	4	20	84
Sebastiani 2007[257]	APRI_F2_high	APRI	F2	1.5	27	96	20	1	55	34
Sebastiani 2007[257]	APRI_F2_low	APRI	F2	0.5	71	87	53	5	22	30
Sebastiani 2007[257]	APRI_F4_high	APRI	F4	2	43	85	9	13	13	75
Sebastiani 2007[257]	Fibrotest_F2	Fibrotest	F2		81	90	61	4	14	32
Sebastiani 2007[257]	Fibrotest_F4	Fibrotest	F4		56	96	12	4	10	84
Sebastiani 2007[257]	Forns_F2_combined	Forns index	F2				11	0	32	27
Sebastiani 2007[257]	Forns_F2_high	Forns index	F2	6.9	15	100	11	0	64	35

continued

TABLE 68 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with chronic HBV (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Cut-off	Sens.	Spec.	TP	FP	FN	TN
Sebastiani 2007[257]	Forns_F2_low	Forns index	F2	4.2	58	78	44	8	32	27
Sebastiani 2007[257]	GUCI_F2	GUCI	F2	0.2	67	96	50	1	25	34
Sebastiani 2007[257]	GUCI_F4	GUCI	F4	1	21	91	5	8	17	80
Sebastiani 2007[257]	Hui_F2	Hui index	F2	0.15	50	91	38	3	38	32
Seto 2011[258]	APGA_F2	APGA	F2	6.7	17	98	13	3	64	157
Seto 2011[258]	APRI_F2_combined	APRI	F2				30	19	8	64
Seto 2011[258]	APRI_F2_high	APRI	F2	1.5	39	88	30	19	47	141
Seto 2011[258]	APRI_F2_low	APRI	F2	0.5	89	40	69	96	8	64
Seto 2011[258]	FIB4_F2_combined	FIB-4	F2				7	2	37	118
Seto 2011[258]	FIB4_F2_high	FIB-4	F2	3.25	9	99	7	2	70	158
Seto 2011[258]	FIB4_F2_low	FIB-4	F2	1.45	52	74	40	42	37	118
Seto 2011[258]	PAPAS_F2	PAPAS	F2	1.67	73	78	56	35	21	125
Shin 2008[259]	APRI_F2_combined	APRI	F2				106	21	4	42
Shin 2008[259]	APRI_F2_high	APRI	F2	1.5	75	83	106	21	35	102
Shin 2008[259]	APRI_F2_low	APRI	F2	0.5	97	34	137	81	4	42
Sinakos 2011[260]	TE_F4	Fibroscan	F4				5	14	0	40
Sokucu 2010[262]	Fibrotest_F3	Fibrotest	F3				0	2	9	14
Sporea 2010[263]	ARFI_F2	ARFI	F2	1.33	71	66	46	2	19	4
Sporea 2010[263]	TE_F2	Fibroscan	F2	7.6	60	83	39	1	26	5
Sporea 2010[200]	TE_F2	Fibroscan	F2		59	70	40	22	27	51
Sporea 2010[200]	TE_F3	Fibroscan	F3		53	85	17	16	16	91
Sporea 2010[200]	TE_F4	Fibroscan	F4		86	99	6	1	1	132
Vigano 2011[264]	TE_F2	Fibroscan	F2		55	95	36	3	30	56

Study ID	Test	Index test assessed	Fibrosis stage assessed	Cut-off	Sens.	Spec.	TP	FP	FN	TN
Vigano 2011[264]	TE_F4	Fibroscan	F4		75	93	15	7	5	98
Wong 2008[267]	TE_F4	Fibroscan	F4		90	79	18	17	2	63
Wong 2010 training[266]	Forns_F3_combined	Forns index	F3				21	7	1	21
Wong 2010 training[266]	Forns_F3_high	Forns index	F3	8.4	28	91	21	7	53	75
Wong 2010 training[266]	Forns_F3_low	Forns index	F3	5.2	99	26	73	61	1	21
Wong 2010 training[266]	TE_F3	Fibroscan	F3		95	58	48	4	34	70
Wong 2010 validation[266]	Forns_F3_combined	Forns index	F3				9	4	0	8
Wong 2010 validation[266]	Forns_F3_high	Forns index	F3	8.4	43	93	9	4	12	57
Wong 2010 validation[266]	Forns_F3_low	Forns index	F3	5.2	100	13	21	53	0	8
Wong 2010 validation[266]	TE_F3	Fibroscan	F3		81	61	37	4	24	17
Wong 2011[268]	TE_F3	Fibroscan	F3		75	47	26	4	29	12
Wong 2011 after Tx[268]	TE_F3	Fibroscan	F3		100	47	24	0	47	20
Wu 2012[269]	APRI_F2_low	APRI	F2	0.6	62	65	167	74	103	138
Wu 2012[269]	FIB4_F2_low	FIB-4	F2	1.57	70	69	189	66	81	146
Zhang 2008[271]	APRI_F1_high	APRI	F1	1.5	35.7	81.6	21	14	36	66
Zhang 2008[271]	APRI_F2_high	APRI	F2	1.5	44.7	84.3	36	9	44	48
Zhu 2011[272]	APRI_F2_low	APRI	F2	0.5	82	83.3	65	16	14	80
Zhu 2011[272]	APRI_F4_low	APRI	F4	1	75.9	69.2	22	45	7	101
Zhu 2011[272]	FIB4_F2_low	FIB-4	F2	1.7	74	84.4	58	15	21	81
Zhu 2011[272]	FIB4_F4_low	FIB-4	F4	1.9	69	75.3	20	37	9	110
Zhu 2011[272]	TE_F2	Fibroscan	F2		88	90.6	70	9	9	87
Zhu 2011[272]	TE_F4	Fibroscan	F4		93.1	91.1	27	13	2	133

AAR, AST–ALT ratio; APGA, AST, platelet count, GGT, α-fetoprotein; API, Age–Platelet Index; CDS, Cirrhosis Discriminant Score; FN, false negative; FP, false positive; PAPAS, age, ALP, α-fetoprotein, AST; sens., sensitivity; spec., specificity; TN, true negative; TP, true positive; Tx, treatment.

TABLE 69 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with ALD

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Forestier 2010[114]	TE_F4	Fibroscan	F4	11.4 kPa	84	91	66	1	13	10
Janssens 2010[273]	APRI_F2_combined	APRI	F2				18	1	15	6
Janssens 2010[273]	APRI_F2_high	APRI	F2	1.5	63	71	26	2	15	6
Janssens 2010[273]	APRI_F2_low	APRI	F2	0.5	43	88	18	1	23	7
Janssens 2010[273]	APRI_F4_high	APRI	F4	2	40	61	8	8	12	13
Janssens 2010[273]	TE_F3	Fibroscan	F3	17 kPa	72	76.5	23	4	9	13
Janssens 2010[273]	TE_F4	Fibroscan	F4	21.1 kPa	75	80	15	6	5	23
Janssens 2010[273]	Forns_F3_high	Forns index	F3	6.9	42.1	85.7	9	2	13	15
Kim 2009[74]	TE_F4	Fibroscan	F4	25.8	90	87	26	2	3	14
Lavallard 2011[275]	CK18_F3	CK18-Total	F3		84	71	49	25	9	60
Melin 2005[276]	TE_F4	Fibroscan	F4	13			34	1	0	0
Mueller 2010[277]	TE_F3	Fibroscan	F3	12.5	96	80	43	11	2	45
Nahon 2008[278]	TE_F3	Fibroscan	F3	11.6	87	89	96	4	14	33
Nahon 2008[278]	TE_F4	Fibroscan	F4	22.7	84	83	66	12	13	56
Naveau 2005[280]	FT_F2_high	Fibrotest	F2	0.7	55	93	77	6	63	75
Naveau 2005[280]	FT_F2_low	Fibrotest	F2	0.3	84	66	118	28	22	53
Naveau 2005[280]	FT_F4_high	Fibrotest	F4	0.7	91	87	62	20	6	133
Naveau 2005[280]	FT_F4_low	Fibrotest	F4	0.3	100	50	68	77	0	77
Naveau 1994[279]	PGAA_F4	PGAA	F4	7	79	89	36	30	10	241
Nguyen-Khac 2008[281]	TE_F1	Fibroscan	F1	5.9	83	86	79	1	16	7
Nguyen-Khac 2008[281]	TE_F2	Fibroscan	F2	7.8	80	91	62	2	15	24

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Nguyen-Khac 2008[281]	TE_F3	Fibroscan	F3	11	87	81	46	10	7	41
Nguyen-Khac 2008[281]	TE_F4	Fibroscan	F4	19.5	86	84	28	11	5	59
Tran 2000[282]	YKL40_F3	YKL-40	F3	330 µg/l	50.8	88.5	30	10	29	77
Vanbiervliet 2005[283]	APRI_F2_combined	APRI	F2				44	19	13	37
Vanbiervliet 2005[283]	APRI_F2_high	APRI	F2	1.5	50	78	44	19	44	68
Vanbiervliet 2005[283]	APRI_F2_low	APRI	F2	0.5	85	43	74	50	13	37

FN, false negative; FP, false positive; sens., sesitivity; spec., specificity; TN, true negative; TP, true positive.

TABLE 70 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with NAFLD

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Adams 2011[284]	APRI_F2	APRI	F2	0.43	71.1	69.6	69	44	28	101
Adams 2011[284]	APRI_F3	APRI	F3	0.54	72	77.1	38	43	15	146
Adams 2011[284]	APRI_F4	APRI	F4	0.54	77.3	70.9	18	64	5	155
Adams 2011[284]	BARD_F2	BARD	F2	2	44.3	70.4	43	43	54	102
Adams 2011[284]	BARD_F3	BARD	F3	2	60.4	71.5	32	54	21	135
Adams 2011[284]	BARD_F4	BARD	F4	3	52.2	83.8	12	35	11	184
Adams 2011[284]	FIB4_F2_low	FIB-4	F2	1.45	54.4	87.5	53	18	44	127
Adams 2011[284]	FIB4_F3_low	FIB-4	F3	1.54	74	86.9	39	25	14	164
Adams 2011[284]	FIB4_F4_low	FIB-4	F4	1.92	72.7	70.9	17	64	6	155
Adams 2011[284]	FT_F2_low	Fibrotest	F2	0.34	58.9	73.2	57	39	40	106
Adams 2011[284]	FT_F3_high	Fibrotest	F3	0.47	60.8	89.8	32	19	21	170
Adams 2011[284]	FT_F4_high	Fibrotest	F4	0.57	72.7	92.1	17	17	6	202
Adams 2011[284]	Hepascore_F2	Hepascore	F2	0.44	50.5	88.3	49	17	48	128
Adams 2011[284]	Hepascore_F3	Hepascore	F3	0.37	75.5	84.1	40	30	13	159
Adams 2011[284]	Hepascore_F4	Hepascore	F4	0.7	87.0	89.0	20	24	3	196
Angulo 2007 – validation[285]	NFS_F3_all	NFS	F3	1.455 and 0.676			32	7	17	127
Angulo 2007 – estimation[285]	NFS_F3_all	NFS	F3	1.455 and 0.676			64	7	23	273
Angulo 2007 – estimation[285]	NFS_F3_high	NFS	F3	> 0.676	51	98	64	7	61	319
Angulo 2007 – estimation[285]	NFS_F3_low	NFS	F3	< –1.455	82	77	103	82	23	273
Angulo 2007 – validation[285]	NFS_F3_high	NFS	F3	> 0.676	43	96	32	7	42	172
Angulo 2007 – validation[285]	NFS_F3_low	NFS	F3	< –1.455	77	71	57	52	17	127
Cales 2009[287]	APRI_F2	APRI	F2	0.5	66.1	90.6	43	16	22	154
Cales 2009[287]	Fibrometer_F2	FibroMeter	F2		78.5	95.9	51	7	14	163

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Cales 2009[287]	NFS_F2_high	NFS	F2		60.9	96.3	40	6	25	164
Dixon 2001[229]	HAIR_NASH	HAIR	NASH	2	80	89	21	9	5	70
Fujii 2009[117]	AAR_F4	AST–ALT ratio	F4		88	72	8	11	1	30
Fujii 2009[117]	AP_F4	Age–Platelet Index	F4		88	84	8	7	1	34
Fujii 2009[117]	APRI_F4	APRI	F4		76	70	7	12	2	29
Fujii 2009[117]	CDS_F4_high	CDS	F4		33	100	3	0	6	41
Fujii 2009[117]	CDS_F4_low	CDS	F4		89	90	8	4	1	37
Fujii 2009[117]	HALT-C_F4_high	HALT-C	F4		22	100	2	0	7	41
Fujii 2009 [17]	HALT-C_F4_low	HALT-C	F4		89	68	8	13	1	28
Gaia 2011[119]	TE F1	Fibroscan	F1		84	57	41	10	8	13
Gaia 2011[119]	TE_F1	Fibroscan	F2		76	80	25	8	8	31
Gaia 2011[119]	TE_F2	Fibroscan	F3		65	80	11	11	6	44
Gaia 2011[119]	TE_F3	Fibroscan	F4		78	96	7	3	2	60
Guajardo-Salinas 2010[289]	TE_F4	FibroSpect II	F2		100	42	80	49	0	35
Guha 2008[290]	FIBROSPECT_F2	ELF	F1	9.8	61	80	69	16	44	63
Guha 2008[290]	ELF_F1	ELF	F2	9.9	70	80	54	23	23	92
Guha 2008[290]	ELF_F2	ELF	F3	10.35	80	90	35	15	9	133
Guha 2008[290]	ELF_F3	Combined panel (NFS + ELF)					68	7	9	41
Guha 2008[290]	NFS_ELF_F1_all	Combined panel (NFS + ELF)	F1		60	91	68	7	45	72
Guha 2008[290]	NFS_ELF_F1_high	Combined panel (NFS + ELF)	F1		92	52	104	38	9	41
Guha 2008[290]	NFS_ELF_F1_low	Combined panel (NFS + ELF)	F2				61	10	8	99
Guha 2008[290]	NFS_ELF_F2_all	Combined panel (NFS + ELF)	F2		79	91	61	10	16	105
Guha 2008[290]	NFS_ELF_F2_high	Combined panel (NFS + ELF)	F2		89	86	69	16	8	99

continued

TABLE 70 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with NAFLD (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Guha 2008[290]	NFS_ELF_F2_low	Combined panel (NFS + ELF)	F3				38	1	4	142
Guha 2008[290]	NFS_ELF_F3_all	Combined panel (NFS + ELF)	F3		86	99	38	1	6	147
Guha 2008[290]	NFS_ELF_F3_high	Combined panel (NFS + ELF)	F3		91	96	40	6	4	142
Guha 2008[290]	NFS_ELF_F3_low	NFS	F1				36	7	9	40
Guha 2008[290]	NFS_F1_all	NFS	F1		32	91	36	7	77	72
Guha 2008[290]	NFS_F1_high	NFS	F1		92	50	104	40	9	40
Guha 2008[290]	NFS_F1_low	NFS	F2				52	13	8	66
Guha 2008[290]	NFS_F2_all	NFS	F2		68	89	52	13	25	102
Guha 2008[290]	NFS_F2_high	NFS	F2		89	57	69	49	8	66
Guha 2008[290]	NFS_F2_low	NFS	F3				34	10	4	87
Guha 2008[290]	NFS_F3_all	NFS	F3		77	93	34	10	10	138
Guha 2008[290]	NFS_F3_high	NFS	F3		91	59	4	61	4	87
Harrison 2008[291]	NFS_F3_low	BARD	F3		95	65	173	227	9	418
Kaneda 2006[292]	BARD_F3	Hyaluronic acid	F3	42	100	89	40	12	0	96
Kaneda 2006[292]	HA_F3	Platelet count	F4		100	95	19	6	0	123
Kaneda 2006[292]	PLT_F4	Type IV collagen	F3		78	87	31	14	9	94
Kayadibi 2009[293]	AST/ALT_NASH	AST–ALT ratio	NASH	1.09	56	58	24	5	19	6
Kelleger 2005[294]	TE_F2	Fibroscan	F2	10	88	72	57	18	8	46
Khosravi 2011[295]	AST/ALT_F3_low	AST–ALT ratio	F3	0.88	87.5	79.7	7	28	1	111
Ledinghen 2009[288]	TE_F2	Fibroscan	F2	7	77	77	61	30	18	99
Ledinghen 2009[288]	TE_F3	Fibroscan	F3	8.7	84	87	37	21	7	143
Ledinghen 2009[288]	TE_F4	Fibroscan	F4	10.3	95	88	19	23	1	165
Lupsor 2010[296]	TE_F1	Fibroscan	F1	5.3	93.5	78.2	44	5	3	20

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Lupsor 2010[296]	TE_F2	Fibroscan	F2	6.8	66.7	84.3	12	8	6	46
Lupsor 2010[296]	TE_F3	Fibroscan	F3	10.4	100	96.8	5	2	0	65
Lydatakis 2006[297]	HA_NASH	Hyaluronic acid	F1	148.5	95.7	96.3	22	1	1	26
Lydatakis 2006[297]	Laminin_NASH	Laminin	F1	292.5	73.9	71.4	17	8	6	19
Mahadeva 2010[298]	TE_F3	Fibroscan	F3	9.4	83	89	5	2	1	17
Manousou 2011[299]	Ferritin_NASH	Serum ferritin	NASH	240	91	70	58	14	6	33
McPherson 2010[301]	APRI_F3	APRI	F3	1	27	89	7	13	20	105
McPherson 2010[301]	AST/ALT_F3_high	AST–ALT ratio	F3	1	52	90	14	12	13	106
McPherson 2010[301]	AST/ALT_F3_low	AST–ALT ratio	F3	0.8	74	78	20	26	7	92
McPherson 2010[301]	BARD_F3	BARD	F3	2	89	44	24	66	3	52
McPherson 2010[301]	FIB4_F3_all	FIB-4	F3				7	2	4	77
McPherson 2010[301]	FIB4_F3_high	FIB-4	F3	3.25	26	98	7	2	20	116
McPherson 2010[301]	FIB4_F3_low	FIB-4	F3	1.3	85	65	23	41	4	77
McPherson 2010[301]	NFS_F3_all	NFS	F3				9	2	6	68
McPherson 2010[301]	NFS_F3_high	NFS	F3	0.676	33	98	9	2	18	116
McPherson 2010[301]	NFS_F3_low	NFS	F3	–1.45	78	58	21	50	6	68
McPherson 2011[300]	AST/ALT_F3_low	AST–ALT ratio	F3	0.8	94	44	54	139	3	109
McPherson 2011[300]	BARD_F3	BARD	F3	2	94	26	54	184	3	64
McPherson 2011[300]	FIB4_F3_low	FIB-4	F3	1.3	82	77	47	57	10	191
McPherson 2011[300]	NFS_F3_low	NFS	F3	–1.45	82	51	47	122	10	126
Oliveira 2005[165]	AST/ALT_F3_high	AST–ALT ratio	F3		25	86	4	10	11	64
Oliveira 2005[165]	HA_F3	Hyaluronic acid	F3		75	75	11	19	4	56
Pais 2011[303]	FT_TE_F2	Fibrotest + Fibroscan	F2	< 0.48 and ≤ 9.6	76	71	36	38	15	119

continued

TABLE 70 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with NAFLD (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Pais 2011[303]	FT_TE_F3	Fibrotest + Fibroscan	F3	> 0.48 and > 7.9	39	96	20	6	31	151
Pais 2011 [303]	NFS_TE_F2	NFS + Fibroscan	F2	< 1.445 and ≤ 9.6	64	65	33	57	18	100
Pais 2011[303]	NFS_TE_F3	NFS + Fibroscan	F3	> 0.66 and 7.9	7	98	4	3	47	154
Palmeri 2011[304]	ARFI_F3	ARFI	F3	4.2	90	90	36	10	4	86
Park 2011[306]	CBT_F4	Caffeine breath test	F4	1.27	90	76	9	9	1	29
Pawitpok 2006[307]	HA_F2	Hyaluronic acid	F2	218.5	78	89	7	2	2	16
Petta 2011[308]	TE_F2	Fibroscan	F2	7.25	69	70	47	23	21	55
Petta 2011[308]	TE_F3	Fibroscan	F3	8.75	76	78	25	25	8	88
Pimentel 2010[309]	NFS_F3_high	NFS	F3	0.676	83	97	20	3	2	133
Pimentel 2010[309]	NFS_F3_all	NFS	F3				20	3	1	109
Pimentel 2010[309]	NFS_F3_low	NFS	F3	–1.455			21	27	1	109
Poynard 2006 Training[310]	NASH_test_NASH	NASH test	NASH		39	92	11	10	17	122
Poynard 2006 Valid[310]	NASH_test_NASH	NASH test	NASH		29	98	10	1	25	61
Qureshi 2010[311]	NFS_F1_all	NFS	F1	> 0.676 and < –1.455			59	8	49	61
Qureshi 2010[311]	NFS_F1_high	NFS	F1	> 0.676 and < –1.457			59	8	158	106
Qureshi 2010[311]	NFS_F1_low	NFS	F1	> 0.676 and < –1.456			164	57	49	61
Qureshi 2010[311]	NFS_F2_all	NFS	F2	> 0.676 and < –1.455			38	29	14	96
Qureshi 2010[311]	NFS_F2_high	NFS		> 0.676 and < –1.459			38	29	51	213
Qureshi 2010[311]	NFS_F2_low	NFS	F2	> 0.676 and < –1.458			79	142	14	96
Qureshi 2010[311]	NFS_F3_all	NFS	F3	> 0.676 and < –1.455			22	45	2	108
Qureshi 2010[311]	NFS_F3_high	NFS	F3	> 0.676 and <–1.461			22	45	28	236
Qureshi 2010[311]	NFS_F3_low	NFS	F3	> 0.676 and < –1.460			48	173	2	108
Raszeja-Wyscomirska 2010[312]	BARD_F3	BARD	F3	2	87	73	13	24	2	64

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Ratziu 2004[313]	FT_F3_all	Fibrotest	F3				6	2	0	58
Ratziu 2004[313]	FT_F3_high	Fibrotest	F3	0.6	60	97	6	2	4	77
Ratziu 2004[313]	FT_F3_low	Fibrotest	F3	0.3	100	73	10	21	0	58
Ratziu 2006 group1[314]	FT_F2_all	Fibrotest	F2				7	3	7	101
Ratziu 2006 group1[314]	FT_F2_high	Fibrotest	F2	0.7	18	98	7	3	33	127
Ratziu 2006 group1[314]	FT_F2_low	Fibrotest	F2	0.3	83	78	33	29	7	101
Ratziu 2006 group1[314]	FT_F3_all	Fibrotest	F3				5	5	1	107
Ratziu 2006 group1[314]	FT_F3_high	Fibrotest	F3	0.7	25	97	5	5	15	146
Ratziu 2006 group1[314]	FT_F3_low	Fibrotest	F3	0.3	95	71	19	44	1	107
Ratziu 2006 group2[314]	FT_F2_all	Fibrotest	F2				4	1	9	49
Ratziu 2006 group2[314]	FT_F2_high	Fibrotest	F2	0.7	13	98	4	1	27	65
Ratziu 2006 group2[314]	FT_F2_low	Fibrotest	F2	0.3	71	74	22	17	9	49
Ratziu 2006 group2[314]	FT_F3_all	Fibrotest	F3				4	1	2	56
Ratziu 2006 group2[314]	FT_F3_high	Fibrotest	F3	0.7	25	99	4	1	12	80
Ratziu 2006 group2[314]	FT_F3_low	Fibrotest	F3	0.3	88	69	14	25	2	56
Ruffilo 2011[315]	BARD_F3	BARD	F3	≥ 2	51	77	19	23	18	78
Ruffilo 2011[315]	NFS_F3_all	NFS	F3	< −1.455 and > 0.676	23	100	5	0	17	74
Ruffilo 2011[315]	NFS_F3_high	NFS	F3	> 0.676			5	0	32	101
Ruffilo 2011[315]	NFS_F3_low	NFS	F3	> −1.455			20	27	17	74
Sakugawa 2005[316]	coll4_F3	Type IV collagen	F3	5	81	71	39	19	9	45
Sakugawa 2005[316]	coll4_NASH	Type IV collagen	NASH	5	70	81	49	8	21	34
Sakugawa 2005[316]	HA_F3	Hyaluronic acid	F3	50	69	83	33	11	15	53
Sakugawa 2005[316]	HA_NASH	Hyaluronic acid	NASH	50	66	90	46	4	24	38

continued

TABLE 70 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with NAFLD (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Santos 2005[317]	coll4_F1	Type IV collagen	F1	145	64	89	7	2	4	17
Santos 2005[317]	HA_F1	Hyaluronic acid	F1	24.6	82	68	9	6	2	13
Santos 2005[317]	laminin_F1	Laminin	F1	282	82	89	9	2	2	17
Santos 2005[317]	coll4_NASH	Type IV collagen	NASH	5	41	95	27	1	39	18
Sumida 2011 training[320]	NAFIC_NASH_high	NAFIC score	NASH	≥ 2	66	91	65	7	33	72
Sumida 2011 training[320]	NAFIC_NASH_low	NAFIC score	NASH	≥ 1	94	48	92	41	6	38
Sumida 2011 total[320]	NAFIC_F2_all	NAFIC score	F2	0 and ≥ 2			127	118	7	153
Sumida 2011 total[320]	NAFIC_F2_high	NAFIC score	F2	≥ 2	84	74	127	118	25	349
Sumida 2011 total[320]	NAFIC_F2_low	NAFIC score	F2	> 0	95	33	145	314	7	153
Sumida 2011 total[320]	NAFIC_F3_all	NAFIC score	F3	< 1 and ≥ 3			56	99	3	371
Sumida 2011 total[320]	NAFIC_F3_high	NAFIC score	F3	≥ 3	84	82	56	99	11	453
Sumida 2011 total[320]	NAFIC_F3_low	NAFIC score	F3	> 1	96	67	64	181	3	371
Sumida 2011 total[320]	NFS_F2_all	NFS	F2	≤ 1.145 and ≥ 0.676			33	16	25	305
Sumida 2011 total[320]	NFS_F2_high	NFS	F2	≥ 0.676	23	96	33	18	112	425
Sumida 2011 total[320]	NFS_F2_low	NFS	F2	> –1.145	86	69	125	136	20	306
Sumida 2011 total[320]	NFS_F3_all	NFS	F3	≤ –1.455 and >0.676			28	21	5	325
Sumida 2011 total[320]	NFS_F3_high	NFS	F3	> 0.676	33	95	21	28	43	496
Sumida 2011 total[320]	NFS_F3_low	NFS	F3	≥ –1455	92	62	59	199	5	325
Sumida 2011 validation[320]	NAFIC_NASH_high	NAFIC score	NASH	≥ 2	60	87	146	26	98	172
Sumida 2011 validation[320]	NAFIC_NASH_low	NAFIC score	NASH	≥ 1	88	43	215	113	29	85
Sumida 2012[319]	Age-PLT_index_F3	Age–Platelet Index	F3	6	66	78	42	113	22	399
Sumida 2012[319]	APRI_F3	APRI	F3	1	67	81	43	97	21	415
Sumida 2012[319]	AST/ALT_F3_high	AST–ALT ratio	F3	1	48	92	31	41	33	471

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Sumida 2012[319]	AST/ALT_F3_low	AST–ALT ratio	F3	0.8	66	76	42	123	22	389
Sumida 2012[319]	BARD_F3	BARD	F3	2	80	65	51	179	13	333
Sumida 2012[319]	FIB4_F3_all	FIB-4	F3	< 1.45 and > 3.25			31	28	6	330
Sumida 2012[319]	FIB4_F3_high	FIB-4	F3	3.25	48	95	31	26	33	486
Sumida 2012[319]	FIB4_F3_low	FIB-4	F3	1.45	90	64	58	184	6	328
Suzuki 2005[321]	HA_F3	Hyaluronic acid	F3	46.1	85	79.7	17	12	3	47
Wong 2008[322]	NFS_F2_all	NFS	F2	< –1.455 and > 0.676			0	2	26	102
Wong 2008[322]	NFS_F2_high	NFS	F2	0.676	0	98	0	2	41	119
Wong 2008[322]	NFS_F2_low	NFS	F2	–1455	37	84	15	19	26	102
Wong 2008[322]	NFS_F3_all	NFS	F3	< –1.455 and > 0.676			0	2	11	117
Wong 2008[322]	NFS_F3_high	NFS	F3	0.676	0	99	0	2	18	142
Wong 2008[322]	NFS_F3_low	NFS	F3	–1455	39	81	7	27	11	117
Wong 2008[323]	TE_F3	Fibroscan	F3	7.5	82	71	14	15	3	38
Wong 2009[324]	TE_F2	Fibroscan	F2	7.0	79	76	80	35	21	110
Wong 2009[324]	TE_F3	Fibroscan	F3	8.7	84	83	47	32	9	158
Wong 2009[324]	TE_F4	Fibroscan	F4	10.3	92	88	23	27	2	194
Yoneda 2008[326]	TE_F1	Fibroscan	F1	5.9	86.1	88.9	68	2	11	16
Yoneda 2008[326]	TE_F2	Fibroscan	F2	6.65	88.2	73.9	45	12	6	34
Yoneda 2008[326]	TE_F3	Fibroscan	F3	9.8	85.2	81.4	14	15	3	65
Yoneda 2008[326]	TE_F4	Fibroscan	F4	17.5	100	96.6	88	0	0	9
Yoneda 2011[325]	PLT_F3	Platelet count	F3	0.774	62.7	76.3	144	293	84	927
Yoneda 2011[325]	PLT_F4	Platelet count	F4	0.918	80.5	88.8	33	111	8	896
Younossi 2011[327]	APRI_F1	APRI	F1	0.5	56	97.4	2	1	37	39

continued

TABLE 70 Diagnostic test accuracy of non-invasive fibrosis tests in individual studies in patients with NAFLD (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Younossi 2011[327]	APRI_F3	APRI	F3	0.5	71	96.7	1	2	15	61
Younossi 2011[327]	M30_NASH	M30	NASH	272.9	72.5	64.1	29	14	11	25
Younossi 2011[327]	NDP_F1	NAFLD diagnostic panel: model predicting any fibrosis	F1	0.4242	60.6	71.8	24	11	15	29
Younossi 2011[327]	NDP_F3	NAFLD diagnostic panel: model predicting severe fibrosis	F3	0.2442	86.7	70.4	14	19	2	44
Younossi 2011[327]	NDP_NASH	NAFLD diagnostic panel: NASH model	NASH	0.36	79.4	73.7	32	11	8	28
Younossi 2011[327]	NFS_F1_low	NFS	F1	–0.1657	84.8	34.2	33	26	6	14

AAR, AST–ALT ratio; CDS, Cirrhosis Discriminant Score; FN, false negative; FP, false positive; HAIR, Hypertension, ALT > 40 and Insulin Resistance Index > 5; HALT-C, Hepatitis C Antiviral Long-Term Treatment Against Cirrhosis; NDP, NAFLD diagnostic panel; PLT, platelet; sens., sensitivity; spec., specificity; TN, true negative; TP, true positive.

TABLE 71 Diagnostic test accuracy in individual studies assessing imaging non-invasive fibrosis tests

Study ID	Test	Index test assessed	Fibrosis stage assessed	Comment	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Aguirre 2006[368]	DEMRI	Double-enhanced MRI (gadolinium and SPIO)	F3	4.76 millisecond TE SPGR		93	87	67	4	5	25
Asbach 2010[328]	MRE_F1	MR elastography	F1		2.84	80	81	58	3	14	13
Asbach 2010[328]	MRE_F2	MR elastography	F2		3.18	77	97	40	1	12	35
Asbach 2010[328]	MRE_F3	MR elastography	F3		3.32	97	87	34	7	1	46
Asbach 2010[328]	MRE_F4	MR elastography	F4		4.21	100	91	19	6	0	63
Aube 1999[329]	US_F4	US	F4	11 parameters				38	23	10	119
Aube 2004[330]	US_F3	US	F3	Five parameters				23	11	9	63
Aube 2004[330]	US_F4	US	F4	Five parameters				22	4	4	76
Awaya 2002[331]	MRI_F4	MRI	F4	Modified caudate lobe ratio	C/RL-r > 0.9	72	77	87	26	34	89
Cardi 1997[332]	US_F4	US	F4	Echotexture, nodularity, spleen size				79	10	15	81
Chen 2008[93]	US_F2	US	F2	Echotexture, nodularity, irregular surface		27	63	23	8	61	14
Chen 2008[93]	US_F3	US	F3	Echotexture, nodularity, irregular surface		14	66	3	29	19	57
Chen 2008[227]	US_F4	US	F4			86	63	84	124	14	211
Cioni 1992[333]	US_F4	US	F4	PV flow				39	1	20	57
Cobbold 2010[98]	CEUS_F4	CEUS	F4	8		71	91	10	5	4	48
Cobbold 2010[98]	CEUS_F2	CEUS	F2	10.25		53	73	20	8	17	22
Colli 1994[334]	US_F4	US	F4	Hepatic vein waveform				12	8	4	28
Colli 2003[335]	US_F4	US	F4	Any one of nodularity, caudate lobe hypertropy, pattern of PV flow		68	79	50	48	23	179

continued

TABLE 71 Diagnostic test accuracy in individual studies assessing imaging non-invasive fibrosis tests (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Comment	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Corradi 2009[100]	US_SAPI_F2	Splenic arterial pulsatity index	F2			78	83	10	4	3	19
Crespo 2010[101]	MRE_F3	MR elastography	F3	3.5		90	71	10	12	1	28
D'Onofrio 2005[336]	US_F3	US	F3	PV, spleen size, parenchyma, liver margins	Any of the four parameters present	68	68	19	25	9	52
Do 2010[337]	DWMRI_F2	DW-MRI	F2	b-values 0, 50, 500	ADC 1.68	67	61	20	2	10	2
Do 2010[337]	DWMRI_F3	DW-MRI	F3	b-values 0, 50, 500	ADC 1.53	56	71	14	3	11	6
Do 2010[337]	DWMRI_F4	DW-MRI	F4	b-values 0, 50, 500	ADC 1.68	76	60	16	5	5	8
Do 2010[337]	DWMRI_F2	DW-MRI	F2	b-values 0, 50, 500 normalised ADC (liver/spleen)	1.41	90	77	27	1	3	3
Do 2010[337]	DWMRI_F3	DW-MRI	F3	b-values 0, 50, 500 normalised ADC (liver/spleen)	1.41	96	71	24	3	1	6
Do 2010[337]	DWMRI_F4	DW-MRI	F4	b-values 0, 50, 500 normalised ADC (liver/spleen)	1.4	95	66	20	4	1	9
Fahmy 2011[110]	US_F2	US	F2	≥ 0.64		69	74	46	11	21	32
Fahmy 2011[110]	US_F4	US	F4	≥ 0.71		87	65	19	31	3	57
Fahmy 2011[110]	US_SAPI_F2	Splenic arterial pulsatity index	F2	≥ 1		66	72	44	12	23	31
Fahmy 2011[110]	US_SAPI_F4	Splenic arterial pulsatity index	F4	≥ 1.06		74	67	16	29	6	59
Ferral 1992[338]	US_F4	US	F4	Nodular surface				28	7	4	31
Friedrich-Rust 2010[367]	CEMRI_F2	MR elastography intravenoud gadolinium-enhanced sequences	F2	Primovist	After 10 minutes	90	76	31	2	3	8
Fujimoto 2011[118]	DWMRI_F1	DW-MRI	F1		1.35	79	83	34	2	9	10

Study ID	Test	Index test assessed	Fibrosis stage assessed	Comment	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Fujimoto 2011[118]	DWMRI_F2	DW-MRI	F2		1.32	85	91	29	2	5	19
Fujimoto 2011[118]	DWMRI_F3	DW-MRI	F3		1.27	87	84	20	5	3	27
Fujimoto 2011[118]	DWMRI_F4	DW-MRI	F4		1.23	75	72	9	12	3	31
Gaia 2009[339]	US_F4	US	F4	Nodularity		63	86	12	6	7	36
Gaiani 1997[340]	US_F4	US	F4	Portal velocity, nodularity				37	32	10	133
Gierblinski 2008[341]	RTE_F1	RTE	F1	Colour coded	> 35.5%	86	84	30	1	5	3
Goyal 1990[342]	US_F4	US	F4	R/L liver lobe ratio	1.3			28	0	10	38
Hu 2010[234]	US_MARS_F1	US MARS	F1	MARS (a novel parameter from sonographic videos)	22.49	82	75	23	2	5	6
Hu 2010[234]	US_MARS_F2	US MARS	F2	MARS (a novel parameter from sonographic videos)	21.81%	85	56	17	7	3	9
Hu 2010[234]	US_MARS_F4	US MARS	F4	MARS (a novel parameter from sonographic videos)	20.32%	100	50	8	14	0	14
Huwart 2008[30]	MRE_F1	MR elastography	F1		2.42	85	91	63	2	11	20
Huwart 2008[30]	MRE_F2	MR elastography	F2		2.49	100	91	52	4	0	40
Huwart 2008[30]	MRE_F3	MR elastography	F3		3.13	91	97	30	2	3	61
Huwart 2008[30]	MRE_F4	MR elastography	F4		4.13	100	96	18	3	0	75
Iacobellis 2005[132]	US_F2	US	F2	Nodular liver surface	Positive	16	97	104	15	544	480
Iacobellis 2005[132]	US_F4	US	F4	Nodular liver surface		46	93	38	74	44	987
Ibrahim 2011[343]	DWMRI_F2	DW-MRI	F2	Mean hepatic ADC at 300 seconds/mm^2 (b-value)	≤ 1.89	90	89	16	2	2	18
Ishibashi 2012[344]	CEUS_F2	CEUS	F2	Sonazoid (second-generation microbubbles)	15-minute phase intensity of difference	88	72	69	10	9	25
Ishibashi 2012[344]	CEUS_F3	CEUS	F3	Sonazoid (second-generation microbubbles)		85	91	44	5	8	56

continued

TABLE 71 Diagnostic test accuracy in individual studies assessing imaging non-invasive fibrosis tests (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Comment	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Ishibashi 2012[344]	CEUS_F4	CEUS	F4	Sonazoid (second-generation microbubbles)		97	90	28	8	1	76
Joseph 1991[345]	US_F1	US	F1	Parenchyma heterogeneity				24	2	7	17
Kim 2011[346]	MRE_F1	MR elastography	F1		2.87	80	90	32	2	8	18
Kim 2011[346]	MRE_F2	MR elastography	F2		3.05	89.7	87.1	26	4	3	27
Kim 2011[346]	MRE_F3	MR elastography	F3		3.57	94.7	90.2	18	4	1	37
Kim 2011[346]	MRE_F4	MR elastography	F4		5.32	100	92.2	9	4	0	47
Kim 2010 EALT[347]	LSPI_F4	Spleen diameter-to-platelet radio index	F4		42	96.3	67.4	77	28	3	58
Kim 2010 EALT[347]	LSPI_F4	Spleen diameter-to-platelet radio index	F4		94	67.5	97.7	54	2	26	84
Kim 2010 NALT[347]	LSPI_F4	Spleen diameter-to-platelet radio index	F4		38	98	69.2	97	20	2	45
Kim 2010 NALT[347]	LSPI_F4	Spleen diameter-to-platelet radio index	F4		62	85.9	93.8	85	4	14	61
Ladenheim 1992[348]	US_F4	US	F4	Nodular surface				1	5	7	37
Lee 2011[145]	MRE_F1	MR elastography	F1		3.81	88	79	6	5	1	20
Lee 2010[349]	US_F4	US	F4		NA	38.9	87.4	14	21	22	146
Lewin 2007[149]	MRE_F3	MR elastography	F3		1.21	87	87	13	5	2	34
Liu 2006[152]	US_SAPI_F2_high	SAPI	F2		1.05	66.7	89.7	14	6	7	52
Liu 2006[152]	US_SAPI_F2_low	SAPI	F2		0.85	97.5	38.8	20	35	1	23
Liu 2006[152]	US_SAPI_F2	SAPI	F2		1	76	80	256	33	81	133
Liu 2006[152]	US_SAPI_F2_high	SAPI	F2		1.1	61	98	206	3	131	163
Liu 2006[152]	US_SAPI_F2_low	SAPI	F2		0.85	94	39	317	101	20	65
Liu 2006[152]	US_SAPI_F4	SAPI	F4		1.2	88	82	74	75	10	344

Study ID	Test	Index test assessed	Fibrosis stage assessed	Comment	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Lutz 2012[351]	HVRI_US_F4	US	F4	HVRI	1.185	90	86	27	13	3	82
Motosugi 2011[369]	CEMRI_F1	MR elastography intravenoud gadolinium-enhanced sequences	F1	Primovist (liver-specific contrast agent)	1.91	87	75	73	4	11	12
Motosugi 2011[369]	CEMRI_F2	MR elastography intravenoud gadolinium-enhanced sequences	F2		1.76	74.6	56	50	15	17	18
Motosugi 2011[369]	CEMRI_F3	MR elastography intravenoud gadolinium-enhanced sequences	F3		1.76	75.4	50	43	22	14	22
Motosugi 2011[369]	CEMRI_F4	MR elastography intravenoud gadolinium-enhanced sequences	F4		1.75	81.6	49.1	29	33	7	31
Nagata 2003[352]	US_F4	US	F4	Liver nodularity assessed – 3.75 standard		73	58	16	23	6	32
Nagata 2003[352]	US_F4	US	F4	Liver nodularity assessed – 7.5 experimental		68	82	15	10	7	45
Nishiura 2005[353]	US_F4	US	F4	Score of three parameters	6.5	100	100	22	0	0	81
Numminen 2005[354]	MRI_F4	MRI	F4	1.5T MRI		87	92	26	2	4	24
Ong 2003[355]	US_F4	US	F4	3.75 MHz US		38	85	6	11	10	61
Orrlachio 2011[166]	CEUS_F3	Contrast-enhanced US	F3	16		77	83	15	5	4	25
Paggi 2008[167]	US_F3	US	F3			73	90	117	27	43	243
Papalavrentios 2011[305]	MRI_F3	MRI	F3		1.16×10^{-3}			8	1	0	9
Ronot 2010[176]	CT_F2	Mean transit time	F2	13.4		71	65	21	8	9	14
Rustogi 2011[356]	MRE_F3	MR elastography	F3		5.9	85	88	27	5	5	35

continued

TABLE 71 Diagnostic test accuracy in individual studies assessing imaging non-invasive fibrosis tests (*continued*)

Study ID	Test	Index test assessed	Fibrosis stage assessed	Comment	Index test cut-off	Sens.	Spec.	TP	FP	FN	TN
Rustogi 2011[356]	MRI_F3	MRI	F3	Morphological criteria		78	75	25	10	7	30
Sandrasegaran 2009[357]	US_F4	DW-MRI	F2		103	73	59	37	11	14	16
Sandrasegaran 2009[357]	US_F4	DW-MRI	F3		98	52	71	21	11	20	26
Schneider 2005[184]	US_F4	US	F4	Reduced PV undulations		77	100	13	0	4	102
Schneider 2006[185]	SPECT_F2	US	F4	PV flow < 12.5 cm/second		87	69	17	20	2	44
Shen 2006	CEMRE_F1	US	F4	Spleen length	12.1	60	75	18	74	12	221
Shiramizy 2006	MRE_F1	Single-photon emission CT	F2	Minimum spleen pixel and right hepatic lobe				19	4	3	20
Venkatesh 2010[359]	US_F4	MR elastography intravenoud gadolinium-enhanced sequences	F1	Contrast enhanced	2.91 kPa	95	100	18	0	1	5
Venkatesh 2010[359]	US_F4	MR elastography	F1		2.83 kPa	95	100	18	0	1	5
Vigano 2005[360]	US_F4	US	F4	Nodularity		53	91	20	6	17	65
Wang 2009[361]	CEUS_F1	US	F4	Liver surface and parenchyma, hepatic vessels and spleen index		74	86	45	36	16	223
Xu 2005[362]	CEUS_F2	US	F4	Scoring system	10	88	97	21	1	3	41
Zhang 2011[270]	CEUS_F3	CEUS	F1		1.02	80	86	61	1	15	9
Zhang 2011[270]	CEUS_F4	CEUS	F2		0.96	75	86	50	3	17	17
Zhang 2011[270]	DWMRI_F2	CEUS	F3		0.83	71	84	31	10	13	52
Zhang 2011[270]	US_F4	CEUS	F4		0.72	76	80	27	10	8	41
Zhu 2008[363]	US_F4	DW-MRI	F2		b-value = 500 s/mm^2	84	80	21	4	4	14

ADC, apparent diffusion coefficient; C/RL–r, caudate/right lobe ratio; CEMRE, contrast-enhanced magnetic resonance elastography; CEUS, contrast-enhanced ultrasound; DEMRI, double-enhanced magnetic resonance imaging; DW-MRI, diffusion-weighted magnetic resonance imaging; EALT, elevated ALT; FN, false negative; FP, false positive; HVRI, hepatic vein resistance index; NA, not applicable; NALT, normal ALT; PV, portal vein; RTE, real time elastography; sens. sensitivity; spec., specificity; SPGR, spoiled gradient echo; SPIO, super paramagnetic iron oxide; TN, true negative; TP, true positive; US, ultrasound.

Appendix 5 Forest plots

This appendix presents forest plots of sensitivity and specificity of non-invasive tests across all fibrosis stages in patients with chronic HBV, chronic HCV, NAFLD and ALD. Plots are presented when there are data available from at least two studies. Studies are represented by their reference number in the report.

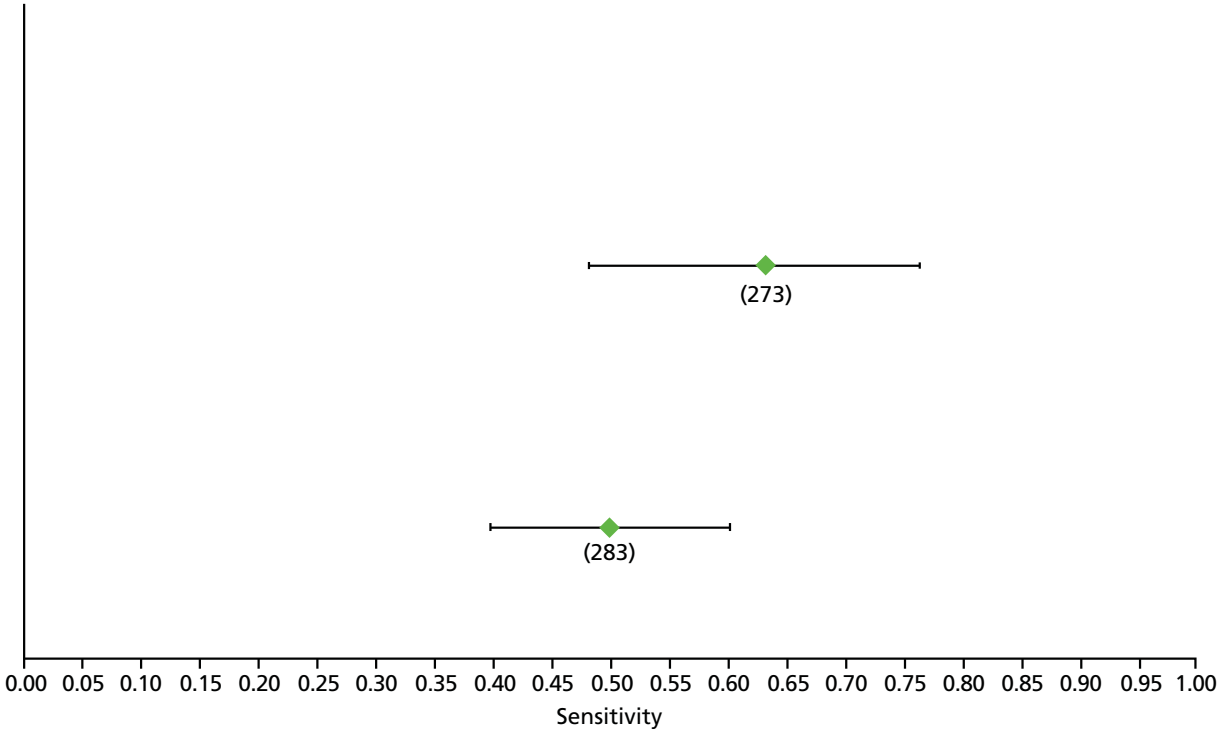

FIGURE 11 Alcoholic liver disease: APRI F2 (high cut-off) – sensitivity.

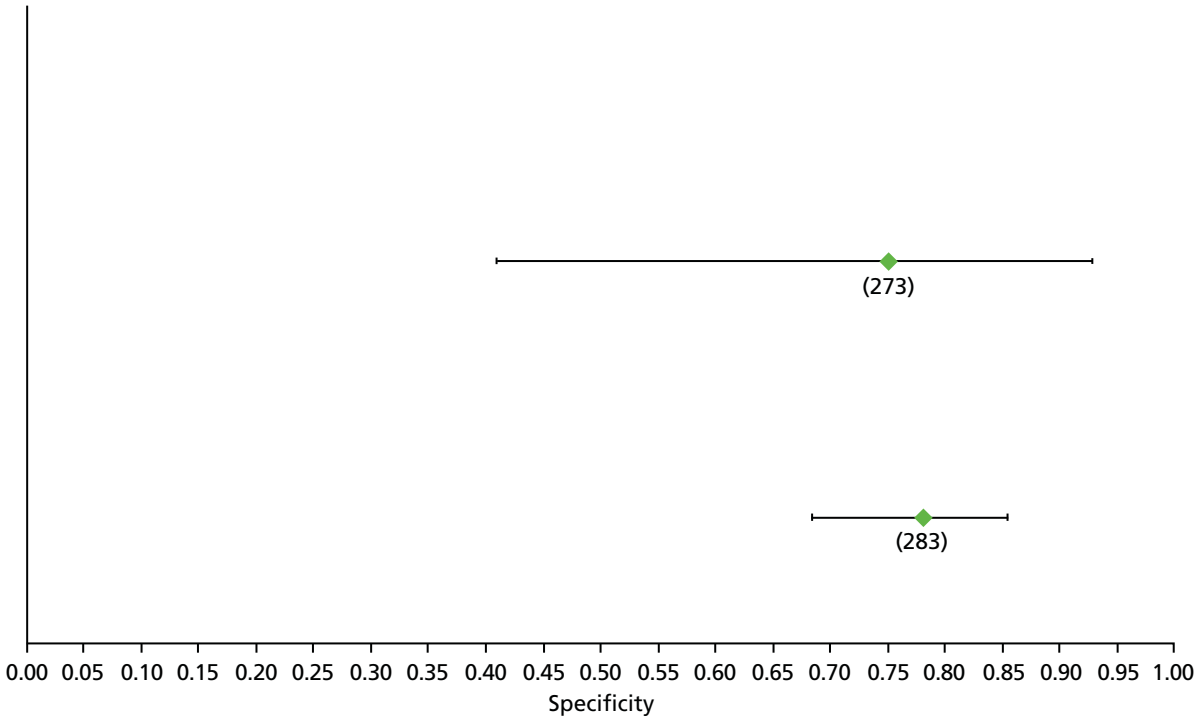

FIGURE 12 Alcoholic liver disease: APRI F2 (high cut-off) – specificity.

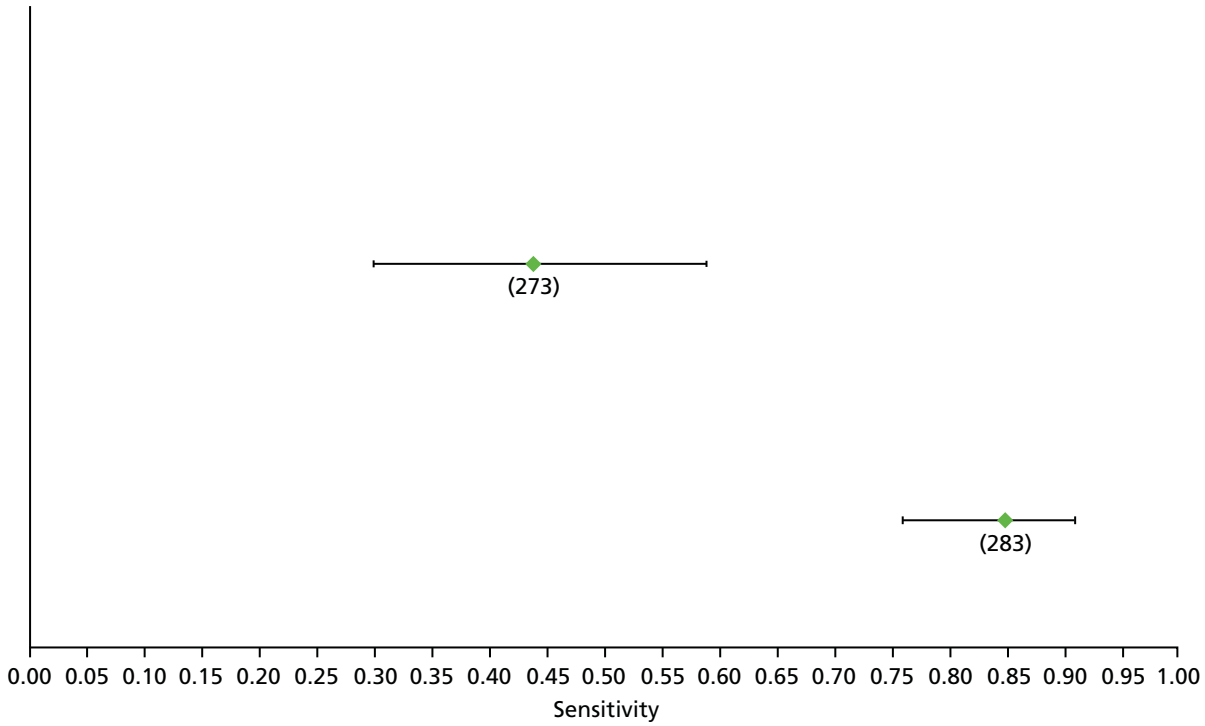

FIGURE 13 Alcoholic liver disease: APRI F2 (low cut-off) – sensitivity.

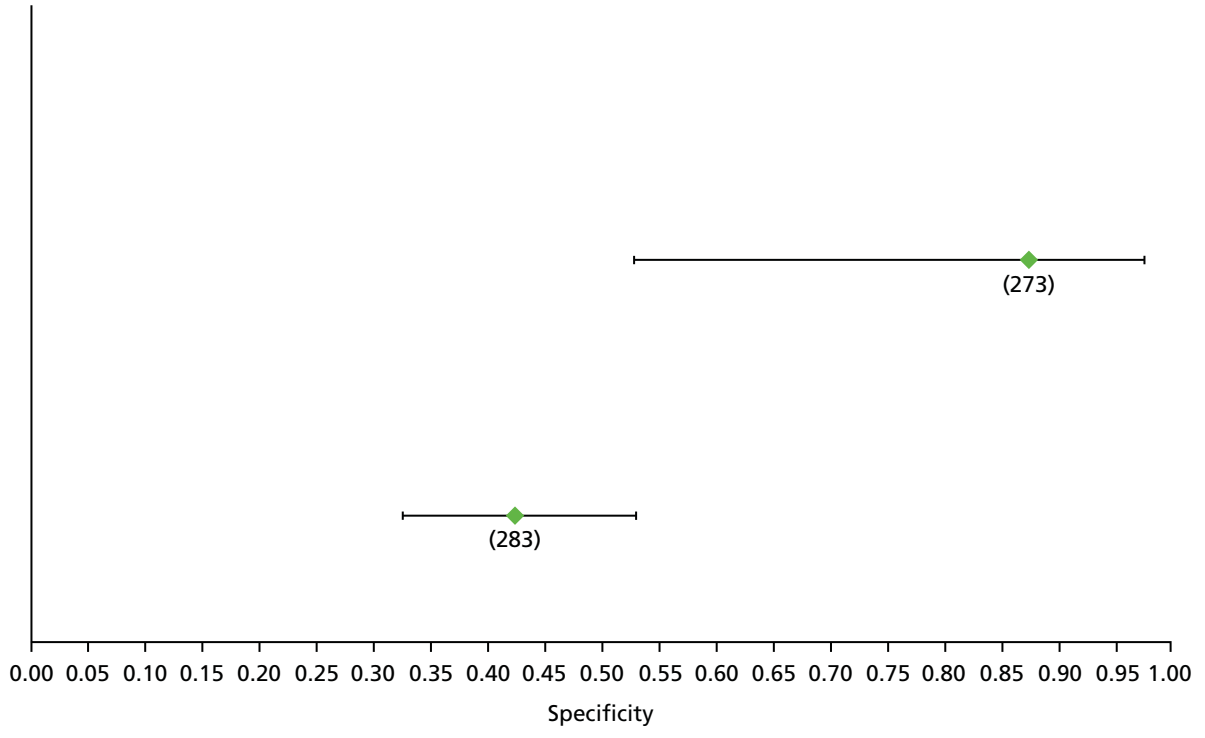

FIGURE 14 Alcoholic liver disease: APRI F2 (low cut-off) – specificity.

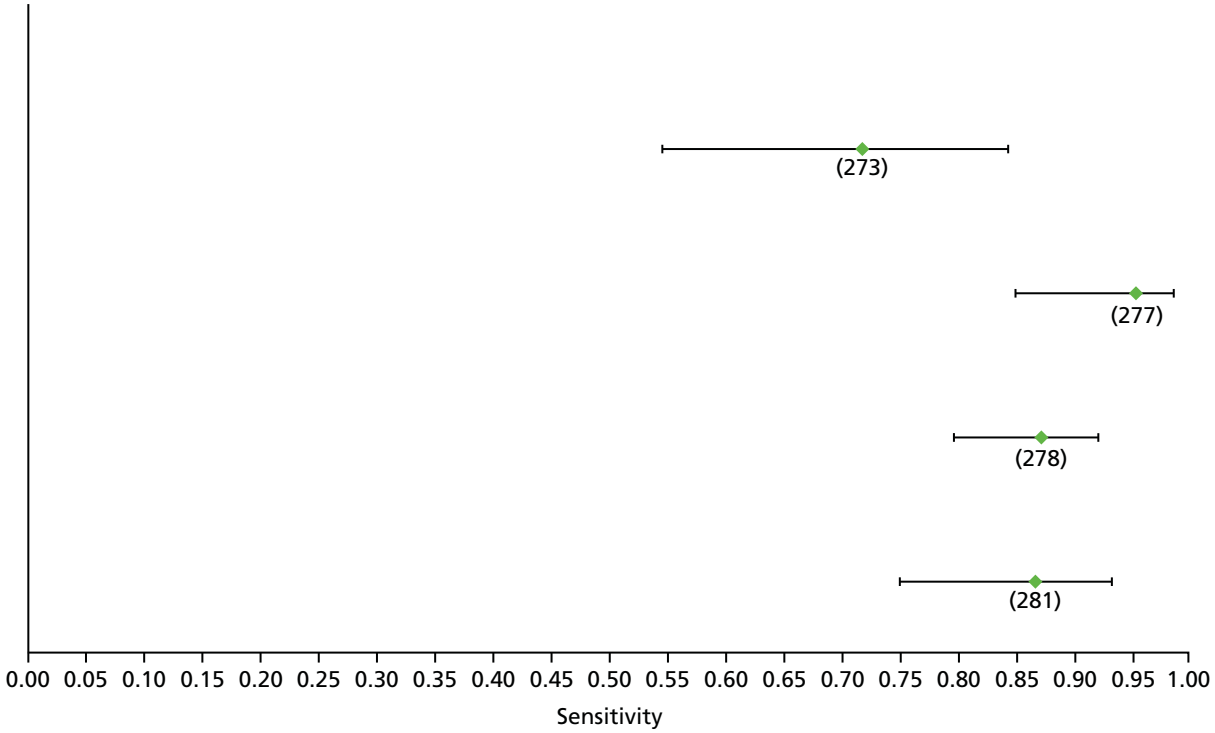

FIGURE 15 Alcoholic liver disease: TE F3 – sensitivity.

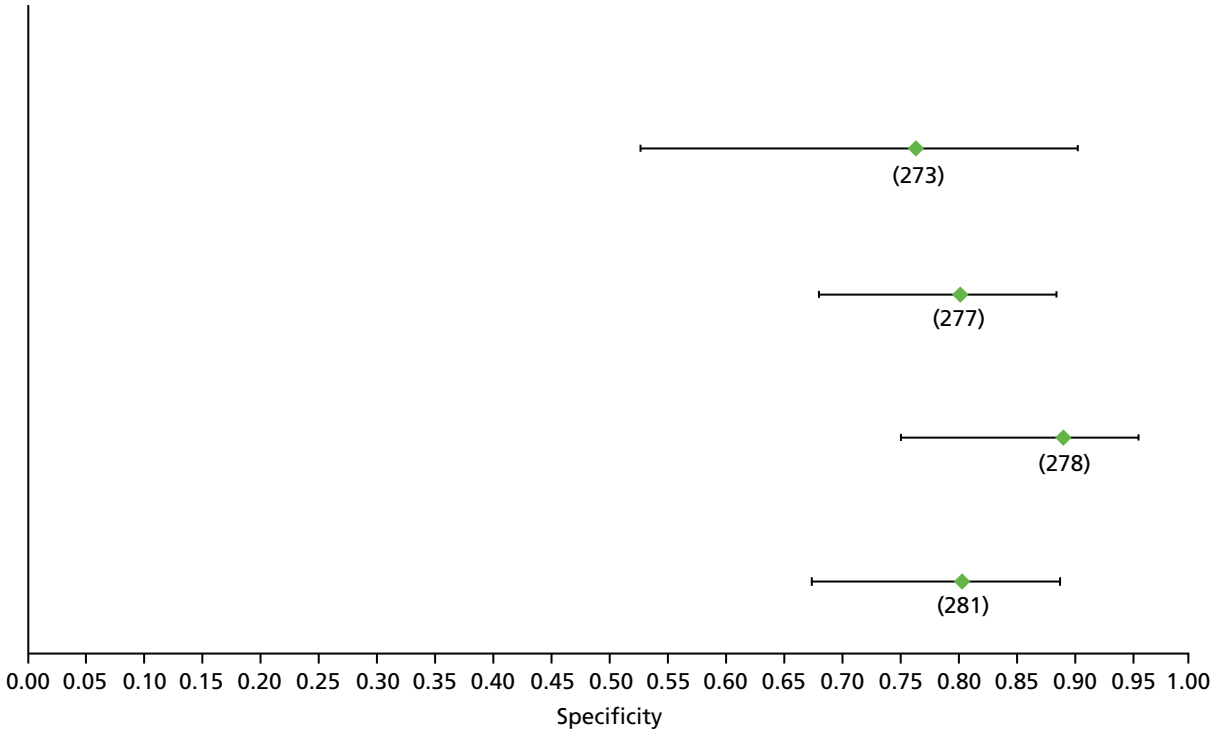

FIGURE 16 Alcoholic liver disease: TE F3 – specificity.

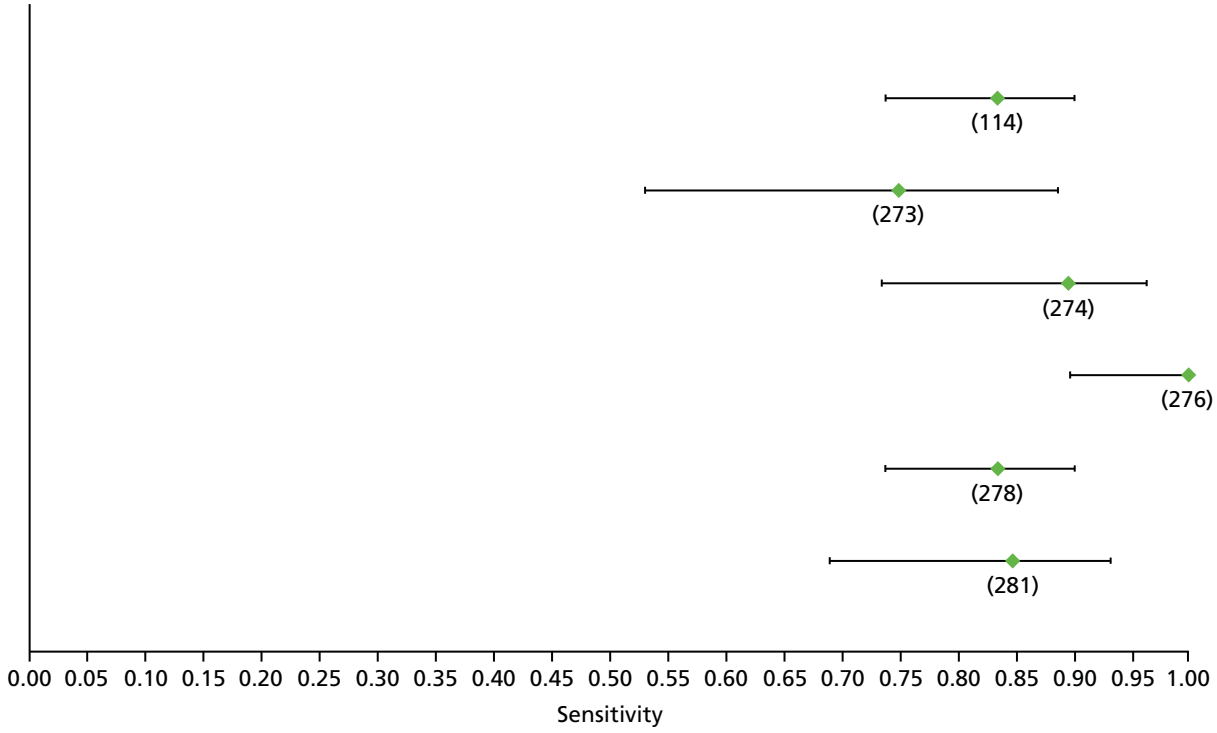

FIGURE 17 Alcoholic liver disease: TE F4 – sensitivity.

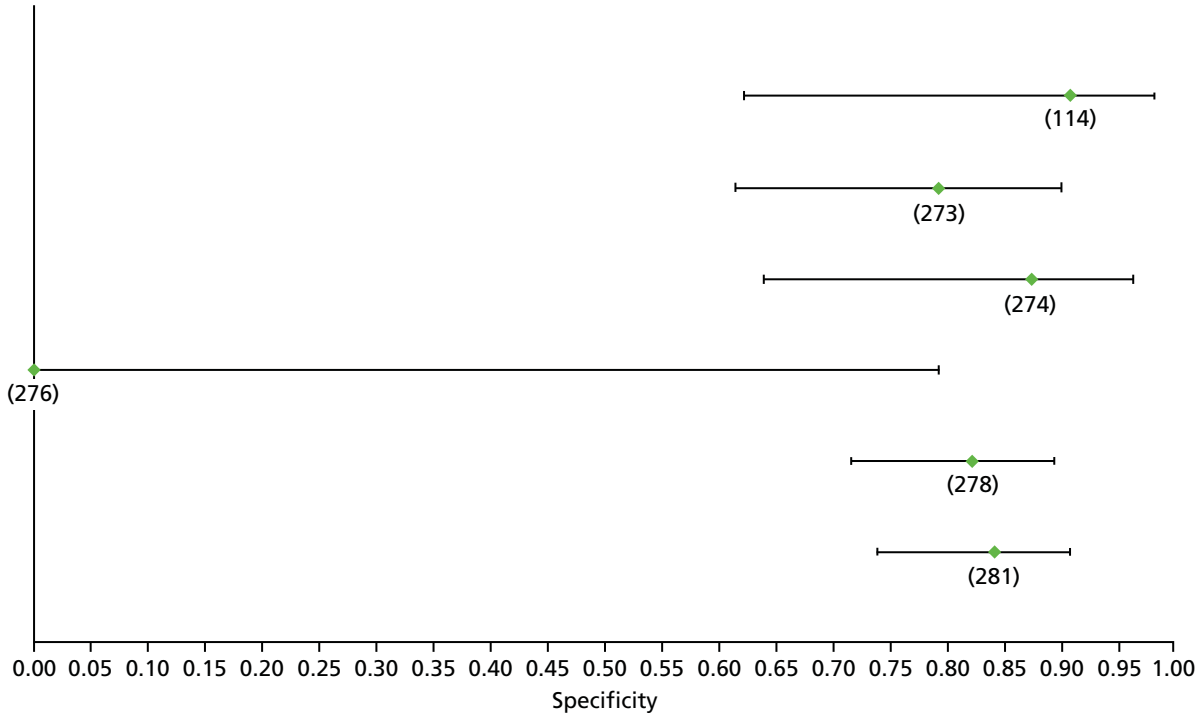

FIGURE 18 Alcoholic liver disease: TE F4 – specificity.

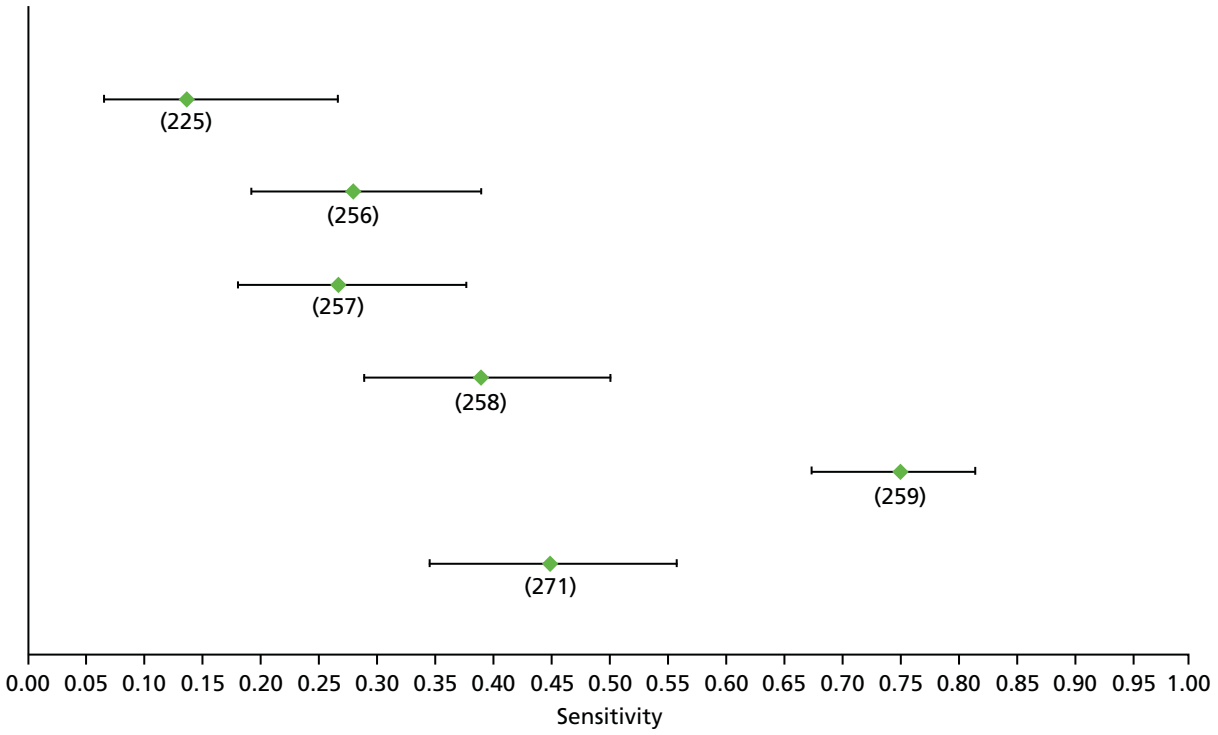

FIGURE 19 Hepatitis B: APRI F2 (high cut-off) – sensitivity.

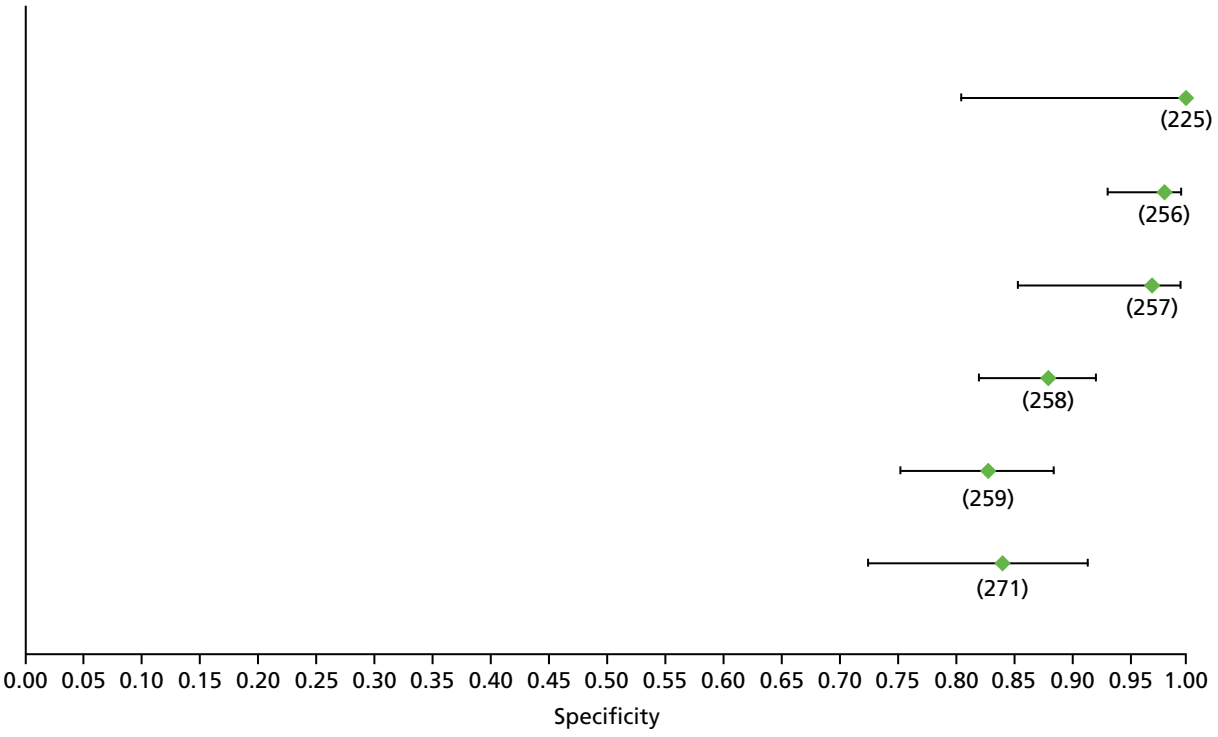

FIGURE 20 Hepatitis B: APRI F2 (high cut-off) – specificity.

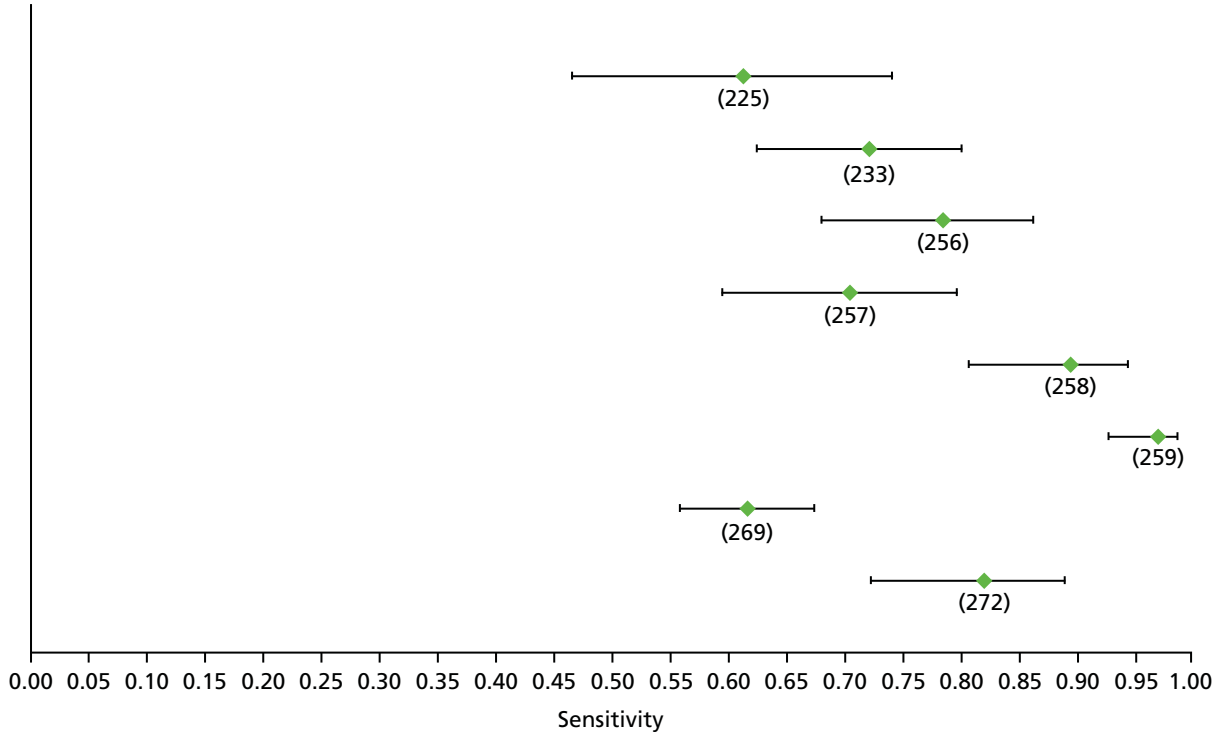

FIGURE 21 Hepatitis B: APRI F2 (low cut-off) – sensitivity.

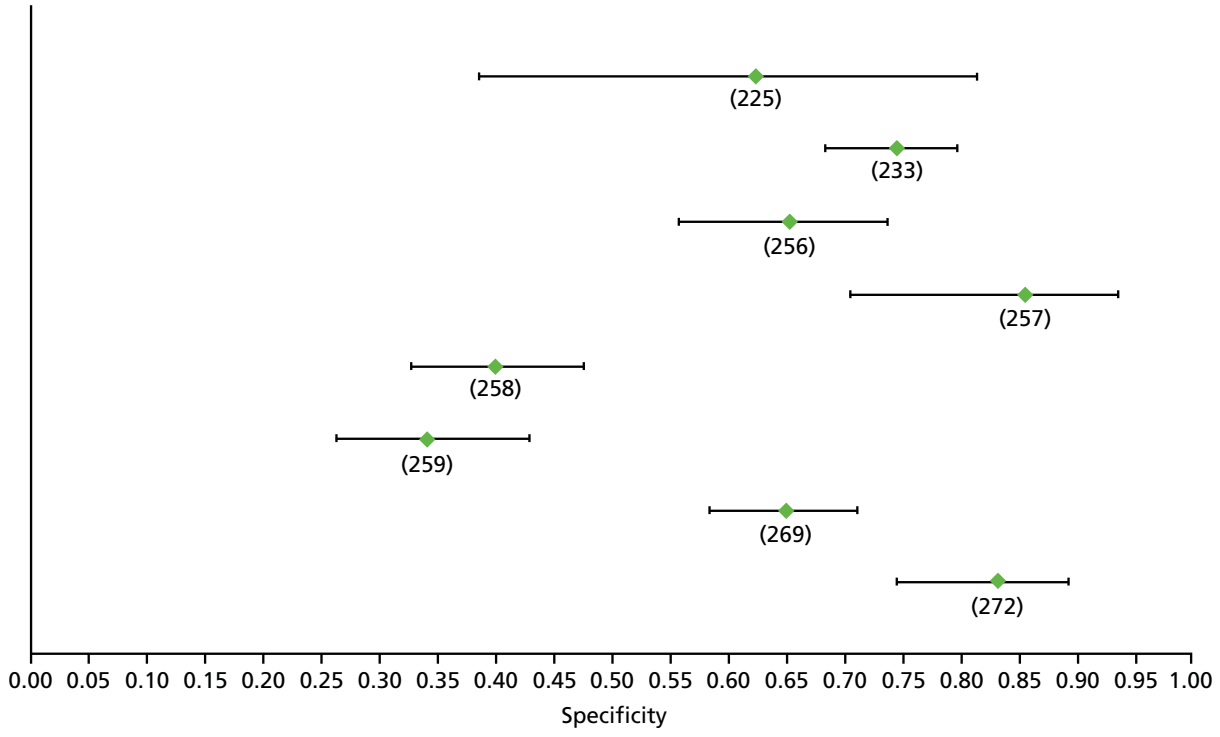

FIGURE 22 Hepatitis B: APRI F2 (low cut-off) – specificity.

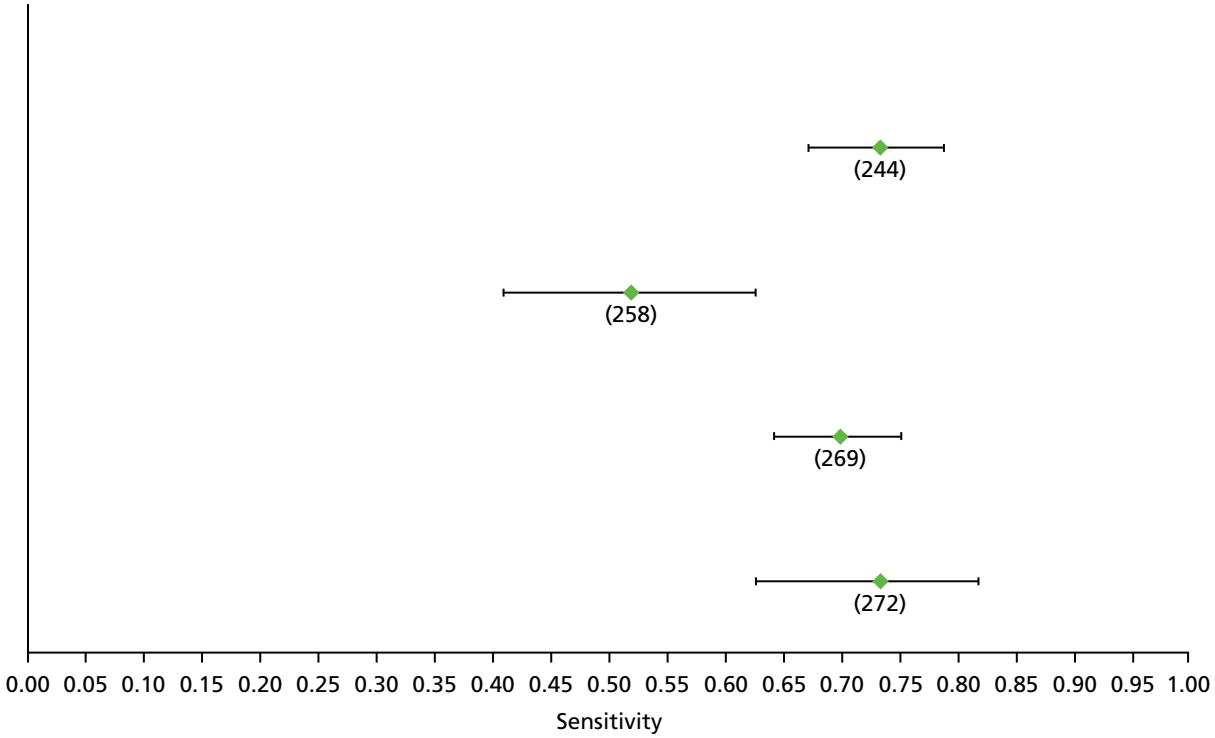

FIGURE 23 Hepatitis B: FIB-4 F2 (low cut-off) – sensitivity.

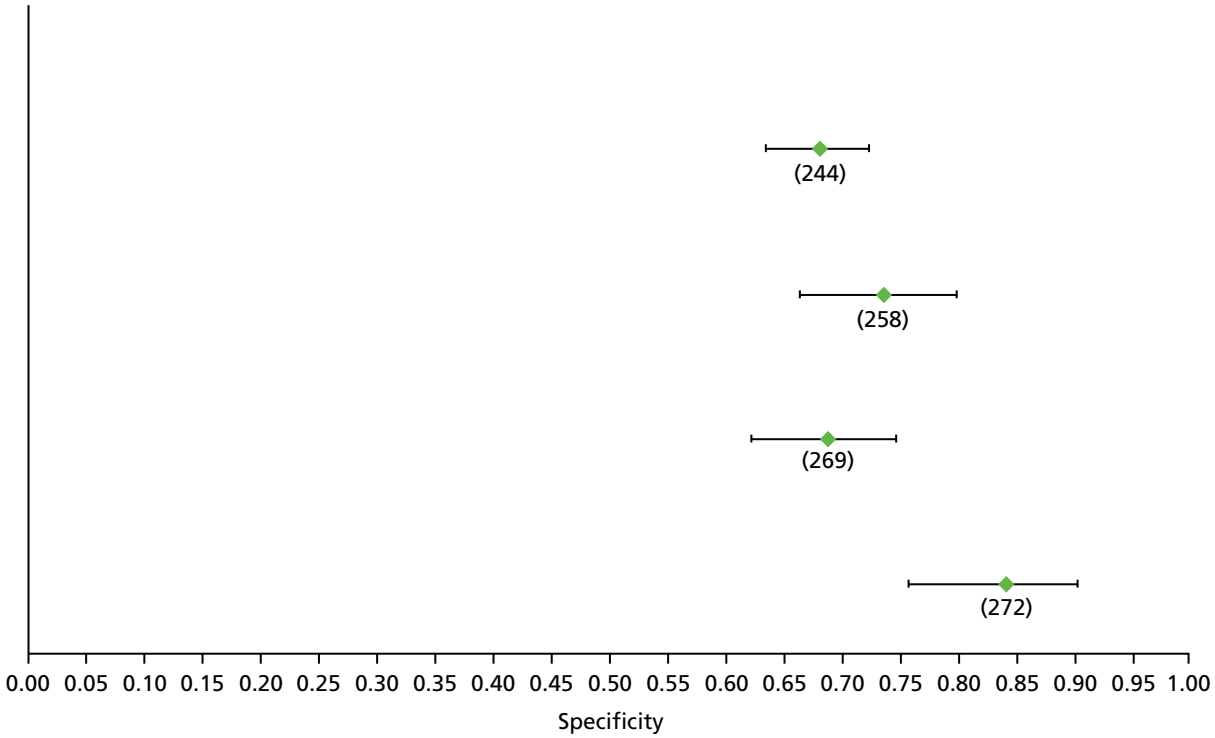

FIGURE 24 Hepatitis B: FIB-4 F2 (low cut-off) – specificity.

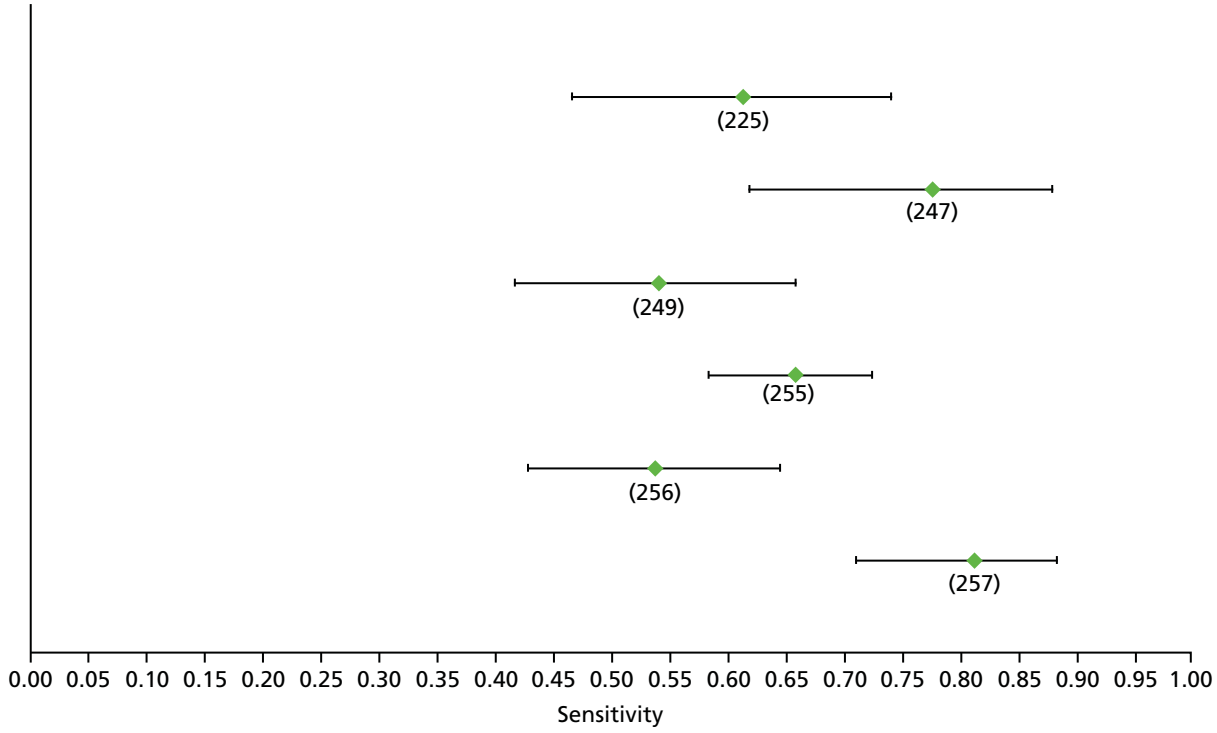

FIGURE 25 Hepatitis B: Fibrotest F2 – sensitivity.

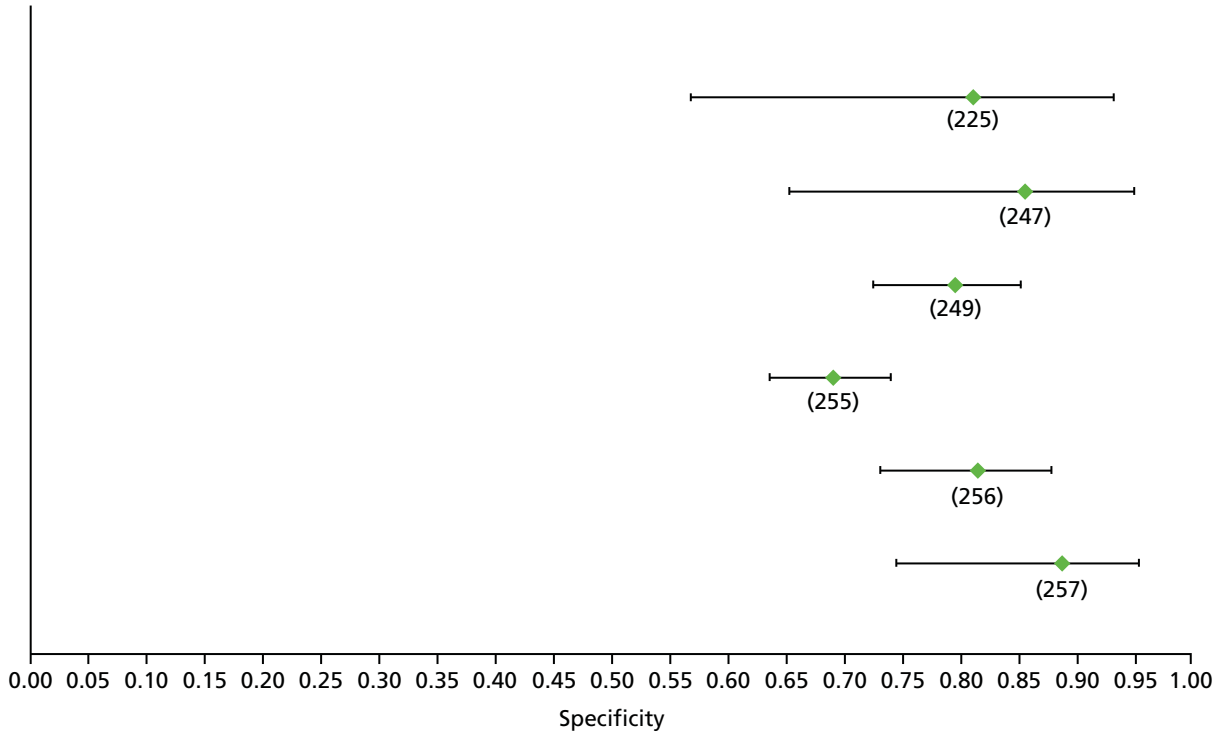

FIGURE 26 Hepatitis B: Fibrotest F2 – specificity.

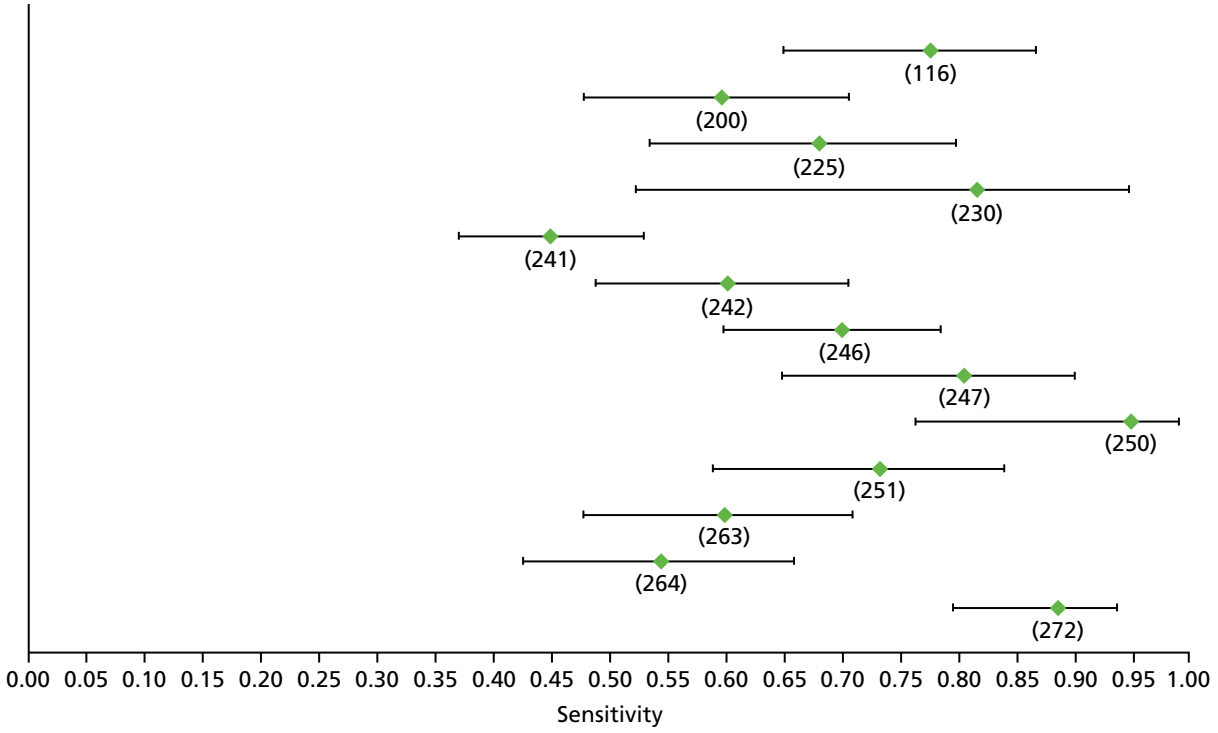

FIGURE 27 Hepatitis B: TE F2 – sensitivity.

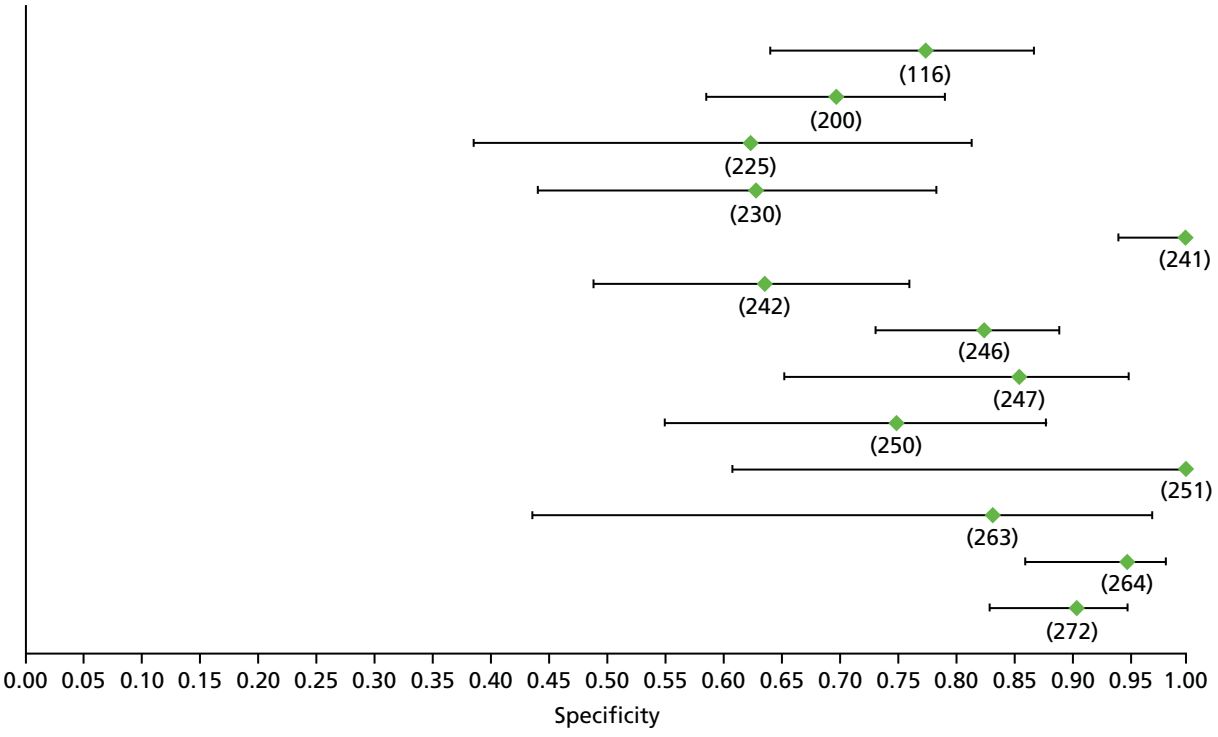

FIGURE 28 Hepatitis B: TE F2 – specificity.

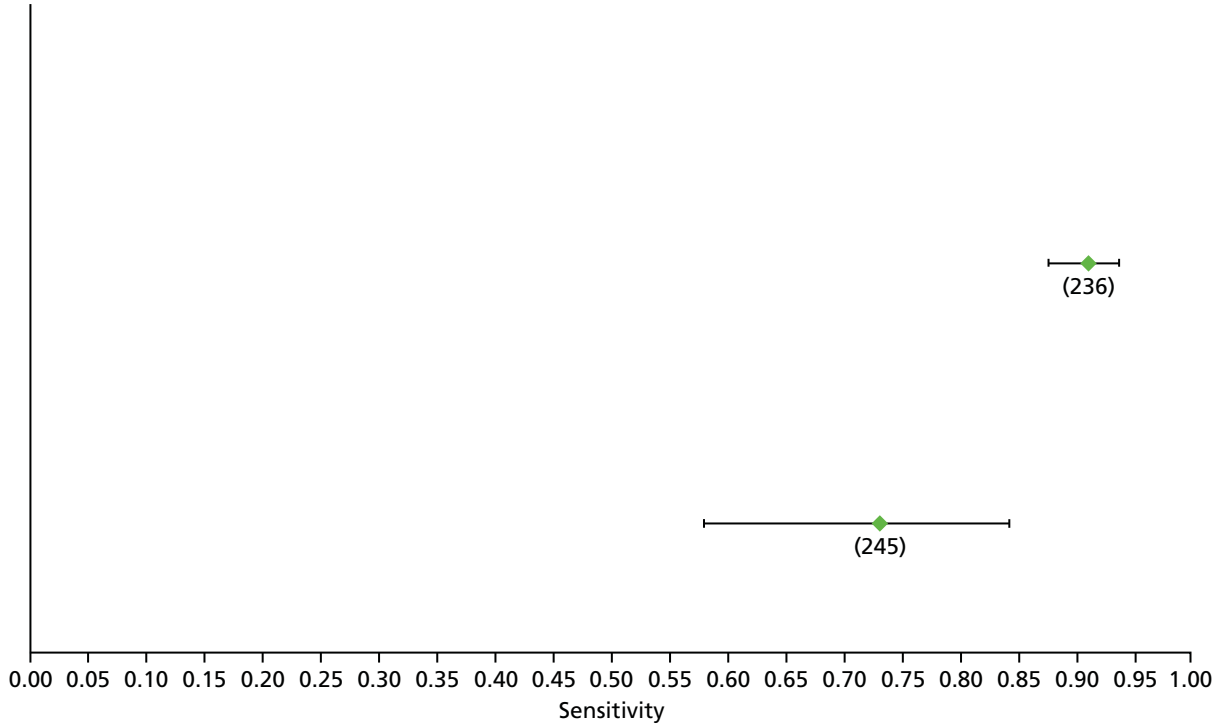

FIGURE 29 Hepatitis B: FIB-4 F3 (low cut-off) – sensitivity.

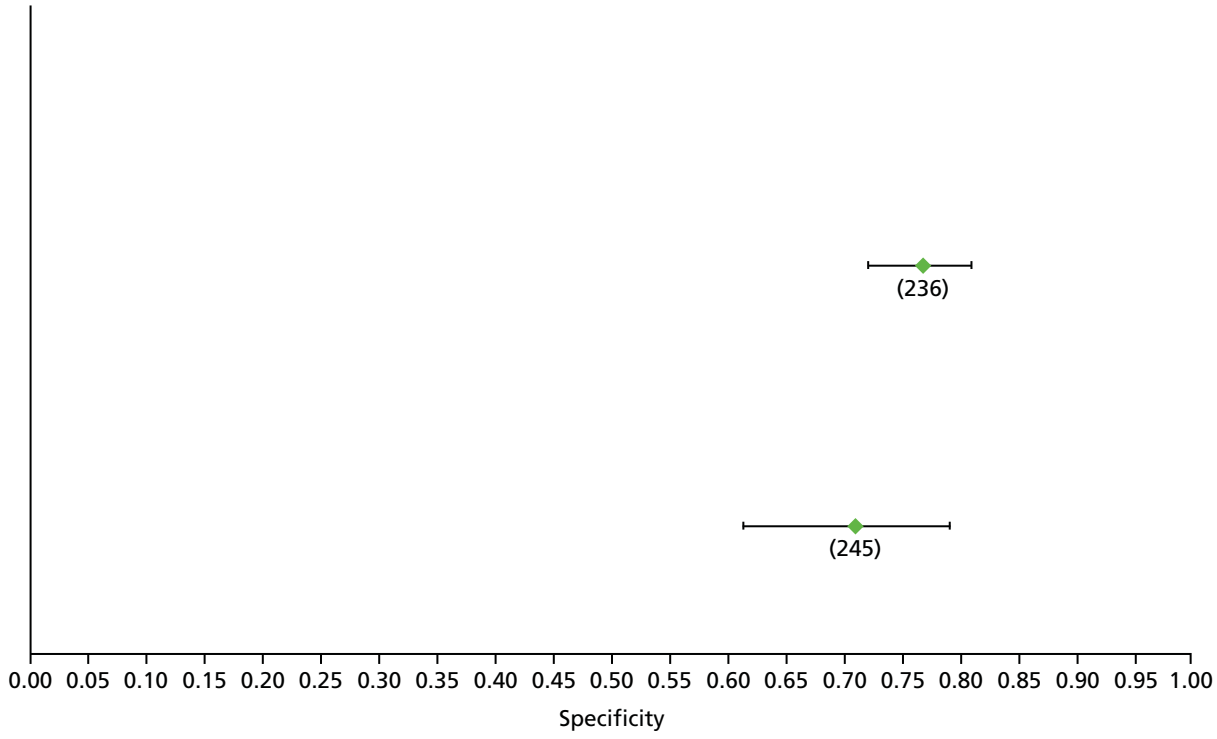

FIGURE 30 Hepatitis B: FIB-4 F3 (low cut-off) – specificity.

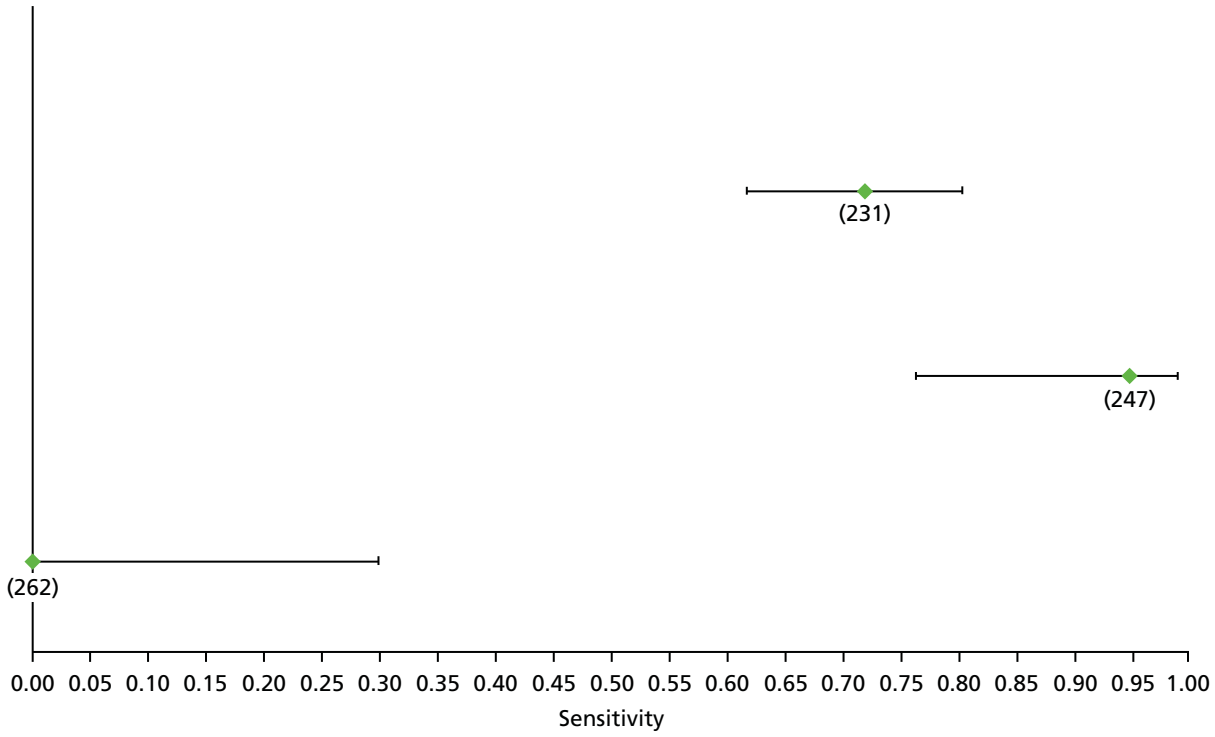

FIGURE 31 Hepatitis B: Fibrotest F3 – sensitivity.

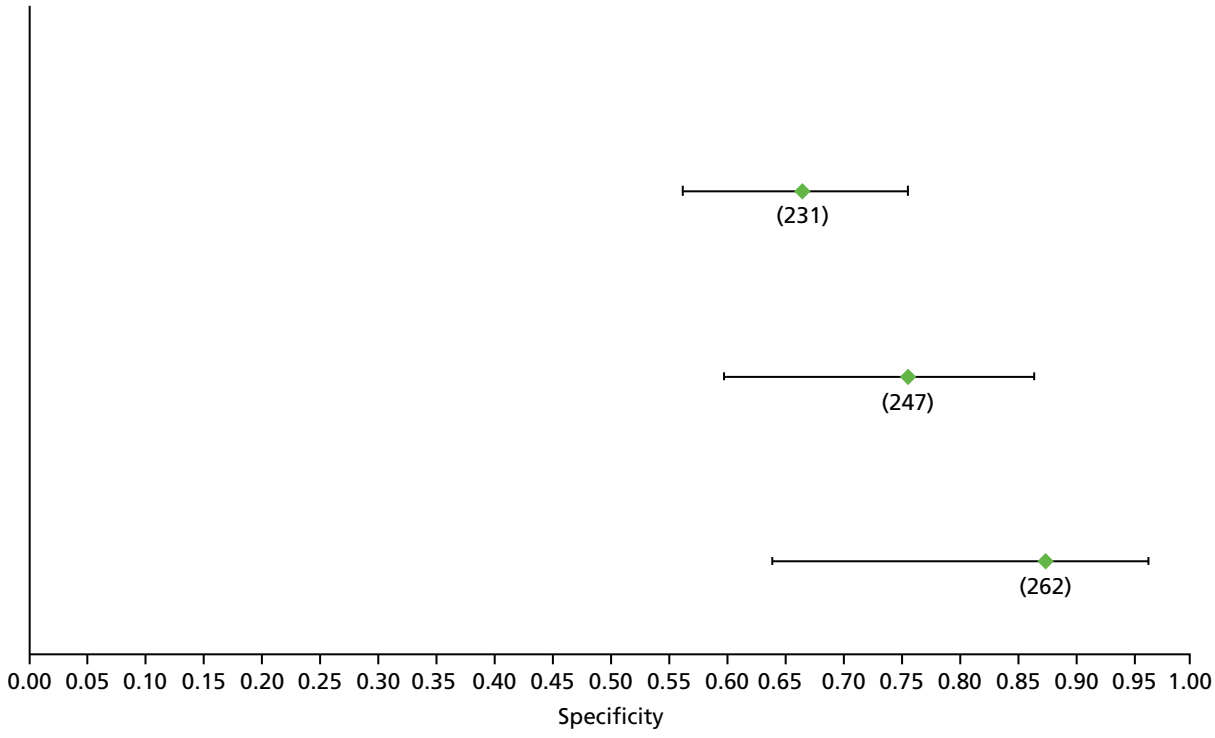

FIGURE 32 Hepatitis B: Fibrotest F3 – specificity.

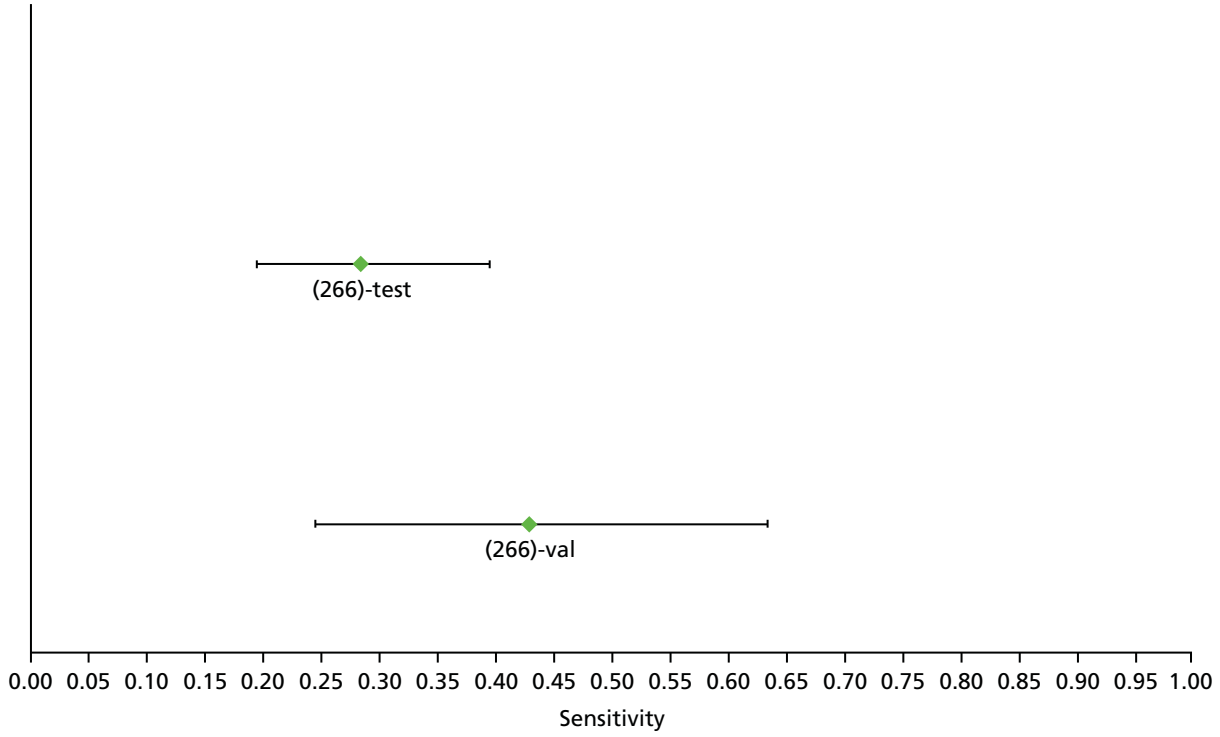

FIGURE 33 Hepatitis B: Forns index F3 (high cut-off) – sensitivity. Val, validation cohort.

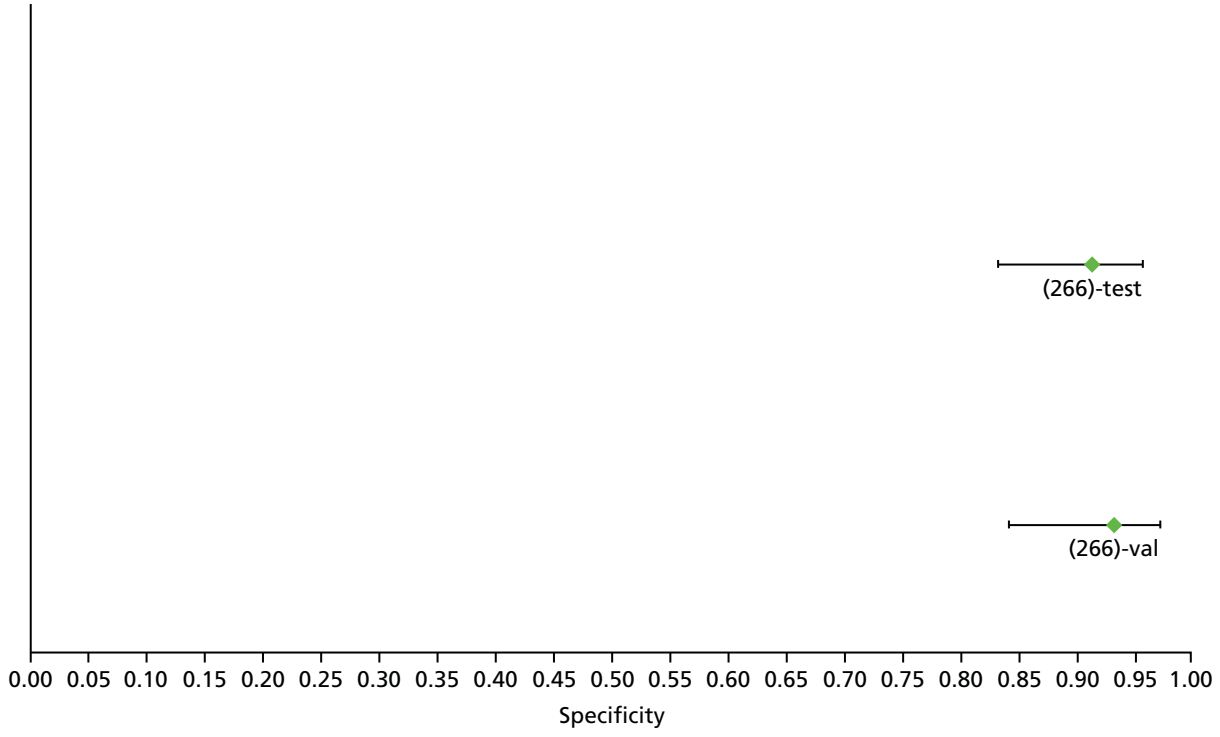

FIGURE 34 Hepatitis B: Forns index F3 (high cut-off) – specificity. Val, validation cohort.

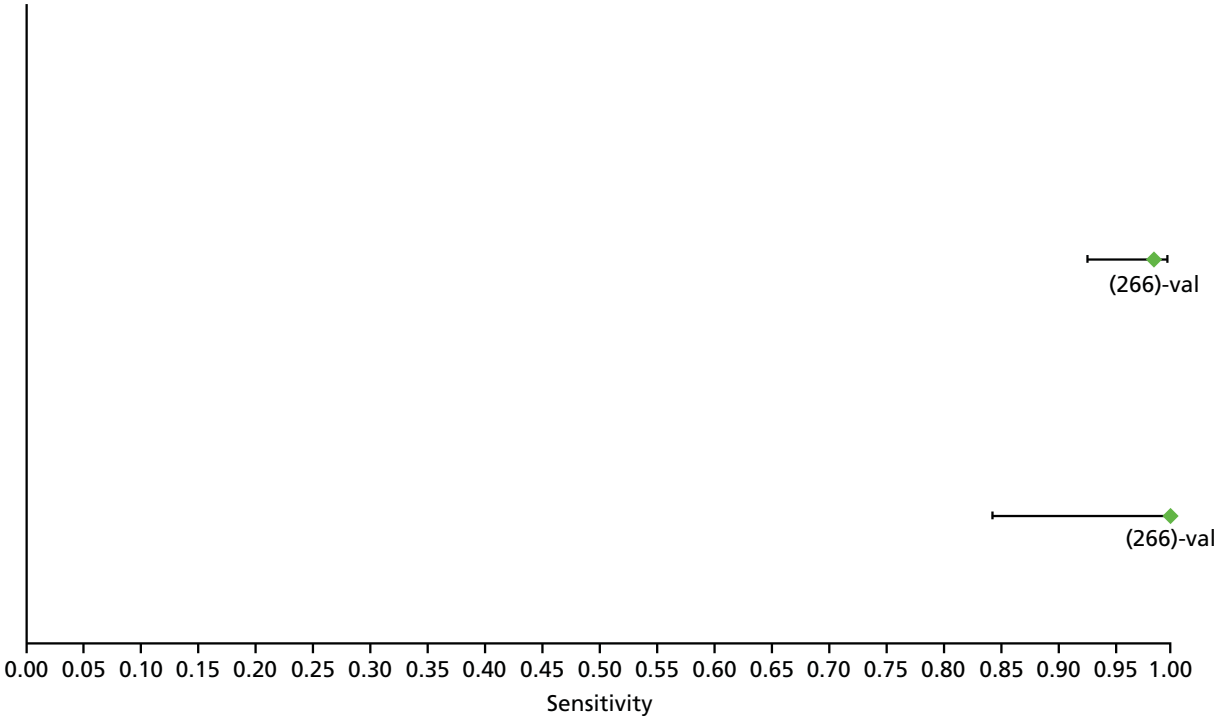

FIGURE 35 Hepatitis B: Forns index F3 (low cut-off) – sensitivity. Val, validation cohort.

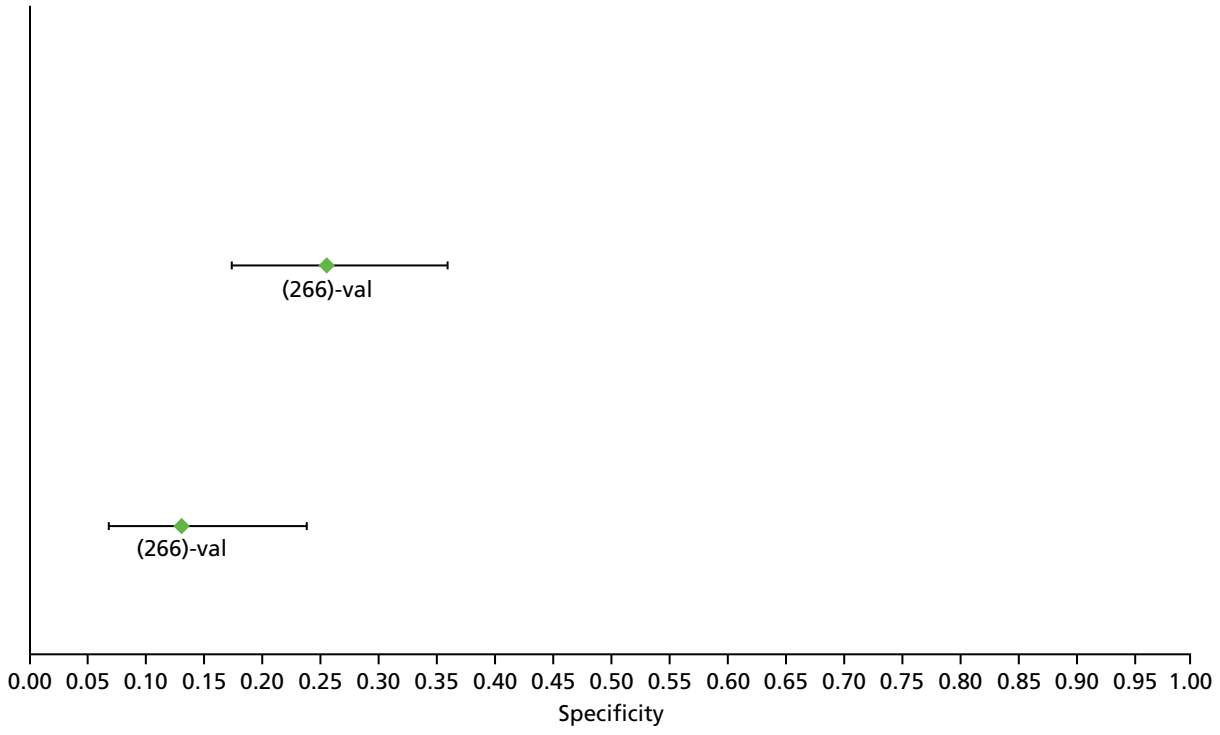

FIGURE 36 Hepatitis B: Forns index F3 (low cut-off) – specificity. Val, validation cohort.

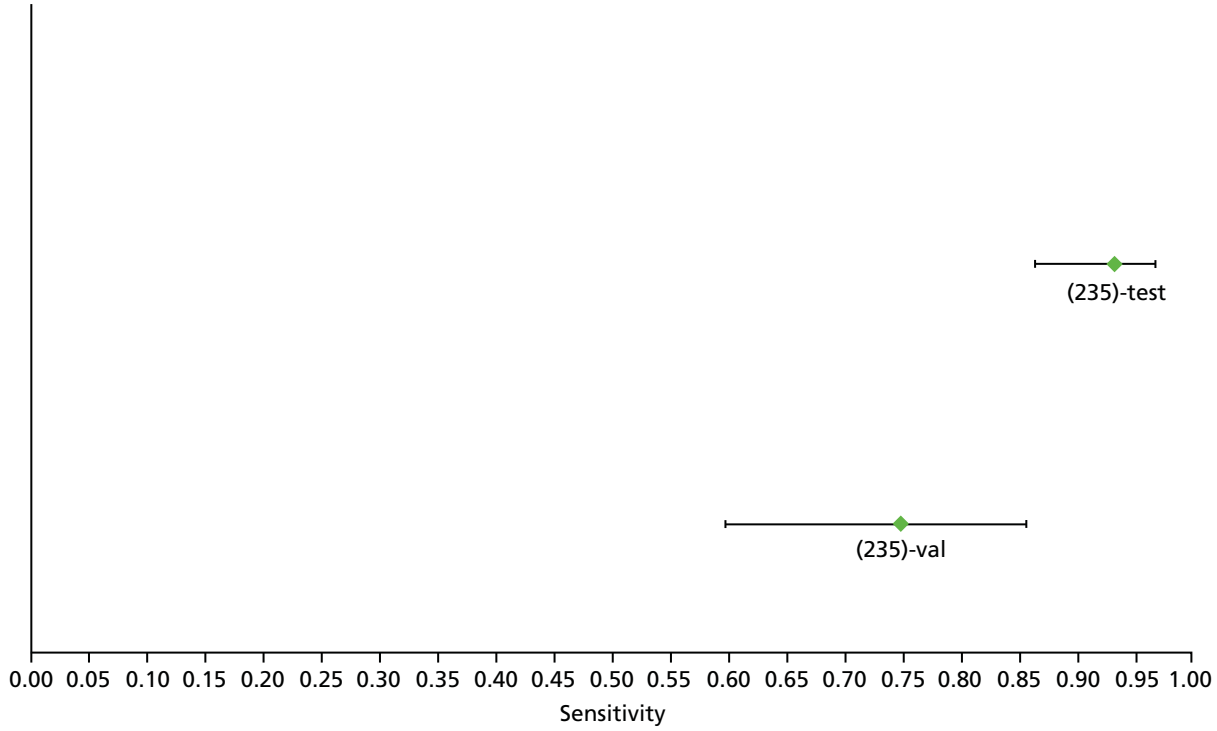

FIGURE 37 Hepatitis B: Hui index F3 – sensitivity. Val, validation cohort.

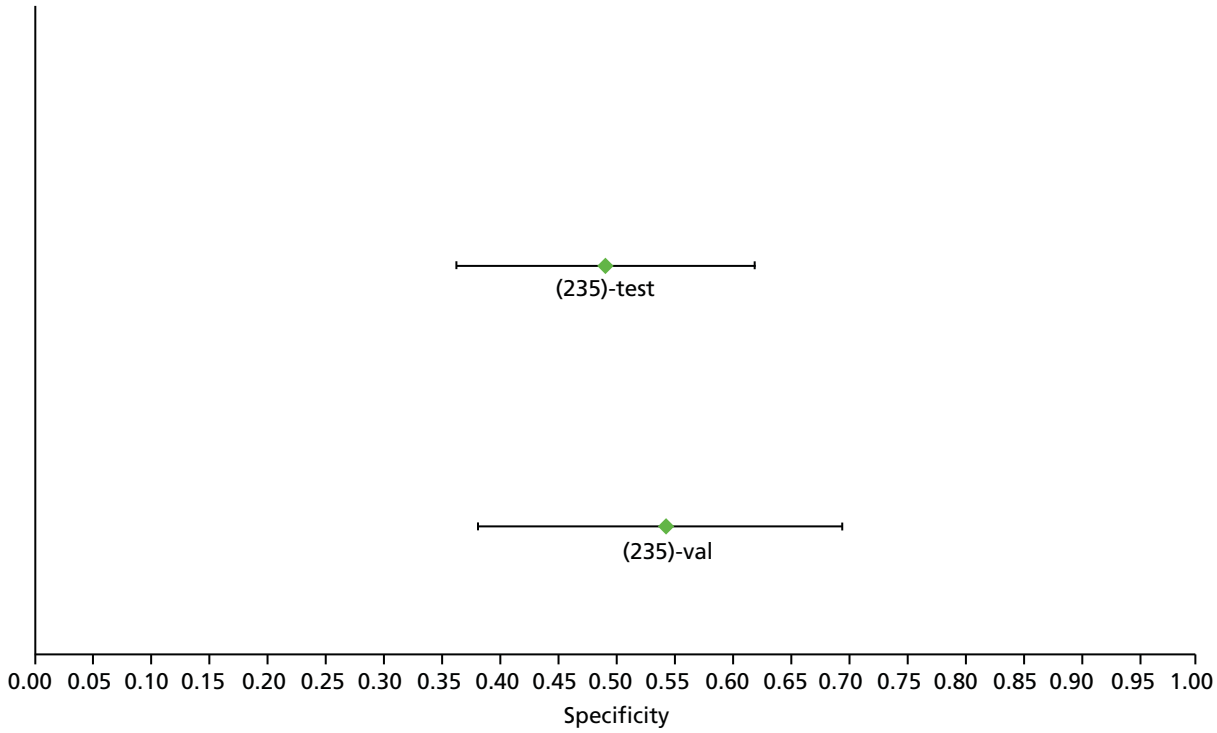

FIGURE 38 Hepatitis B: Hui index F3 – specificity. Val, validation cohort.

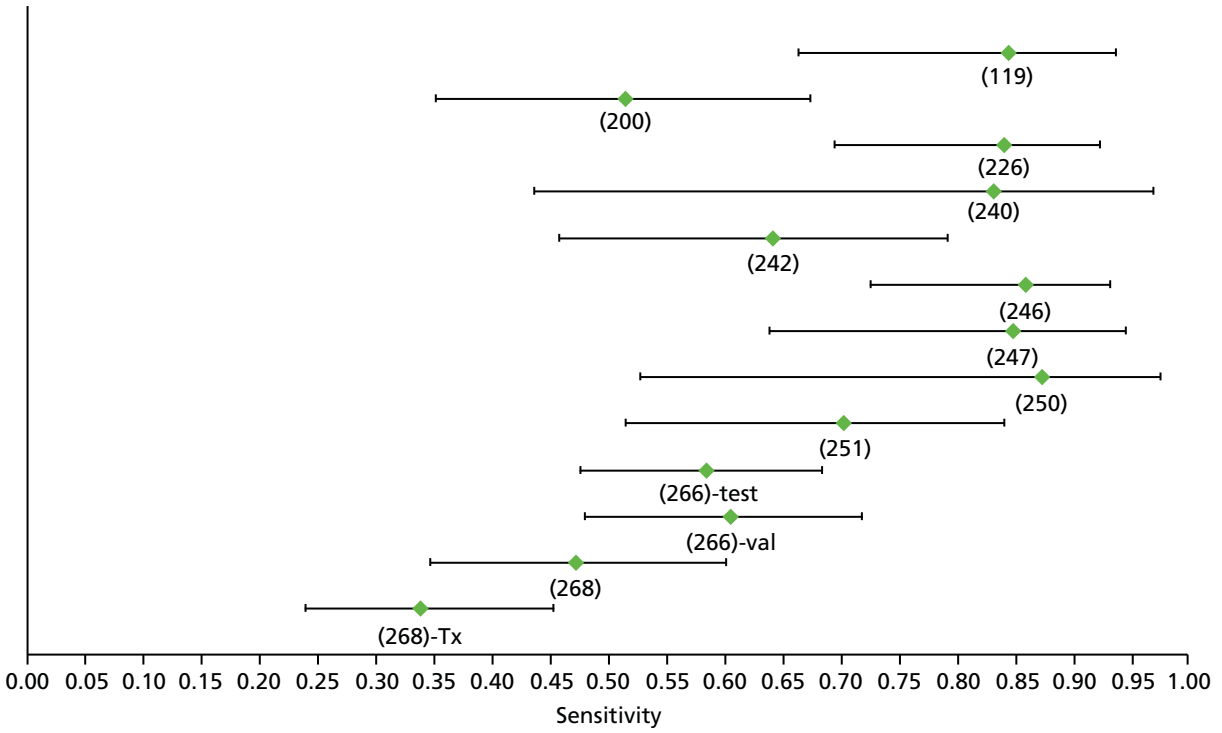

FIGURE 39 Hepatitis B: TE F3 – sensitivity. Tx, treatment; val, validation cohort.

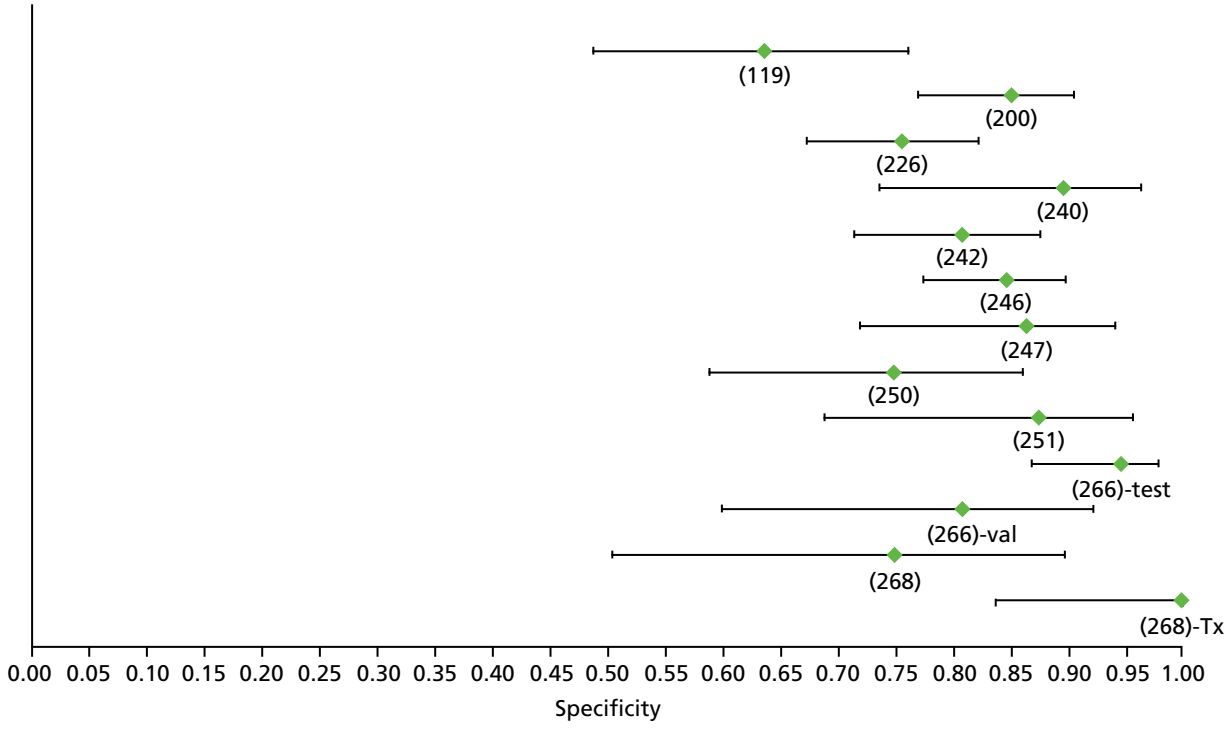

FIGURE 40 Hepatitis B: TE F3 – specificity. Tx, treatment; val, validation cohort.

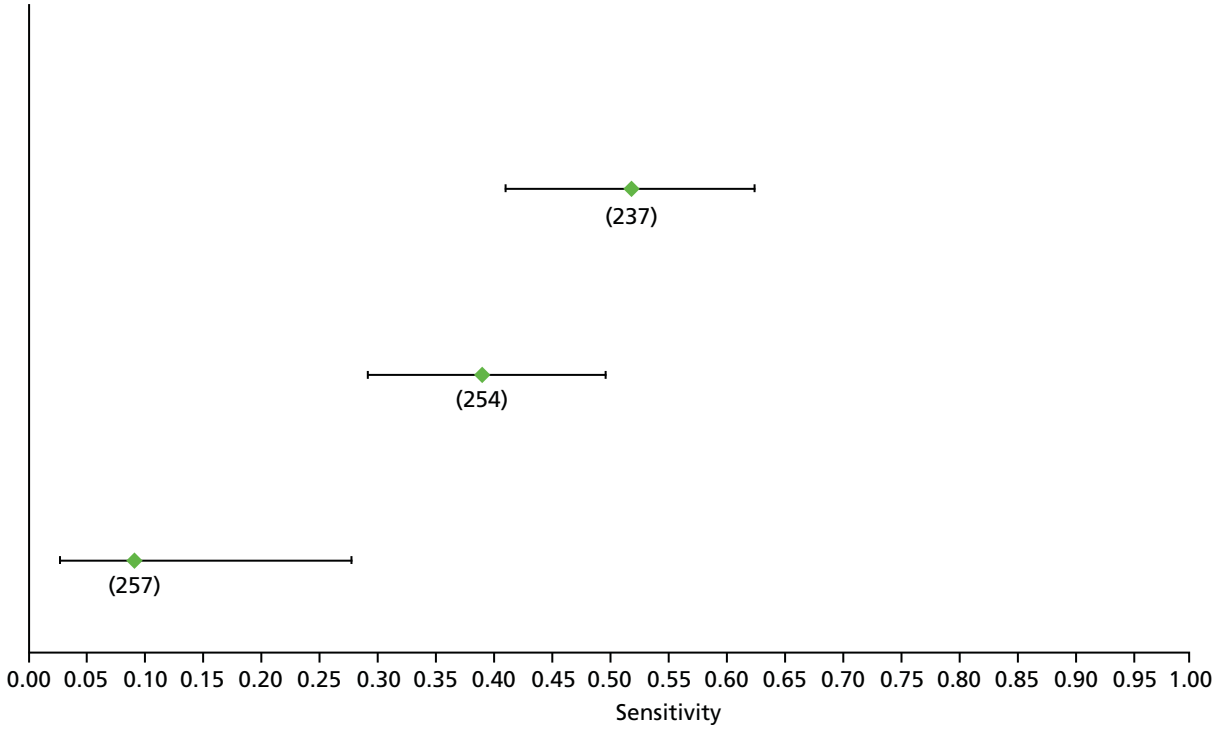

FIGURE 41 Hepatitis B: AST–ALT ratio F4 – sensitivity.

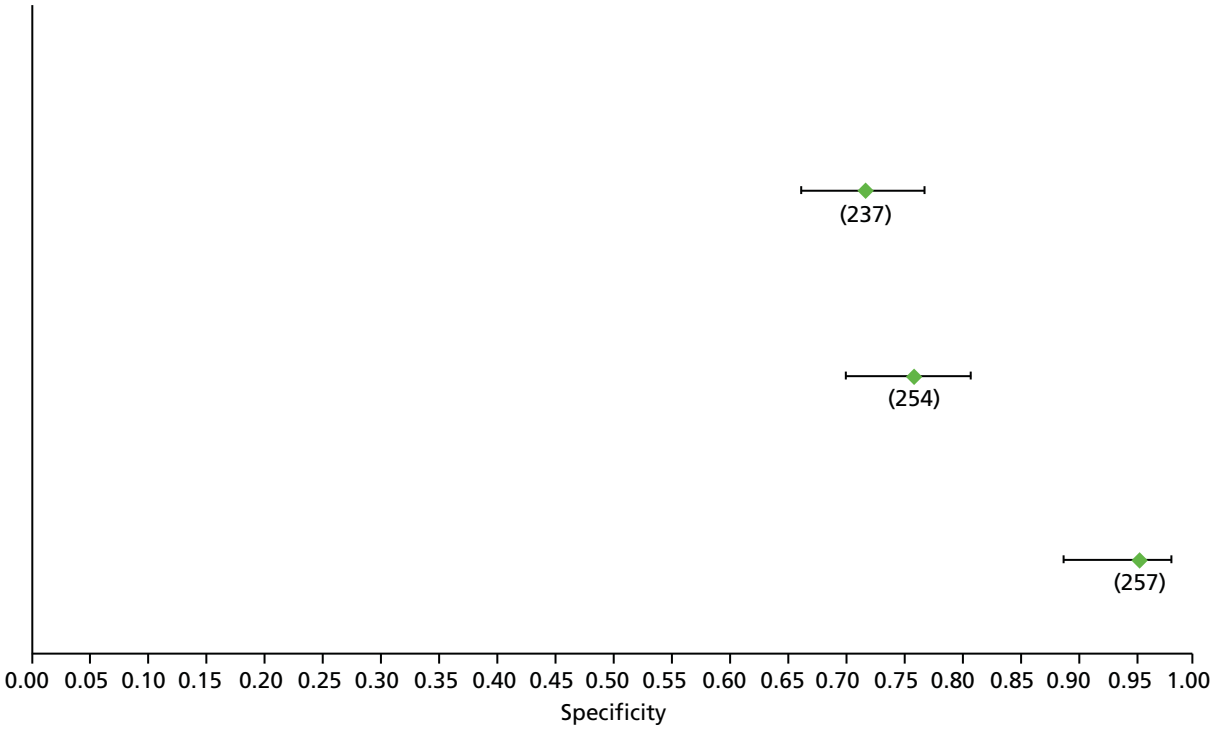

FIGURE 42 Hepatitis B: AST–ALT F4 – specificity.

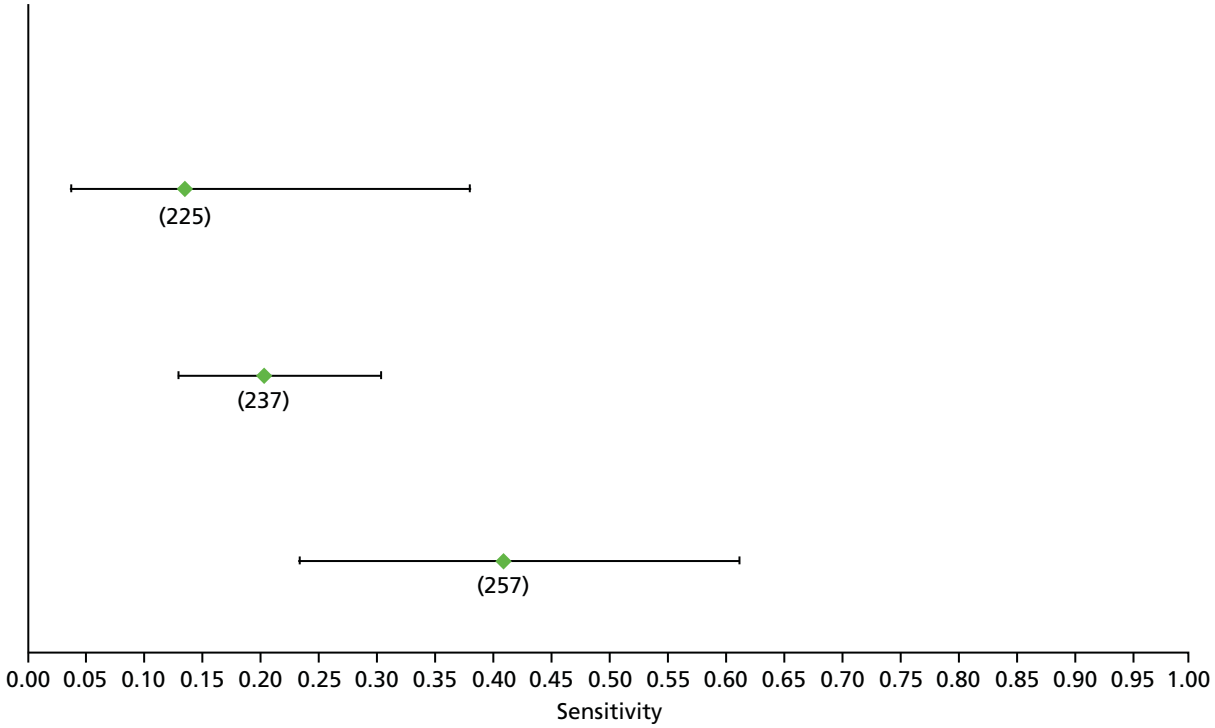

FIGURE 43 Hepatitis B: APRI F4 (high cut-off) – sensitivity.

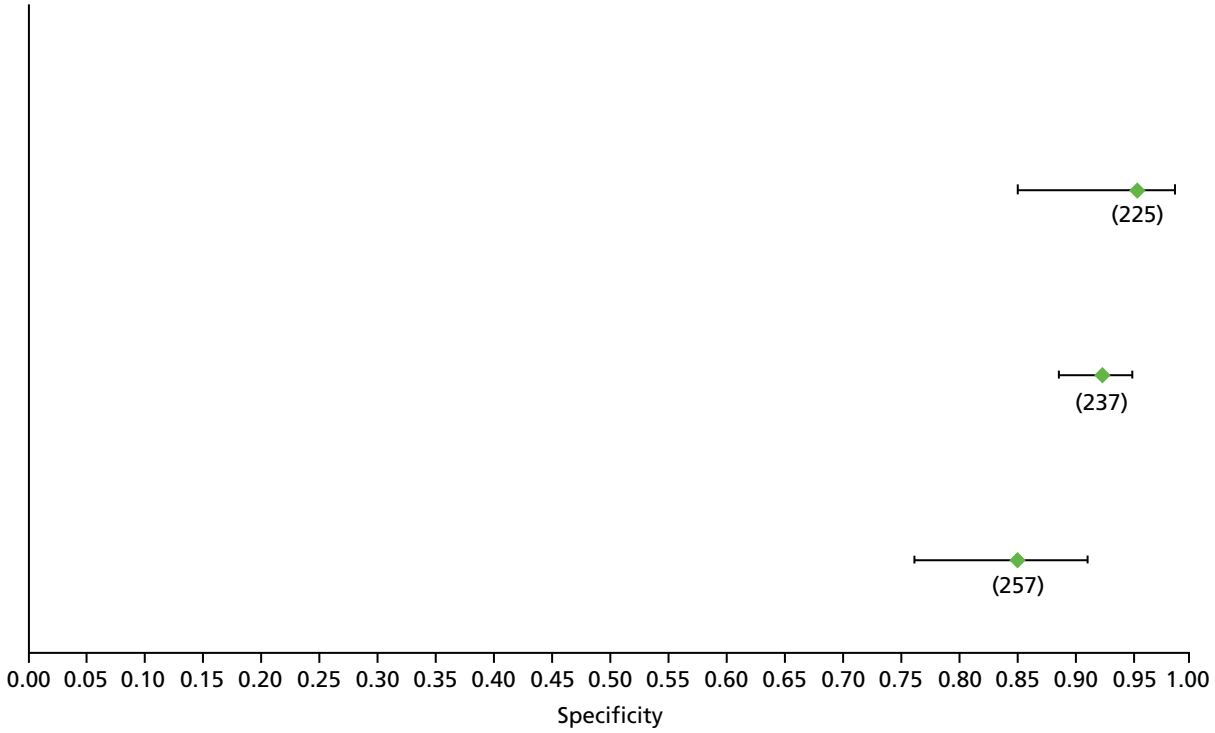

FIGURE 44 Hepatitis B: APRI F4 (high cut-off) – specificity.

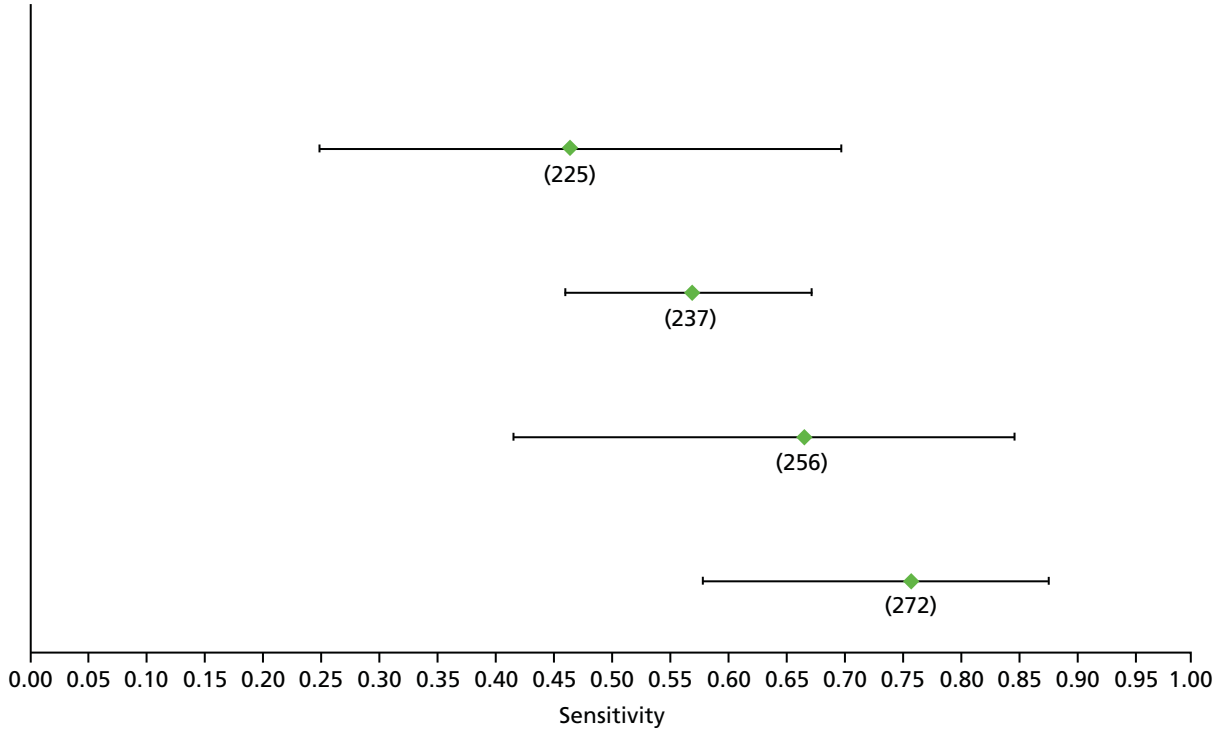

FIGURE 45 Hepatitis B: APRI F4 (low cut-off) – sensitivity.

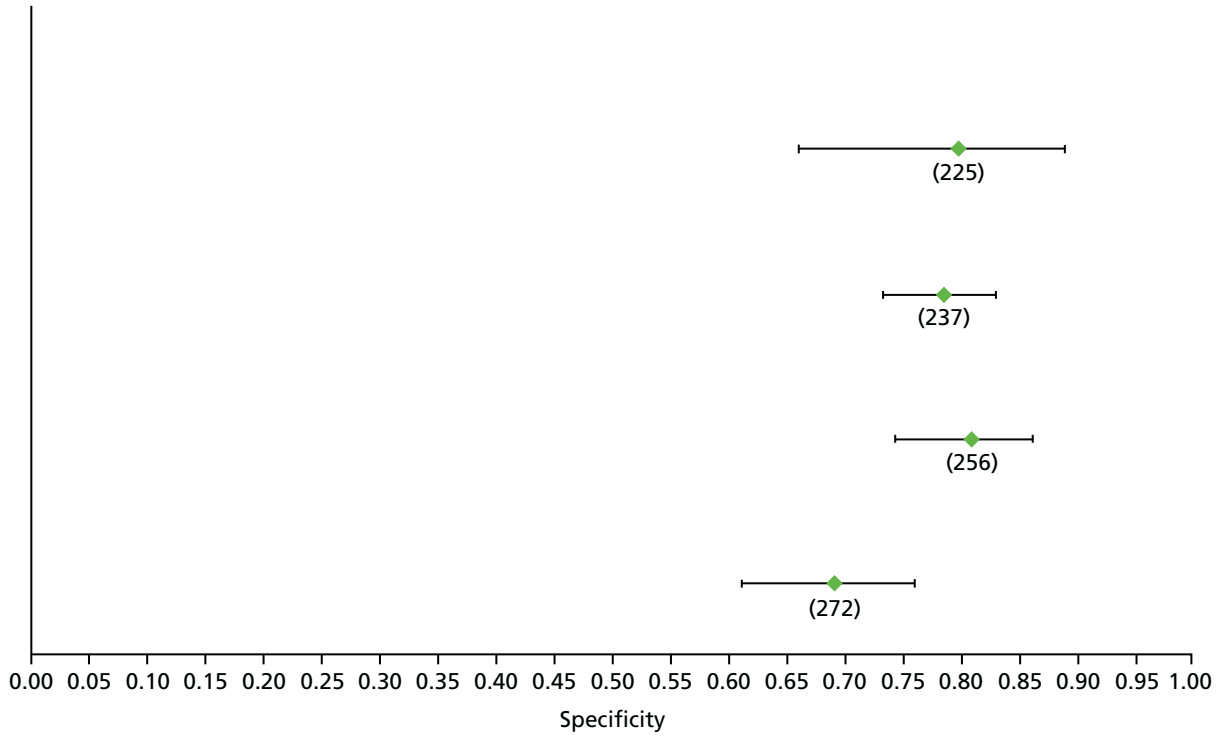

FIGURE 46 Hepatitis B: APRI F4 (low cut-off) – specificity.

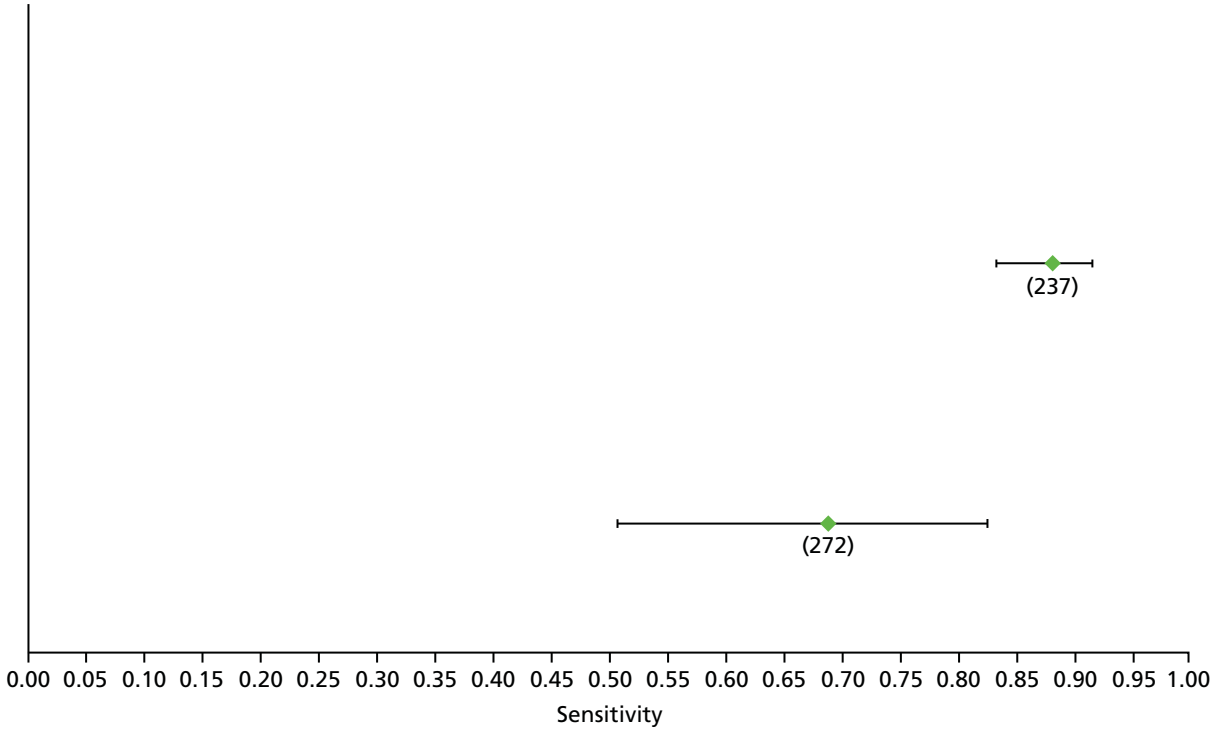

FIGURE 47 Hepatitis B: FIB-4 F4 (low cut-off) – sensitivity.

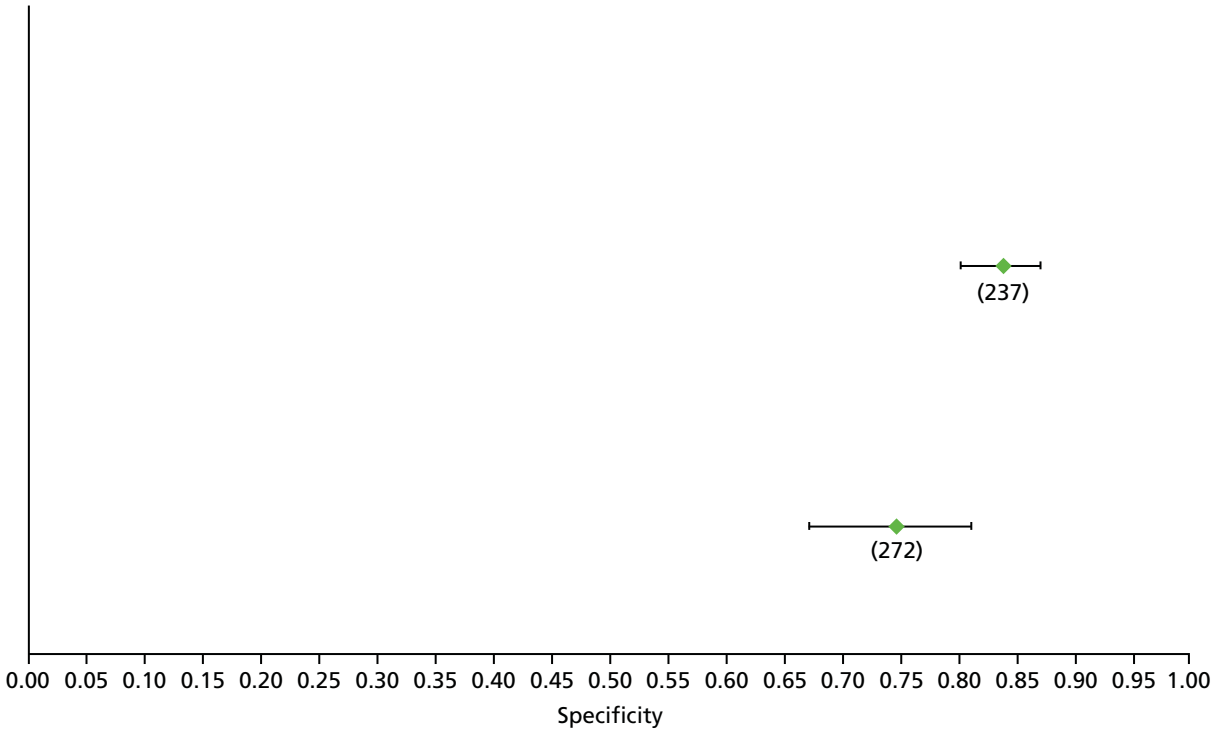

FIGURE 48 Hepatitis B: FIB-4 F4 (low cut-off) – specificity.

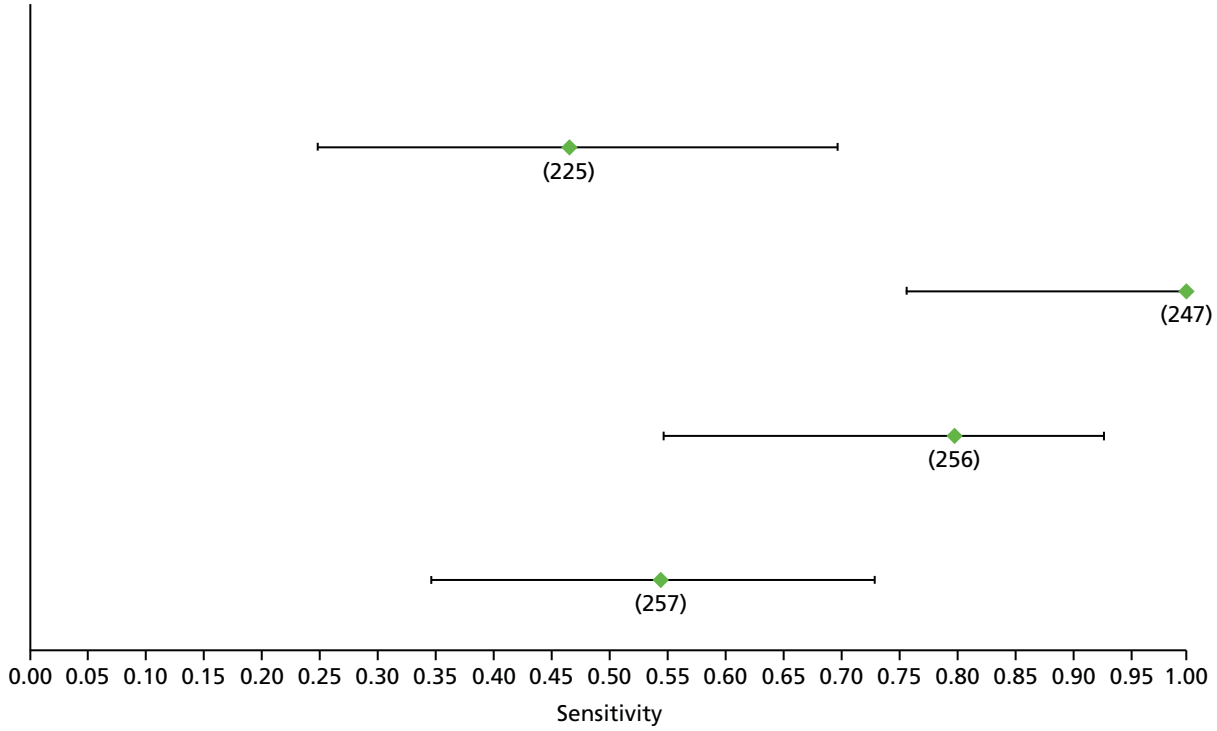

FIGURE 49 Hepatitis B: Fibrotest F4 – sensitivity.

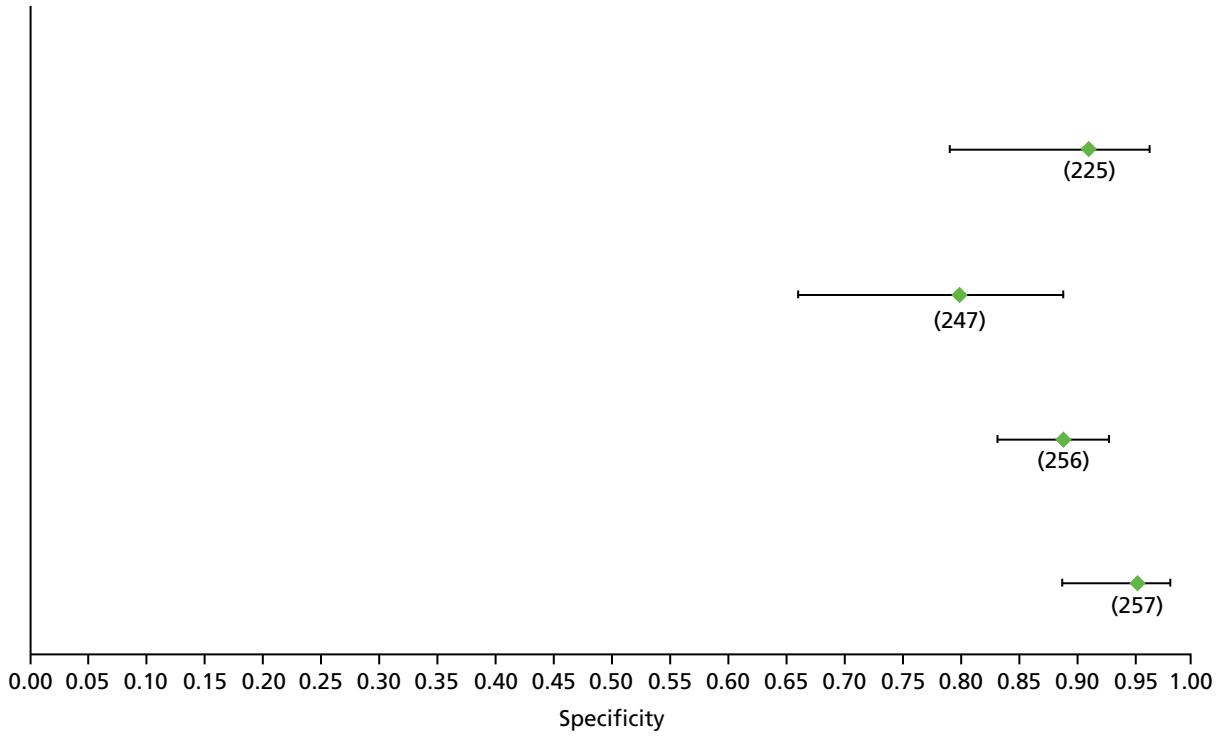

FIGURE 50 Hepatitis B: Fibrotest F4 – specificity.

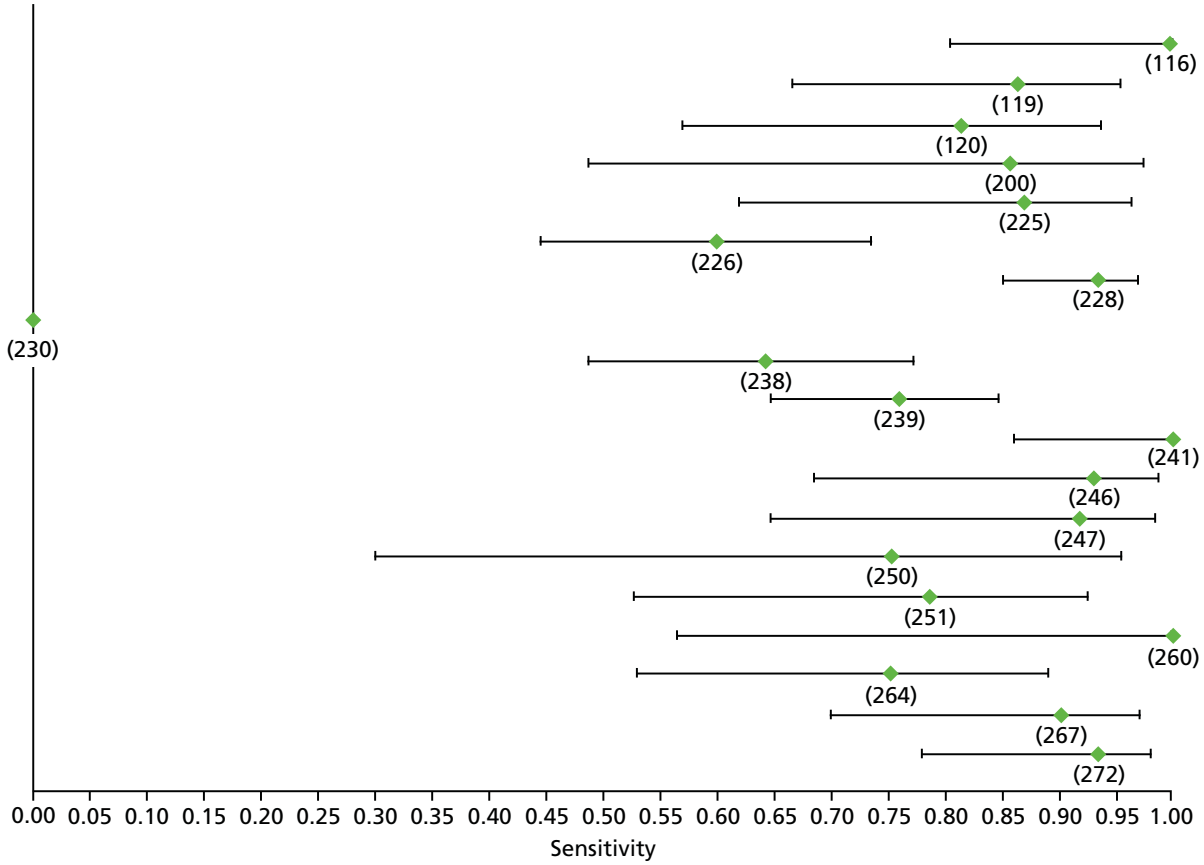

FIGURE 51 Hepatitis B: TE F4 – sensitivity.

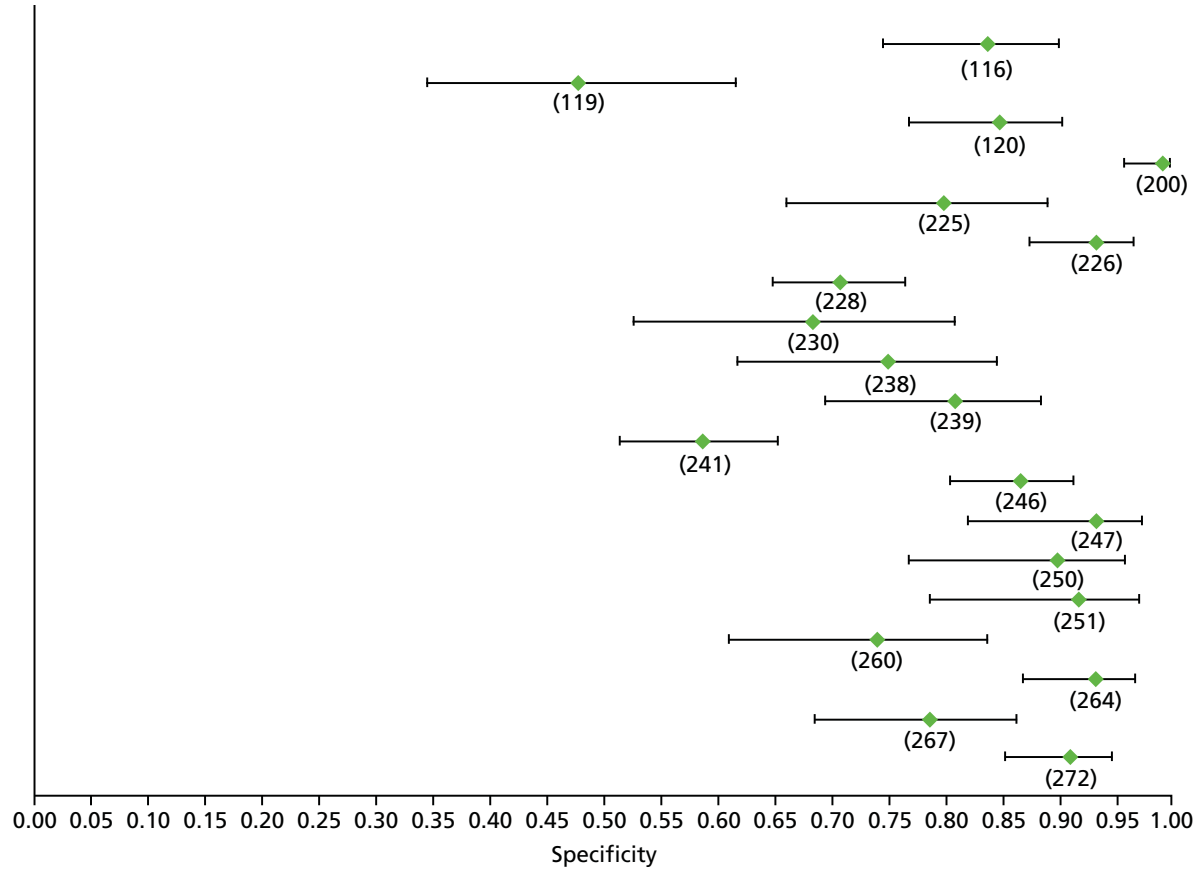

FIGURE 52 Hepatitis B: TE F4 – specificity.

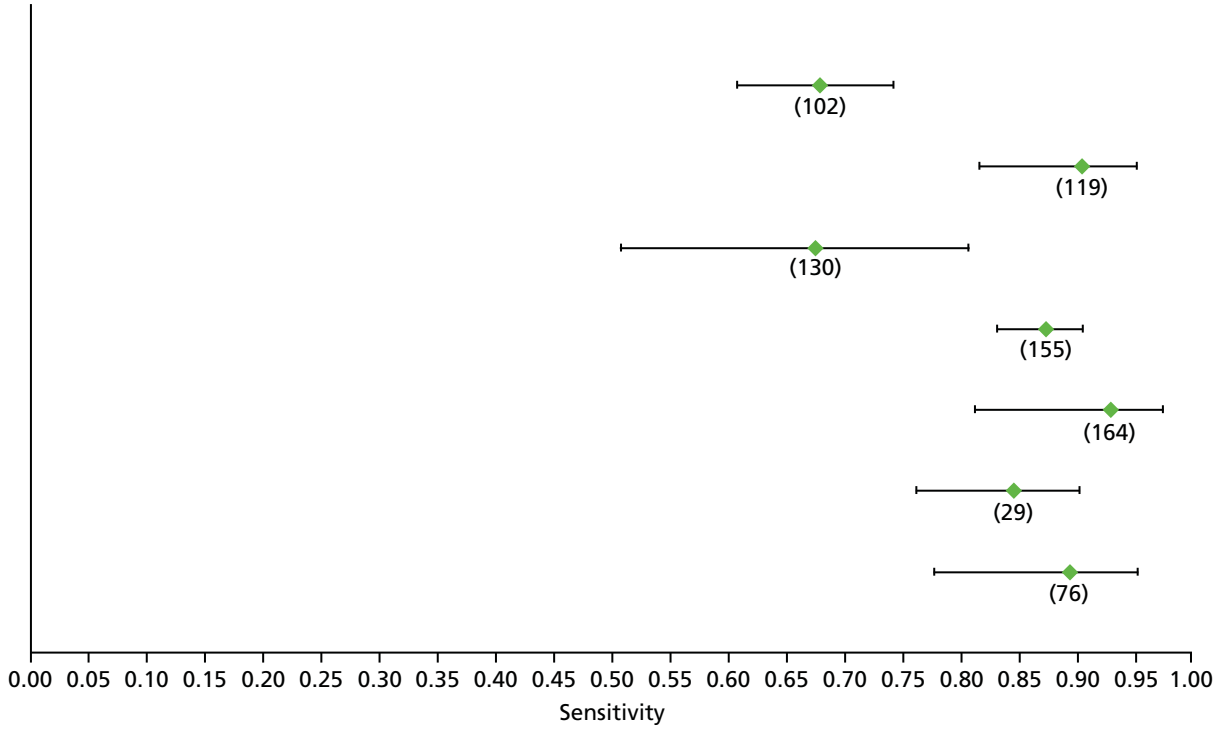

FIGURE 53 Hepatitis C: TE F1 – sensitivity.

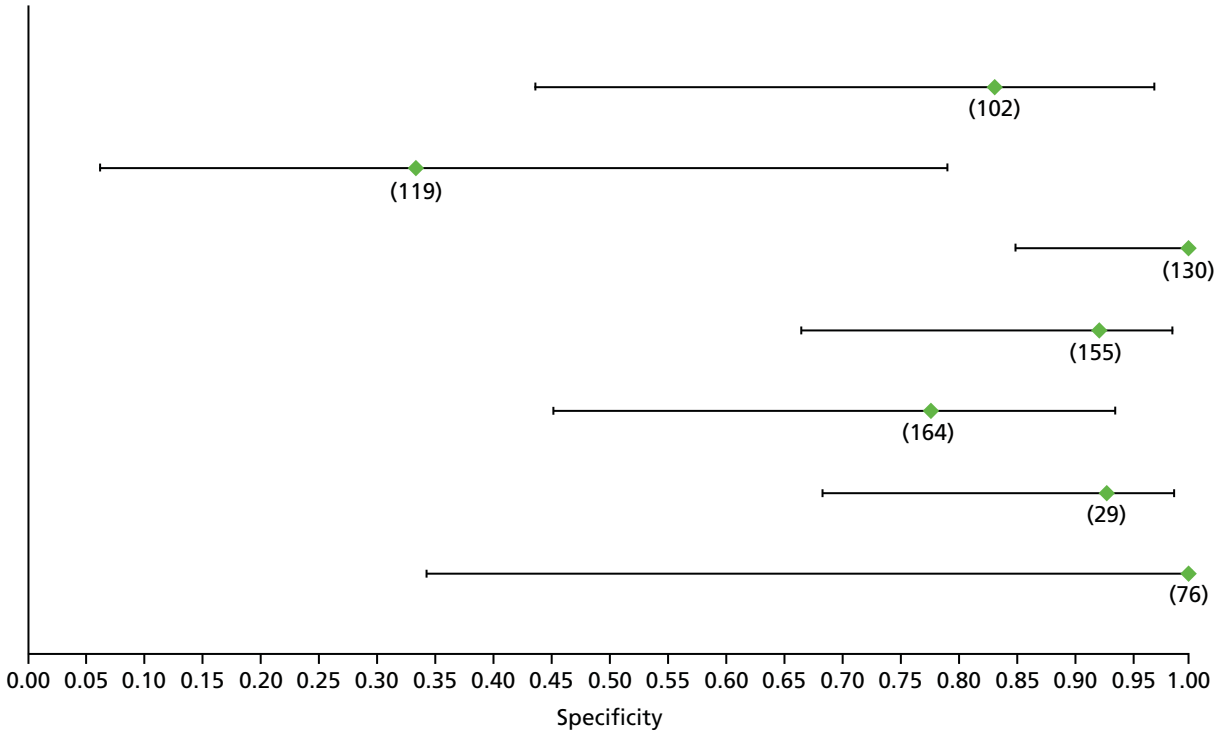

FIGURE 54 Hepatitis C: TE F1 – specificity.

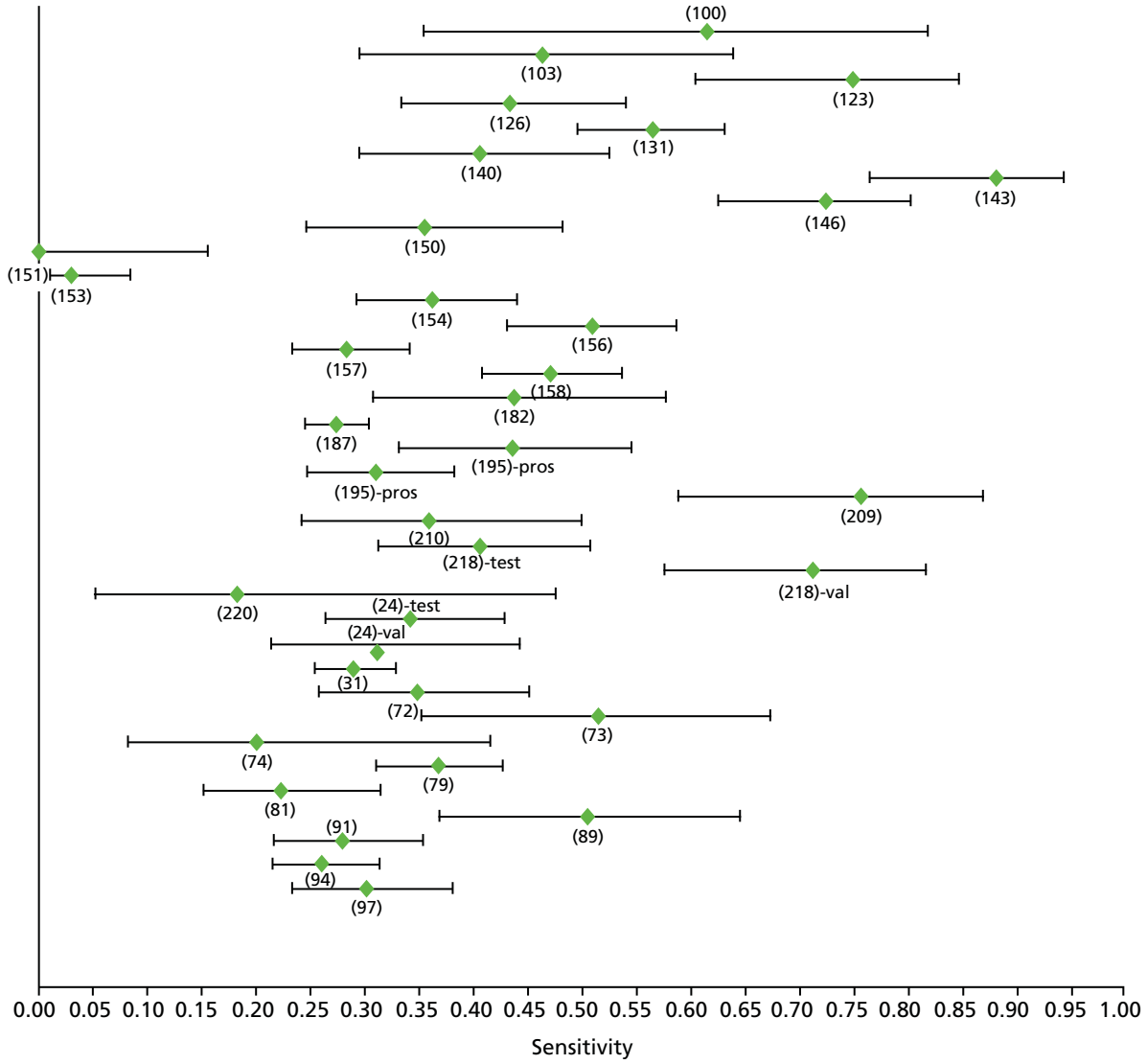

FIGURE 55 Hepatitis C: APRI F2 (high cut-off) – sensitivity. Pros, prospective; val, validation cohort.

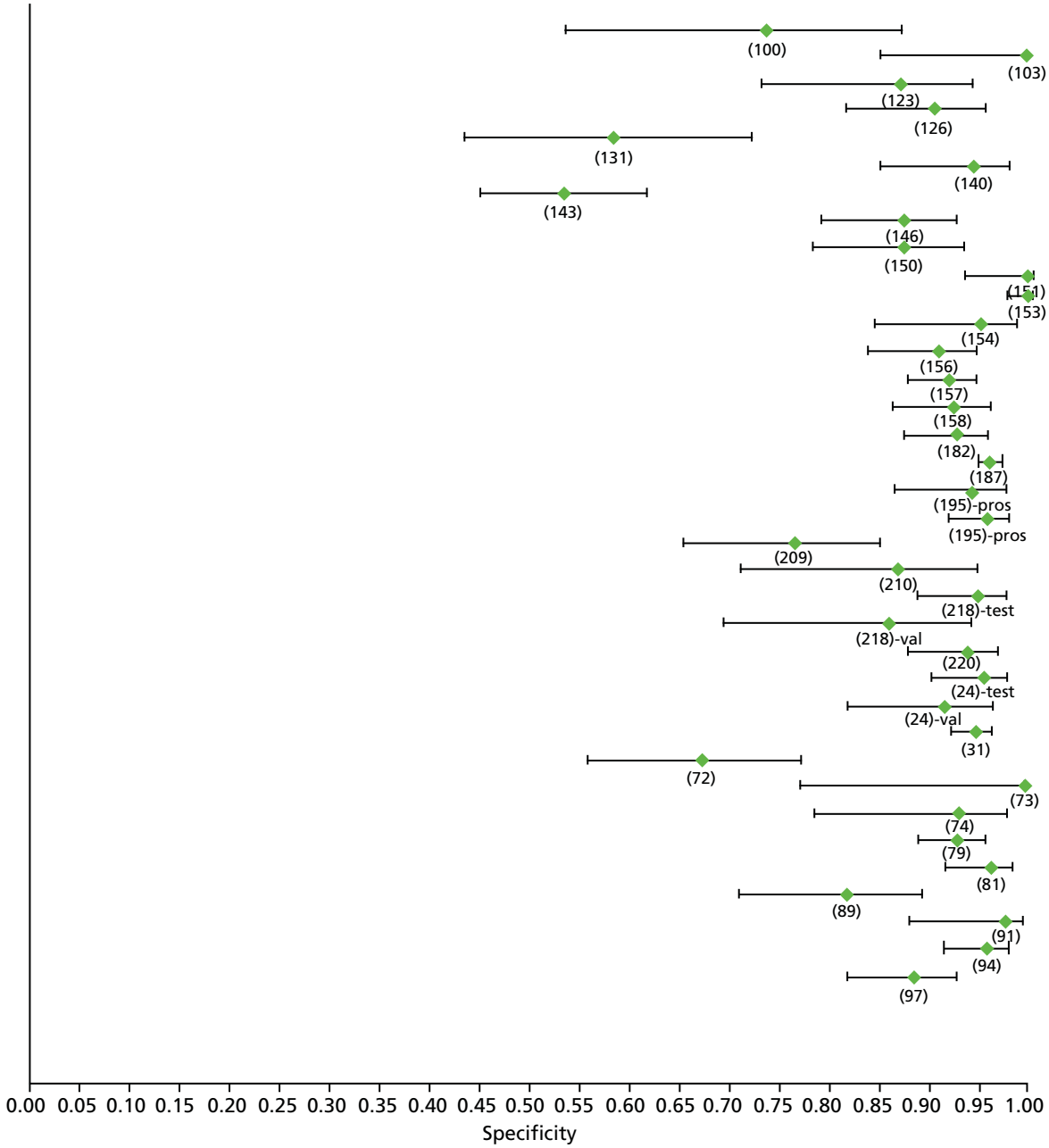

FIGURE 56 Hepatitis C: APRI F2 (high cut-off) – specificity. Pros, prospective; val, validation cohort.

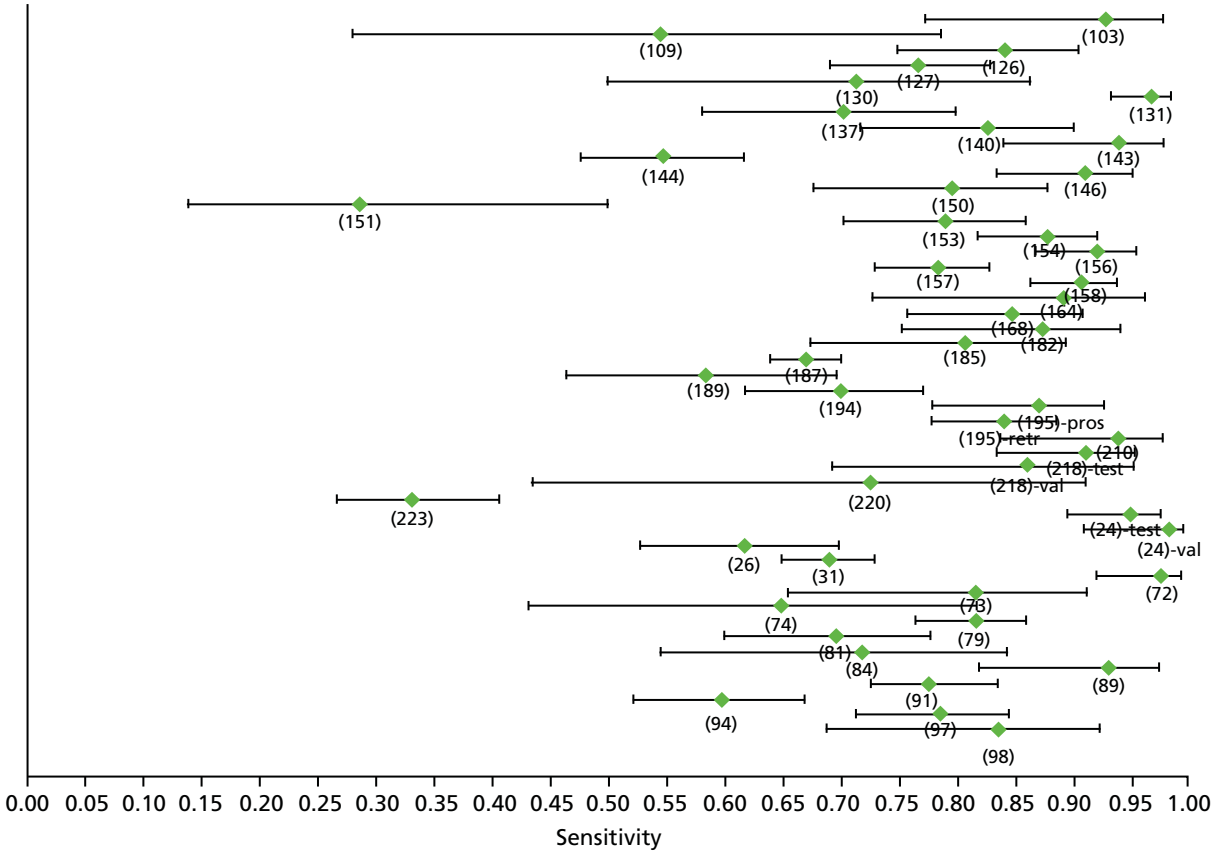

FIGURE 57 Hepatitis C: APRI F2 (low cut-off) – sensitivity. Pros, prospective; retr, retrospective; val, validation cohort.

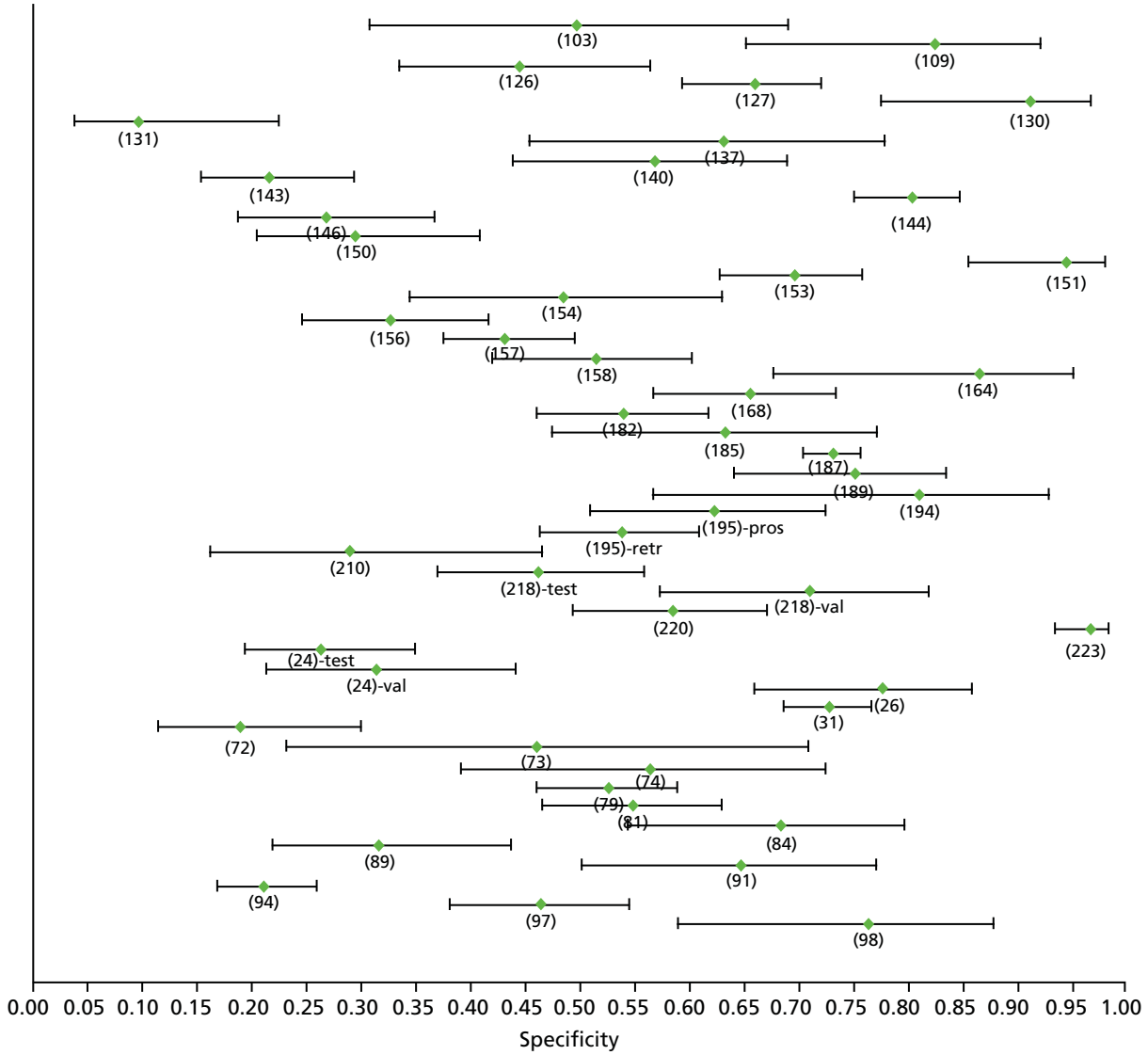

FIGURE 58 Hepatitis C: APRI F2 (low cut-off) – specificity. Pros, prospective; retr, retrospective; val, validation cohort.

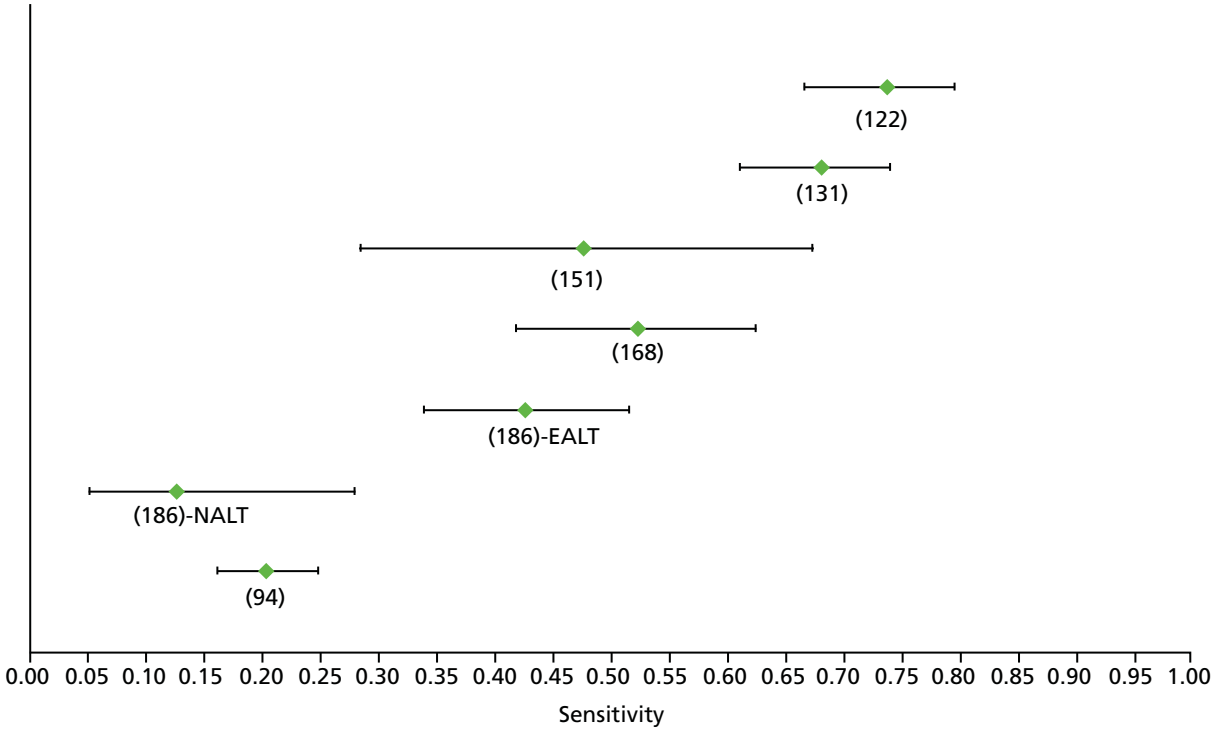

FIGURE 59 Hepatitis C: AST–ALT (F2) – sensitivity. EALT, elevated ALT; NALT, normal ALT.

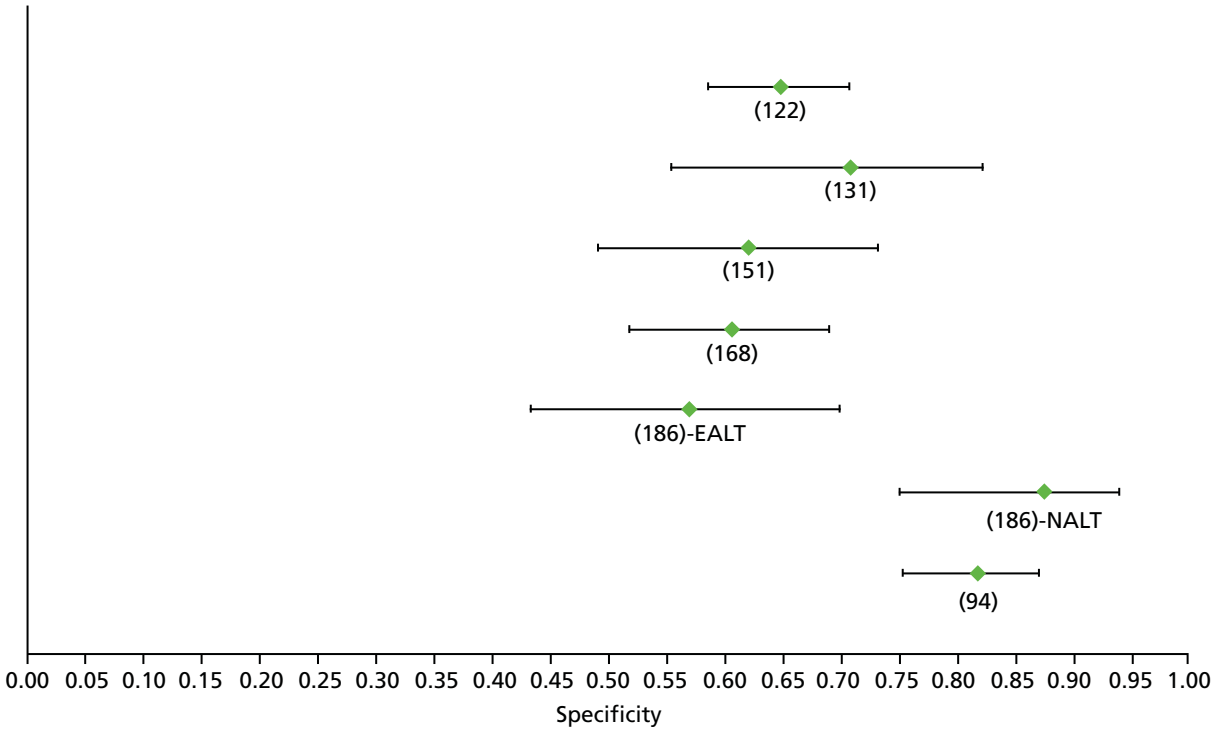

FIGURE 60 Hepatitis C: AST–ALT (F2) – specificity. EALT, elevated ALT; NALT, normal ALT.

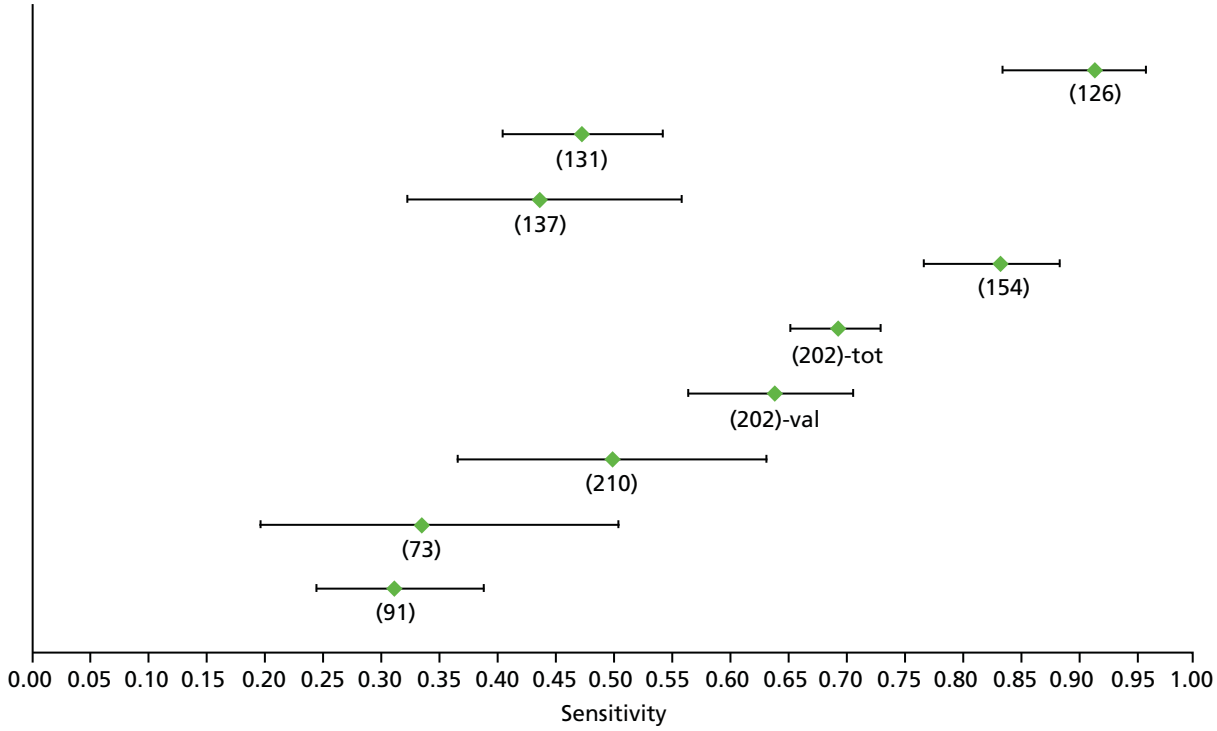

FIGURE 61 Hepatitis C: FIB-4 F2 (high cut-off) – sensitivity. Tot, total; val, validation cohort.

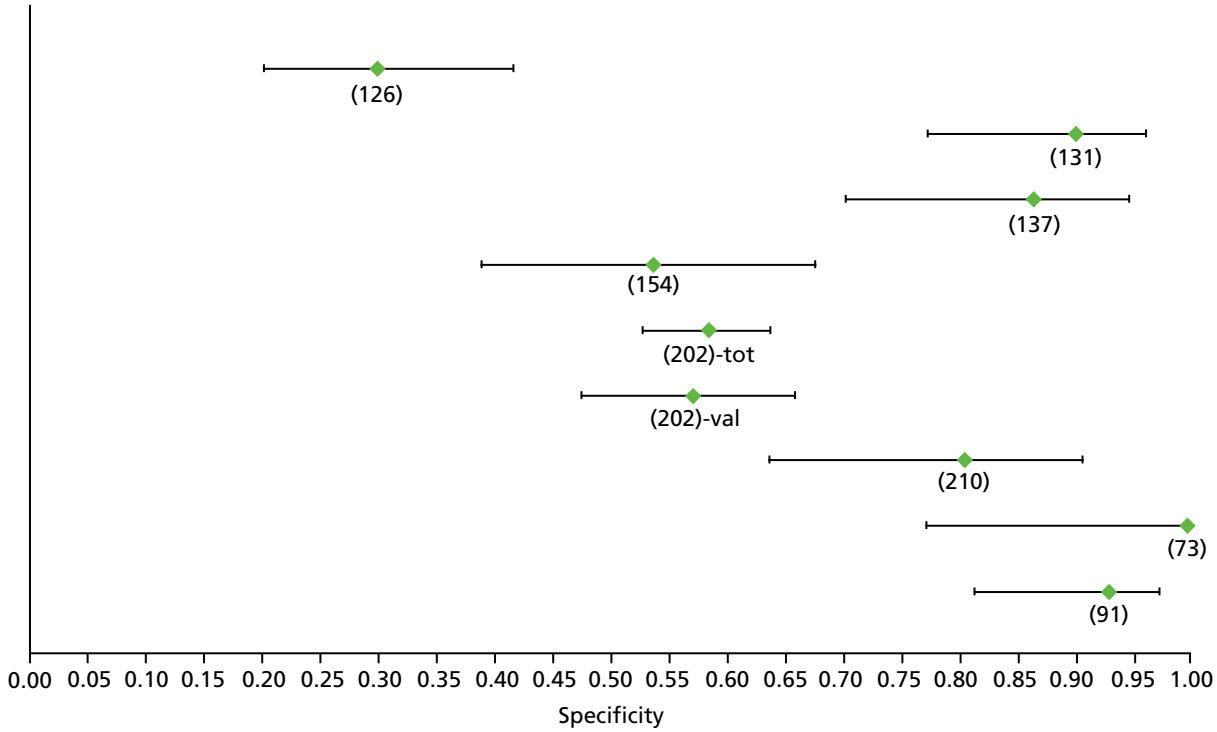

FIGURE 62 Hepatitis C: FIB-4 F2 (high cut-off) – specificity. Tot, total; val, validation cohort.

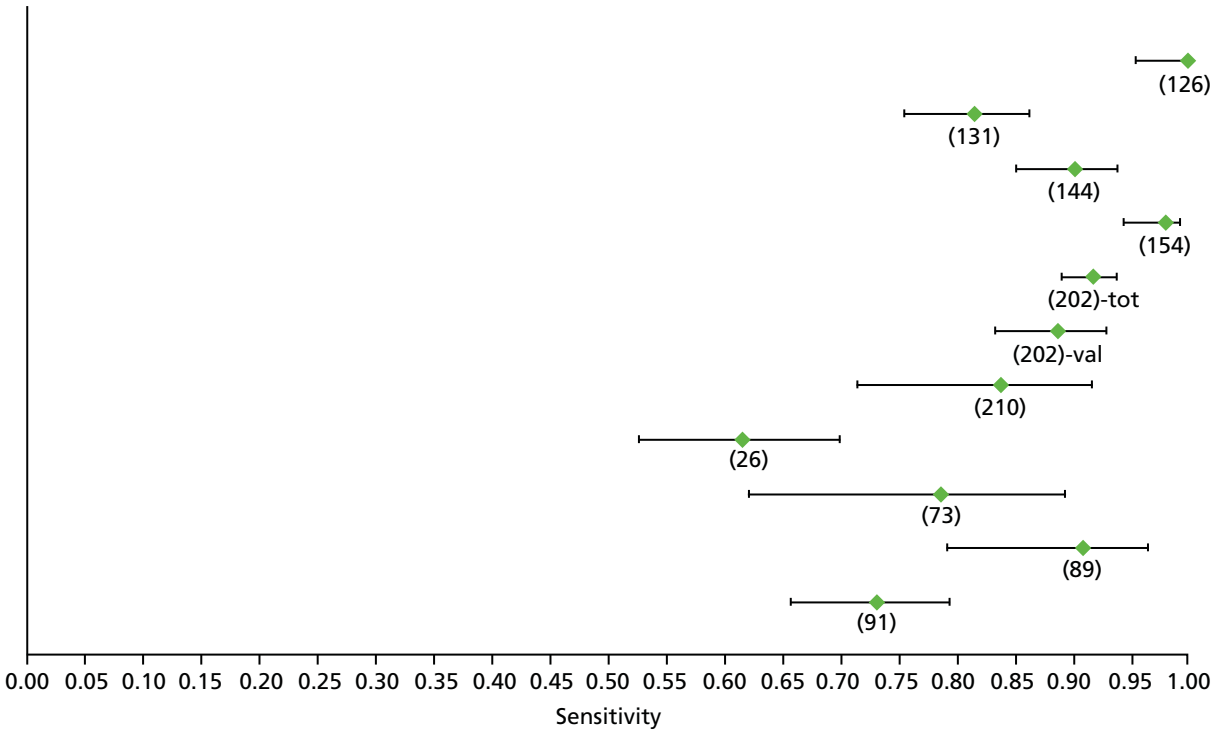

FIGURE 63 Hepatitis C: FIB-4 F2 (low cut-off) – sensitivity. Tot, total; val, validation cohort.

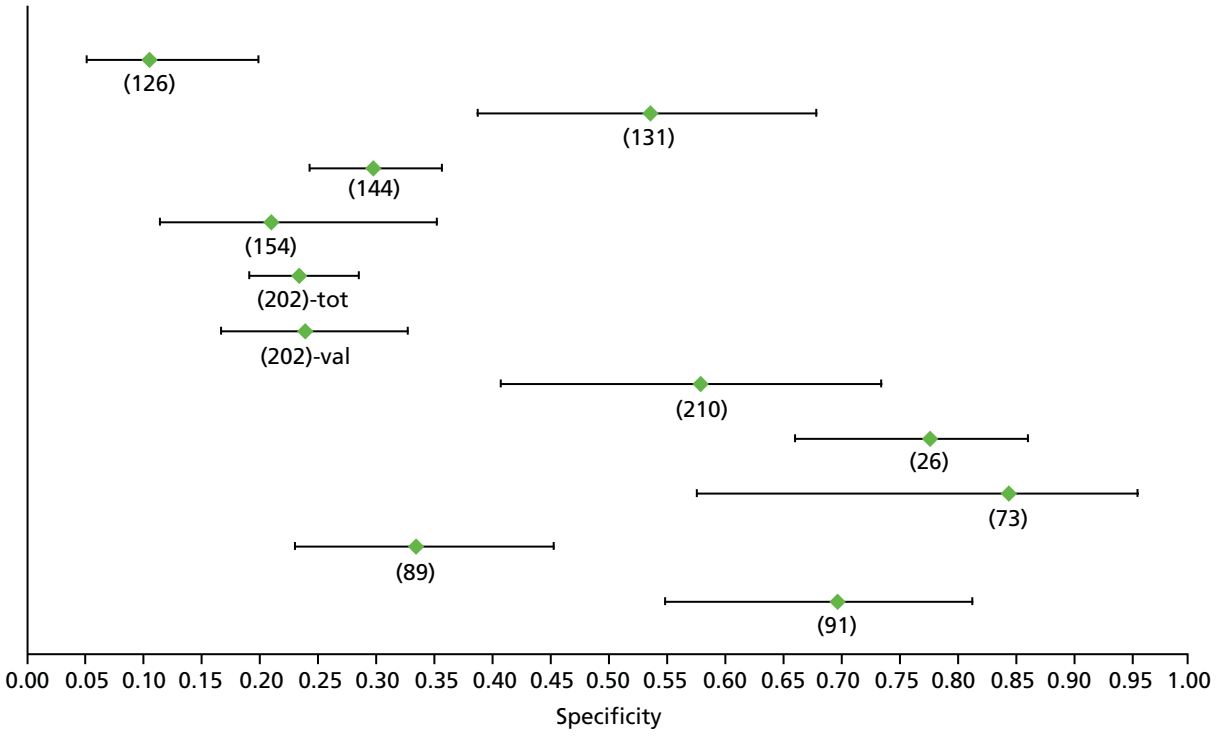

FIGURE 64 Hepatitis C: FIB-4 F2 (low cut-off) – specificity. Tot, total; val, validation cohort.

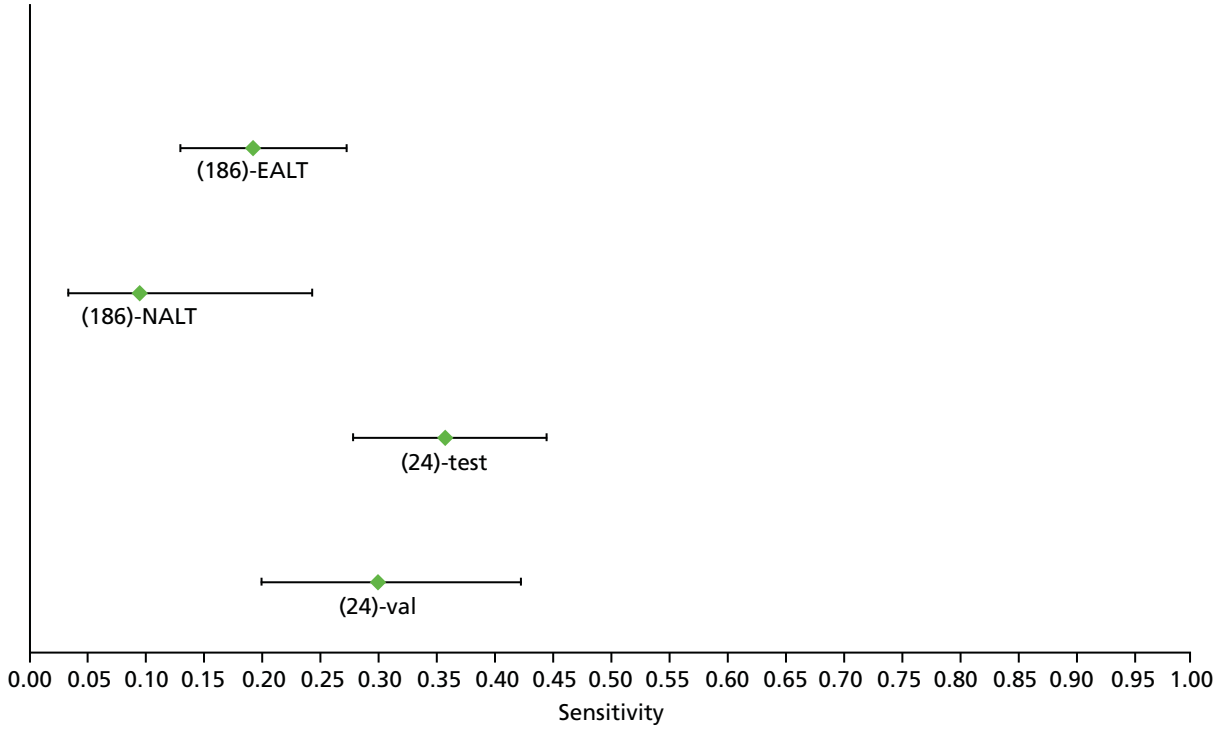

FIGURE 65 Hepatitis C: Fibroindex F2 (high cut-off) – sensitivity. EALT, elevated ALT; NALT, normal ALT; test, derivation cohort; val, validation cohort.

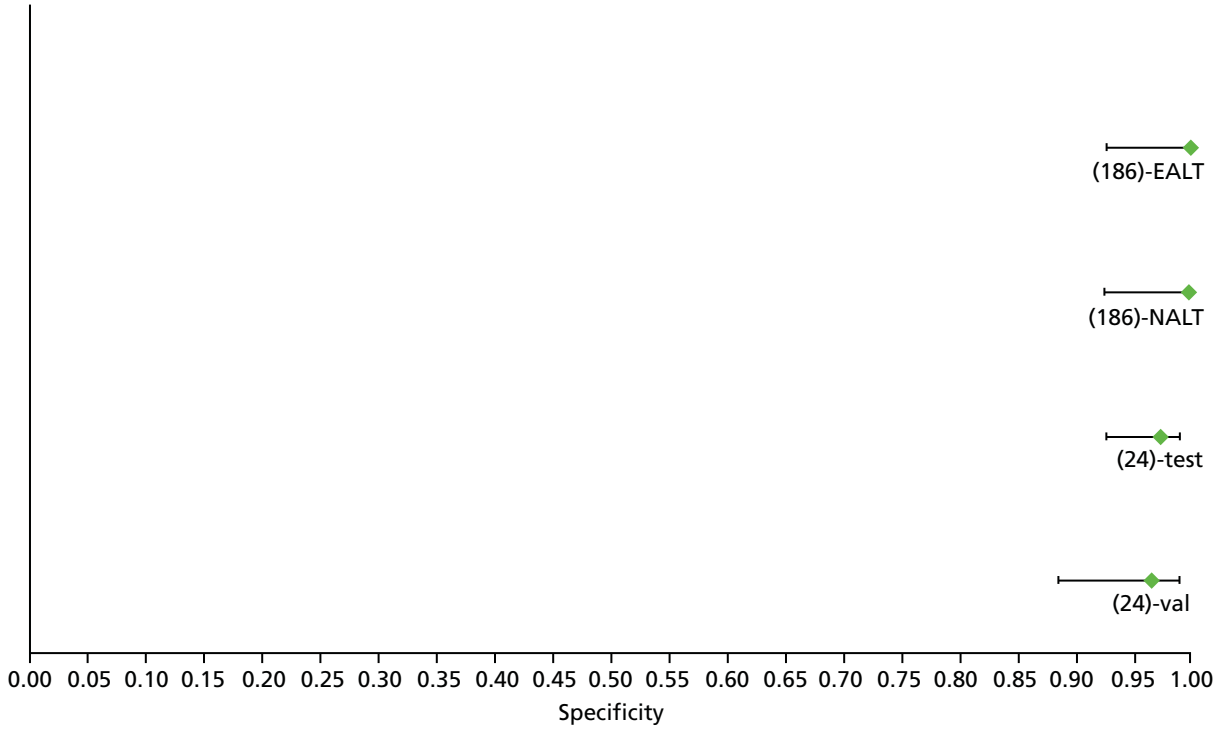

FIGURE 66 Hepatitis C: Fibroindex F2 (high cut-off) – specificity. EALT, elevated ALT; NALT, normal ALT; test, derivation cohort; val, validation cohort.

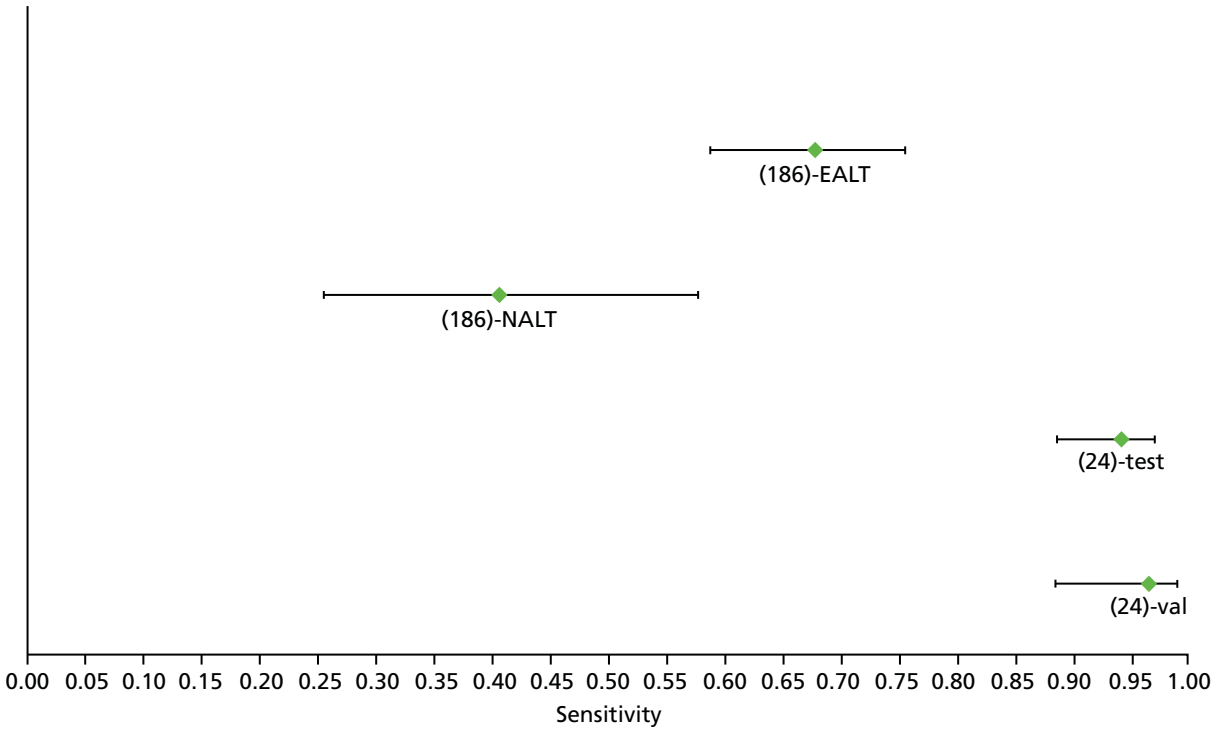

FIGURE 67 Hepatitis C: Fibroindex F2 (low cut-off) – sensitivity. EALT, elevated ALT; NALT, normal ALT; test, derivation cohort; val, validation cohort.

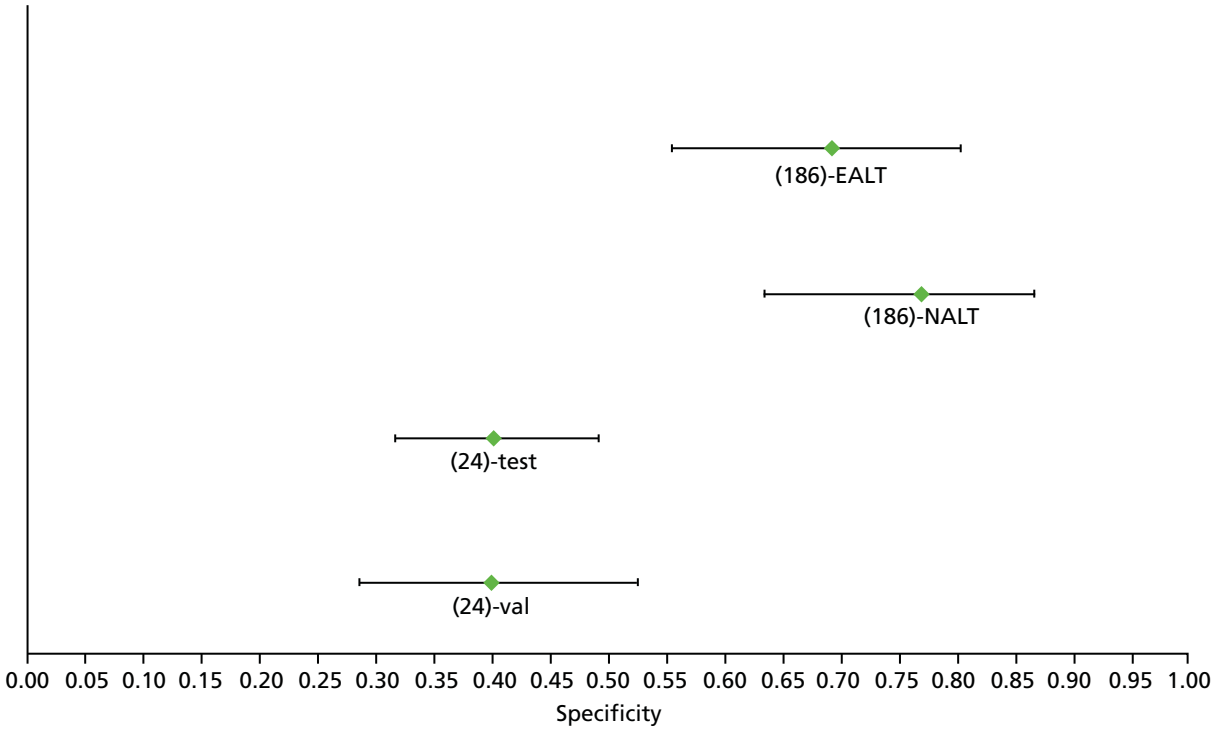

FIGURE 68 Hepatitis C: Fibroindex F2 (low cut-off) – specificity. EALT, elevated ALT; NALT, normal ALT; test, derivation cohort; val, validation cohort.

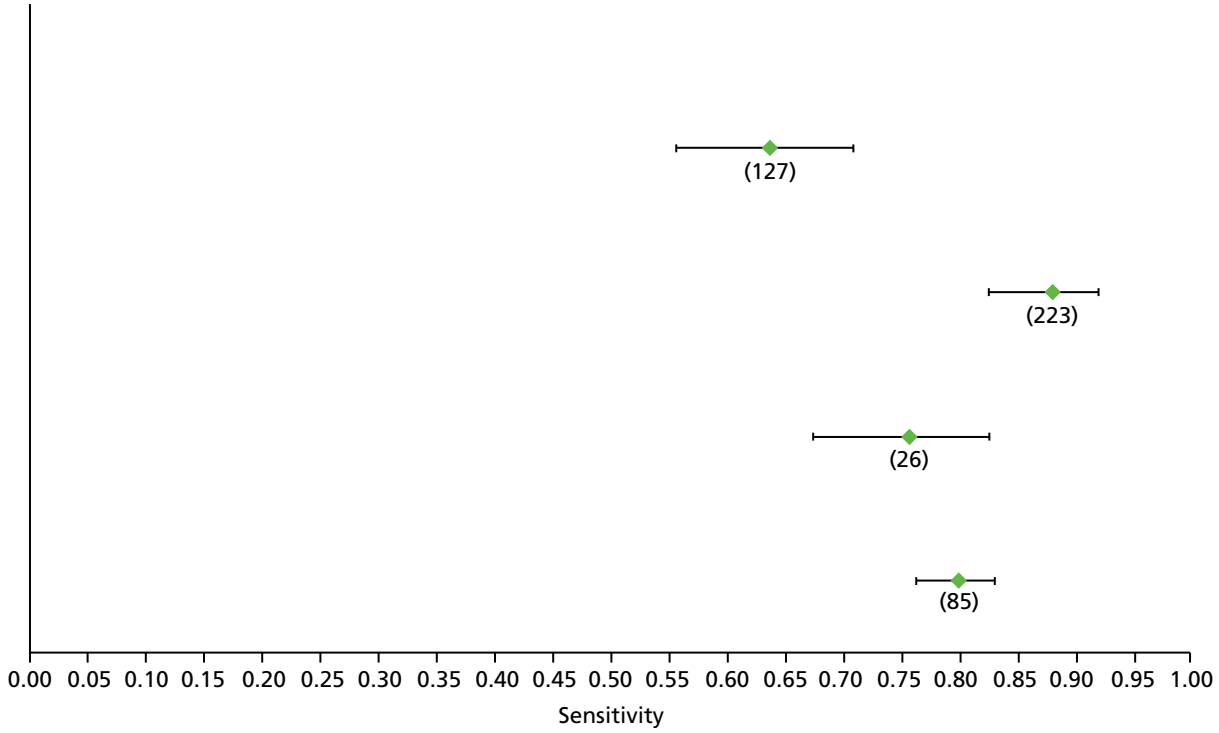

FIGURE 69 Hepatitis C: Fibrometer F2 – sensitivity.

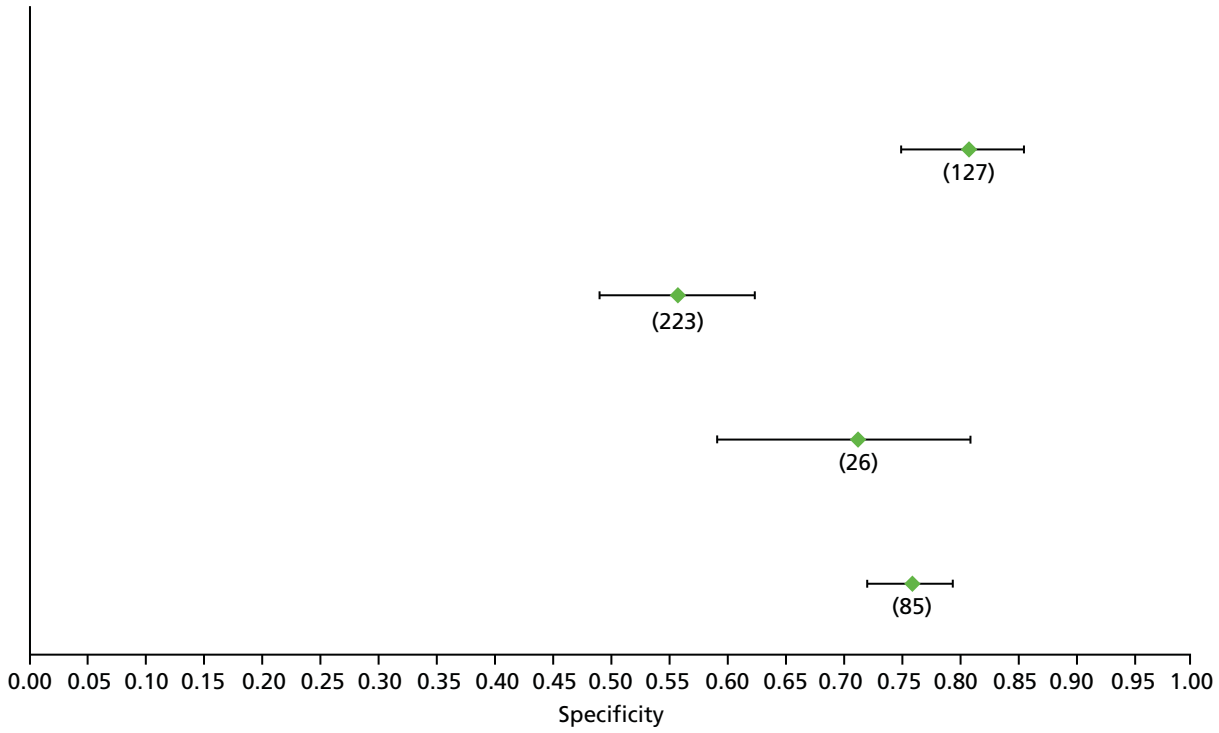

FIGURE 70 Hepatitis C: Fibrometer F2 – specificity.

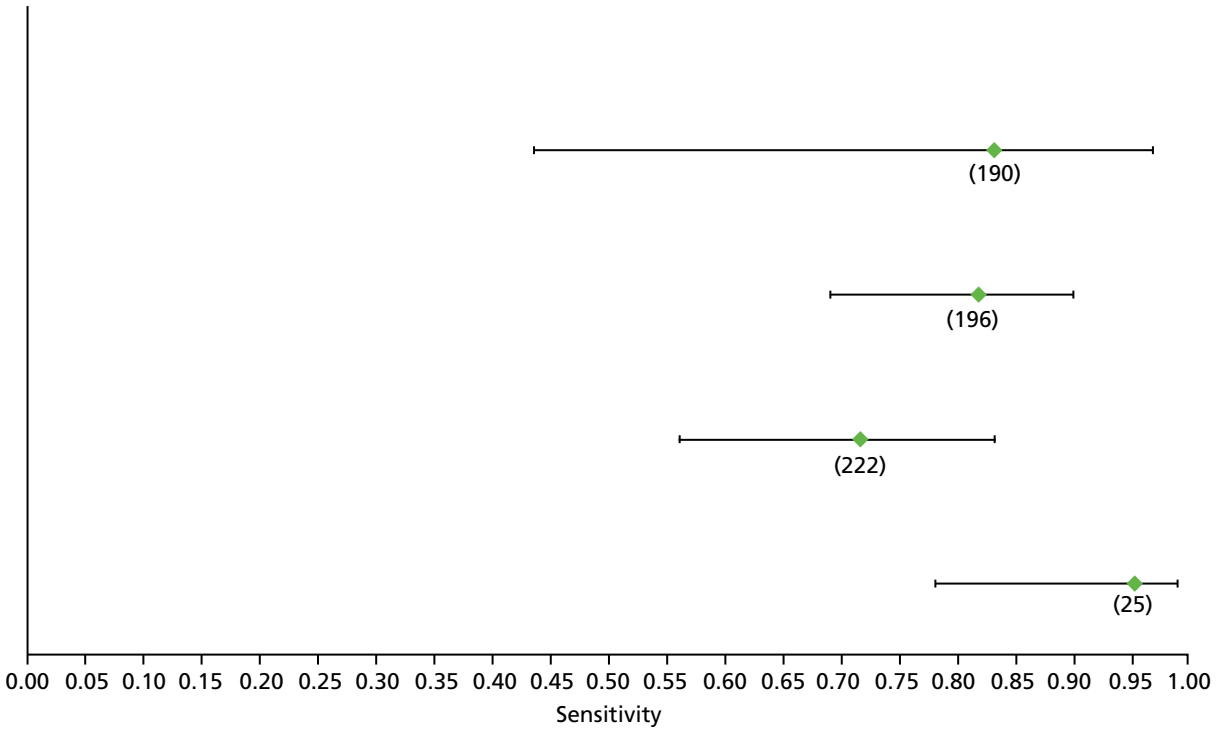

FIGURE 71 Hepatitis C: FibroSpect II F2 – sensitivity.

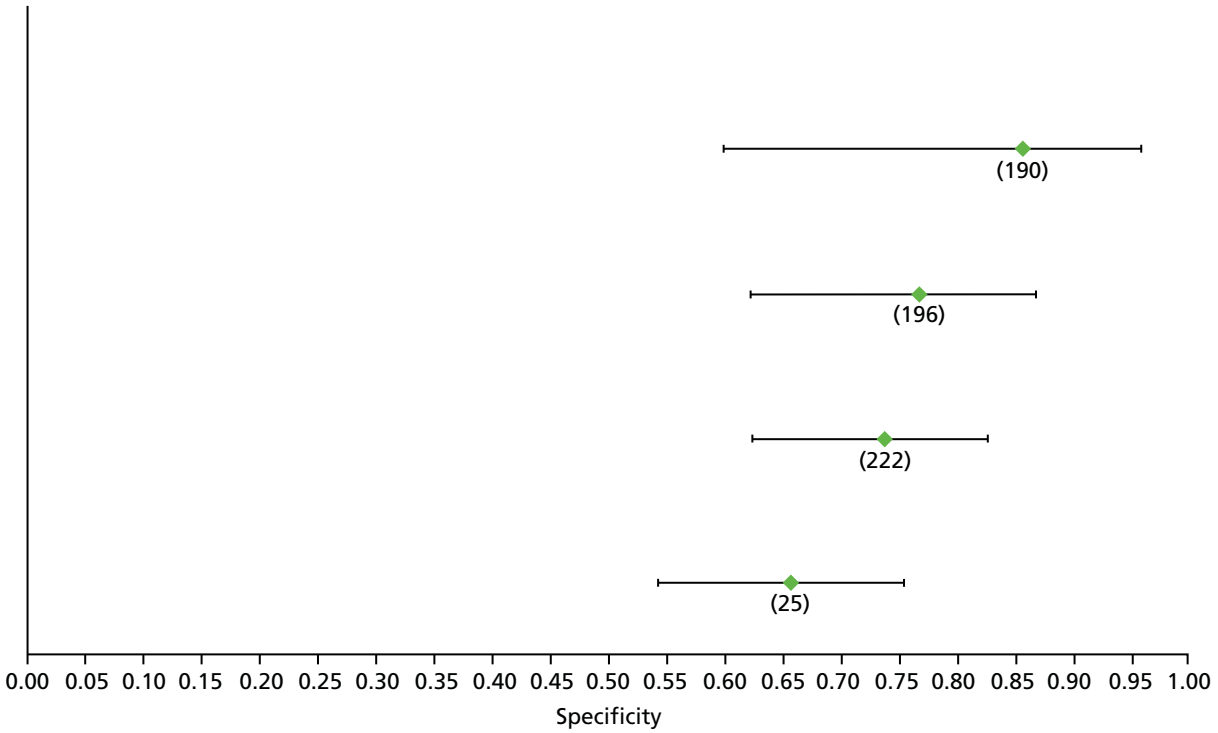

FIGURE 72 Hepatitis C: FibroSpect II F2 – specificity.

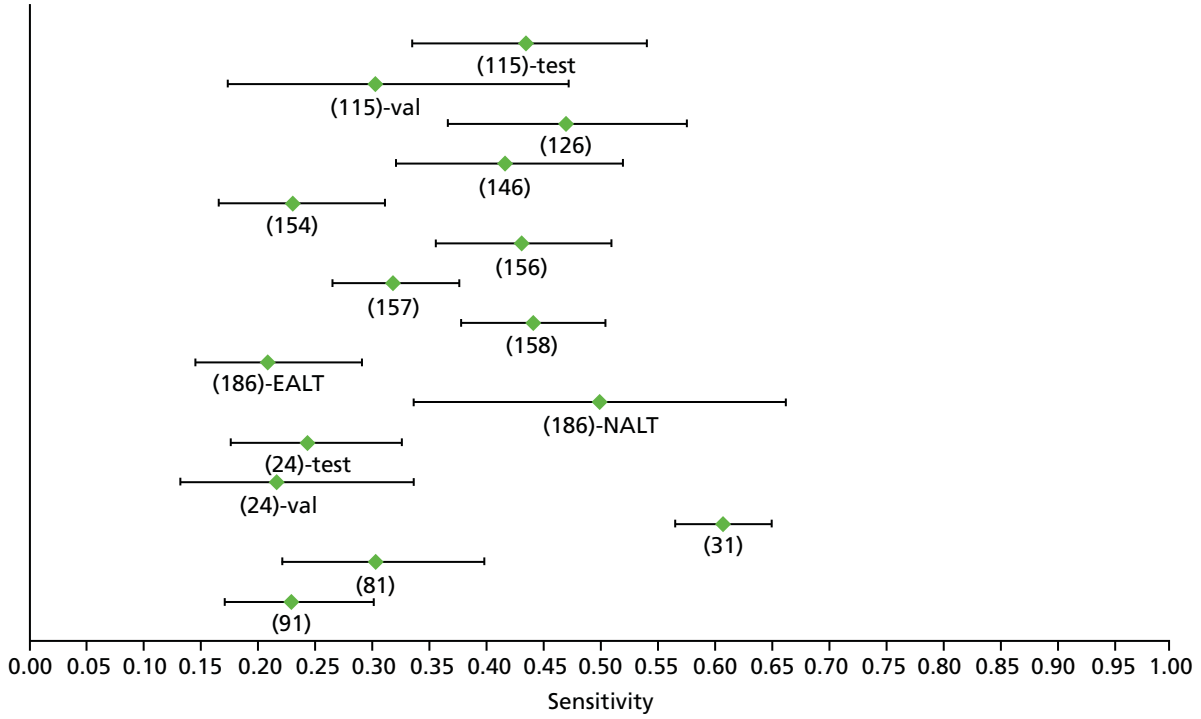

FIGURE 73 Hepatitis C: Forns index F2 (high cut-off) – sensitivity. EALT, elevated ALT; NALT, normal ALT; test, derivation cohort; val, validation cohort.

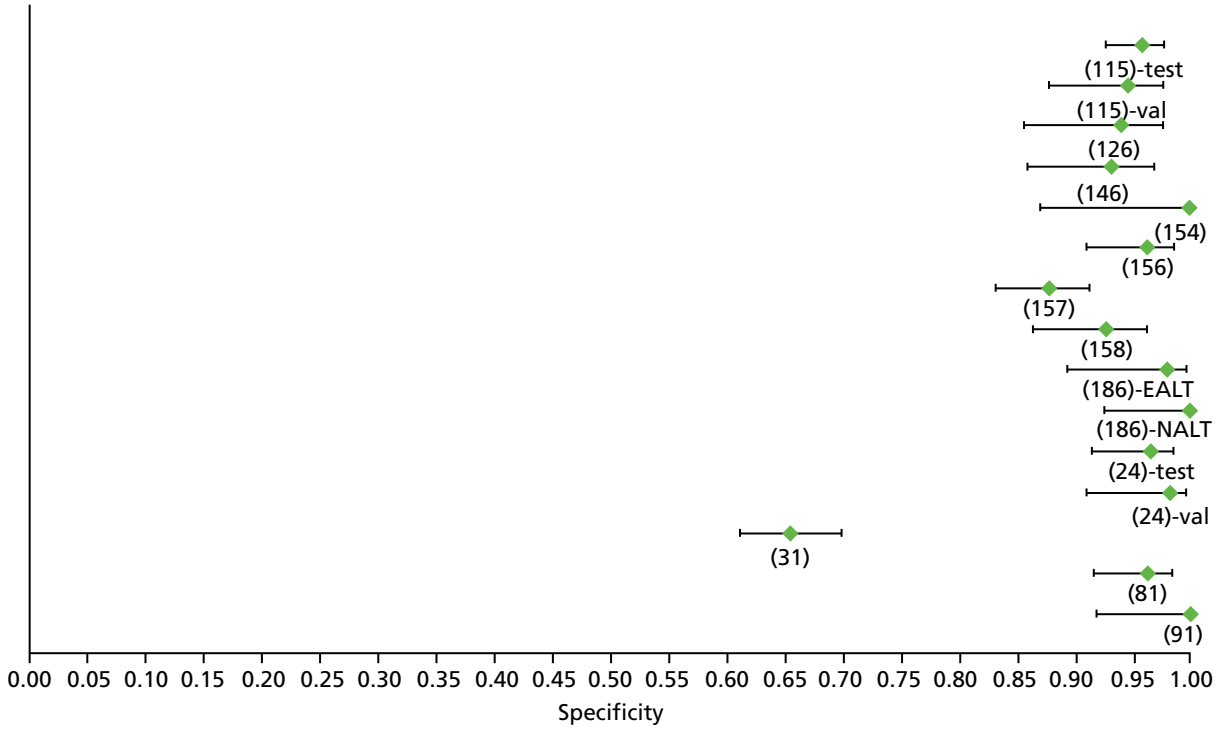

FIGURE 74 Hepatitis C: Forns index F2 (high cut-off) – specificity. EALT, elevated ALT; NALT, normal ALT; test, derivation cohort; val, validation cohort.

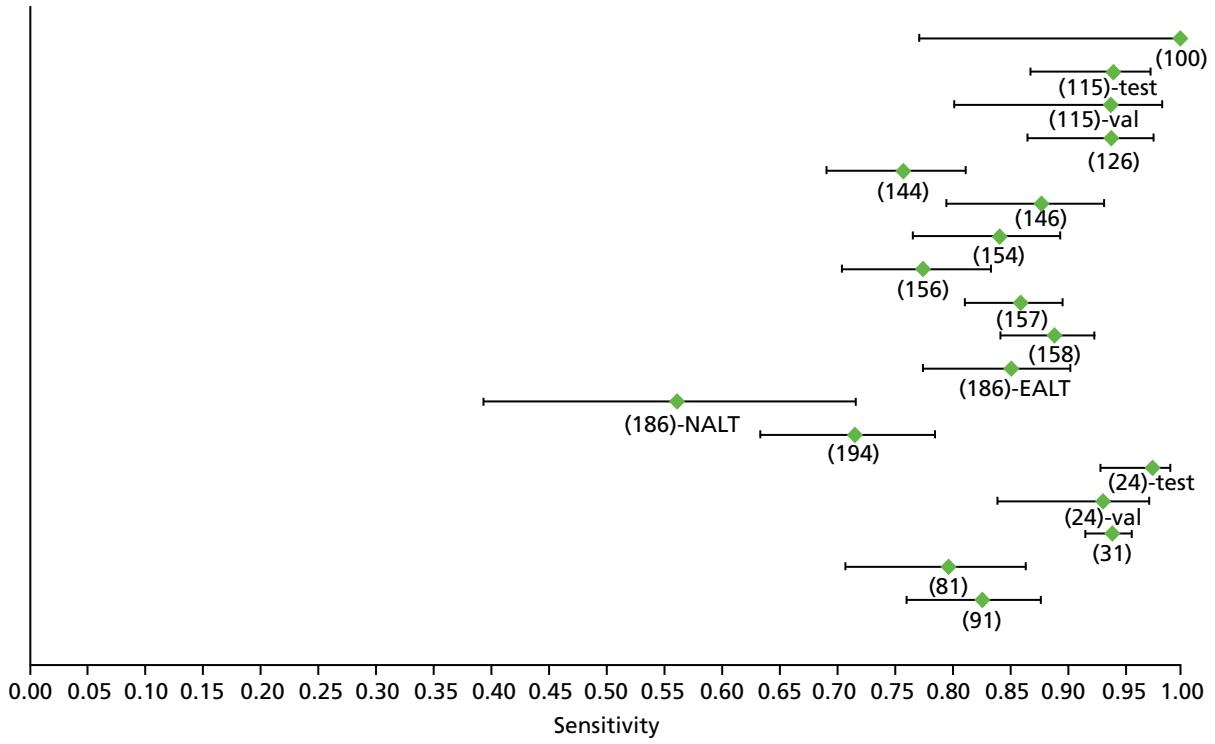

FIGURE 75 Hepatitis C: Forns index F2 (low cut-off) – sensitivity. EALT, elevated ALT; NALT, normal ALT; test, derivation cohort; val, validation cohort.

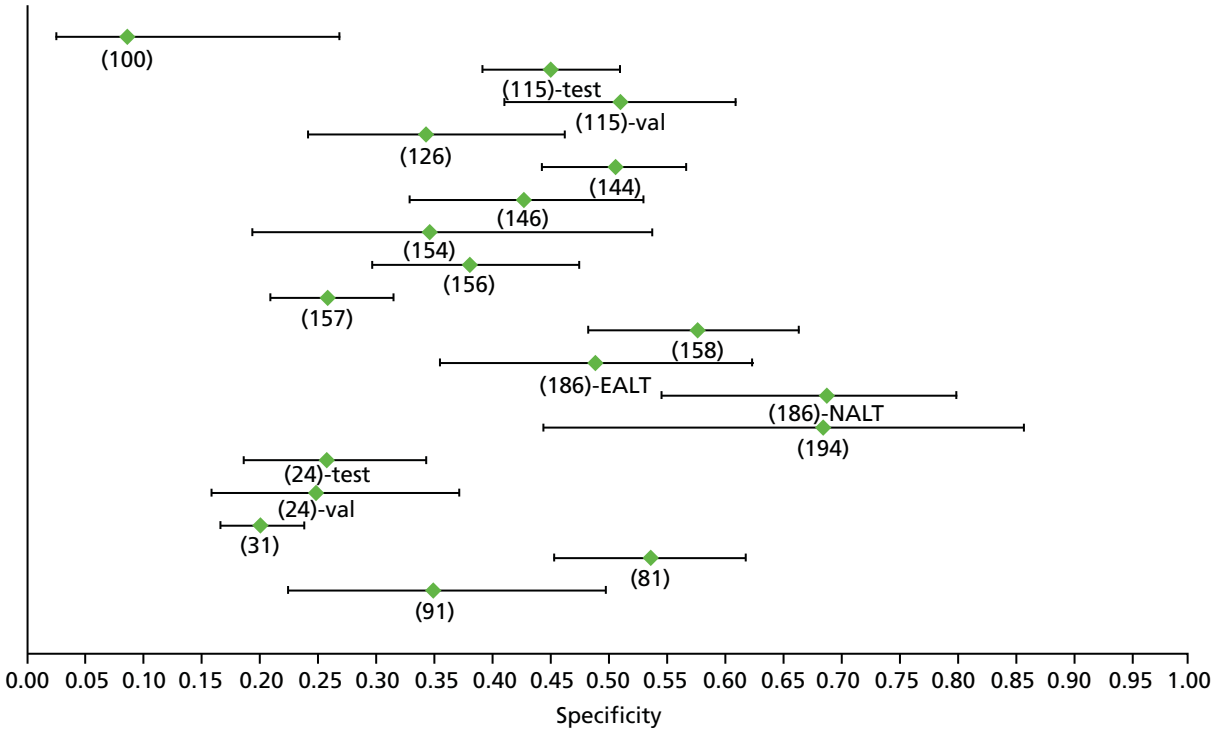

FIGURE 76 Hepatitis C: Forns index F2 (low cut-off) – specificity. EALT, elevated ALT; NALT, normal ALT; test, derivation cohort; val, validation cohort.

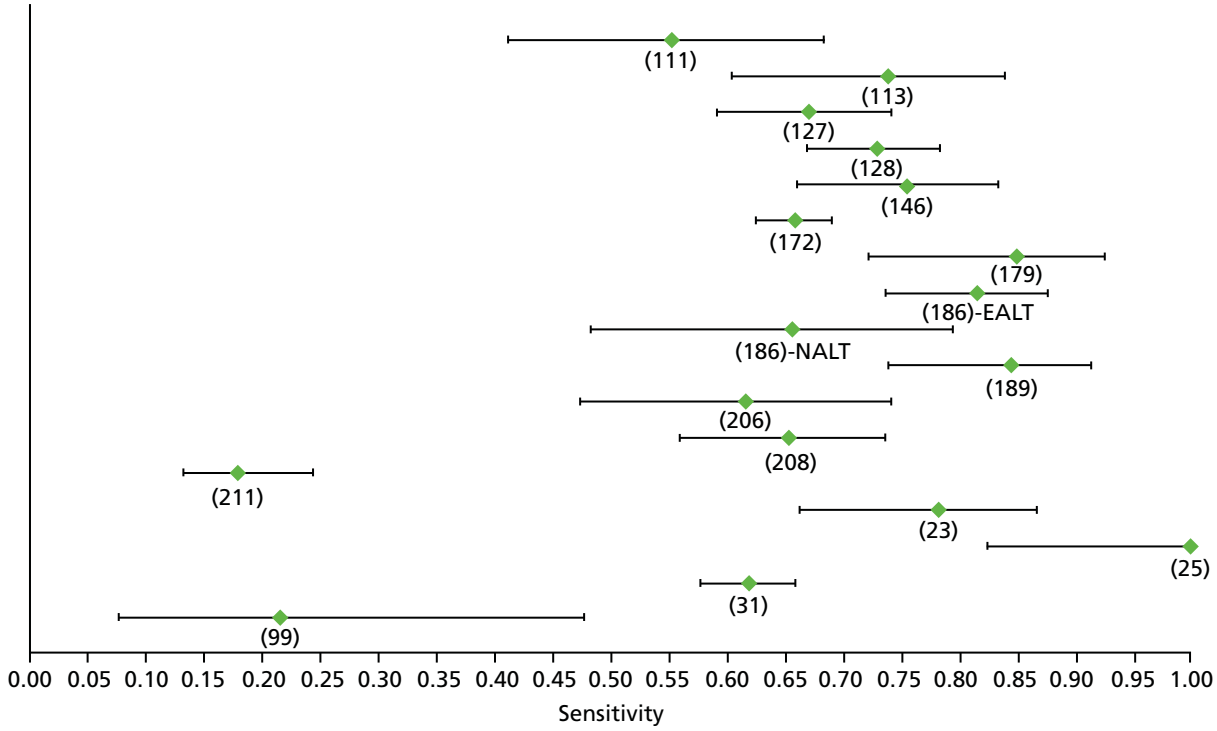

FIGURE 77 Hepatitis C: Fibrotest F2 – sensitivity. EALT, elevated ALT; NALT, normal ALT.

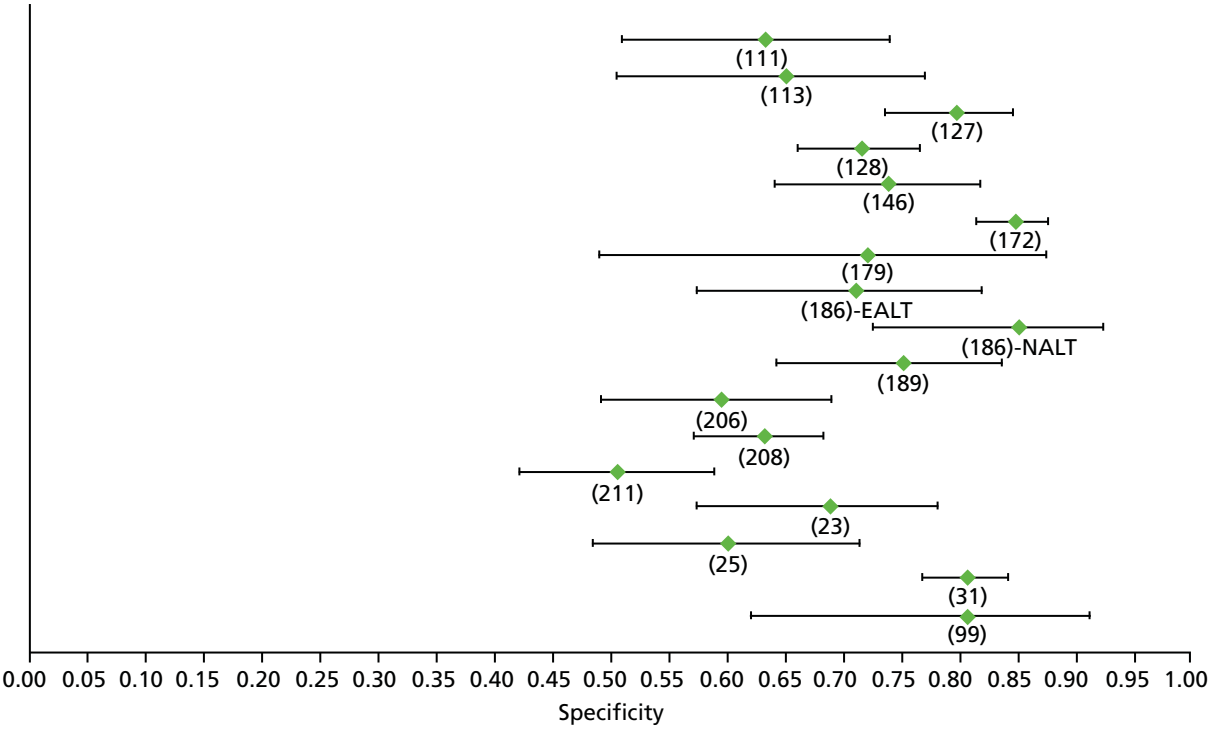

FIGURE 78 Hepatitis C: Fibrotest F2 – specificity. EALT, elevated ALT; NALT, normal ALT;

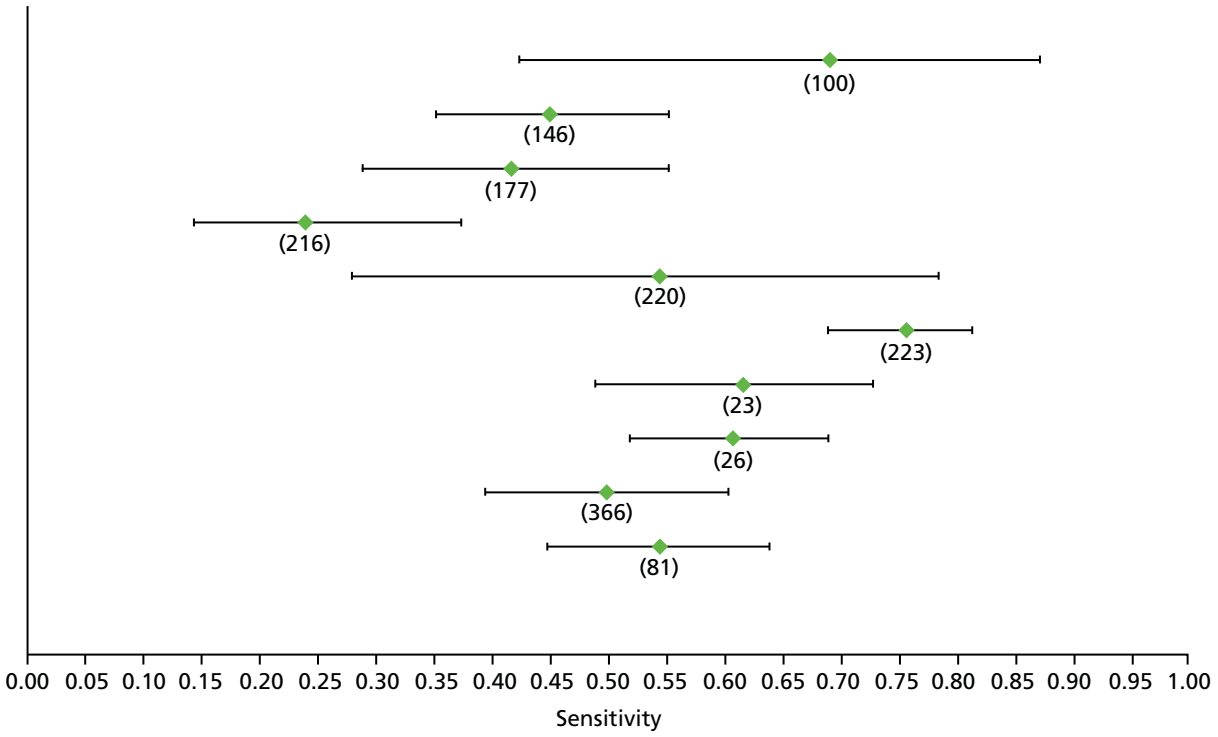

FIGURE 79 Hepatitis C: Fibrotest F2 (high cut-off) – sensitivity.

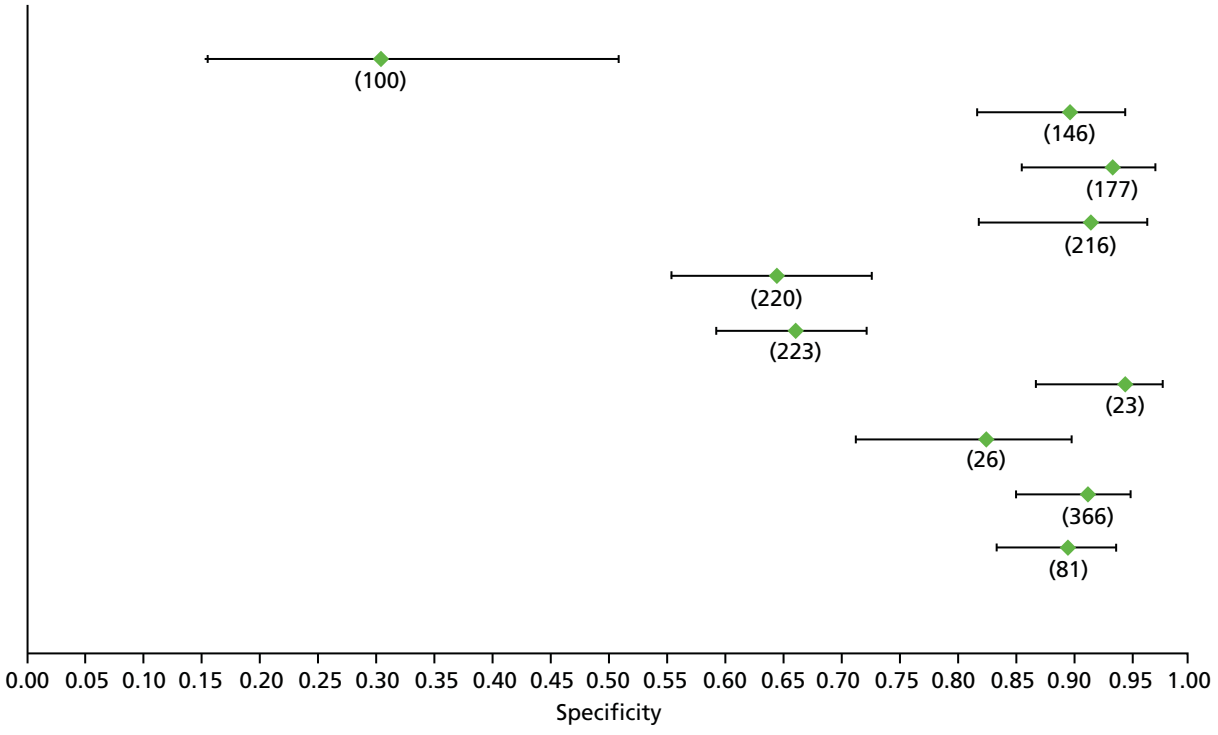

FIGURE 80 Hepatitis C: Fibrotest F2 (high cut-off) – specificity.

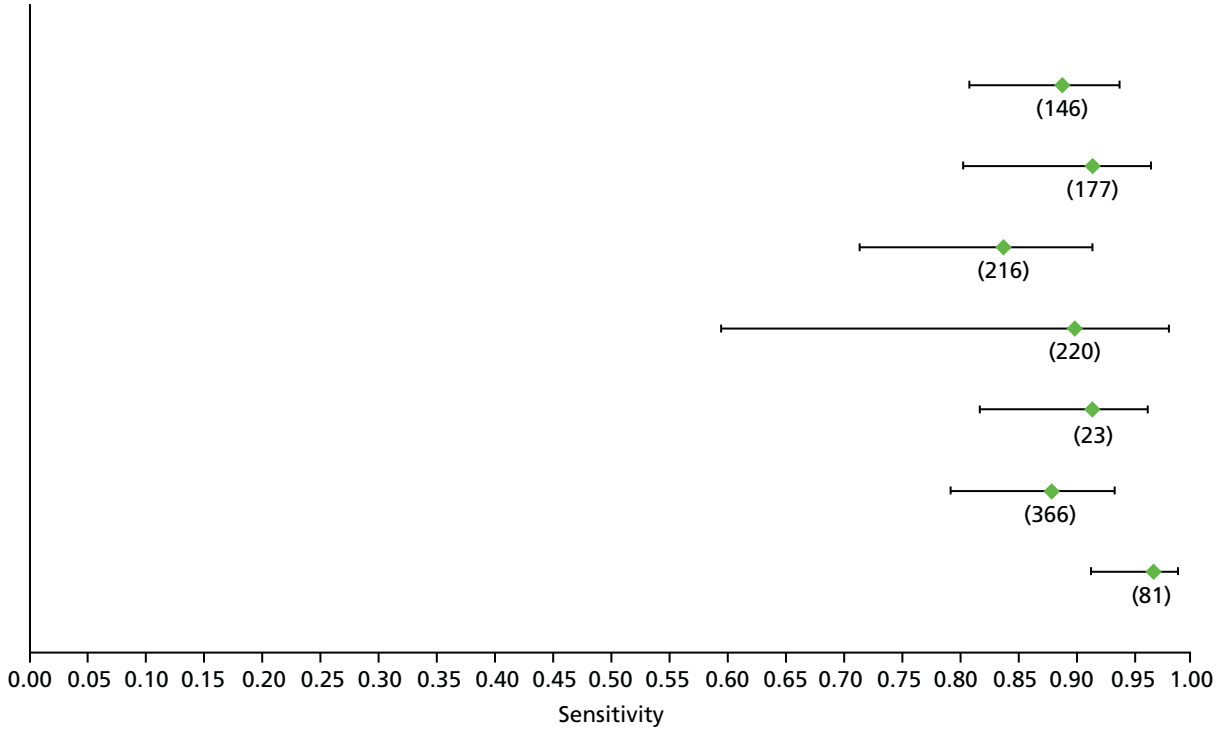

FIGURE 81 Hepatitis C: Fibrotest F2 (low cut-off) – sensitivity.

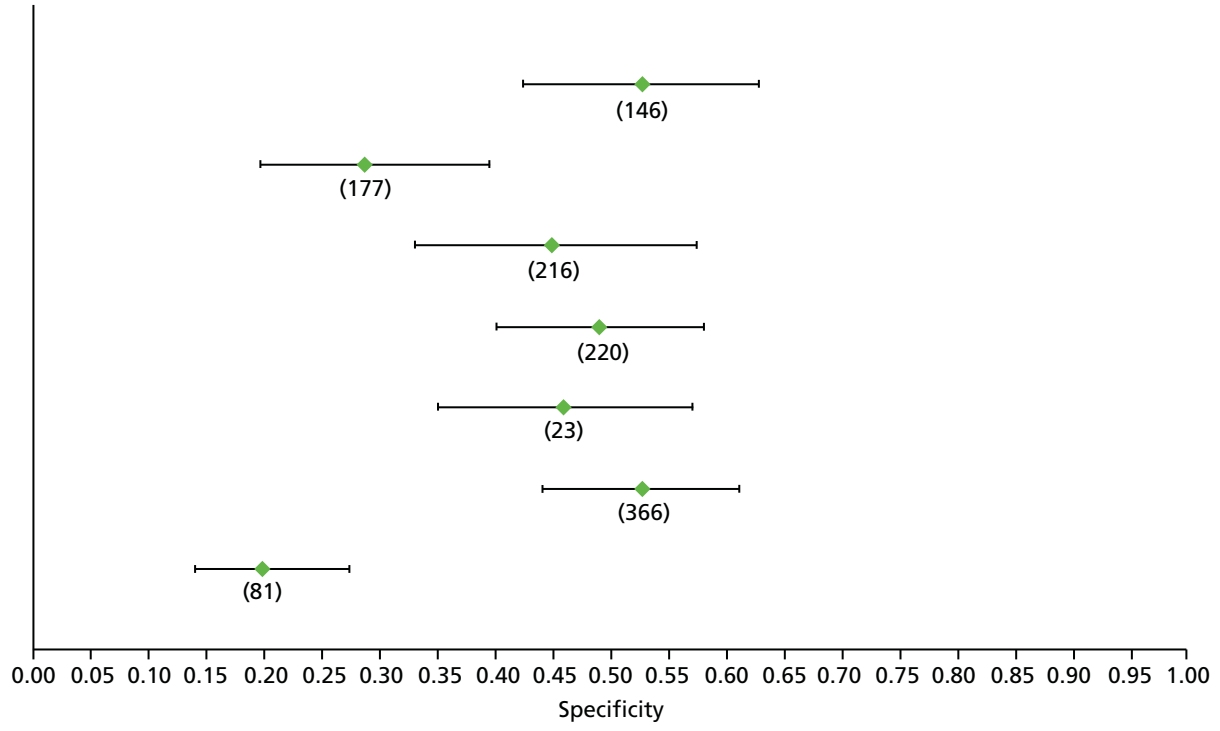

FIGURE 82 Hepatitis C: Fibrotest F2 (low cut-off) – specificity.

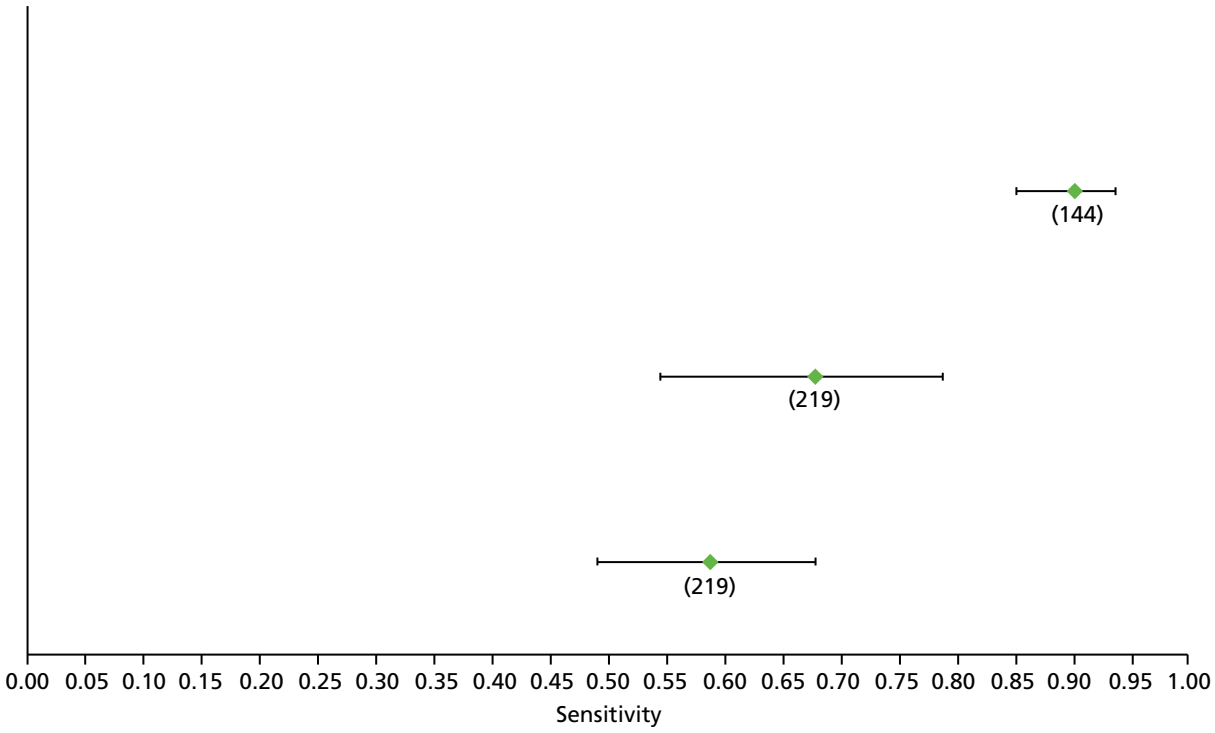

FIGURE 83 Hepatitis C: GUCI F2 – sensitivity.

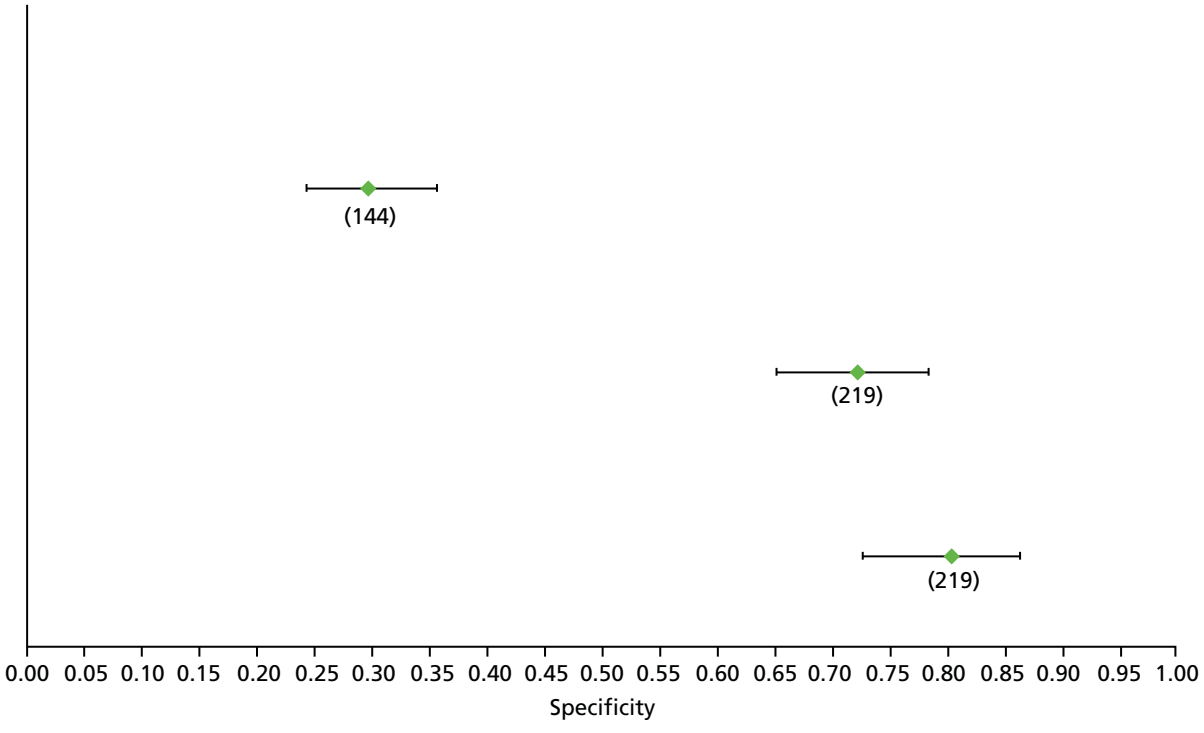

FIGURE 84 Hepatitis C: GUCI F2 – specificity.

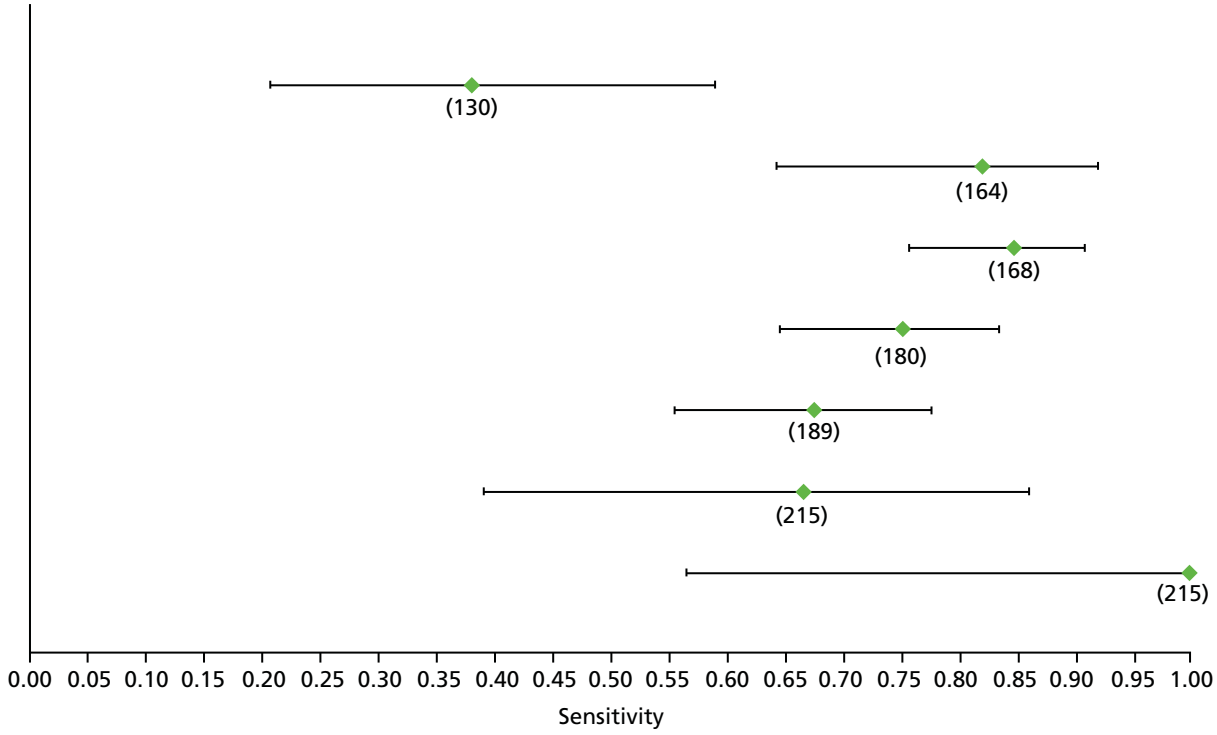

FIGURE 85 Hepatitis C: hyaluronic acid F2 – sensitivity.

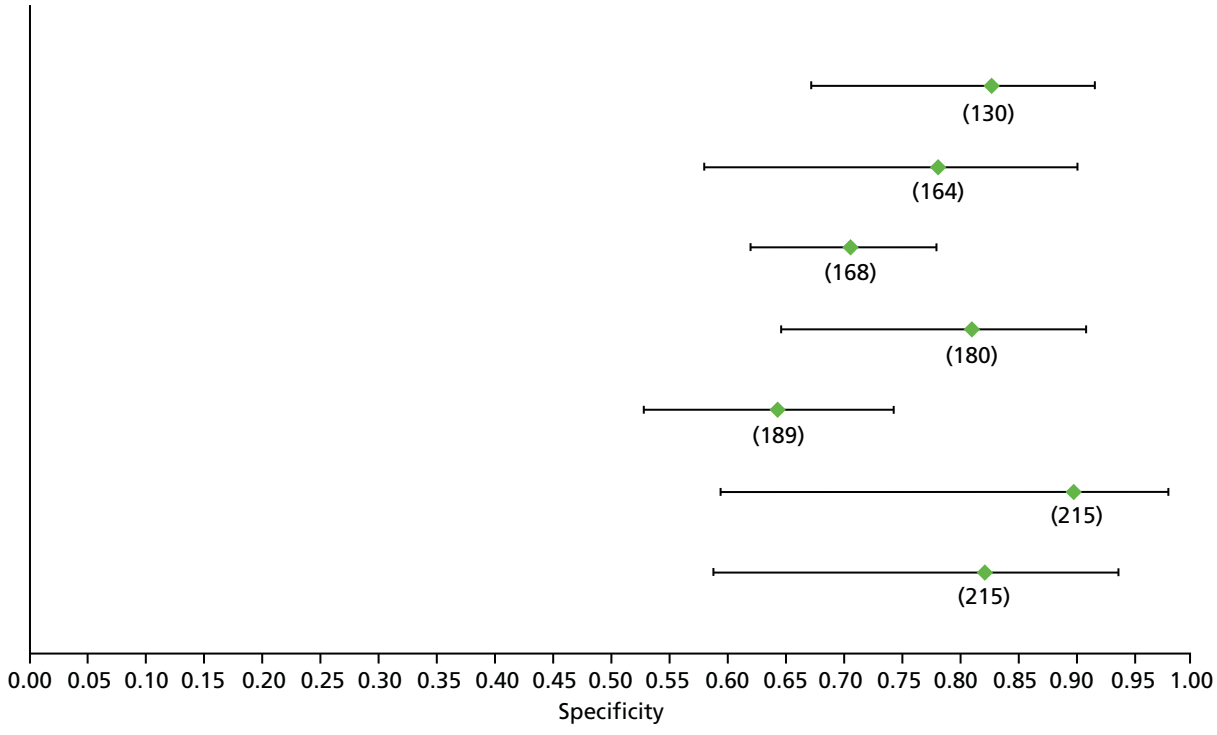

FIGURE 86 Hepatitis C: hyaluronic acid F2 – specificity.

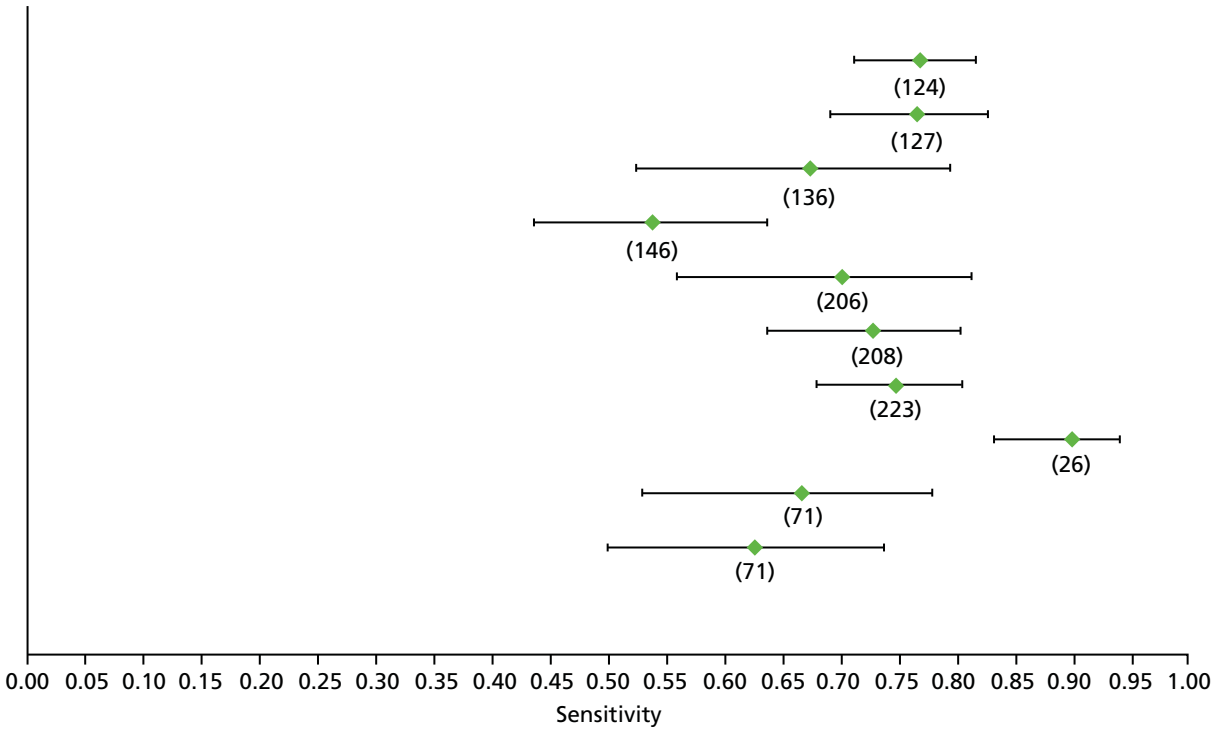

FIGURE 87 Hepatitis C: Hepascore F2 – sensitivity.

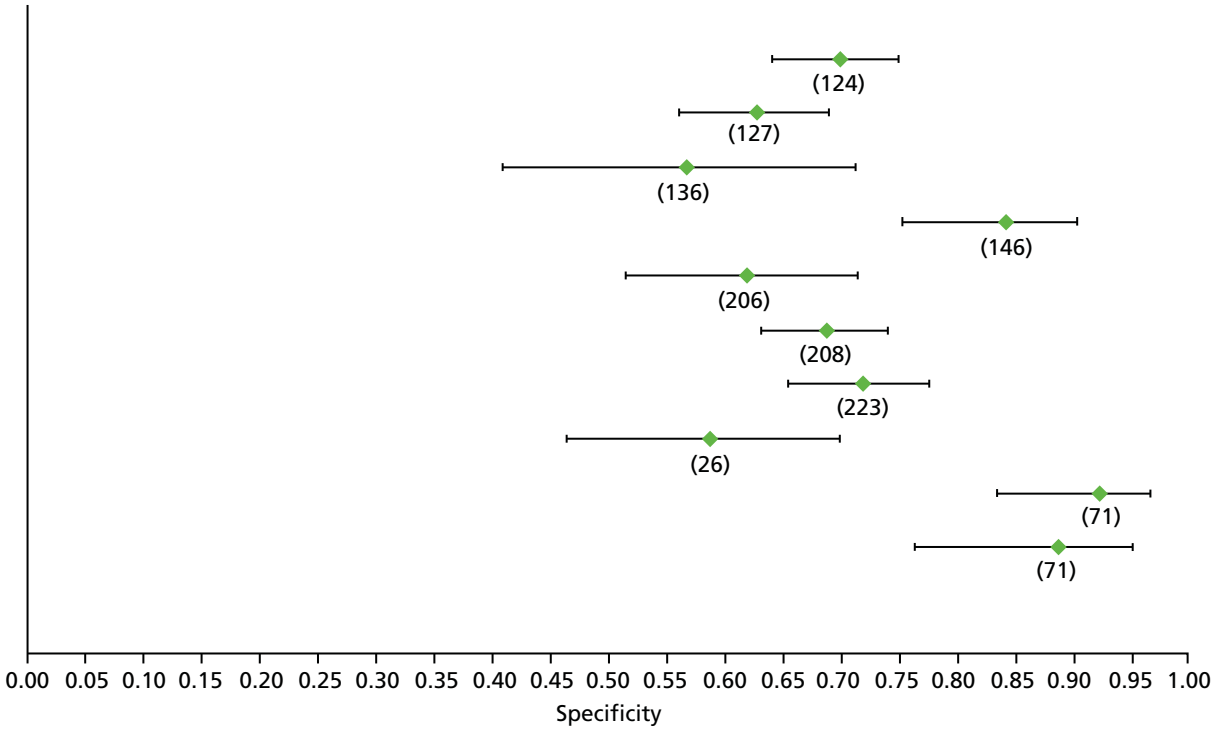

FIGURE 88 Hepatitis C: Hepascore F2 – specificity.

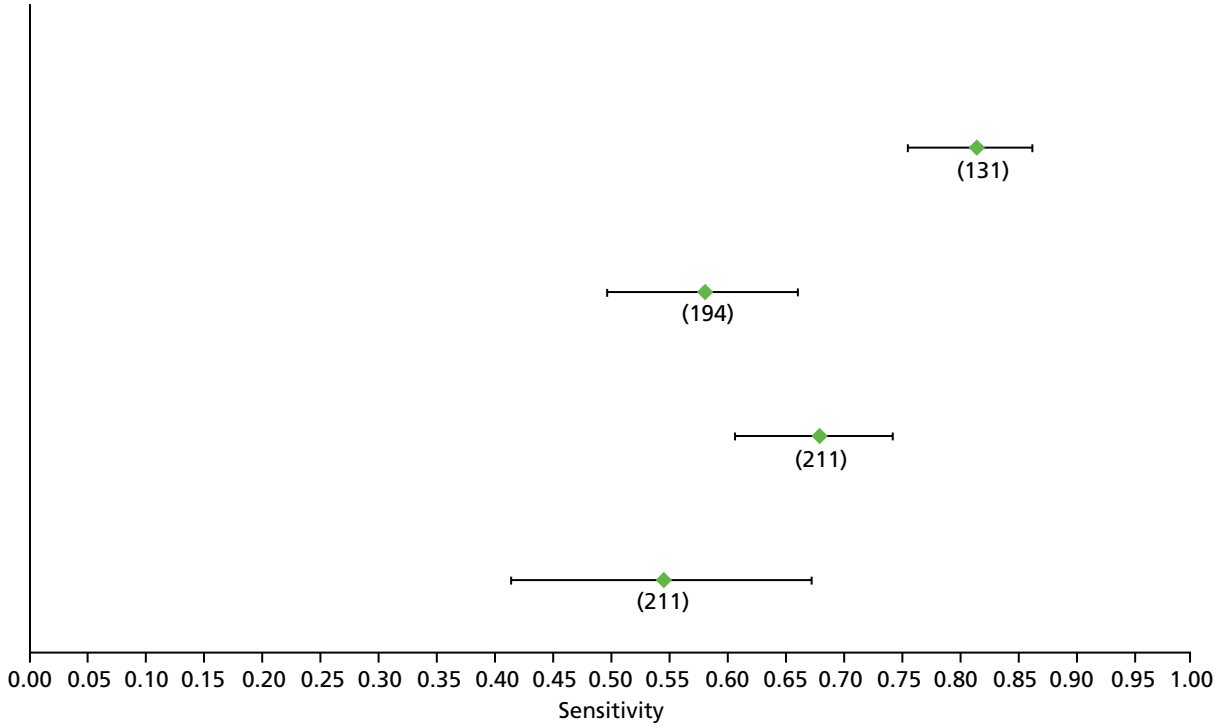

FIGURE 89 Hepatitis C: Lok's index F2 – sensitivity.

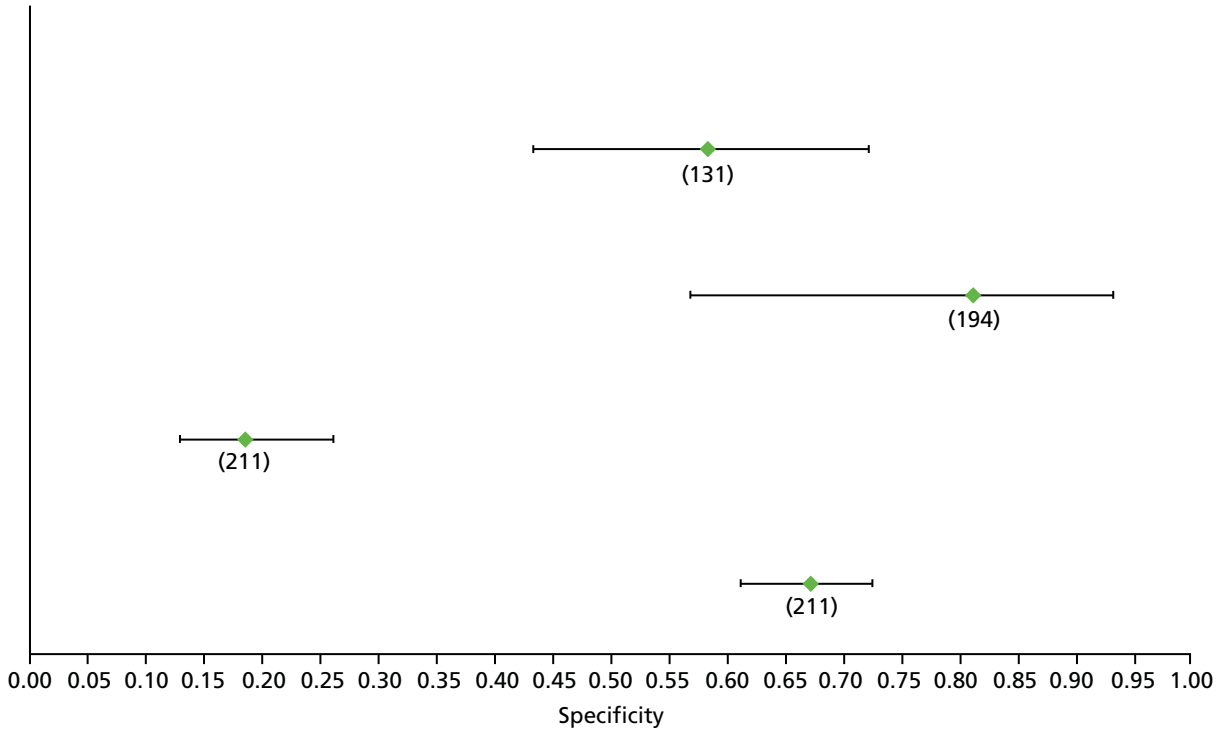

FIGURE 90 Hepatitis C: Lok's index F2 – specificity.

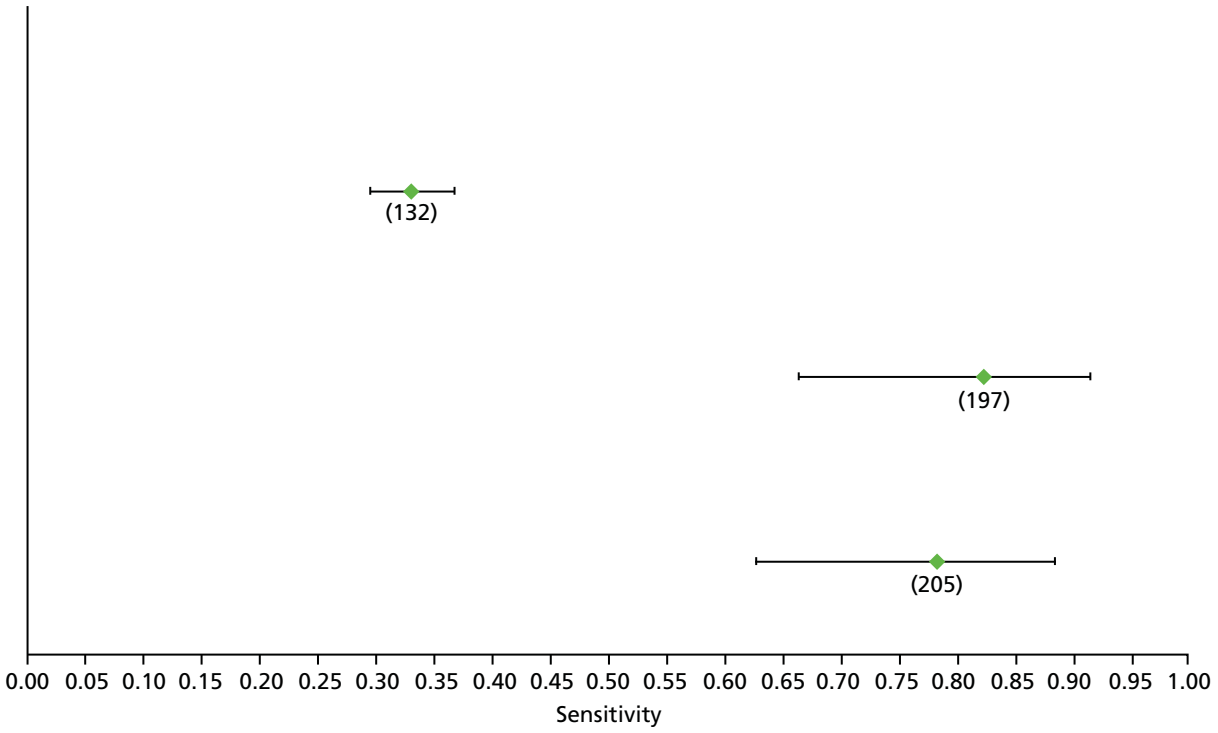

FIGURE 91 Hepatitis C: platelet–spleen F2 – sensitivity.

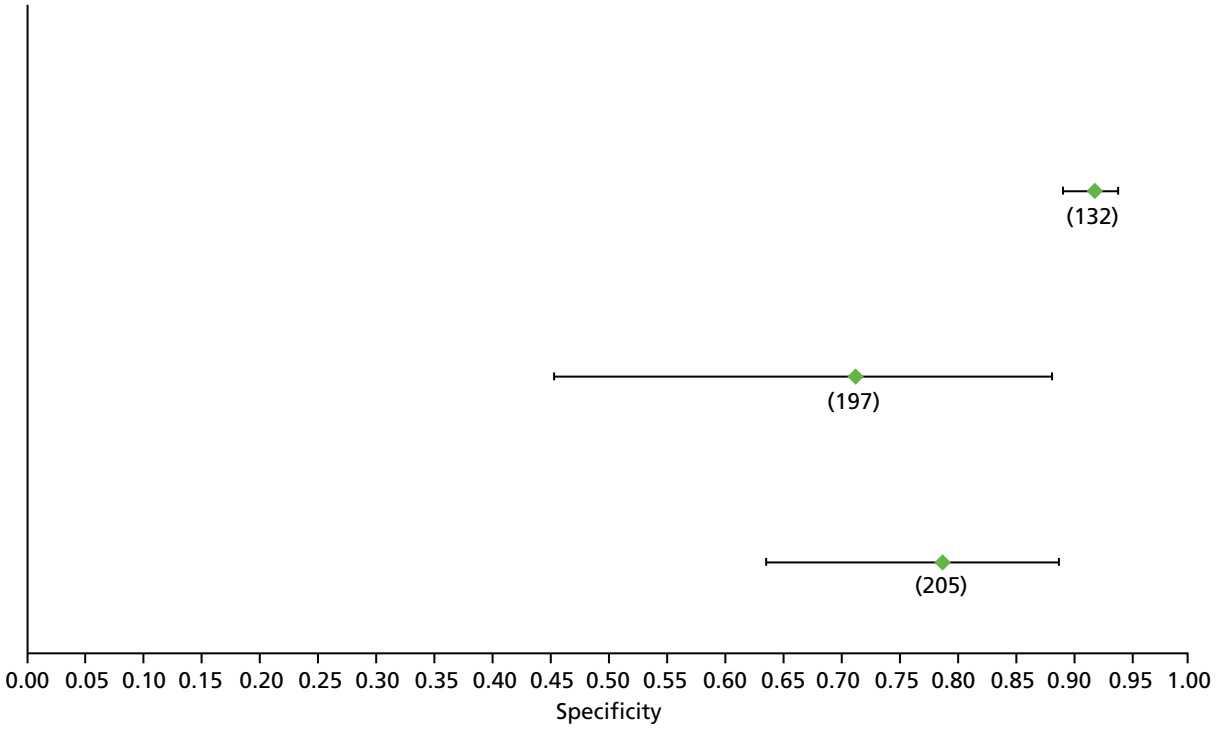

FIGURE 92 Hepatitis C: platelet–spleen F2 – specificity.

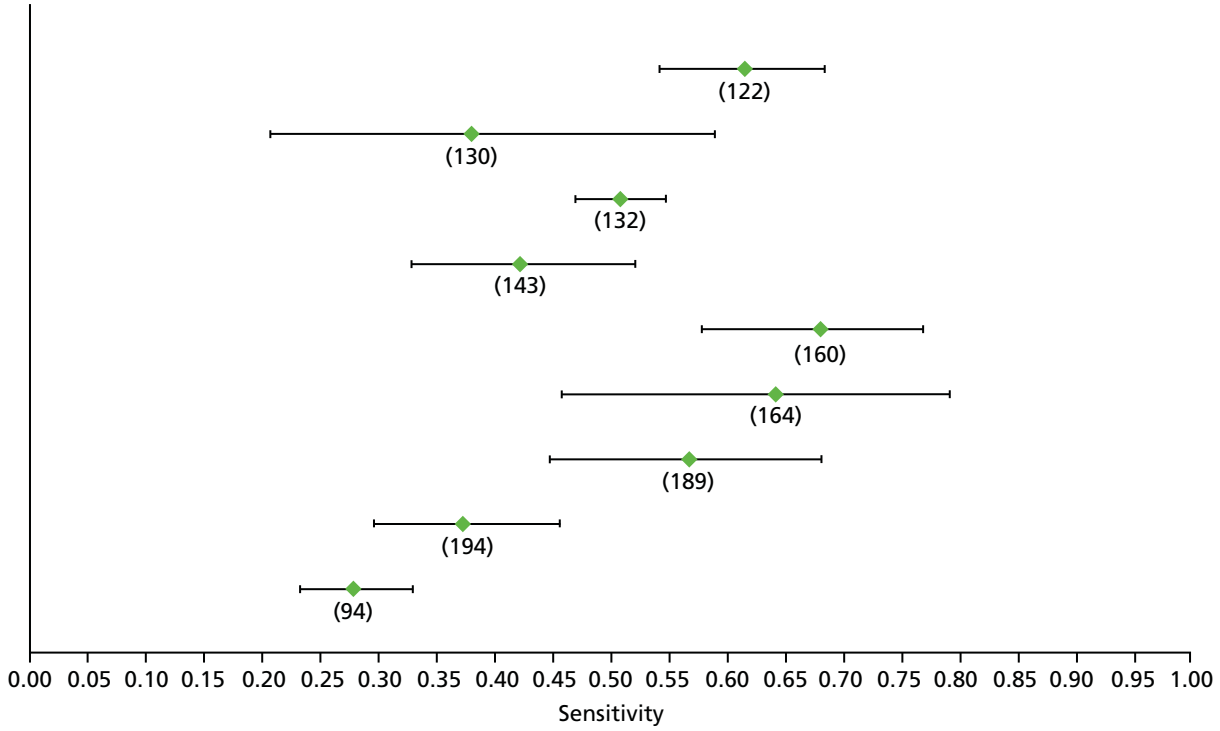

FIGURE 93 Hepatitis C: platelet F2 – sensitivity.

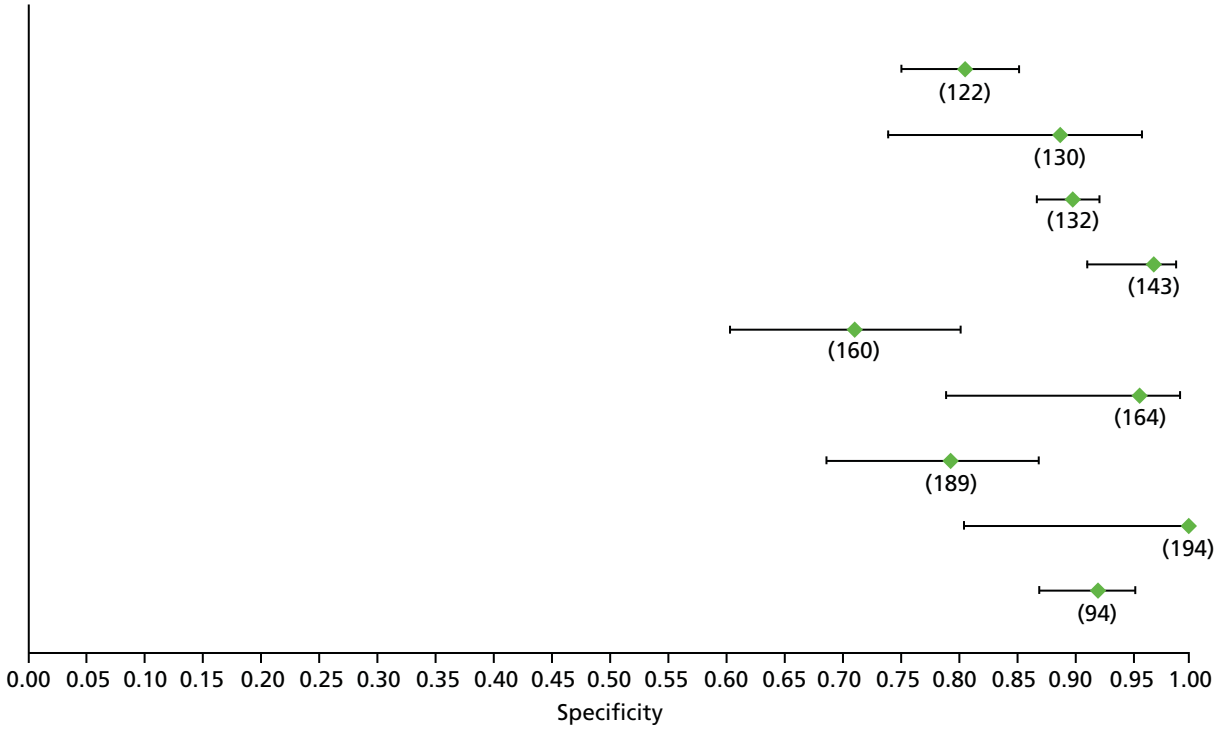

FIGURE 94 Hepatitis C: platelet F2 – specificity.

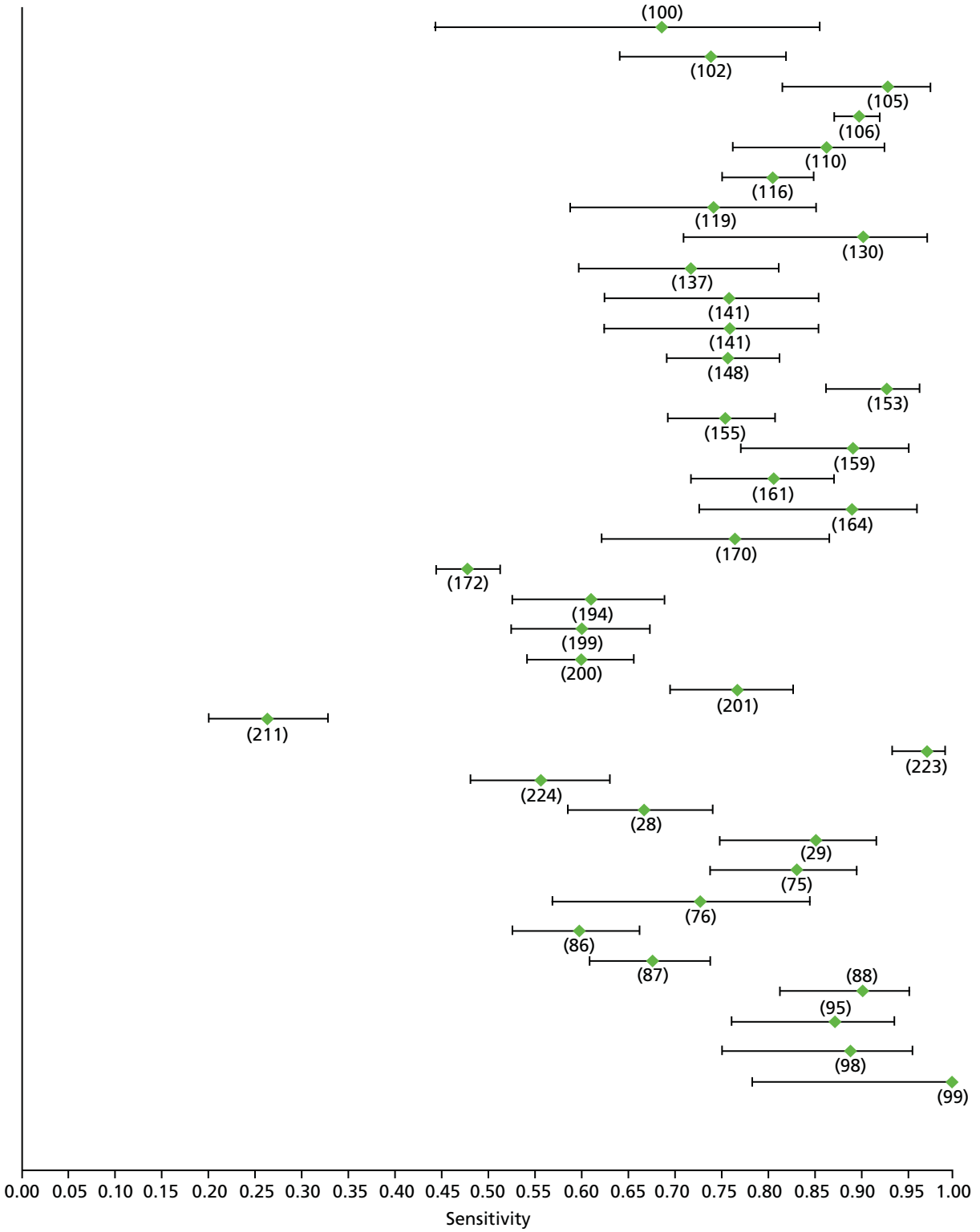

FIGURE 95 Hepatitis C: Fibroscan F2 – sensitivity.

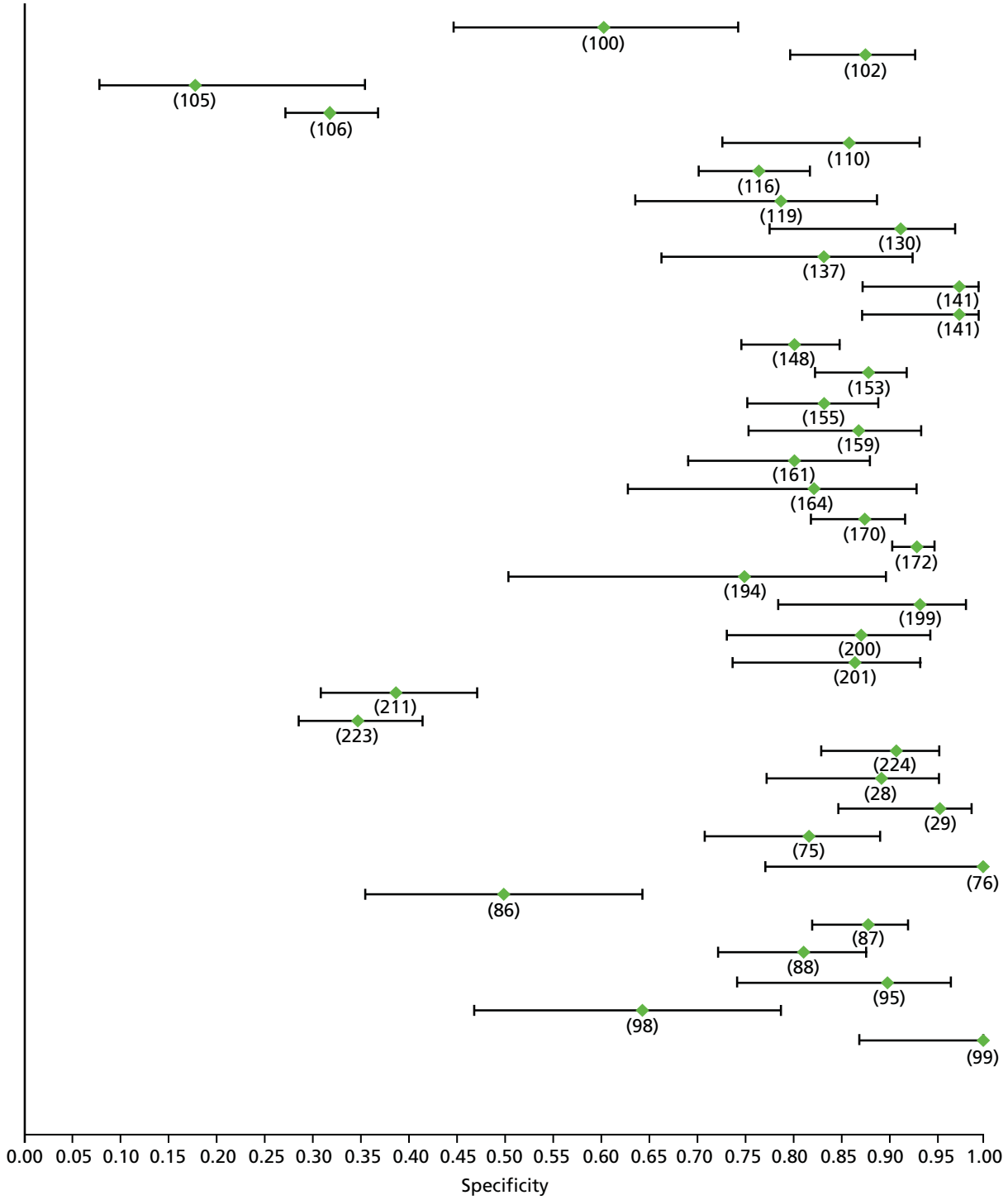

FIGURE 96 Hepatitis C: Fibroscan F2 – specificity.

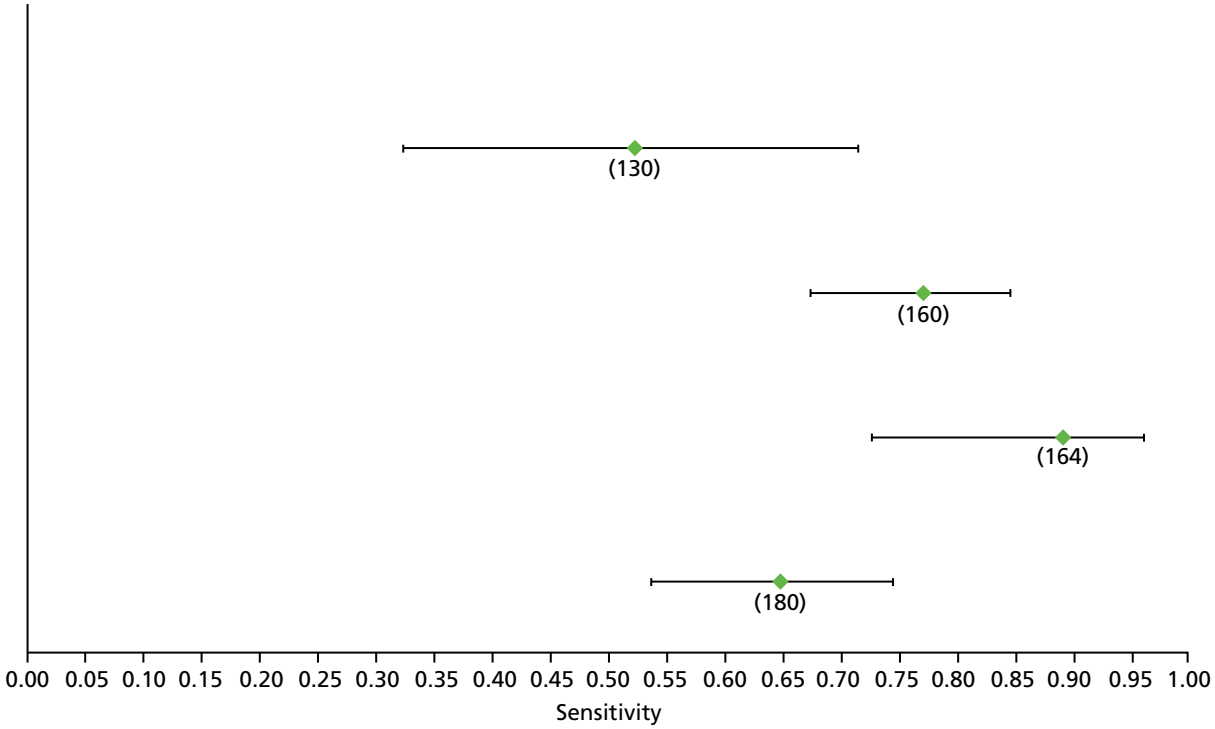

FIGURE 97 Hepatitis C: type IV collagen F2 – sensitivity.

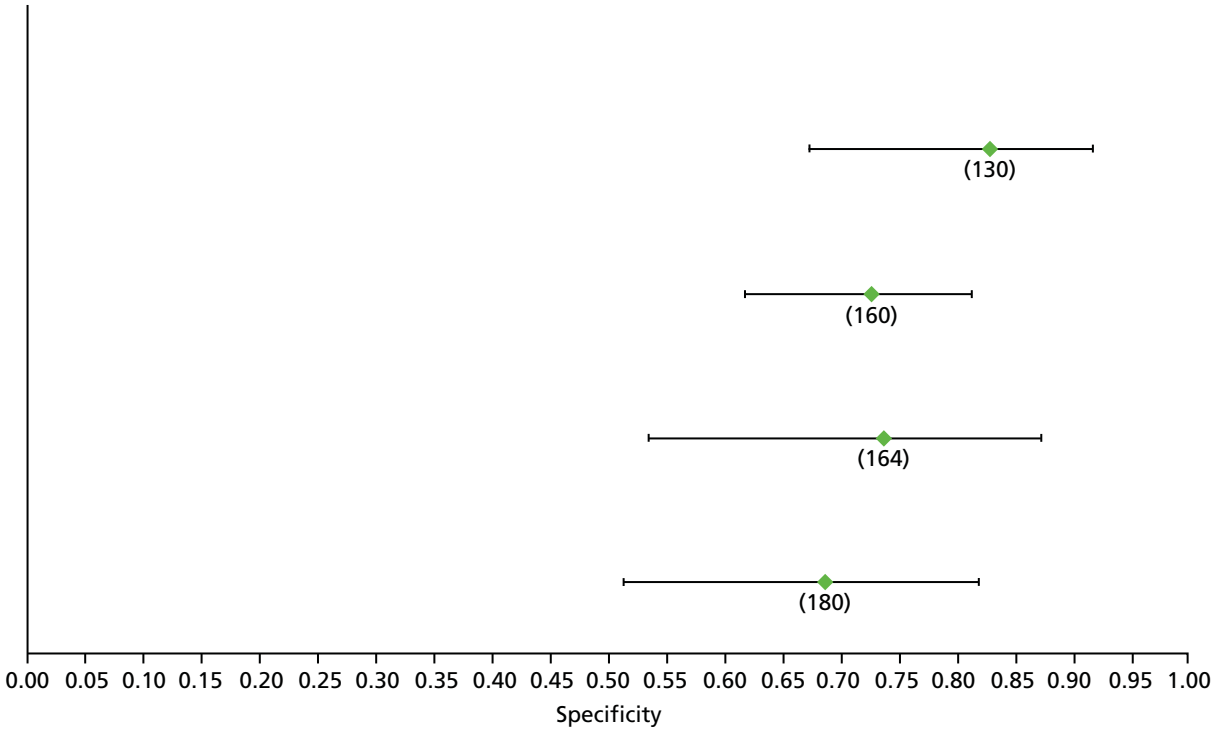

FIGURE 98 Hepatitis C: type IV collagen F2 – specificity.

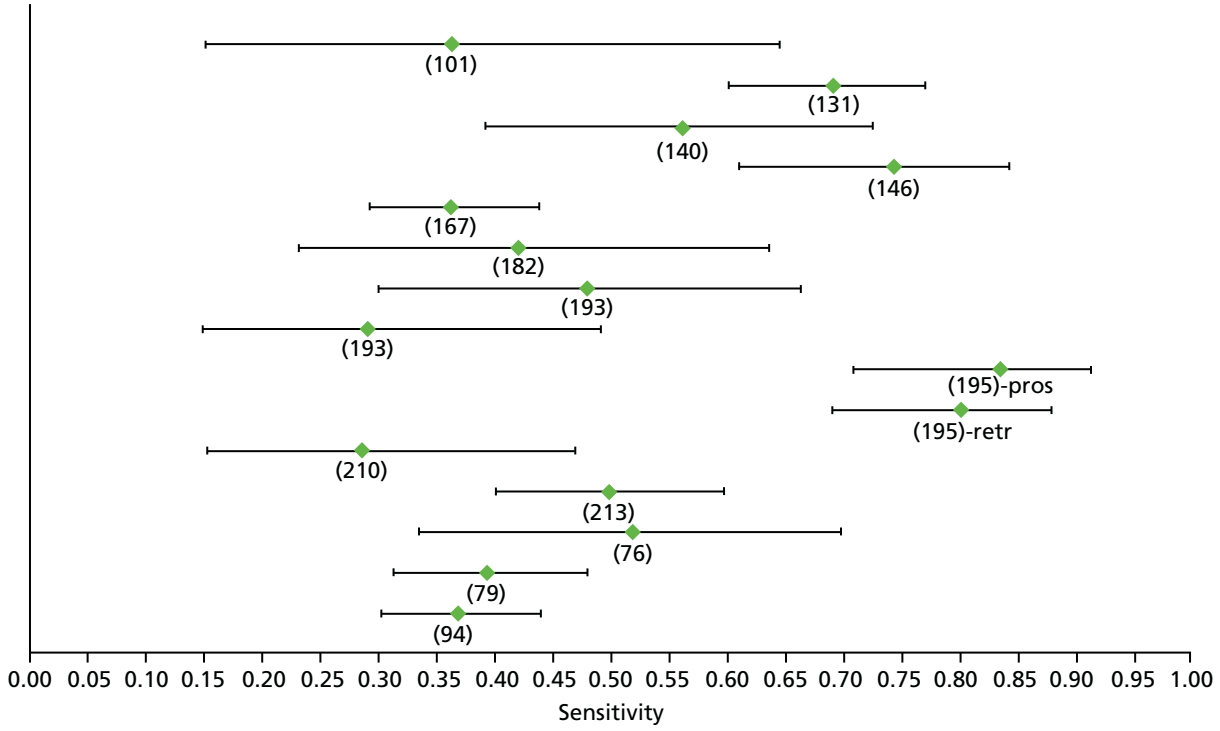

FIGURE 99 Hepatitis C: APRI F3 (high cut-off) – sensitivity. Pros, prospective; retr, retrospective.

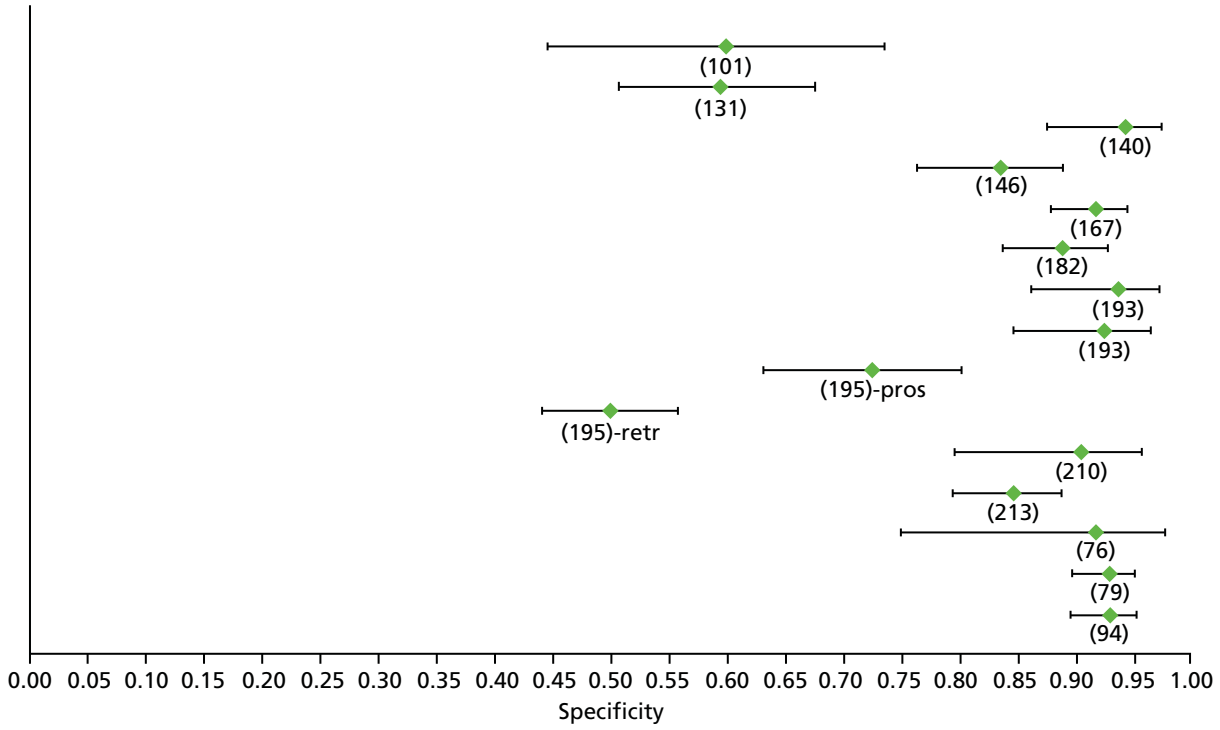

FIGURE 100 Hepatitis C: APRI F3 (high cut-off) – specificity. Pros, prospective; retr, retrospective.

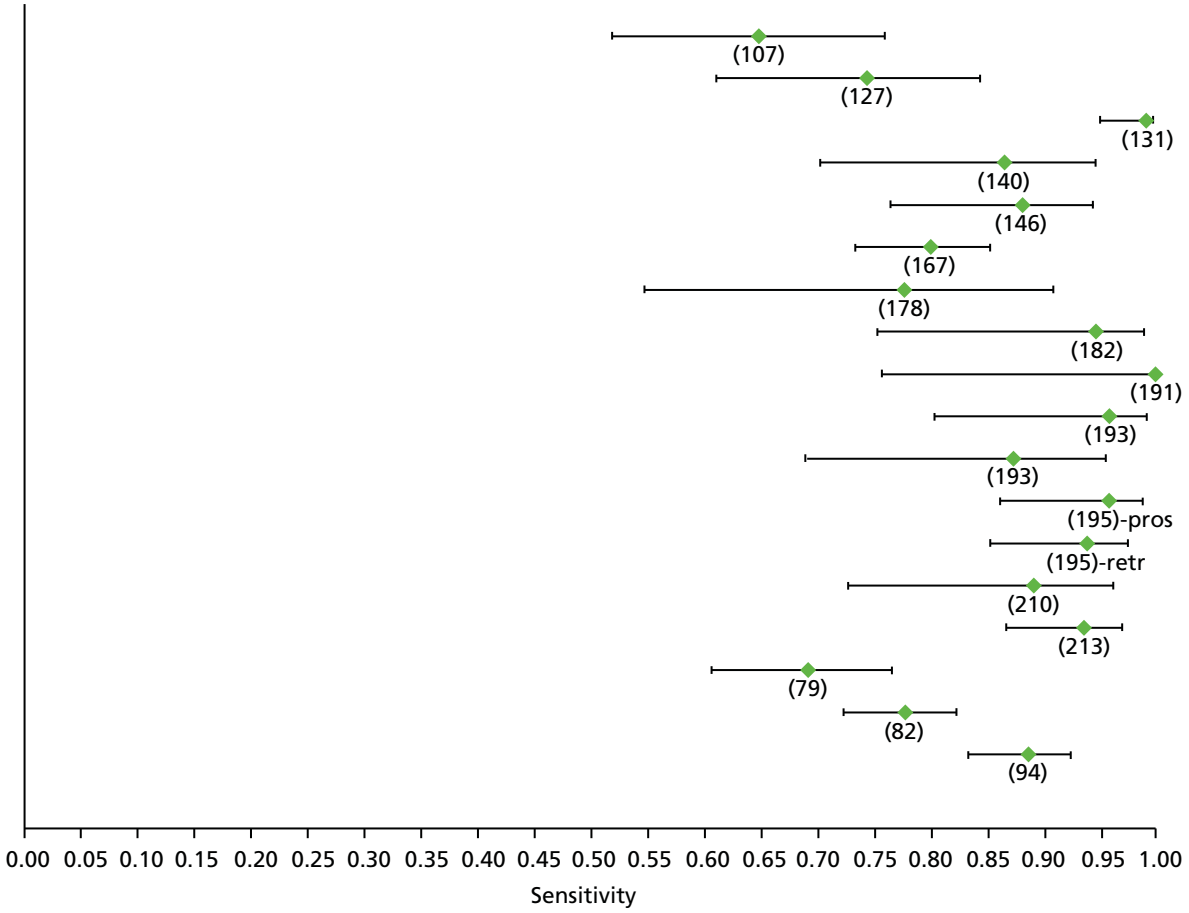

FIGURE 101 Hepatitis C: APRI F3 (low cut-off) – sensitivity. Pros, prospective; retr, retrospective.

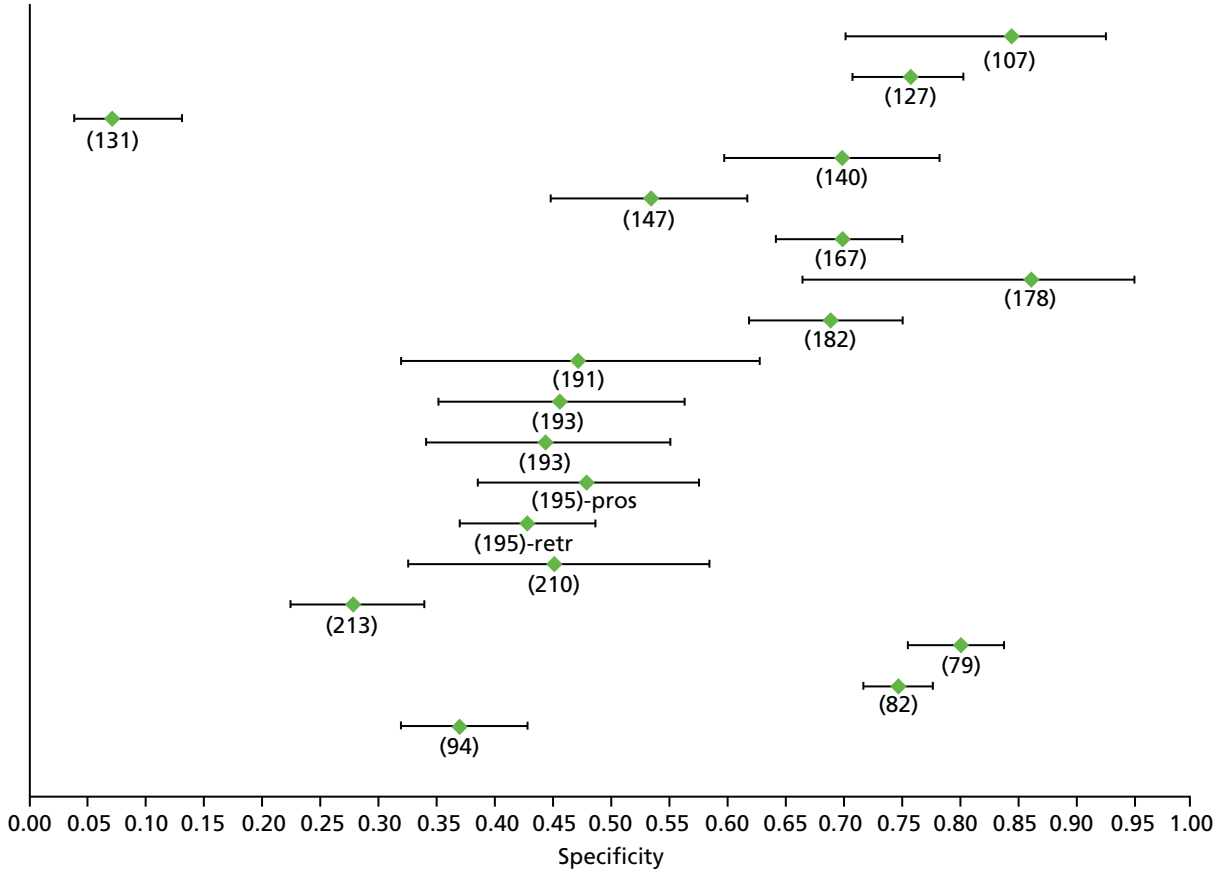

FIGURE 102 Hepatitis C: APRI F3 (low cut-off) – specificity. Pros, prospective; retr, retrospective.

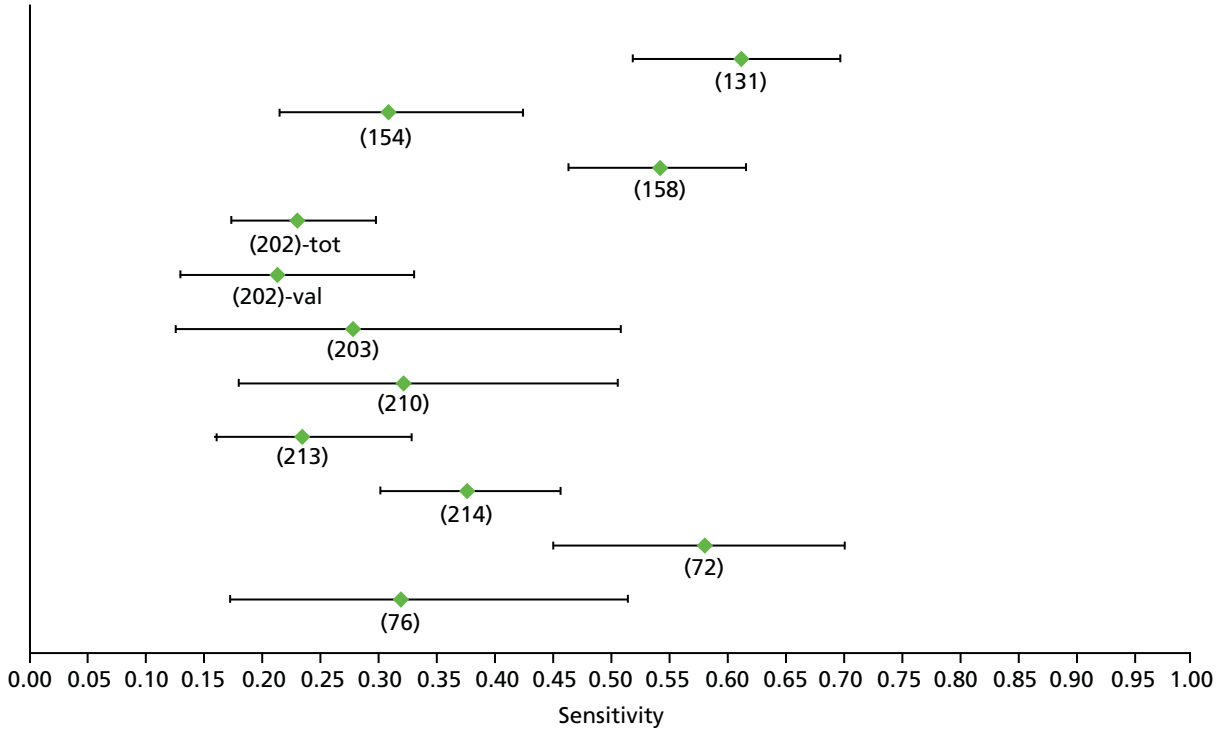

FIGURE 103 Hepatitis C: FIB-4 F3 (high cut-off) – sensitivity. Tot, total; val, validation cohort.

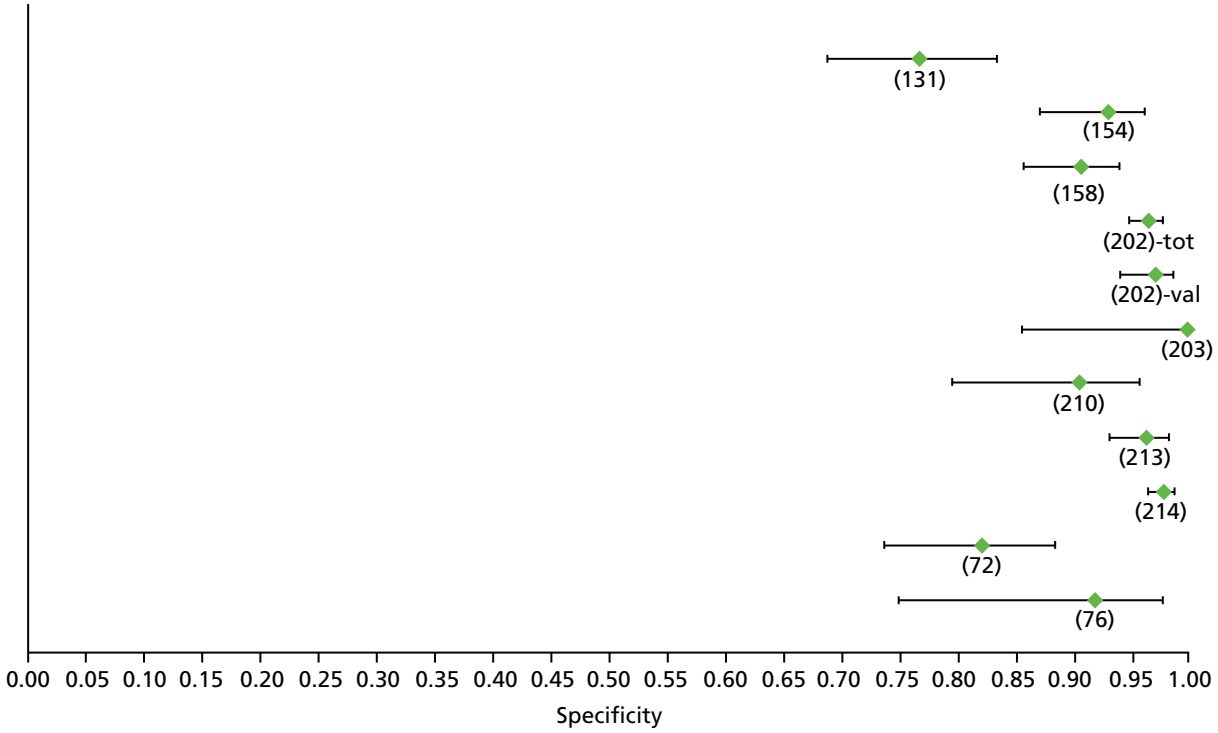

FIGURE 104 Hepatitis C: FIB-4 F3 (high cut-off) – specificity. Tot, total; val, validation cohort.

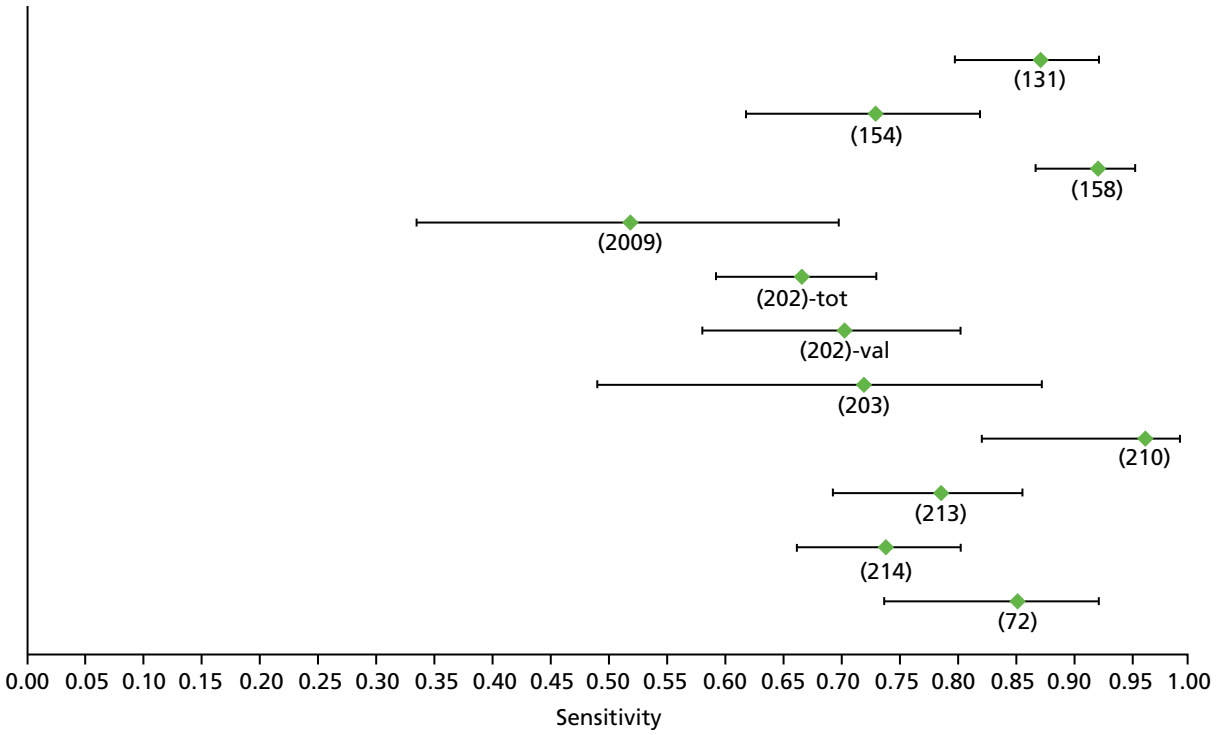

FIGURE 105 Hepatitis C: FIB-4 F3 (low cut-off) – sensitivity. Tot, total; val, validation cohort.

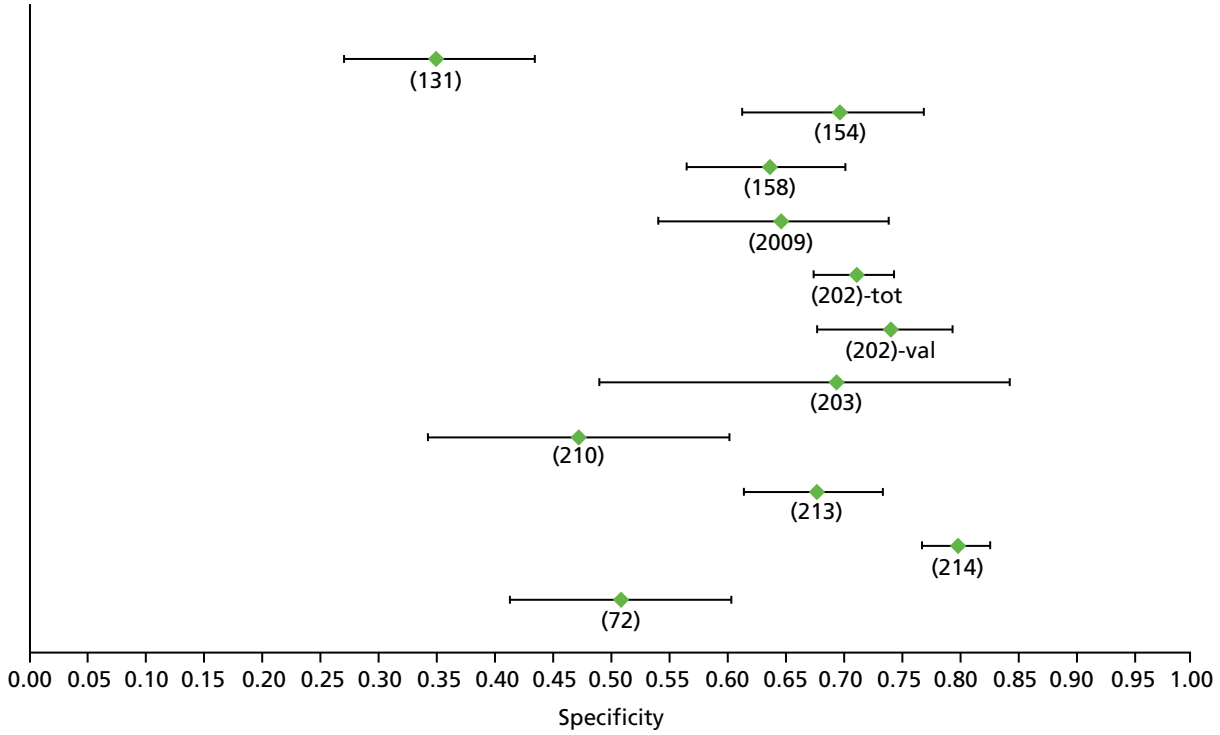

FIGURE 106 Hepatitis C: FIB-4 F3 (low cut-off) – specificity. Tot, total; val, validation cohort.

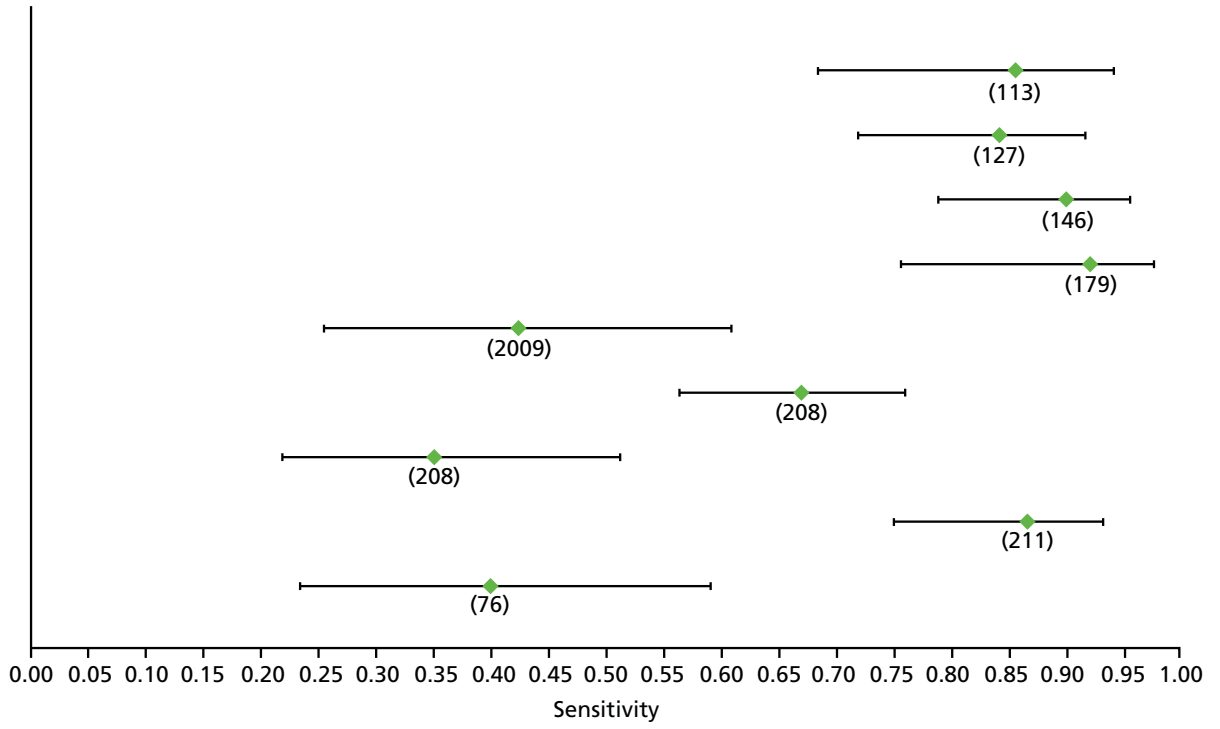

FIGURE 107 Hepatitis C: Fibrotest F3 – sensitivity.

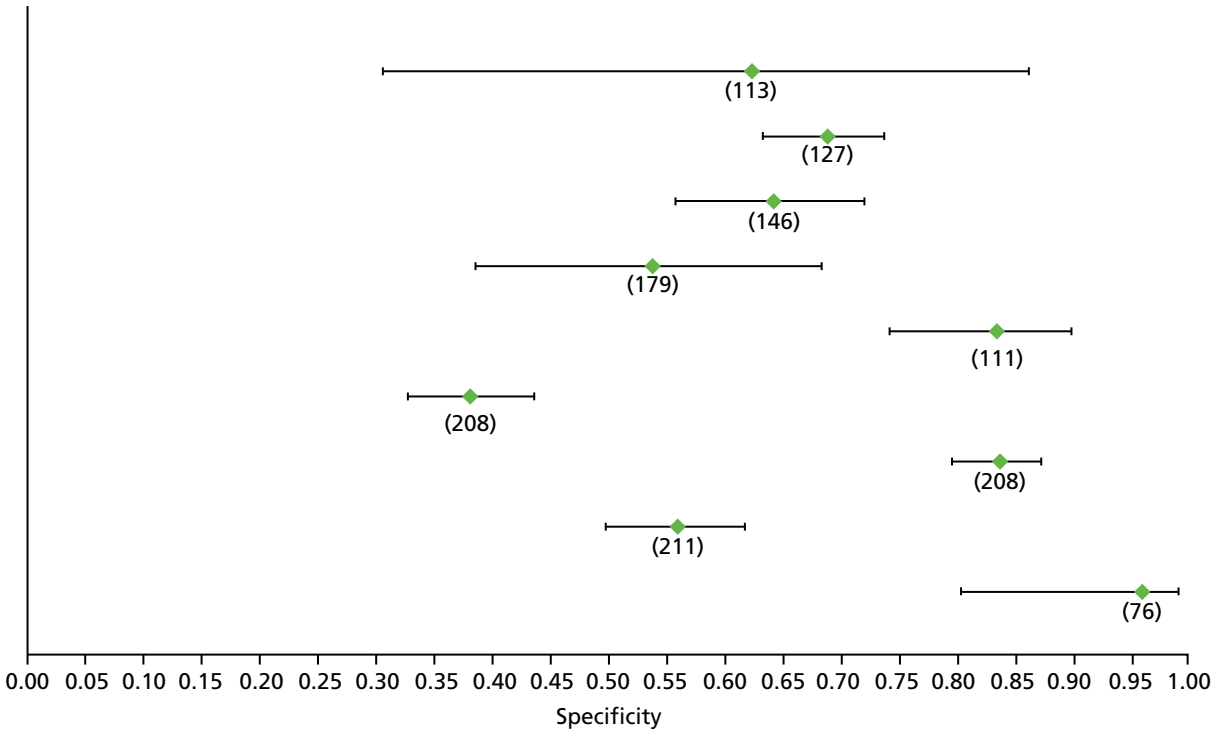

FIGURE 108 Hepatitis C: Fibrotest F3 – specificity.

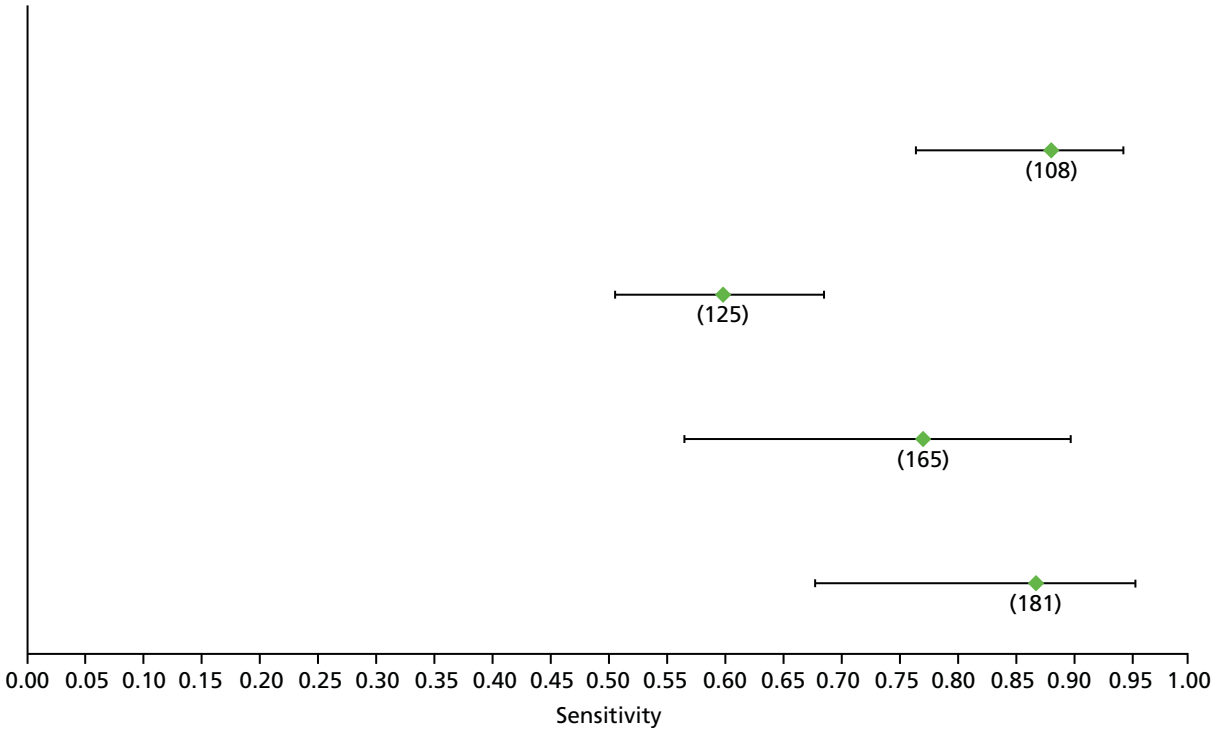

FIGURE 109 Hepatitis C: hyaluronic acid F3 – sensitivity.

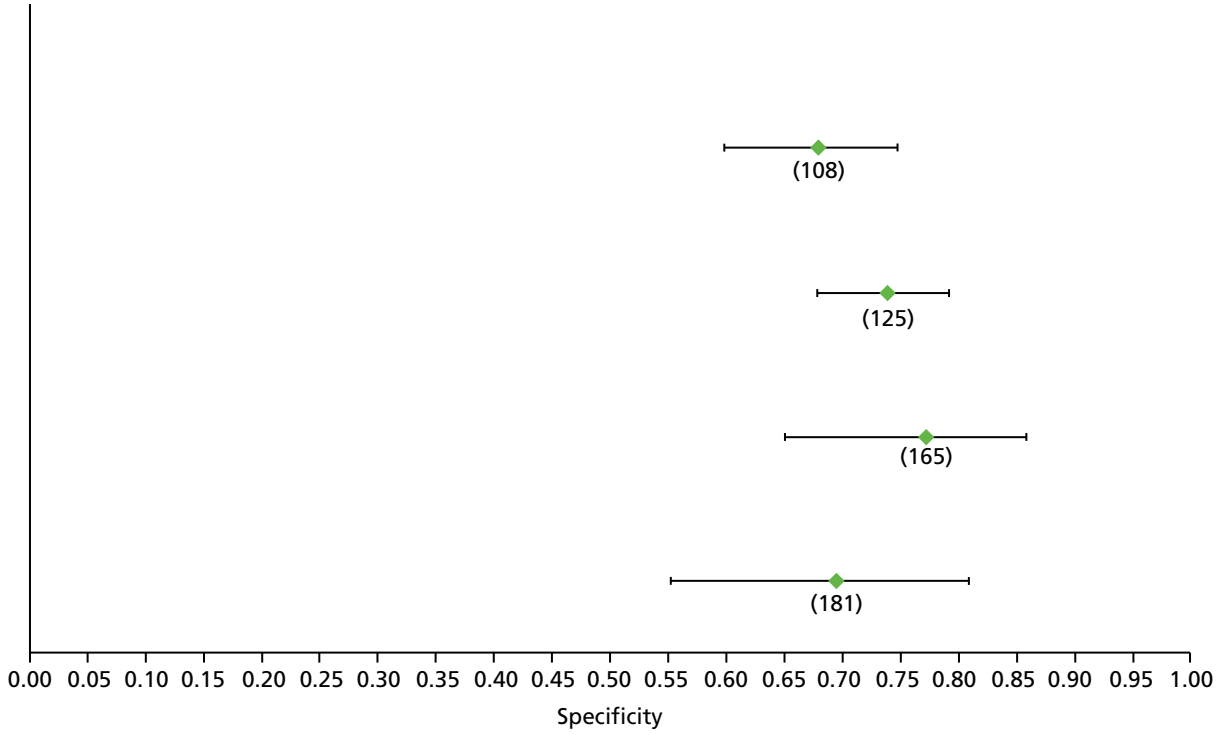

FIGURE 110 Hepatitis C: hyaluronic acid F3 – specificity.

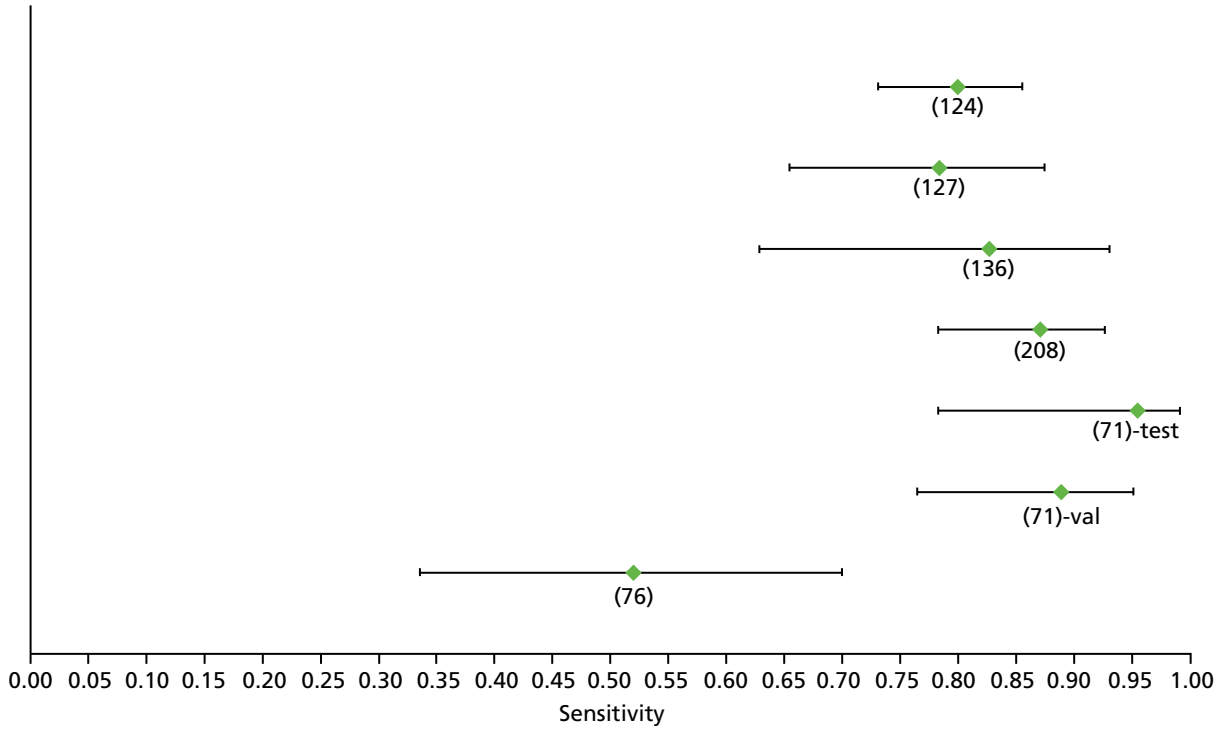

FIGURE 111 Hepatitis C: Hepascore F3 – sensitivity.

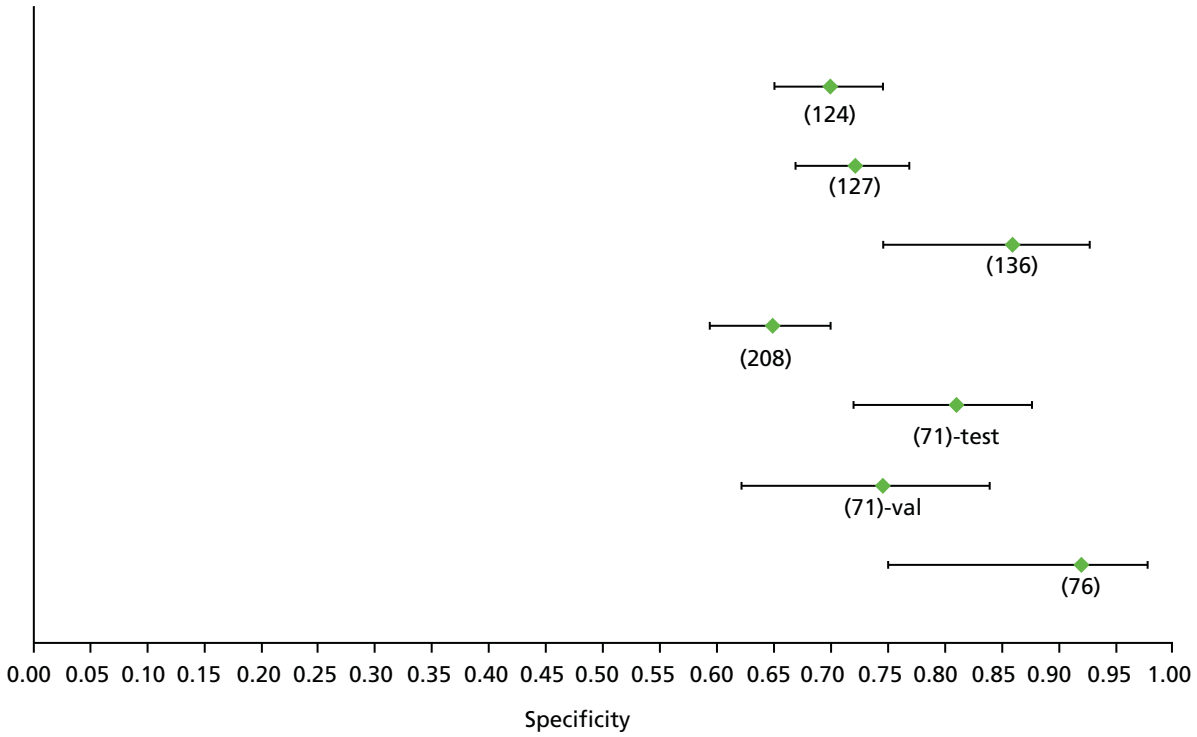

FIGURE 112 Hepatitis C: Hepascore F3 – specificity.

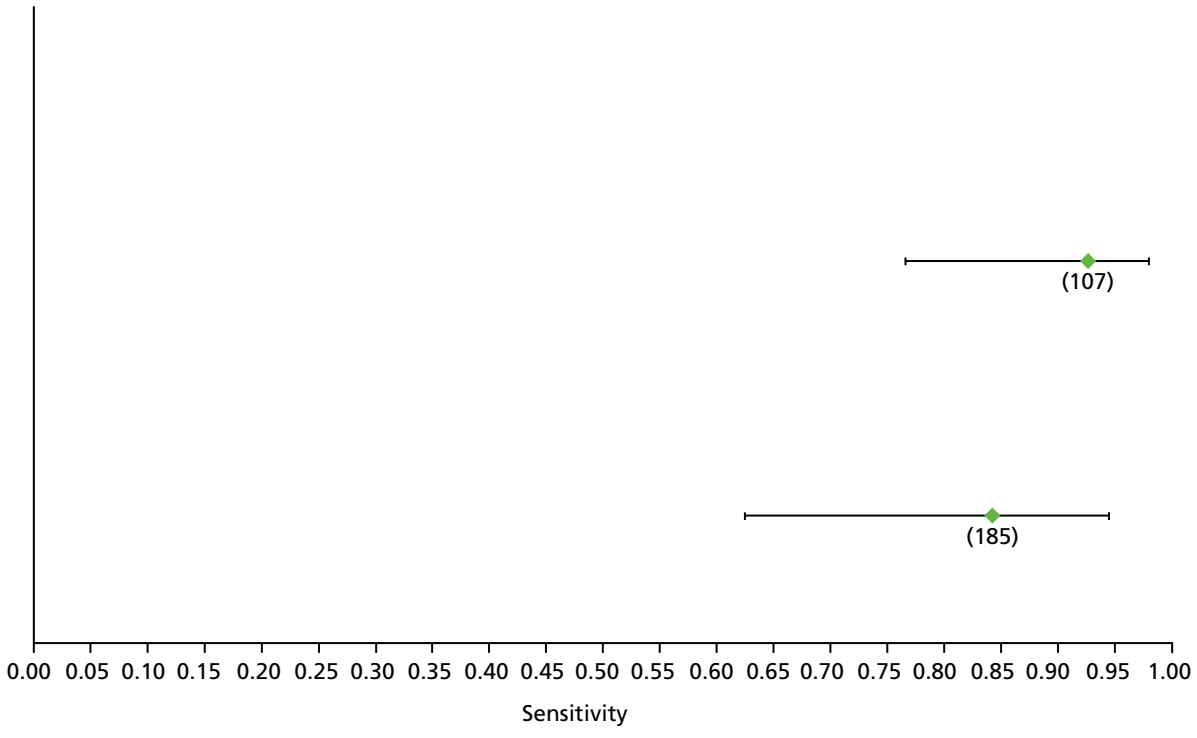

FIGURE 113 Hepatitis C: 13C-caffeine breath test F4 – sensitivity.

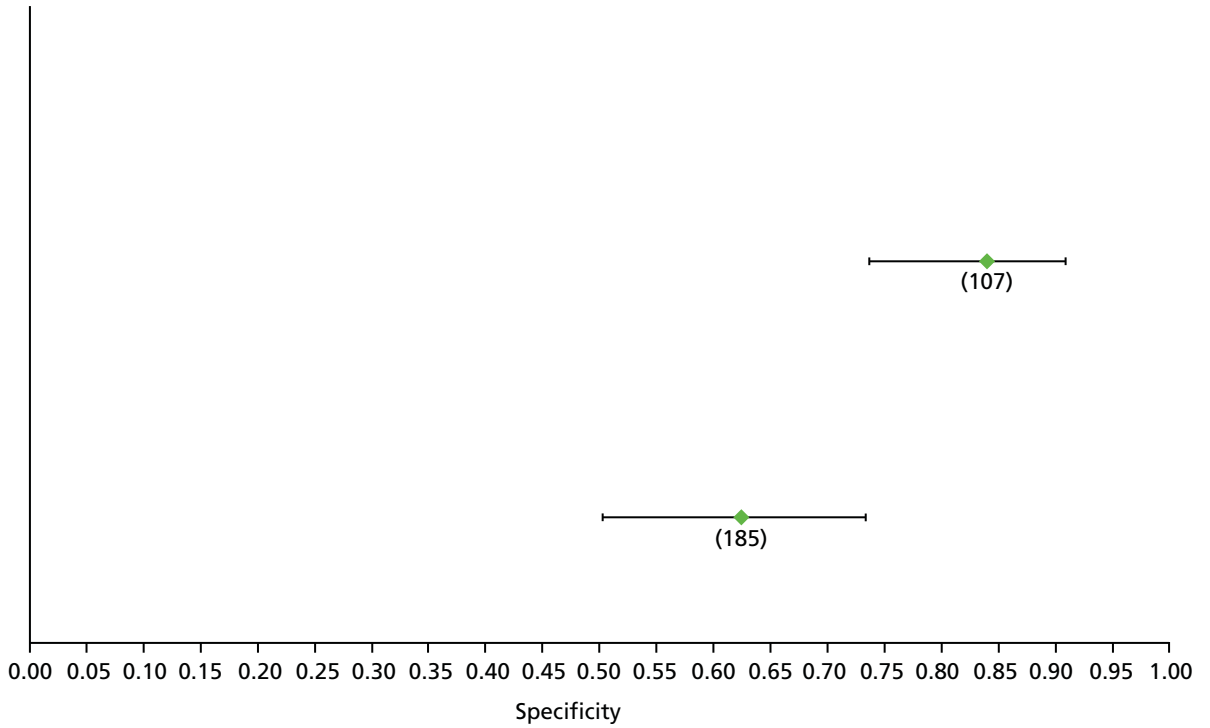

FIGURE 114 Hepatitis C: 13C-caffeine breath test F4 – specificity.

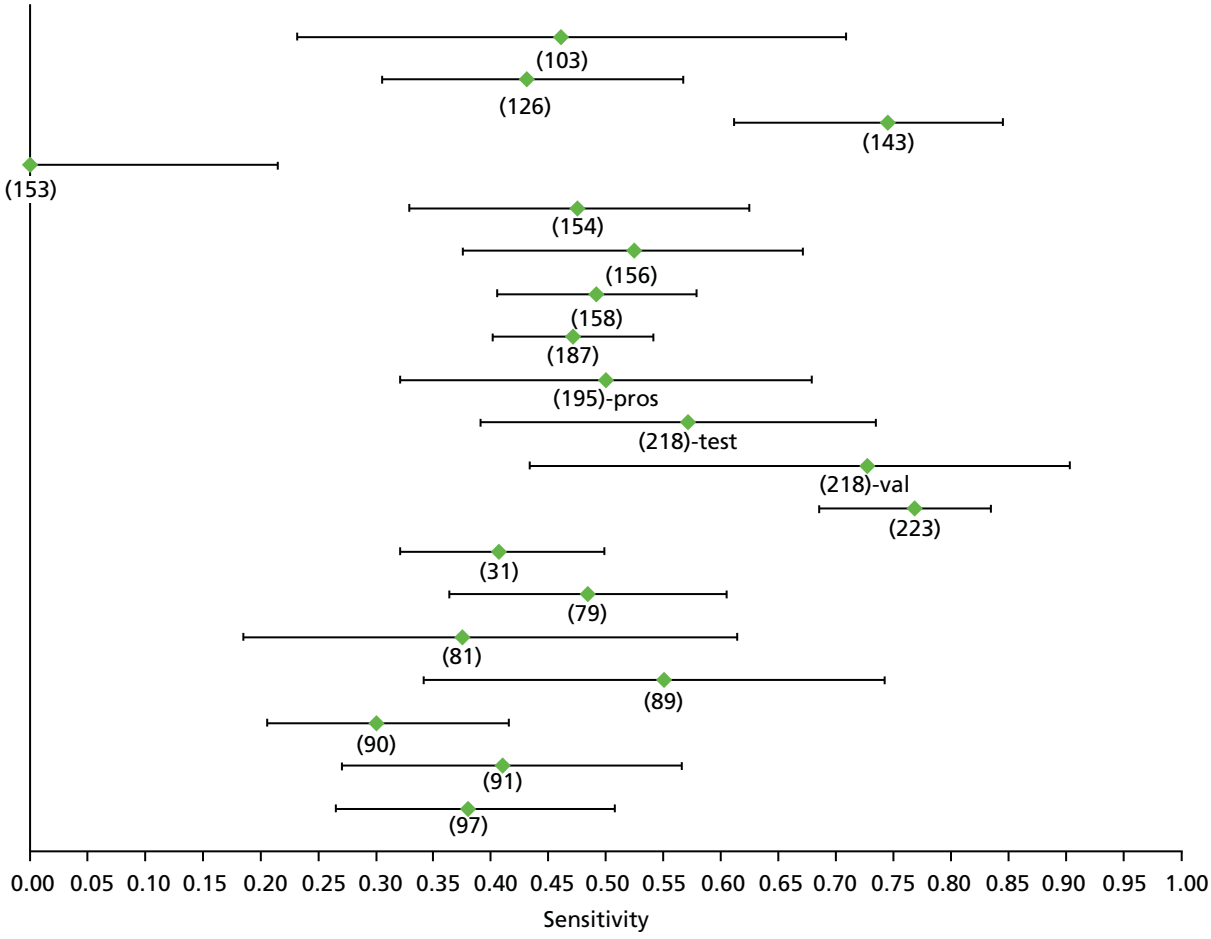

FIGURE 115 Hepatitis C: APRI F4 (high cut-off) – sensitivity.

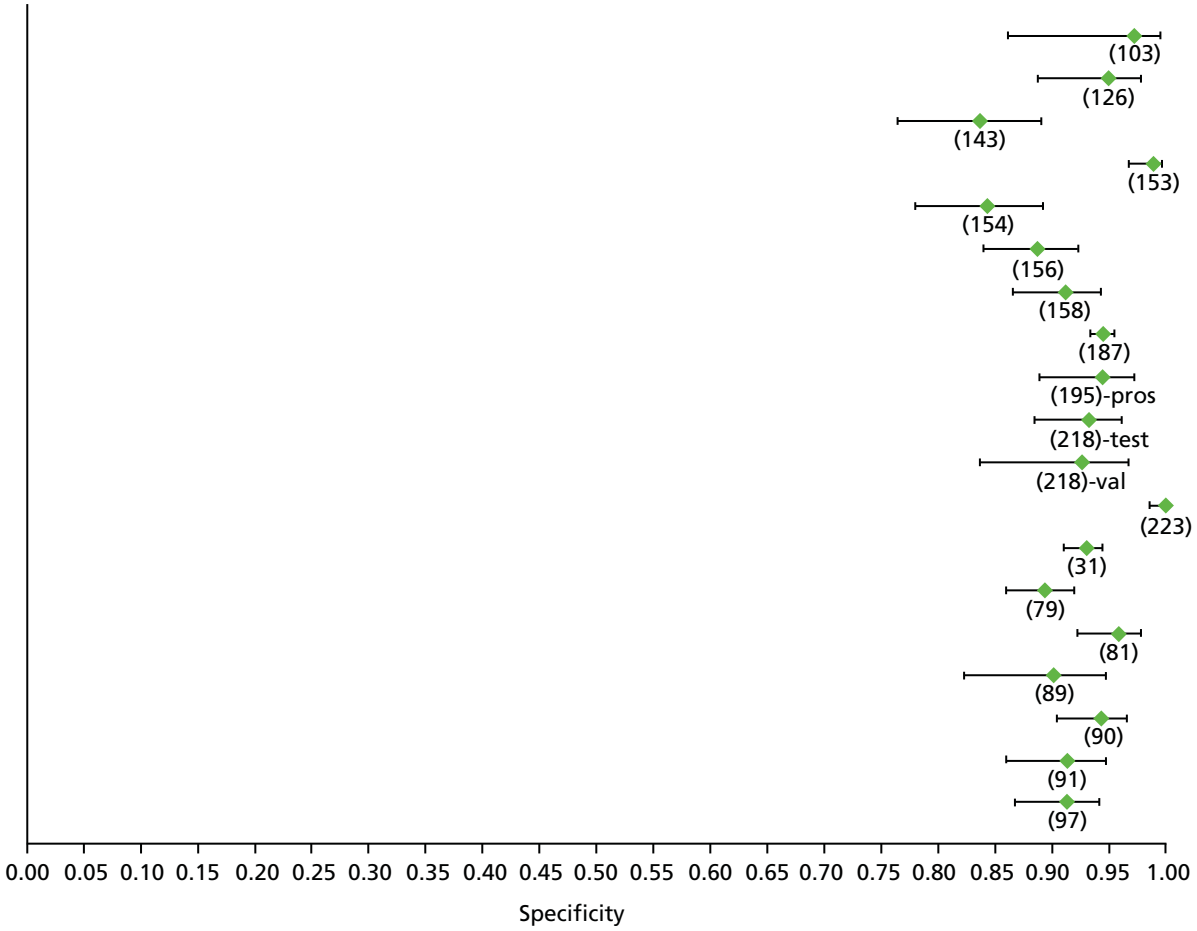

FIGURE 116 Hepatitis C: APRI F4 (high cut-off) – specificity.

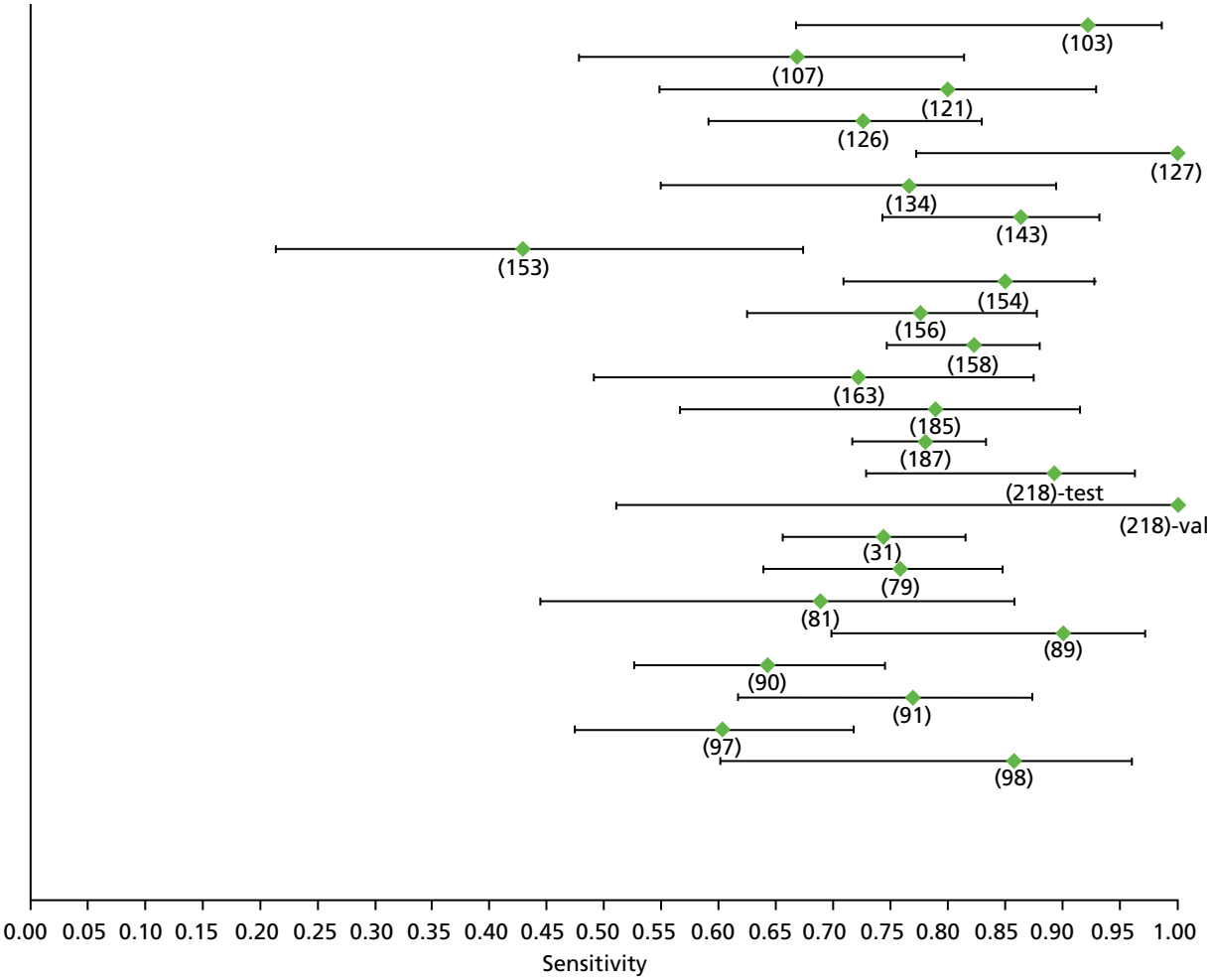

FIGURE 117 Hepatitis C: APRI F4 (low cut-off) – sensitivity.

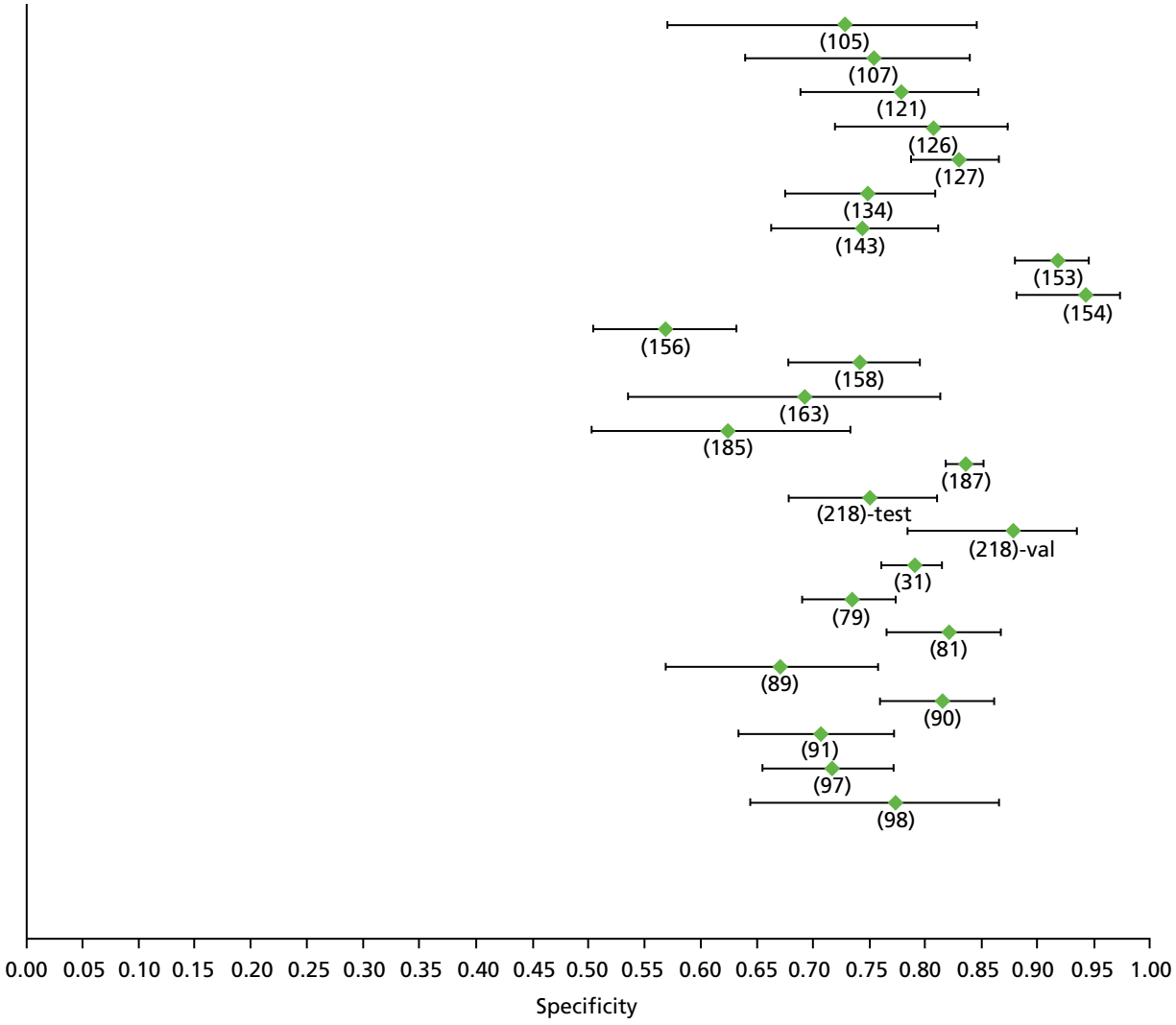

FIGURE 118 Hepatitis C: APRI F4 (low cut-off) – specificity.

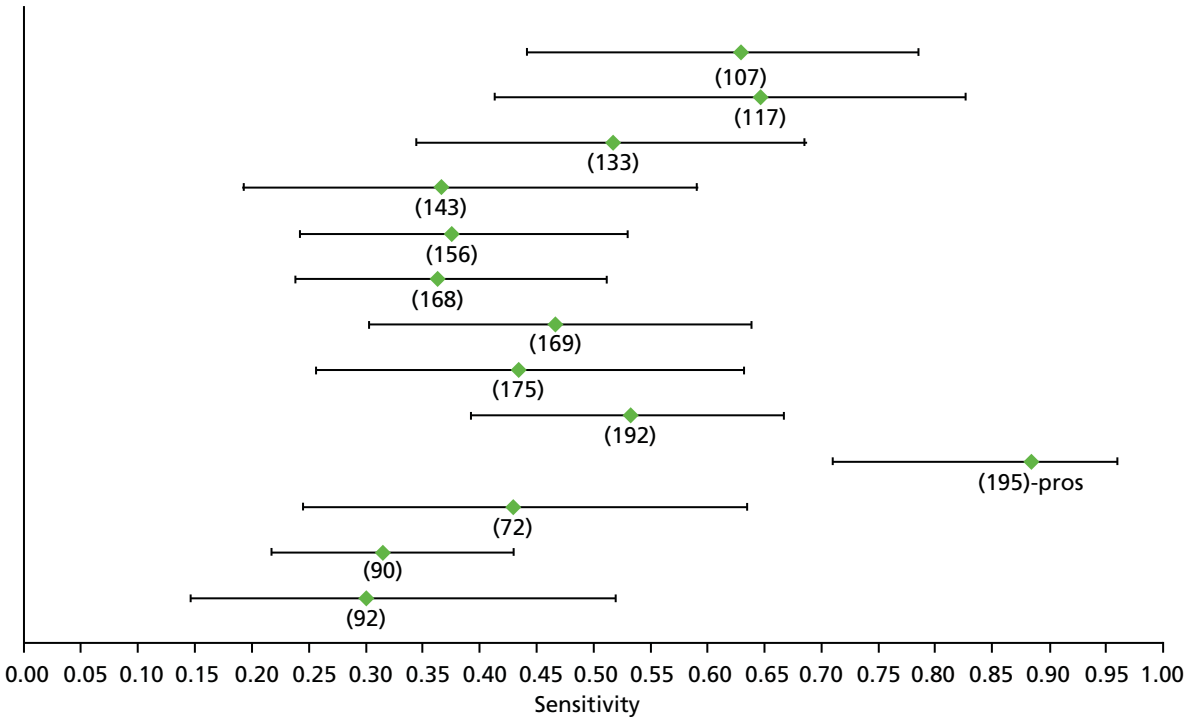

FIGURE 119 Hepatitis C: AST–ALT ratio F4 – sensitivity.

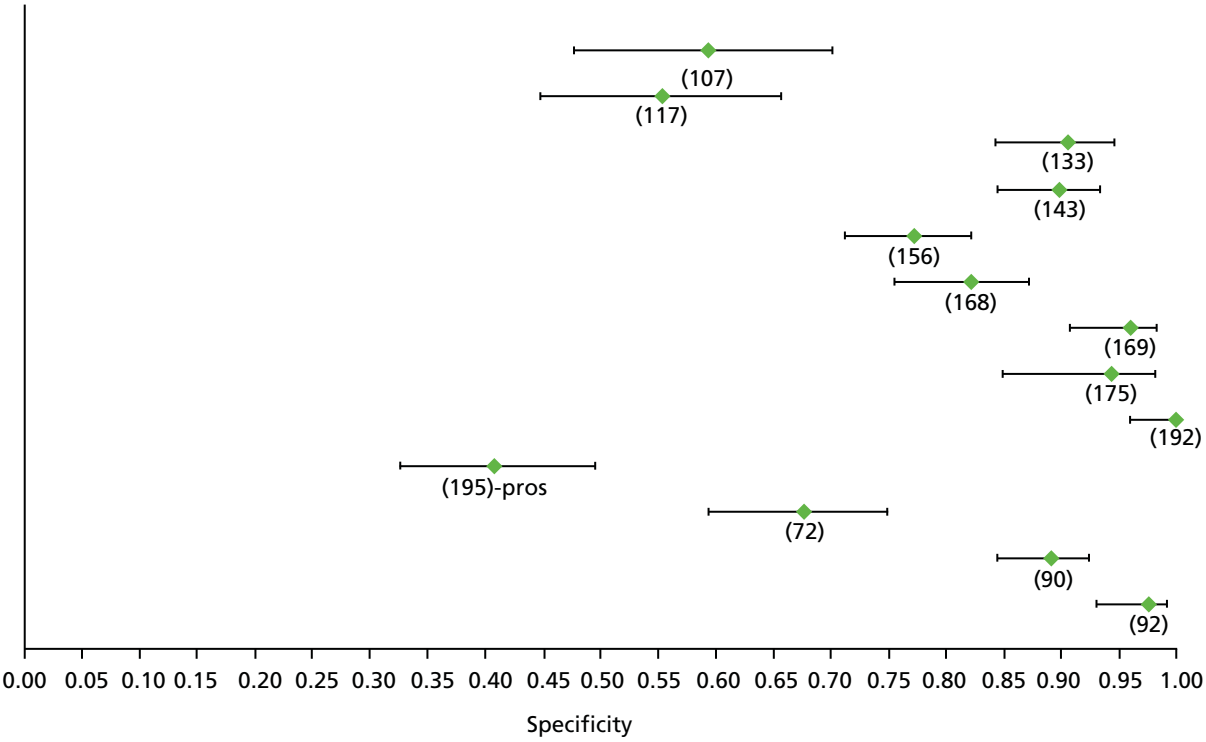

FIGURE 120 Hepatitis C: AST–ALT ratio F4 – specificity.

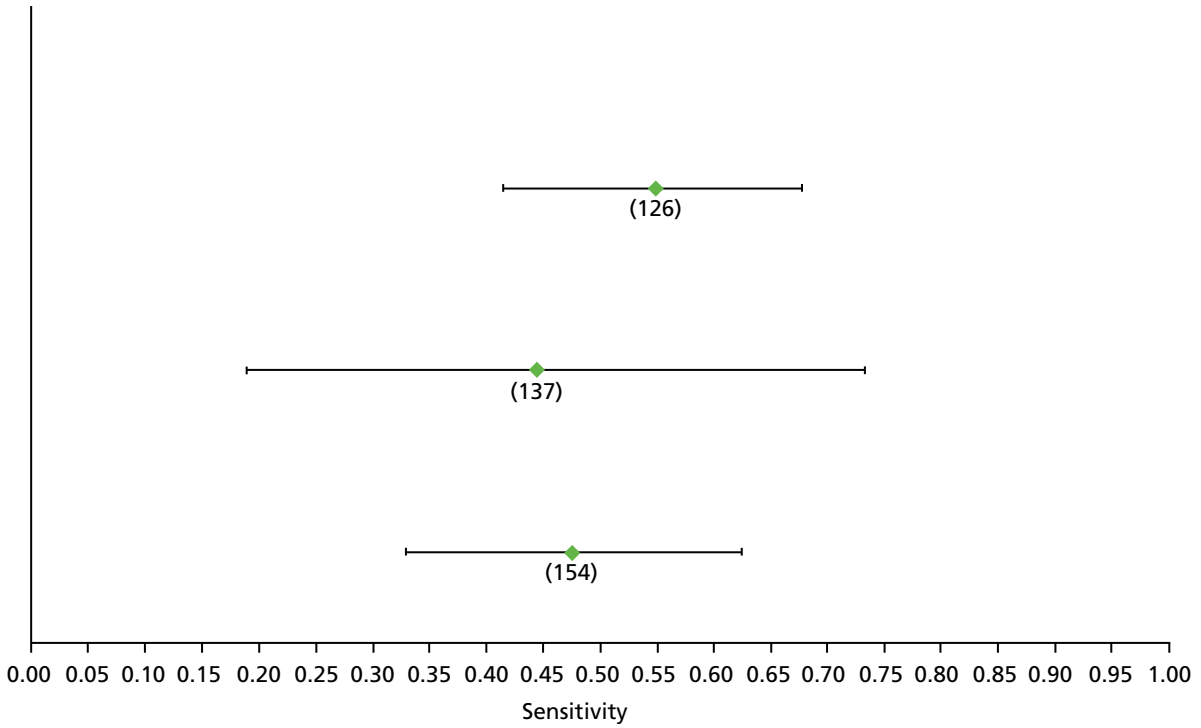

FIGURE 121 Hepatitis C: FIB-4 F4 (high cut-off) – sensitivity.

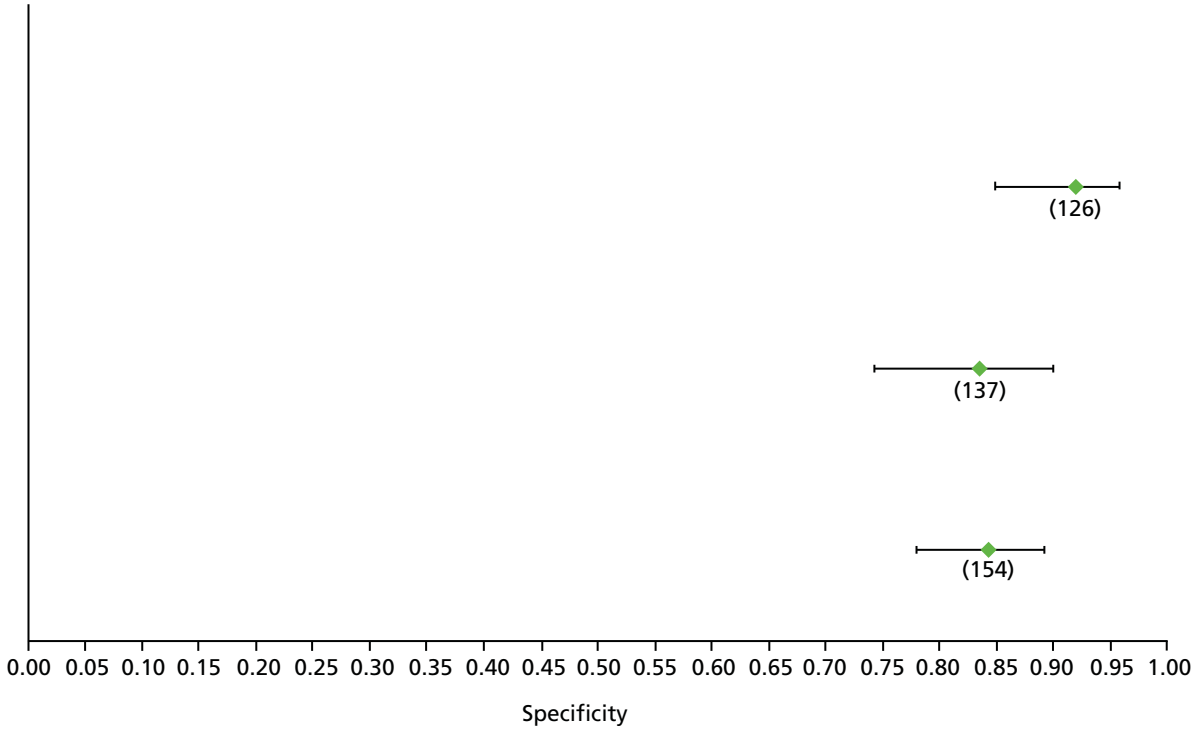

FIGURE 122 Hepatitis C: FIB-4 F4 (high cut-off) – specificity.

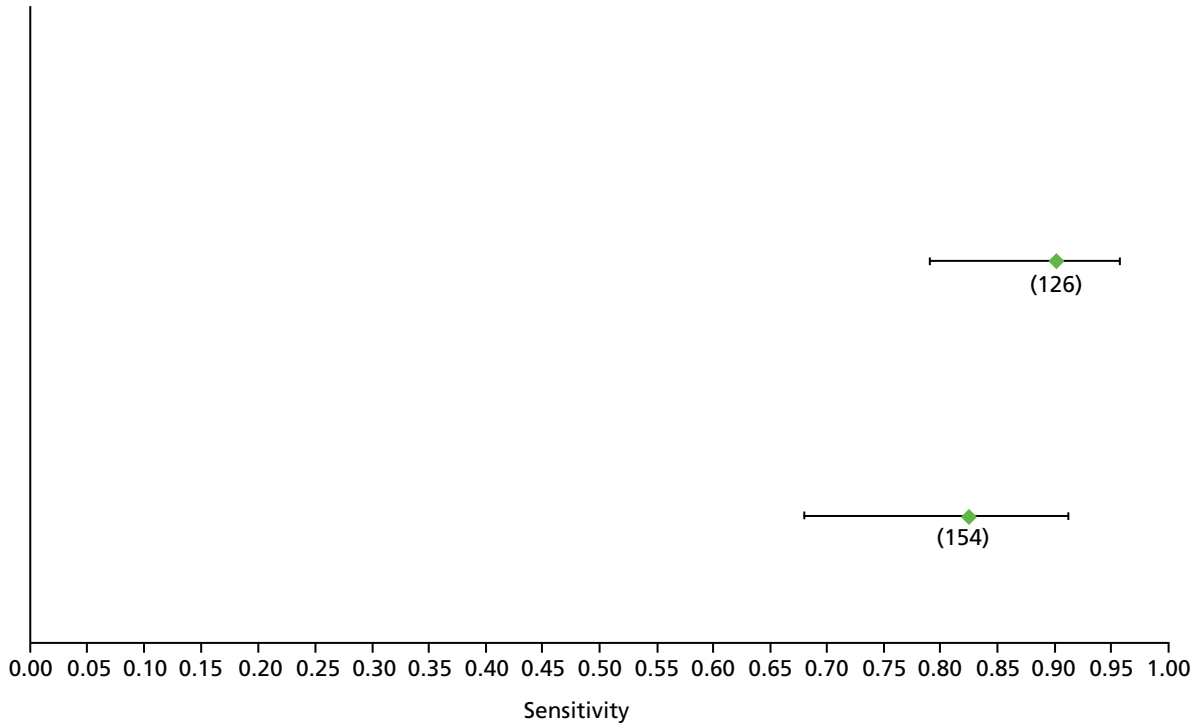

FIGURE 123 Hepatitis C: FIB-4 F4 (low cut-off) – sensitivity.

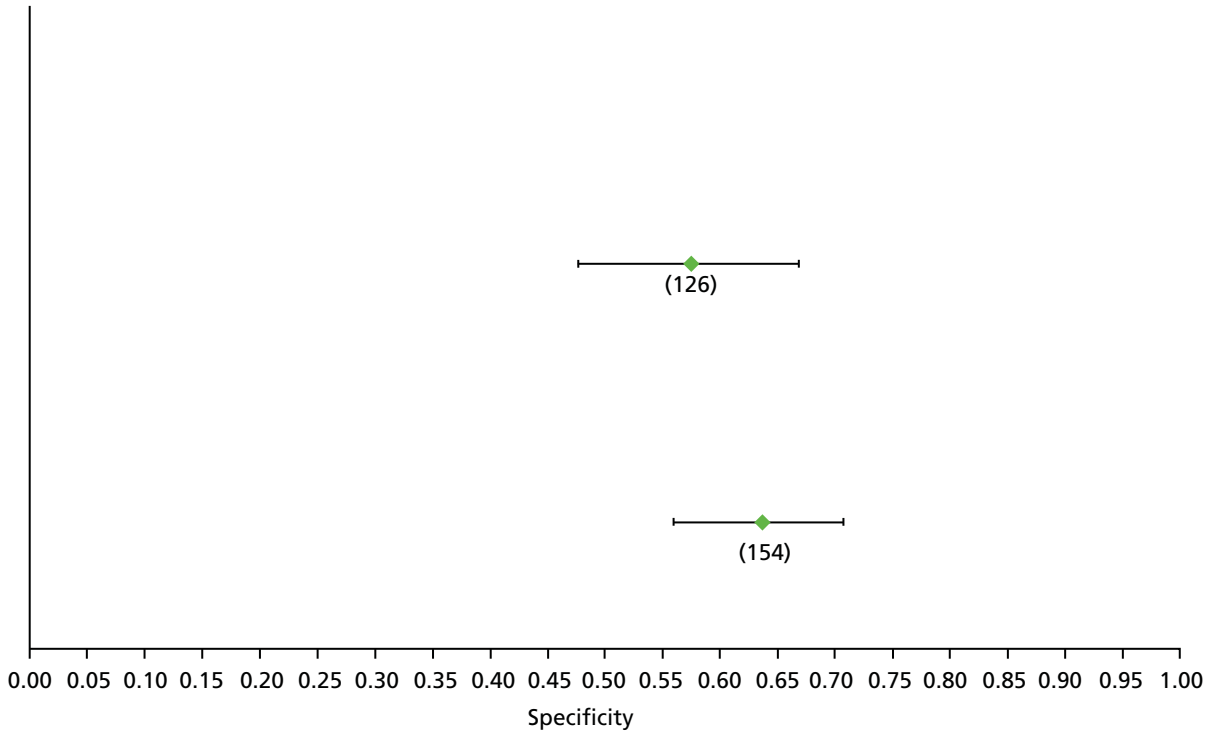

FIGURE 124 Hepatitis C: FIB-4 F4 (low cut-off) – specificity.

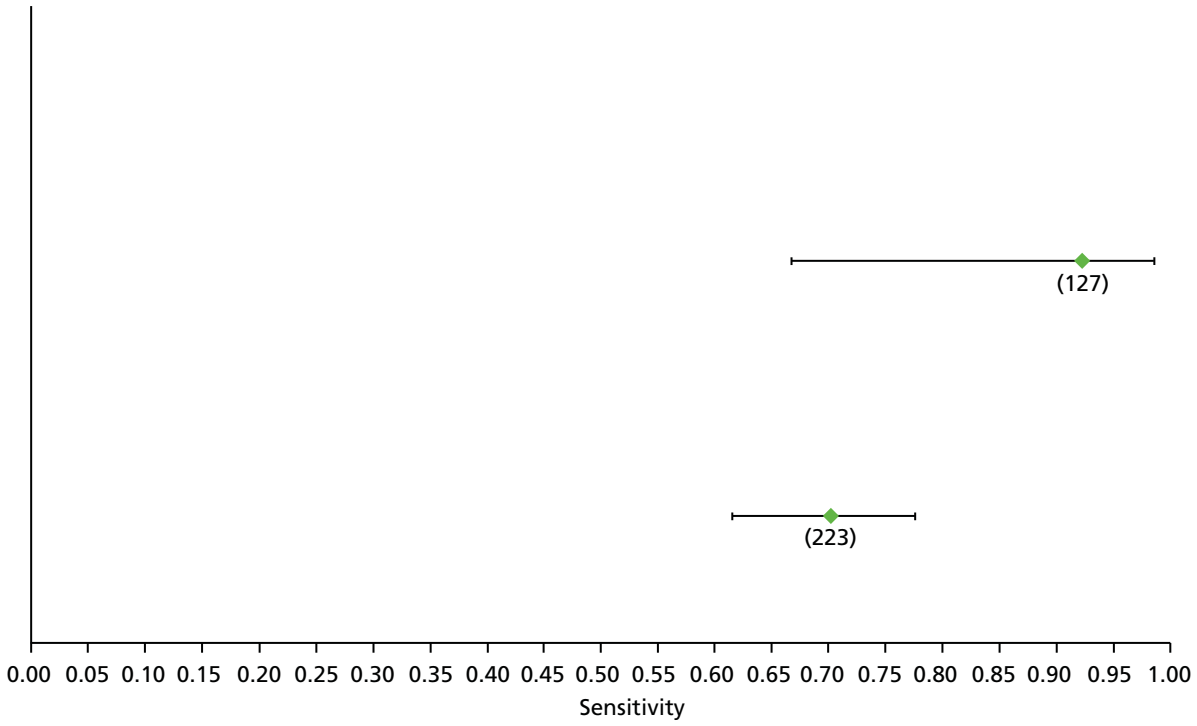

FIGURE 125 Hepatitis C: Fibrometer F4 – sensitivity.

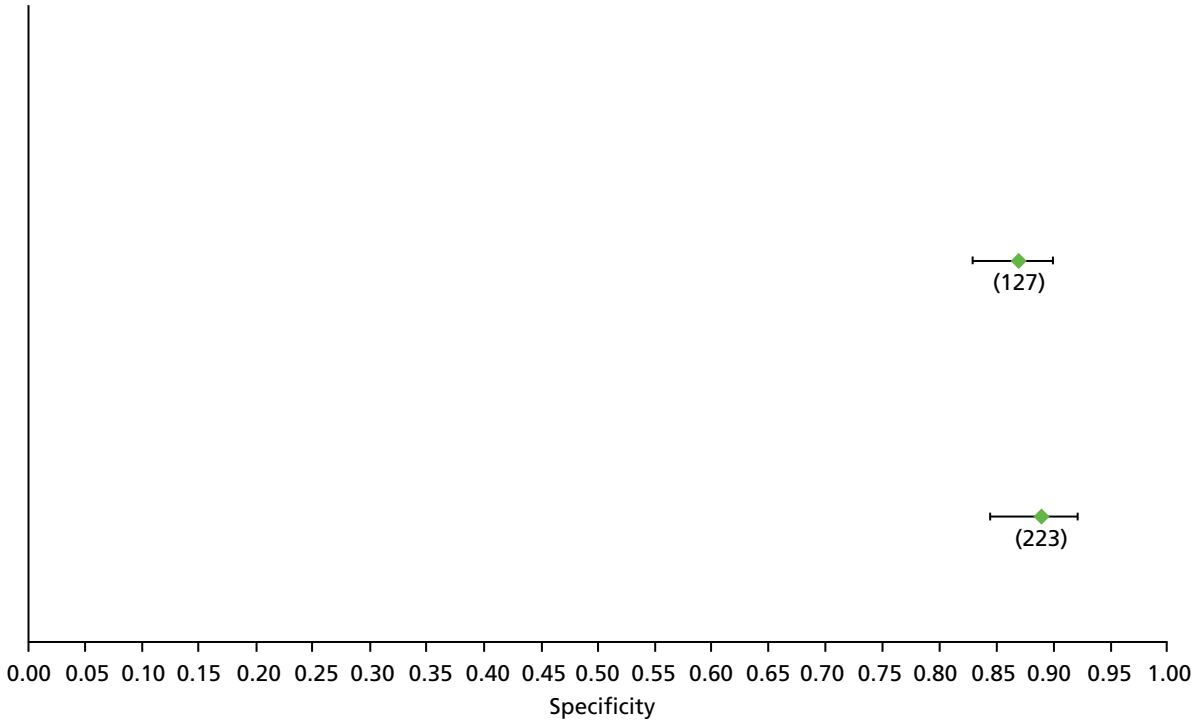

FIGURE 126 Hepatitis C: Fibrometer F4 – specificity.

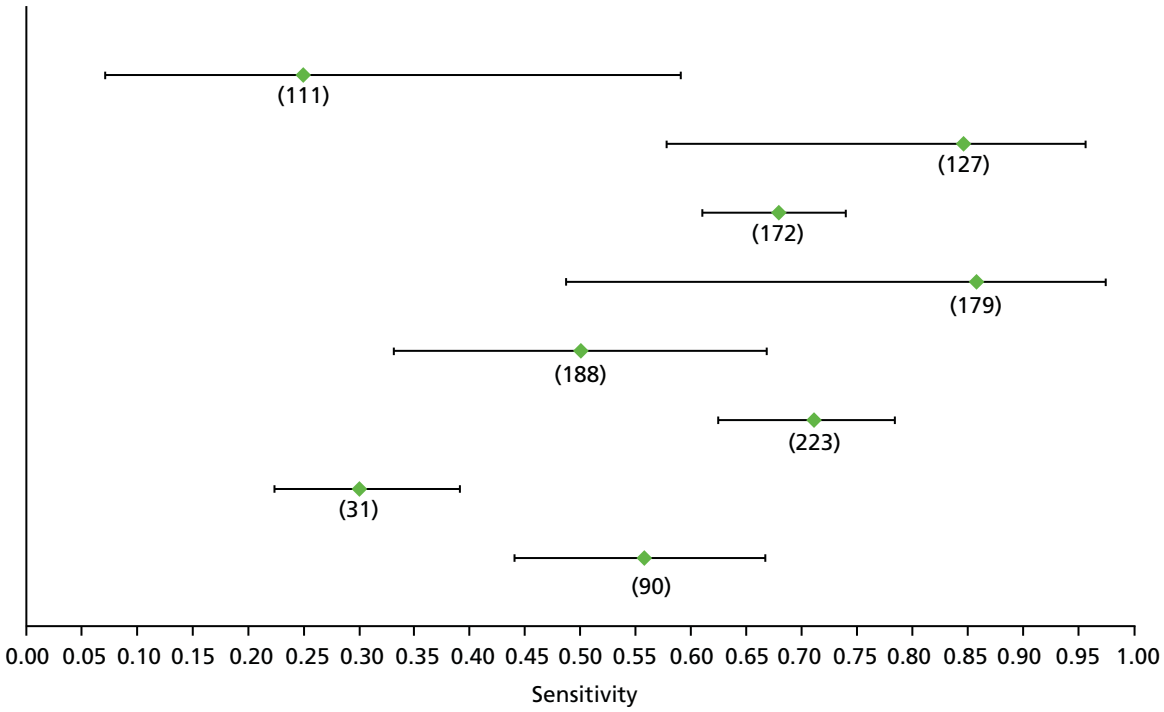

FIGURE 127 Hepatitis C: Fibrotest F4 – sensitivity.

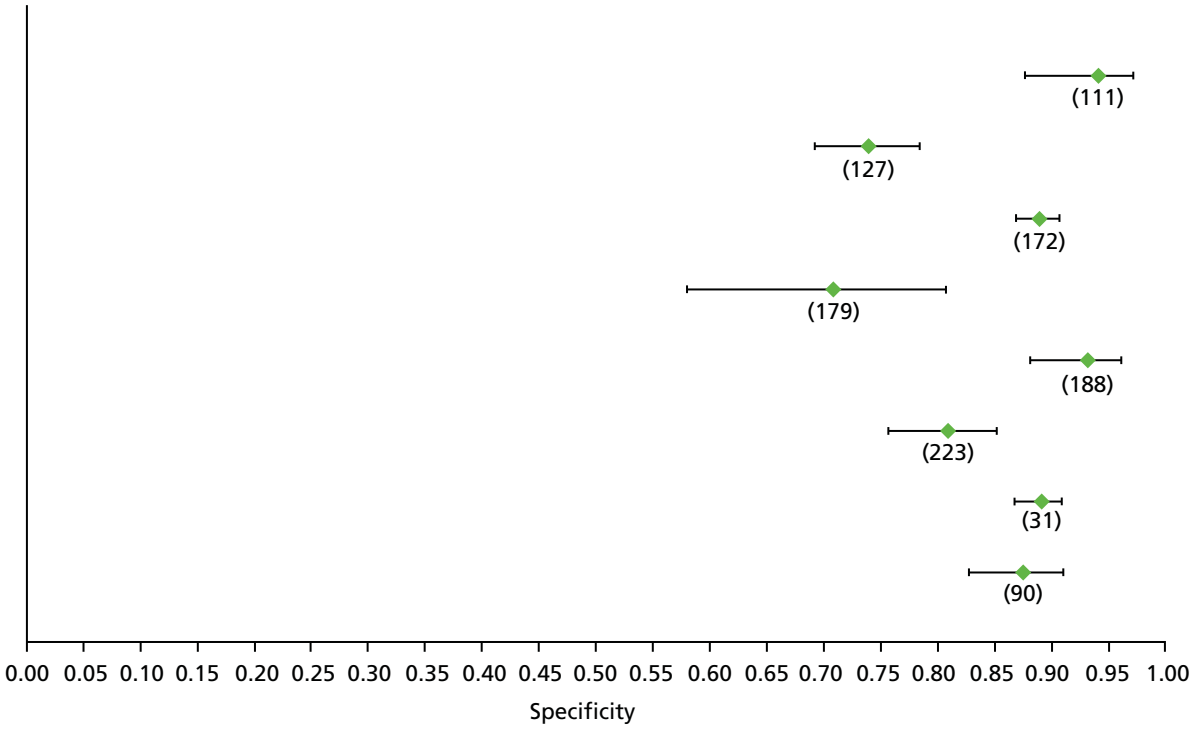

FIGURE 128 Hepatitis C: Fibrotest F4 – specificity.

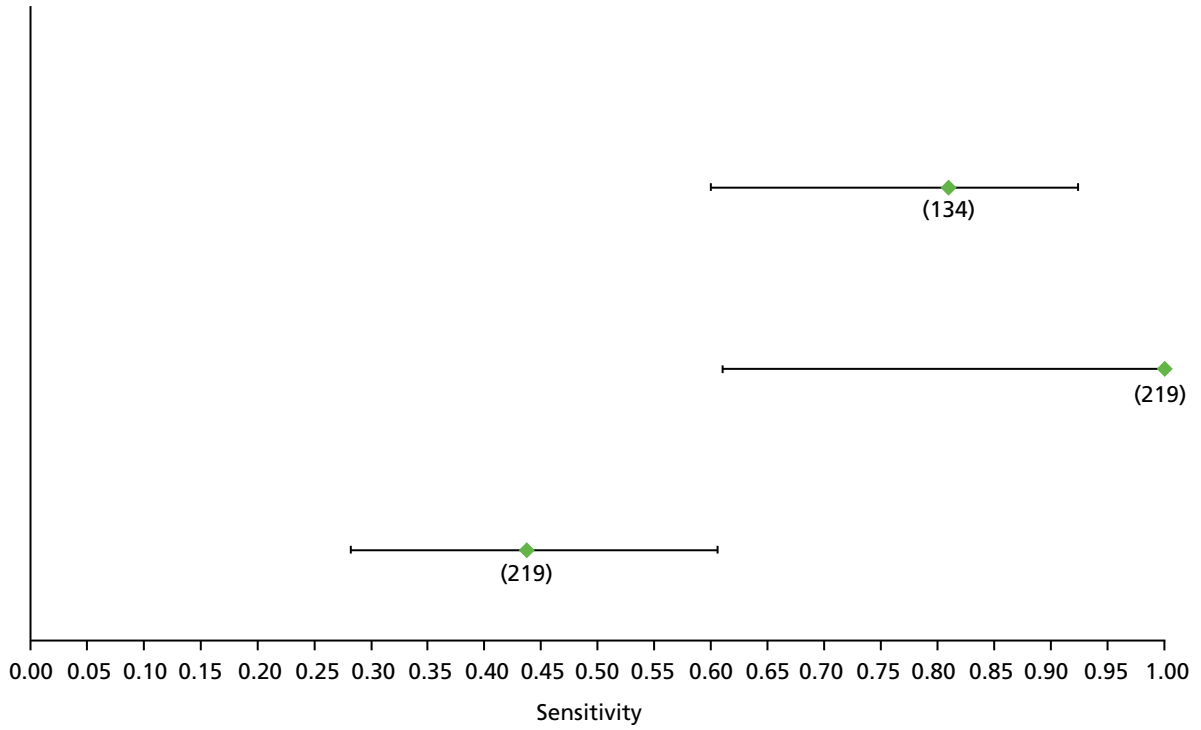

FIGURE 129 Hepatitis C: GUCI F4 – sensitivity.

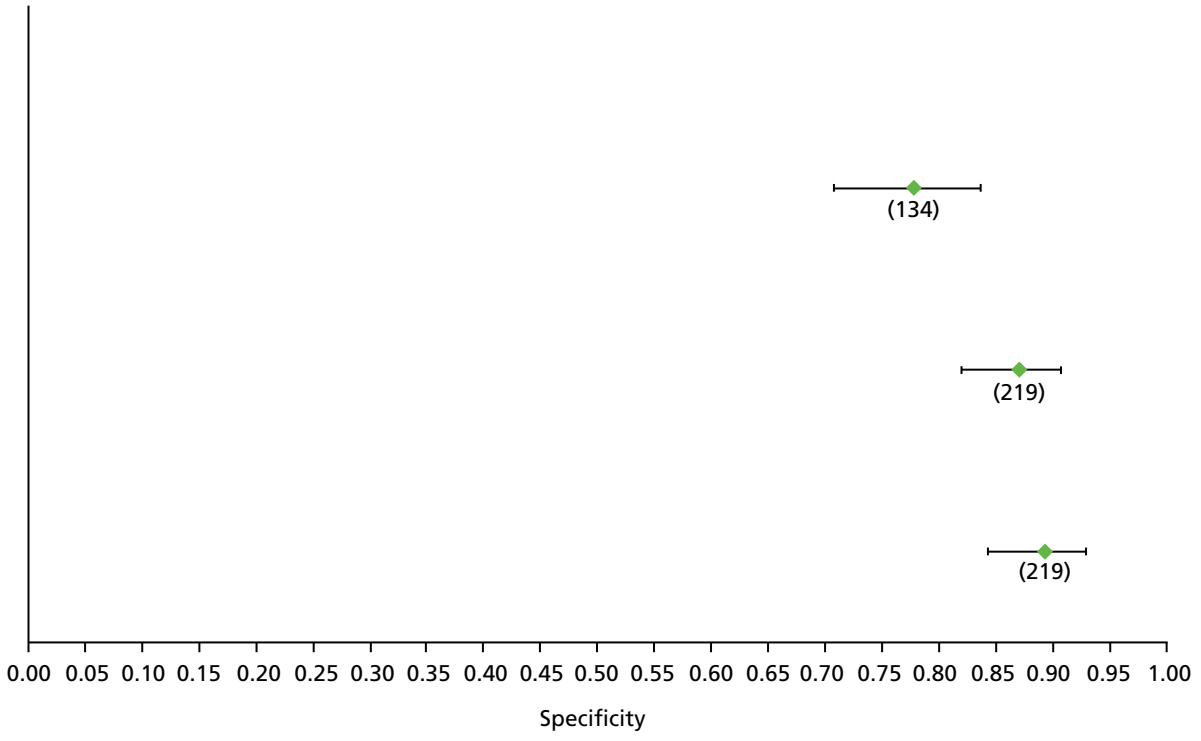

FIGURE 130 Hepatitis C: GUCI F4 – specificity.

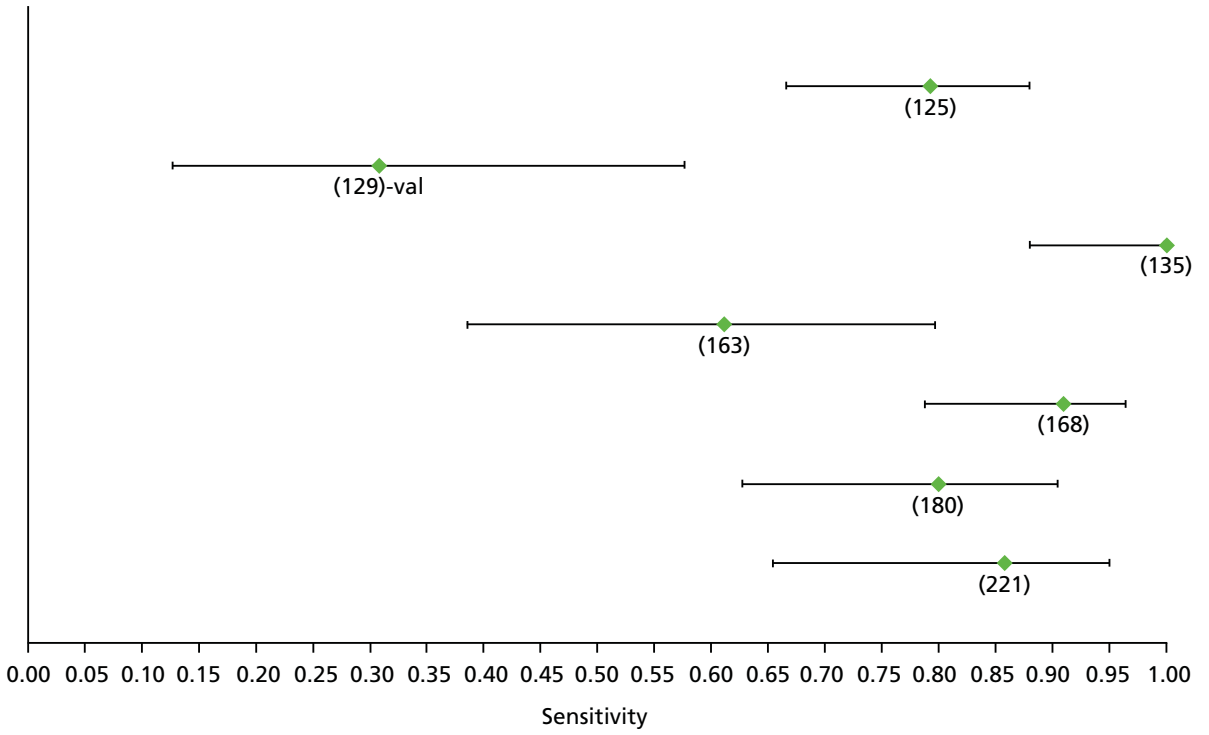

FIGURE 131 Hepatitis C: hyaluronic acid F4 – sensitivity.

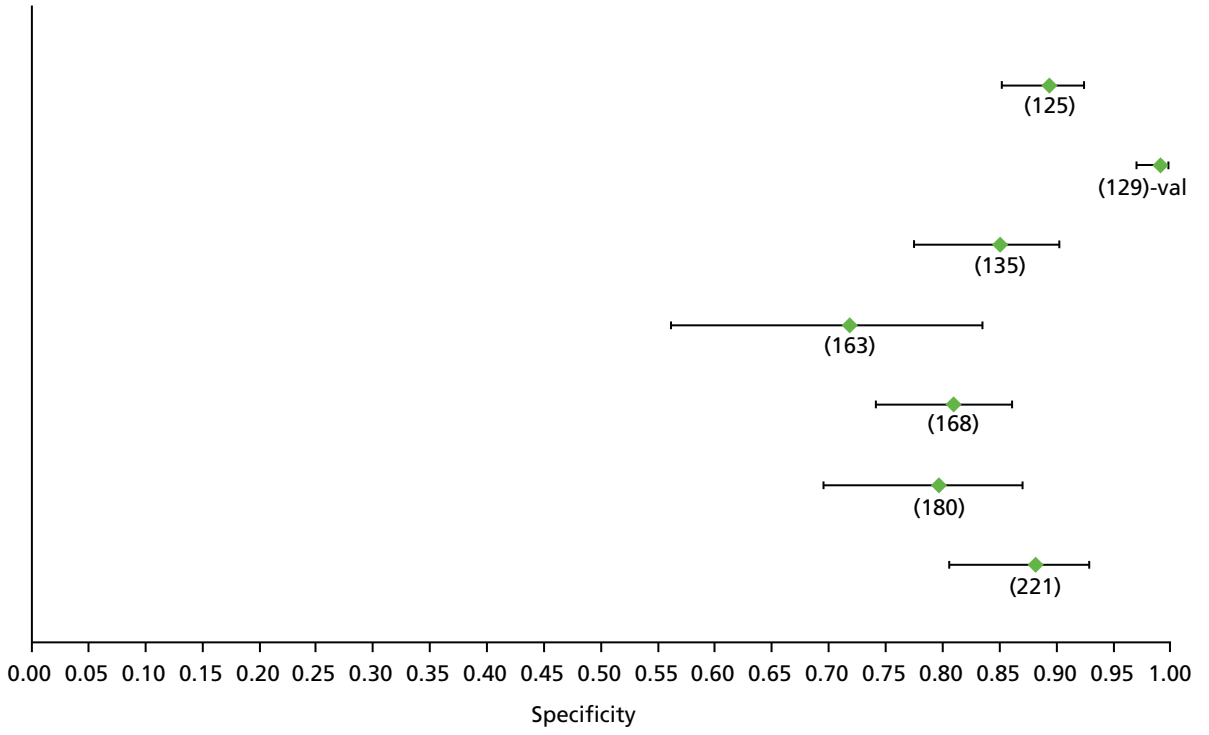

FIGURE 132 Hepatitis C: hyaluronic acid F4 – specificity.

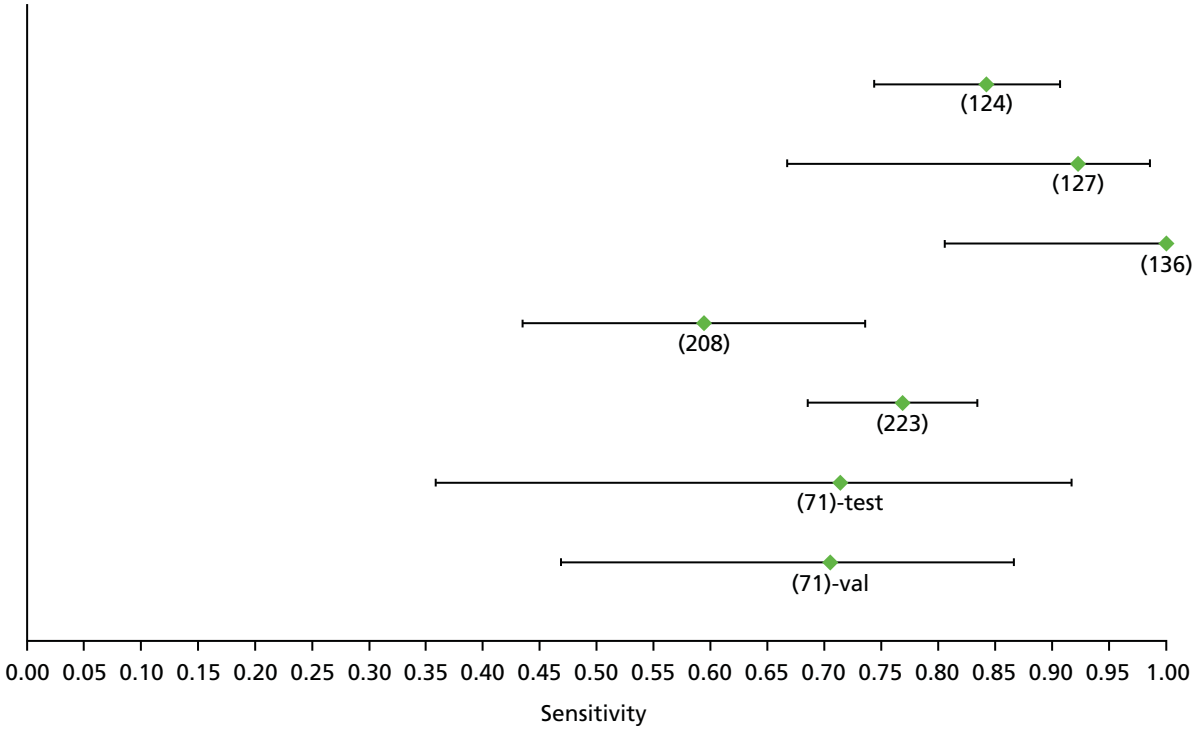

FIGURE 133 Hepatitis C: Hepascore F4 – sensitivity. Val, validation cohort.

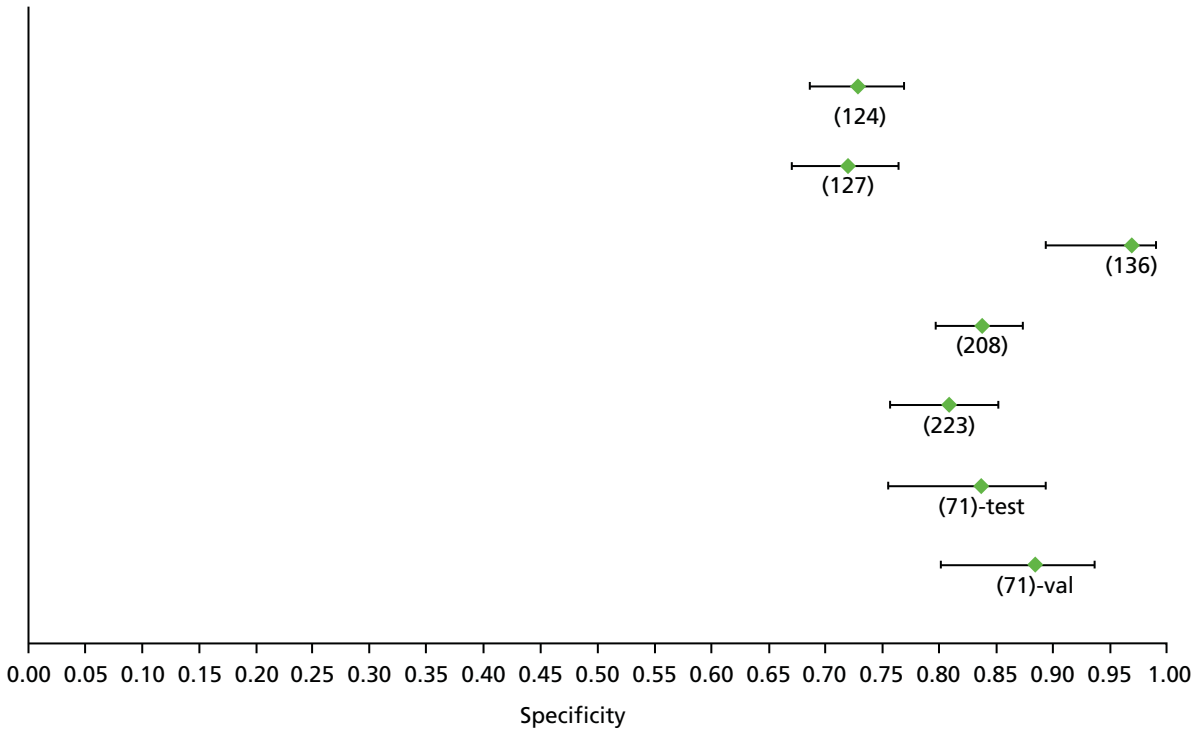

FIGURE 134 Hepatitis C: Hepascore F4 – specificity. Val, validation cohort.

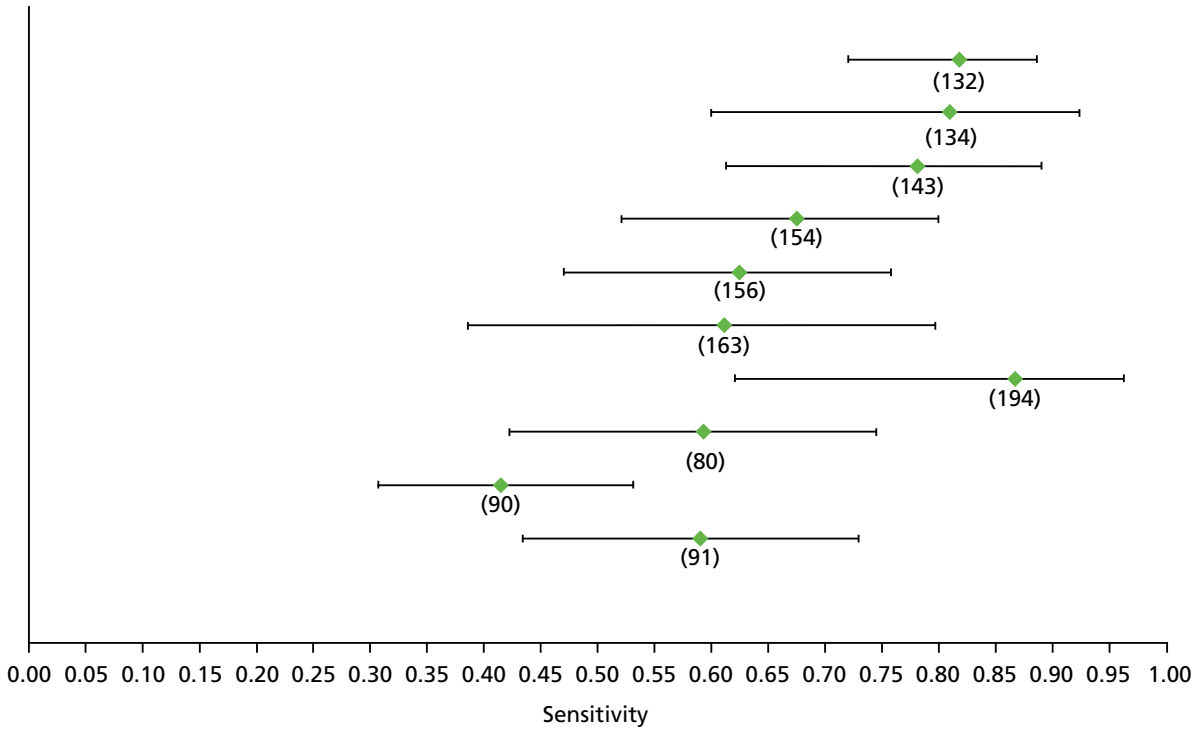

FIGURE 135 Hepatitis C: platelet F4 – sensitivity.

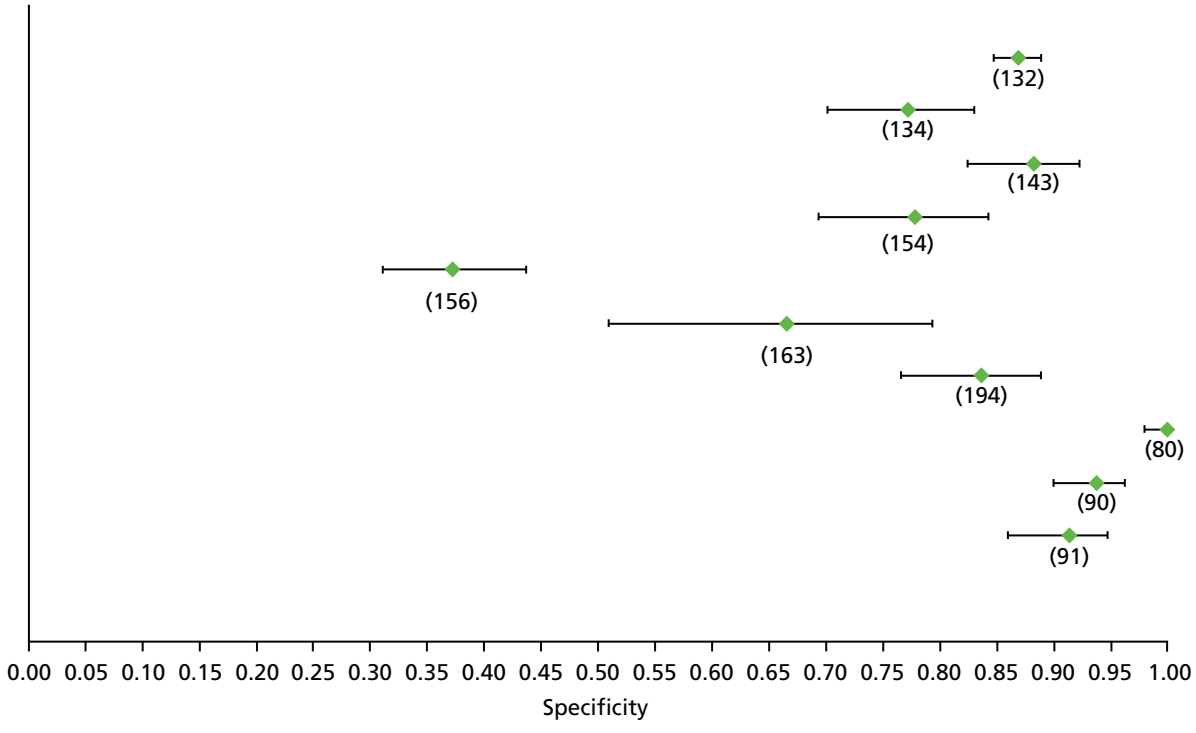

FIGURE 136 Hepatitis C: platelet F4 – specificity.

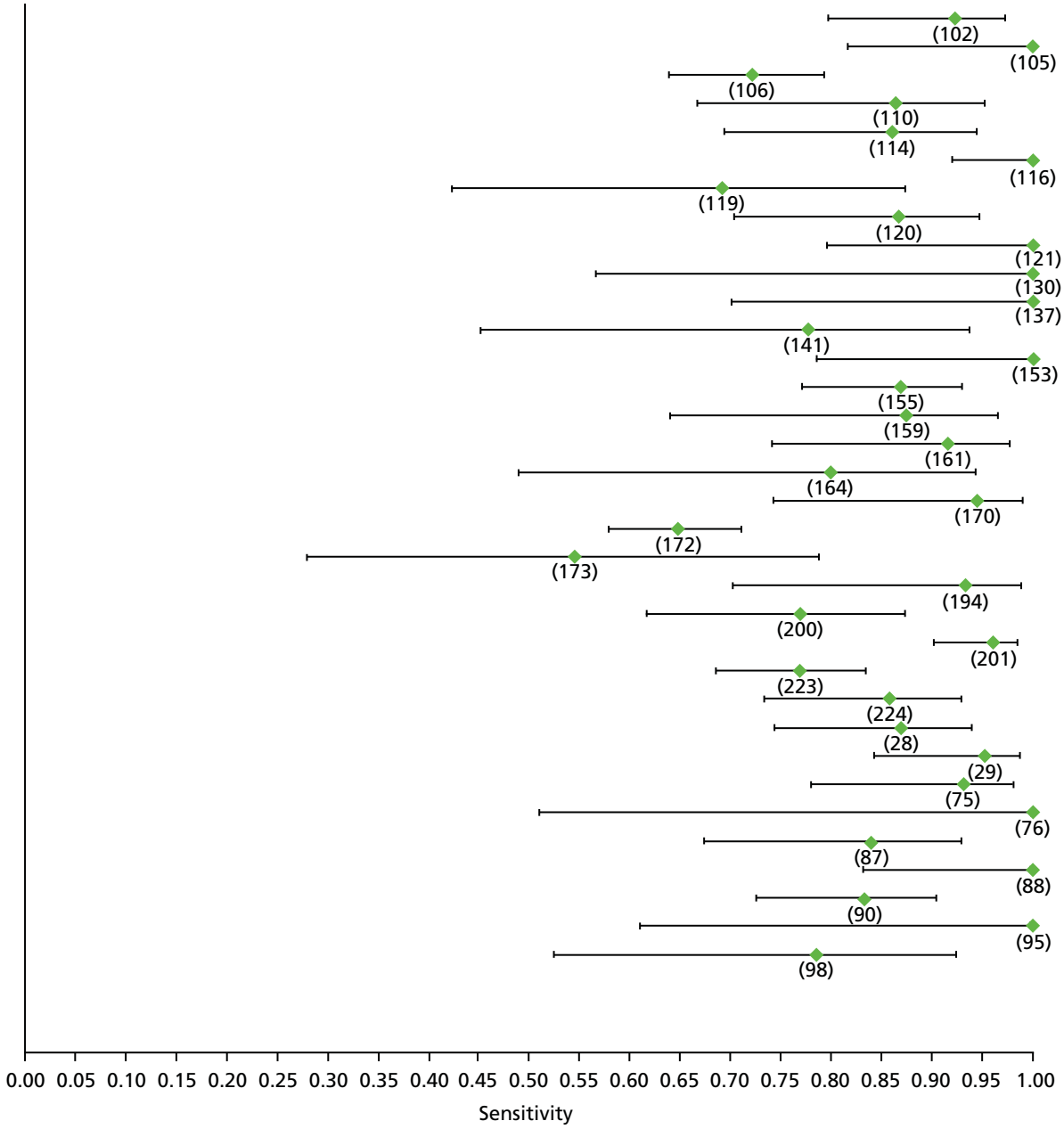

FIGURE 137 Hepatitis C: TE F4 – sensitivity.

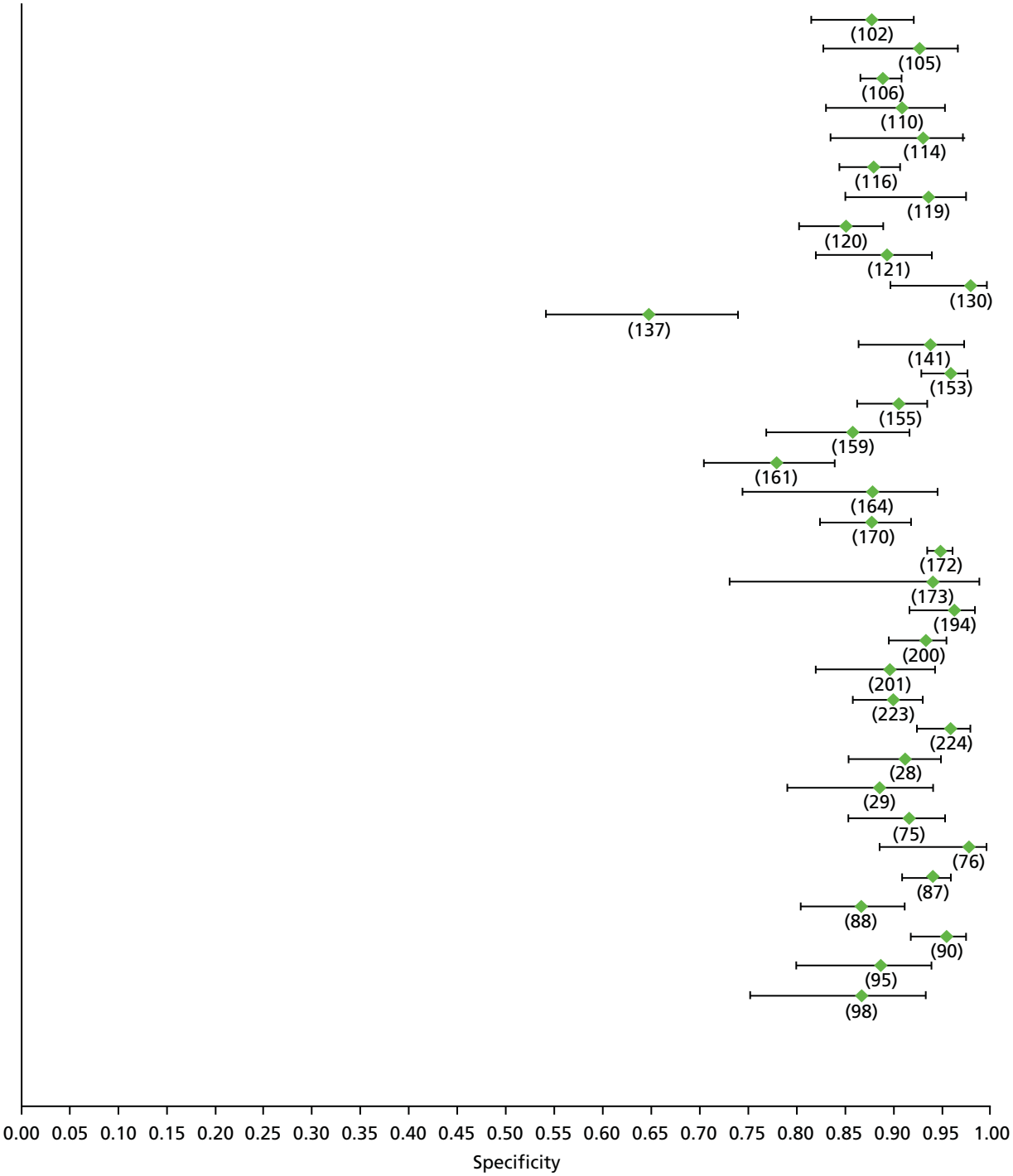

FIGURE 138 Hepatitis C: TE F4 – specificity.

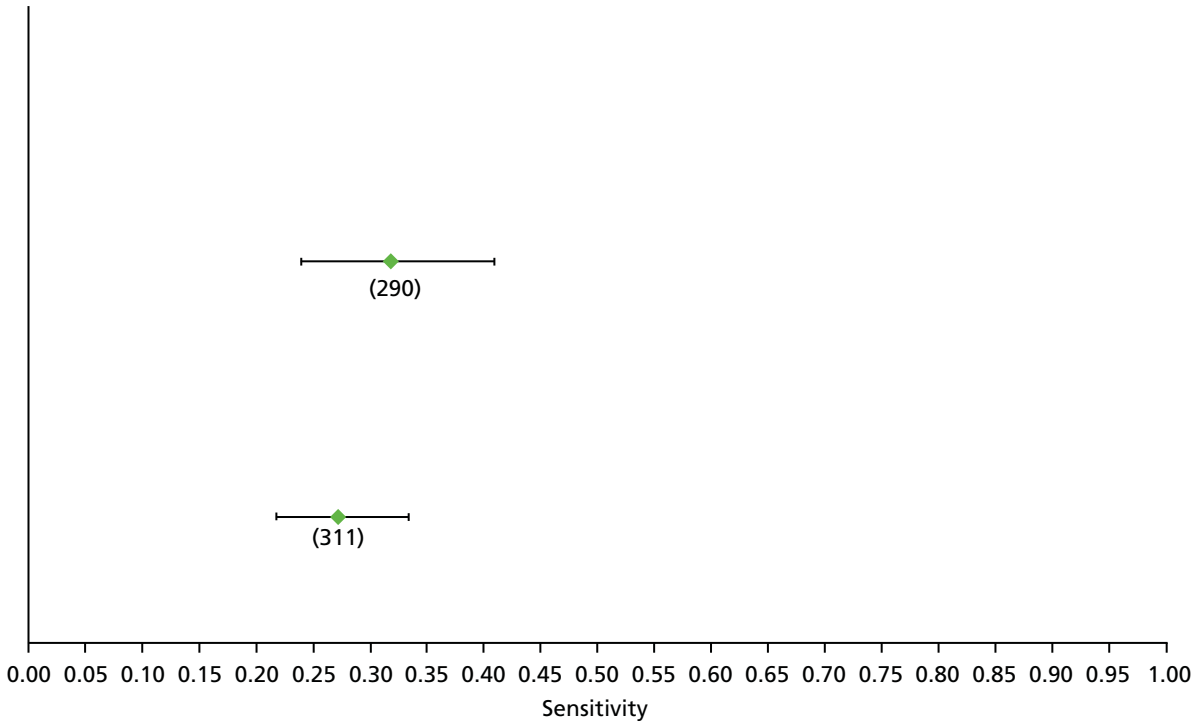

FIGURE 139 Non-alcoholic fatty liver disease: NFS F1 (high cut-off) – sensitivity.

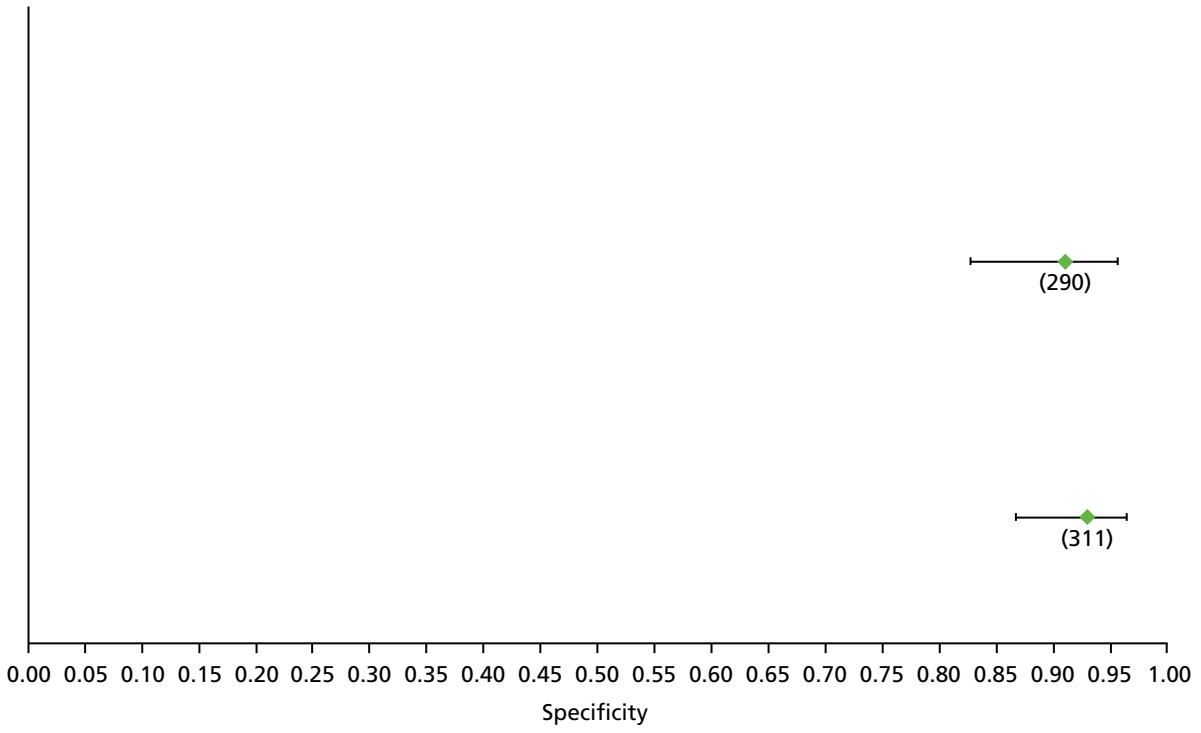

FIGURE 140 Non-alcoholic fatty liver disease: NFS F1 (high cut-off) – specificity.

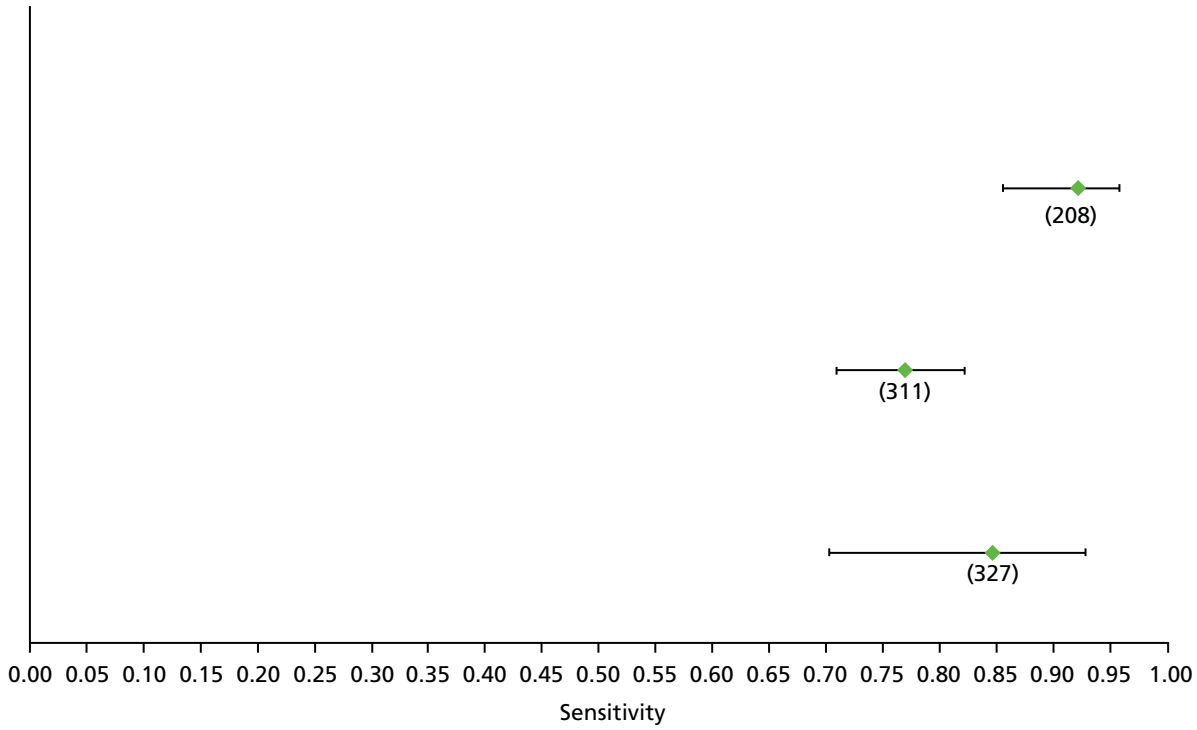

FIGURE 141 Non-alcoholic fatty liver disease: NFS F1 (low cut-off) – sensitivity.

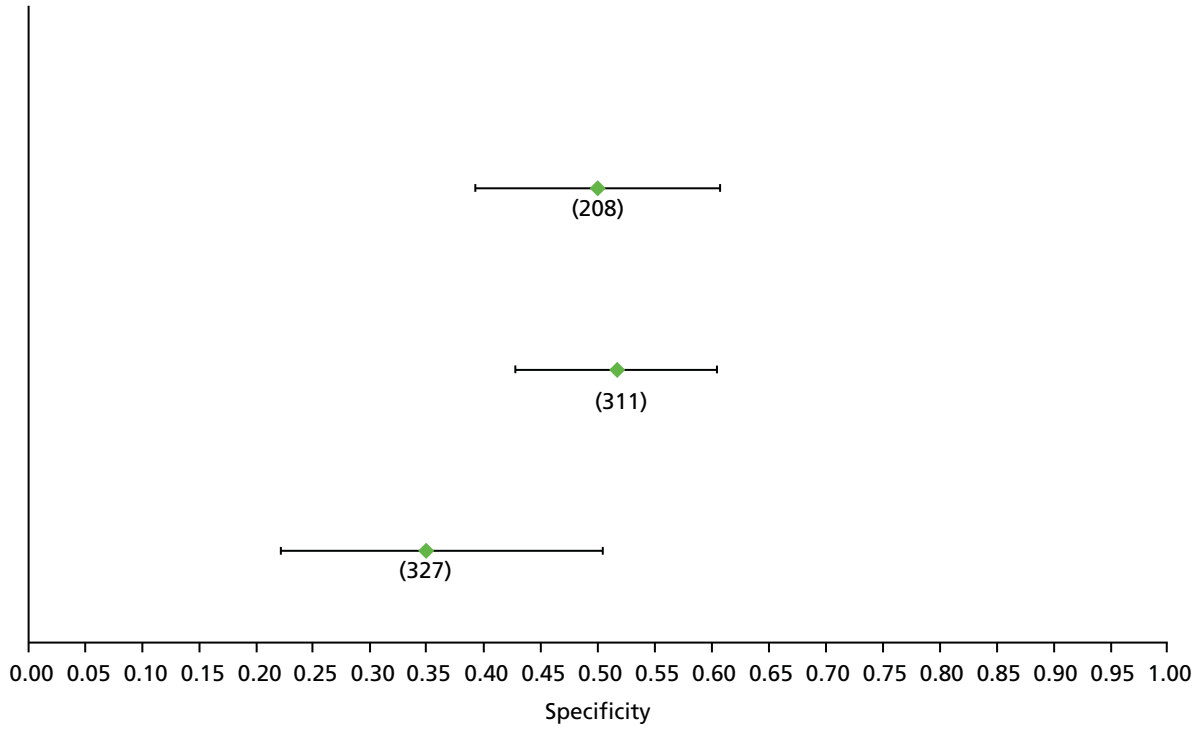

FIGURE 142 Non-alcoholic fatty liver disease: NFS F1 (low cut-off) – specificity.

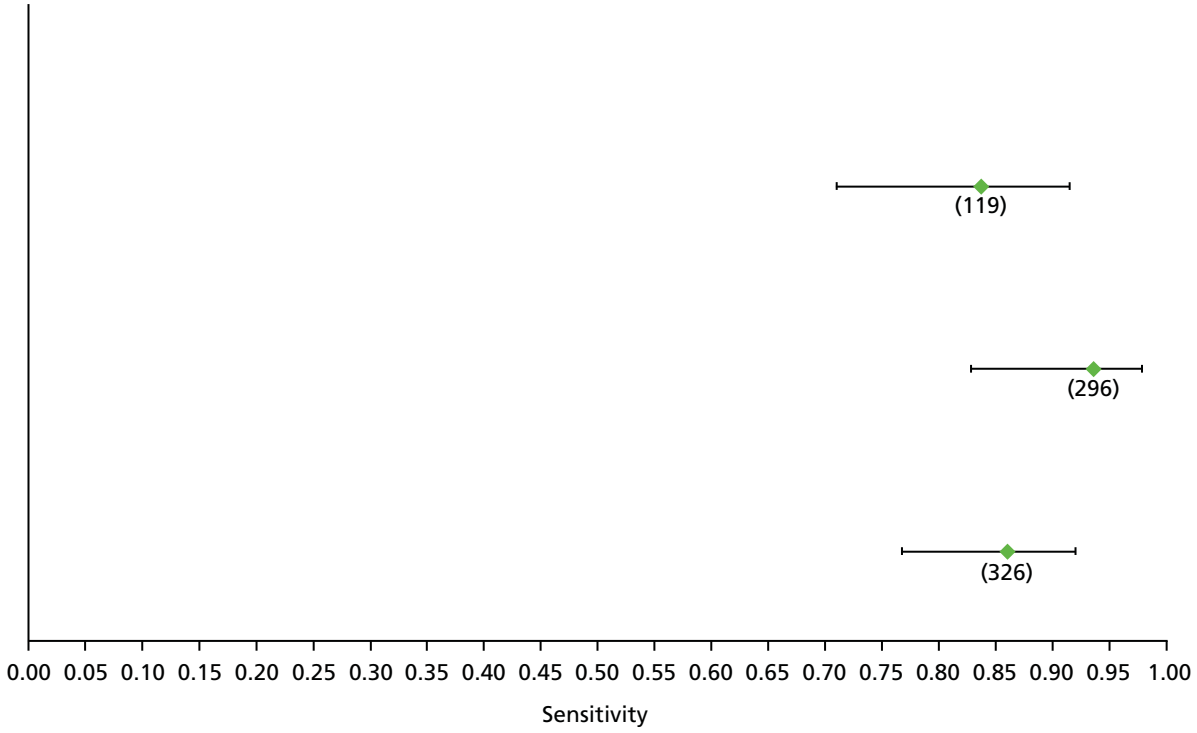

FIGURE 143 Non-alcoholic fatty liver disease: TE F1 – sensitivity.

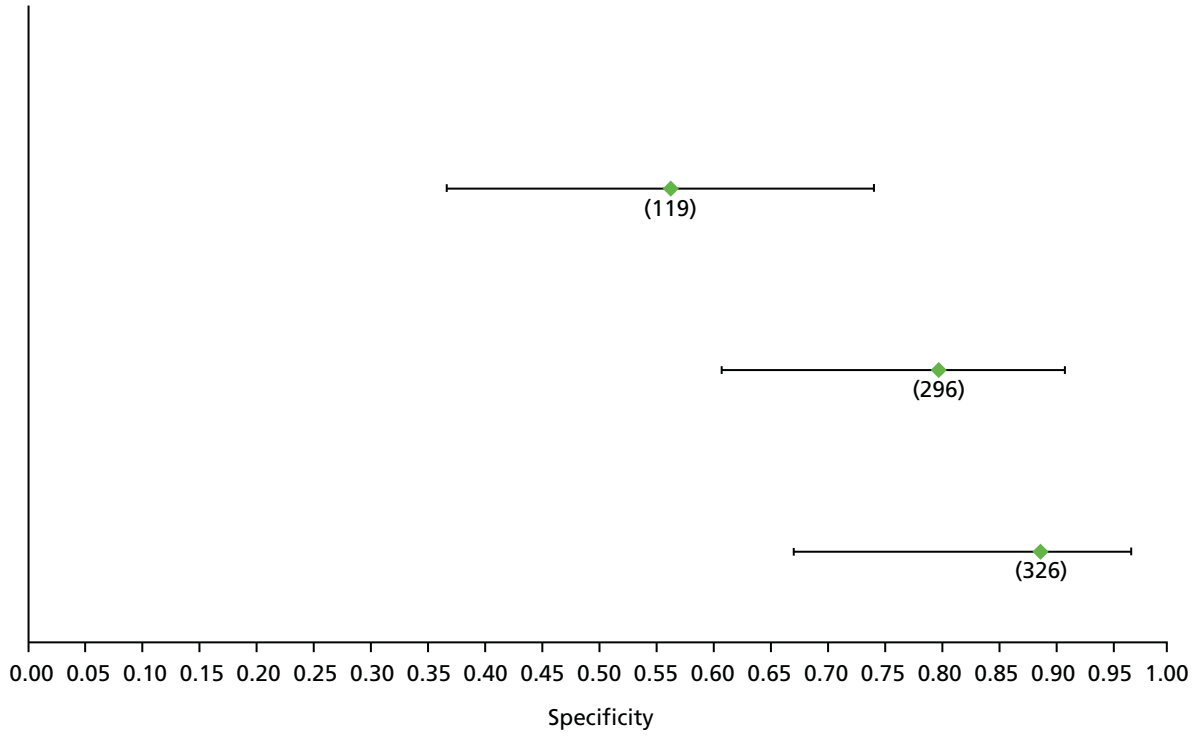

FIGURE 144 Non-alcoholic fatty liver disease: TE F1 – specificity.

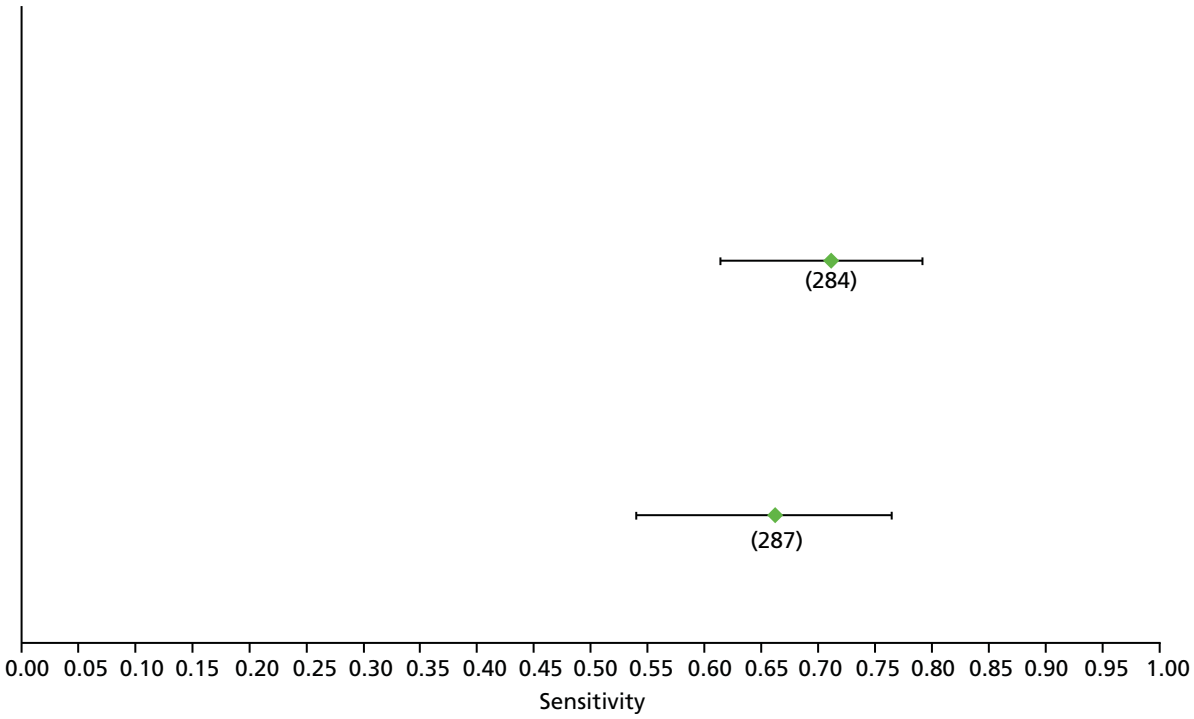

FIGURE 145 Non-alcoholic fatty liver disease: APRI F2 – sensitivity.

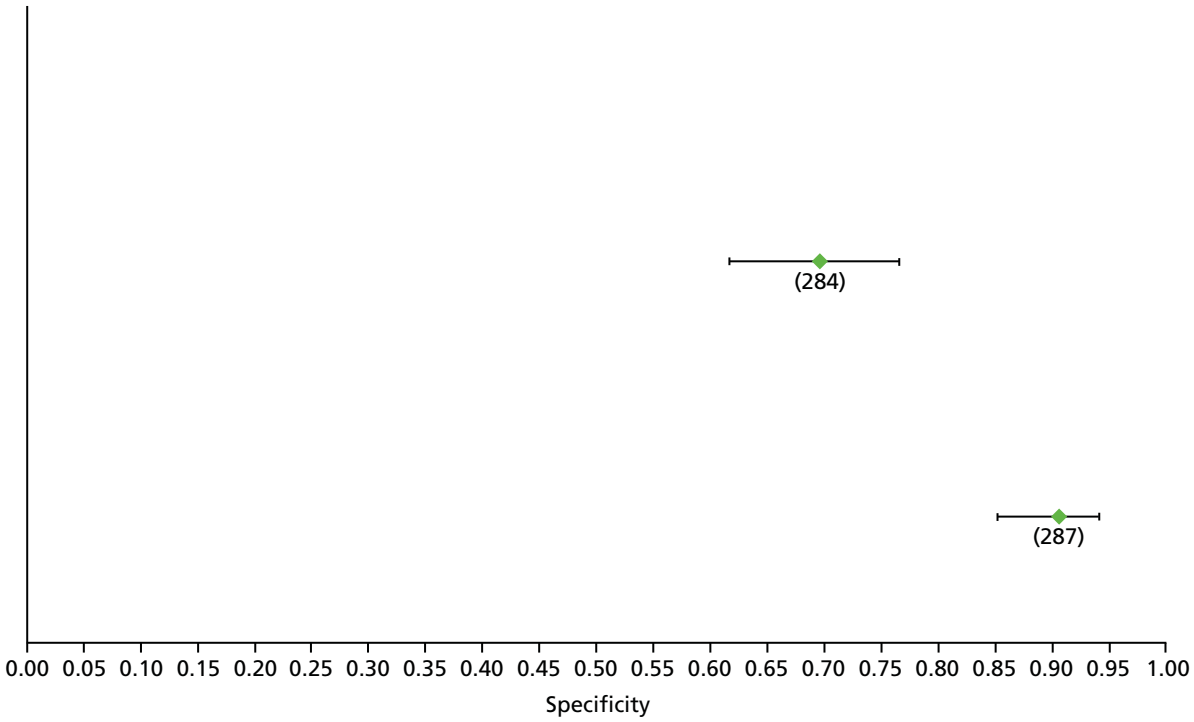

FIGURE 146 Non-alcoholic fatty liver disease: APRI F2 – specificity.

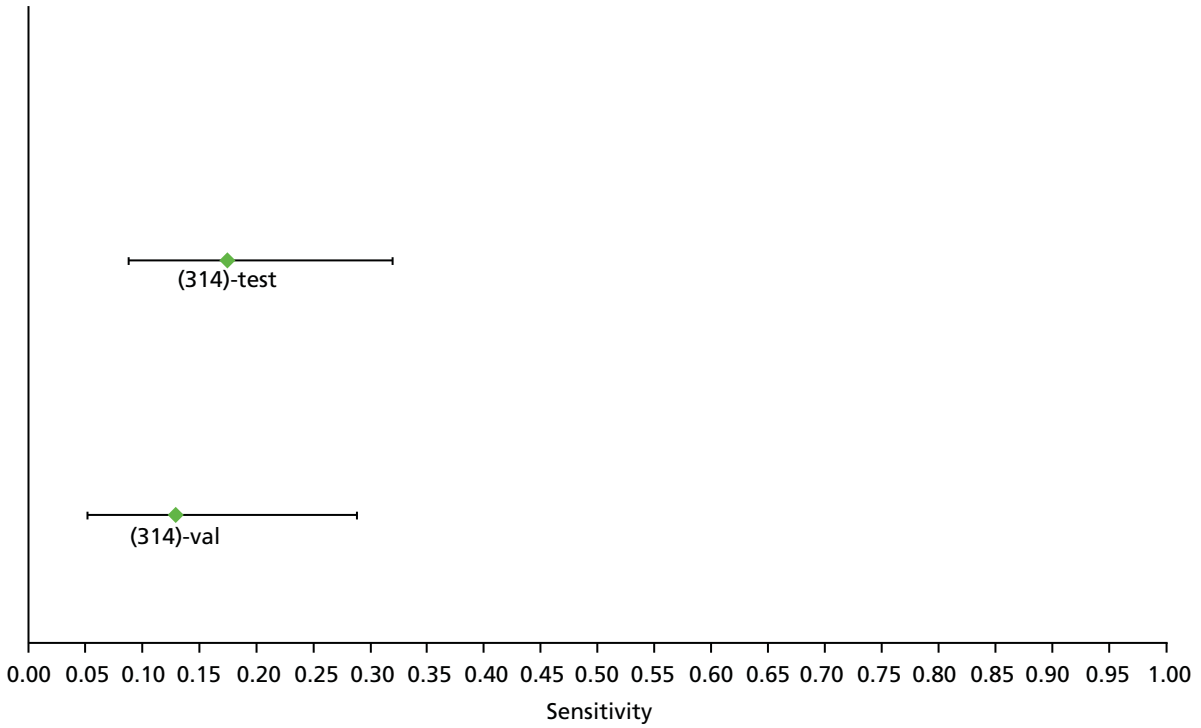

FIGURE 147 Non-alcoholic fatty liver disease: Fibrotest F2 (high cut-off) – sensitivity. Val, validation cohort.

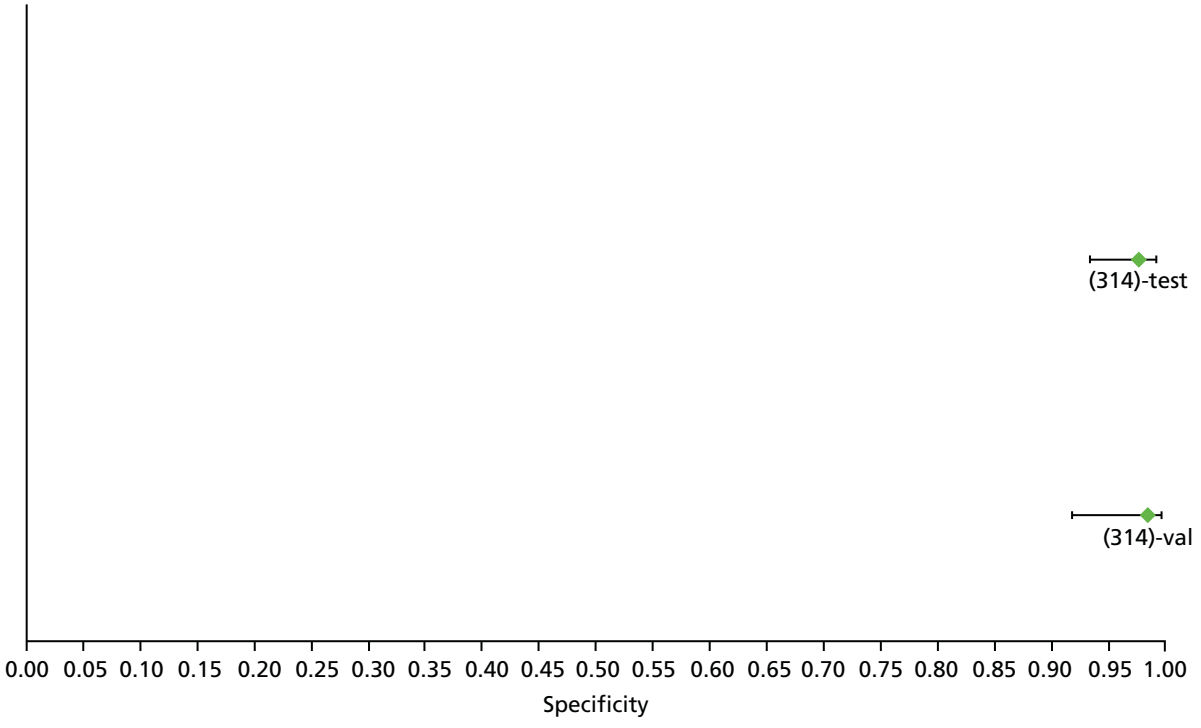

FIGURE 148 Non-alcoholic fatty liver disease: Fibrotest F2 (high cut-off) – specificity. Val, validation cohort.

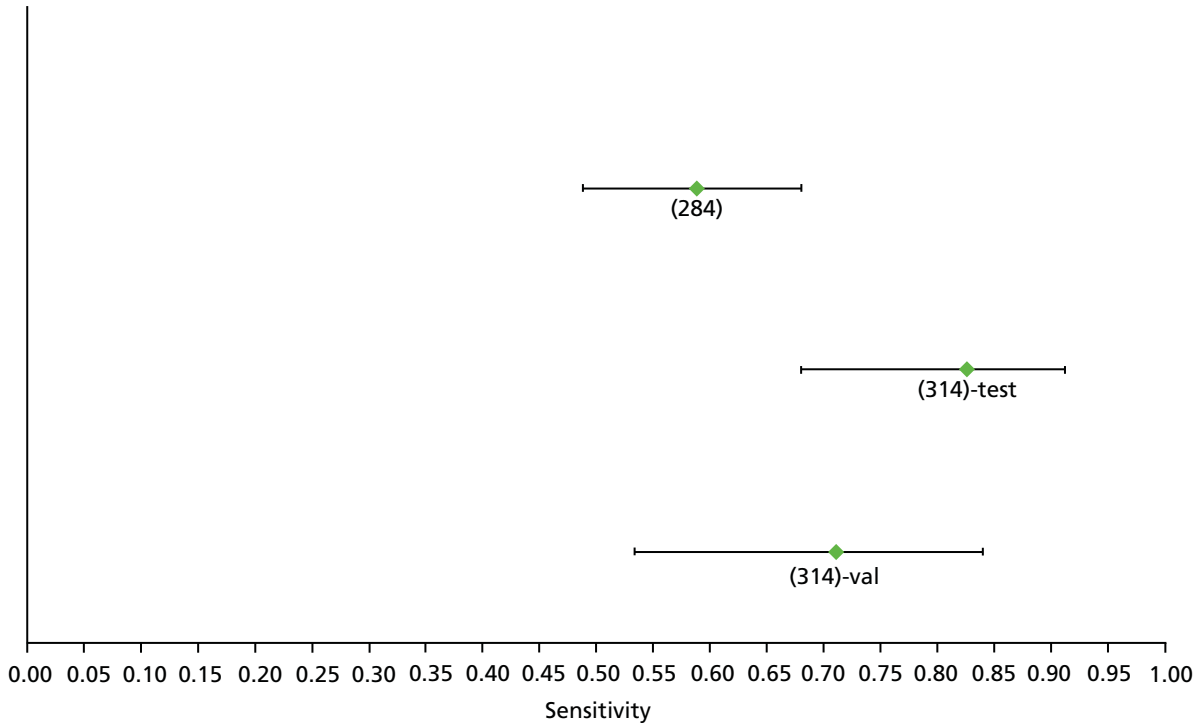

FIGURE 149 Non-alcoholic fatty liver disease: Fibrotest F2 (low cut-off) – sensitivity. Val, validation cohort.

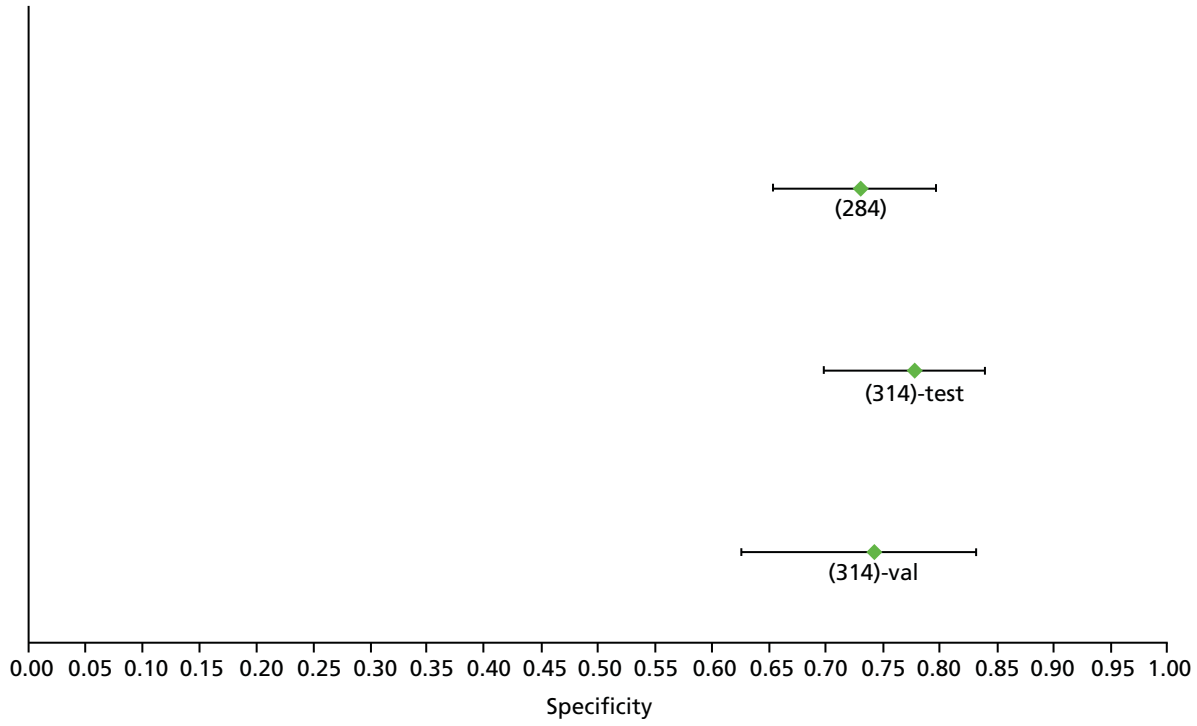

FIGURE 150 Non-alcoholic fatty liver disease: Fibrotest F2 (low cut-off) – specificity. Val, validation cohort.

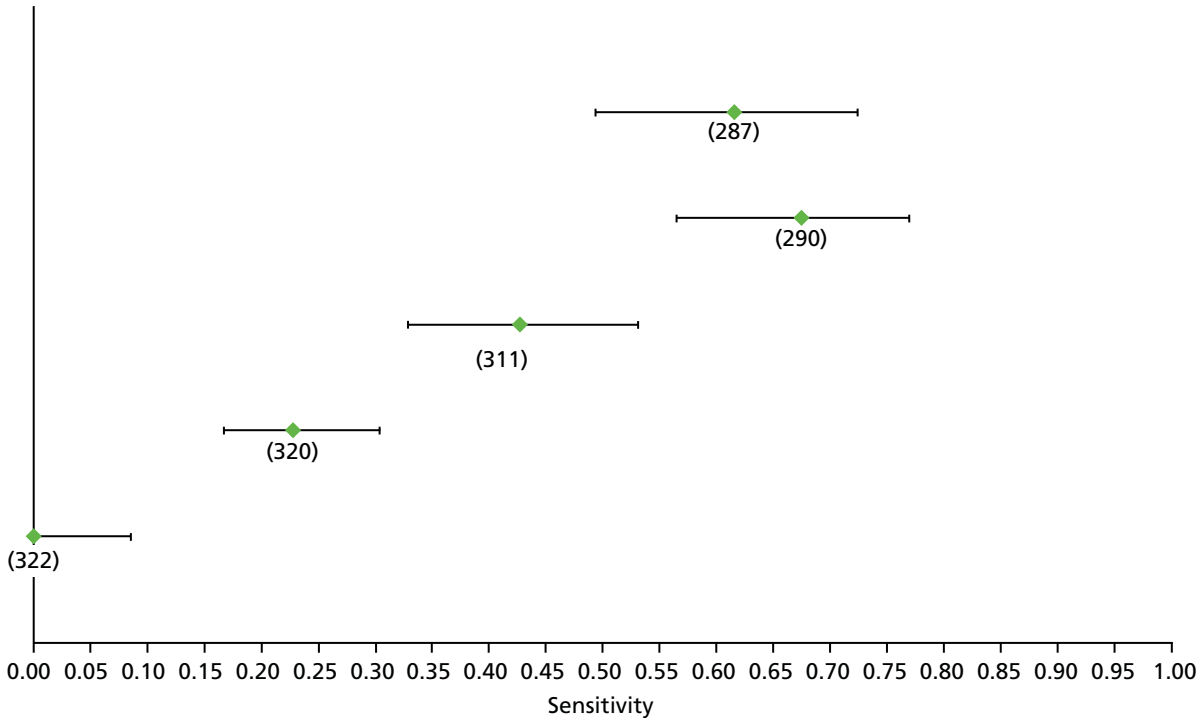

FIGURE 151 Non-alcoholic fatty liver disease: NFS F2 (high cut-off) – sensitivity.

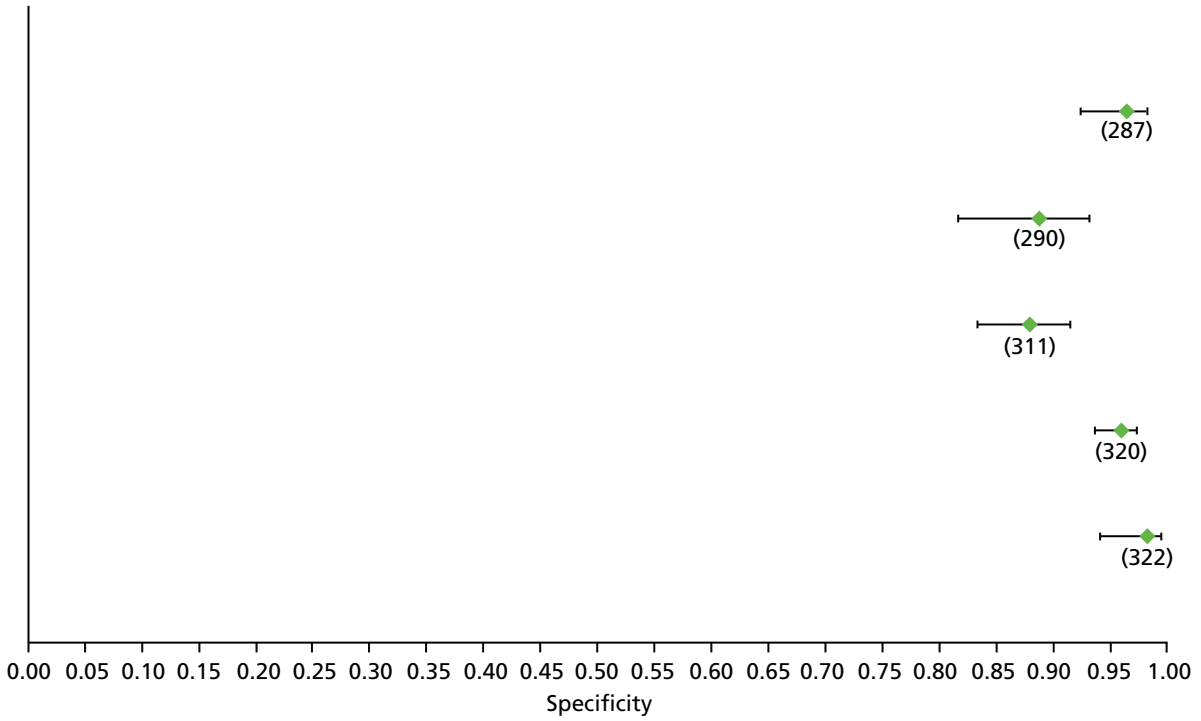

FIGURE 152 Non-alcoholic fatty liver disease: NFS F2 (high cut-off) – specificity.

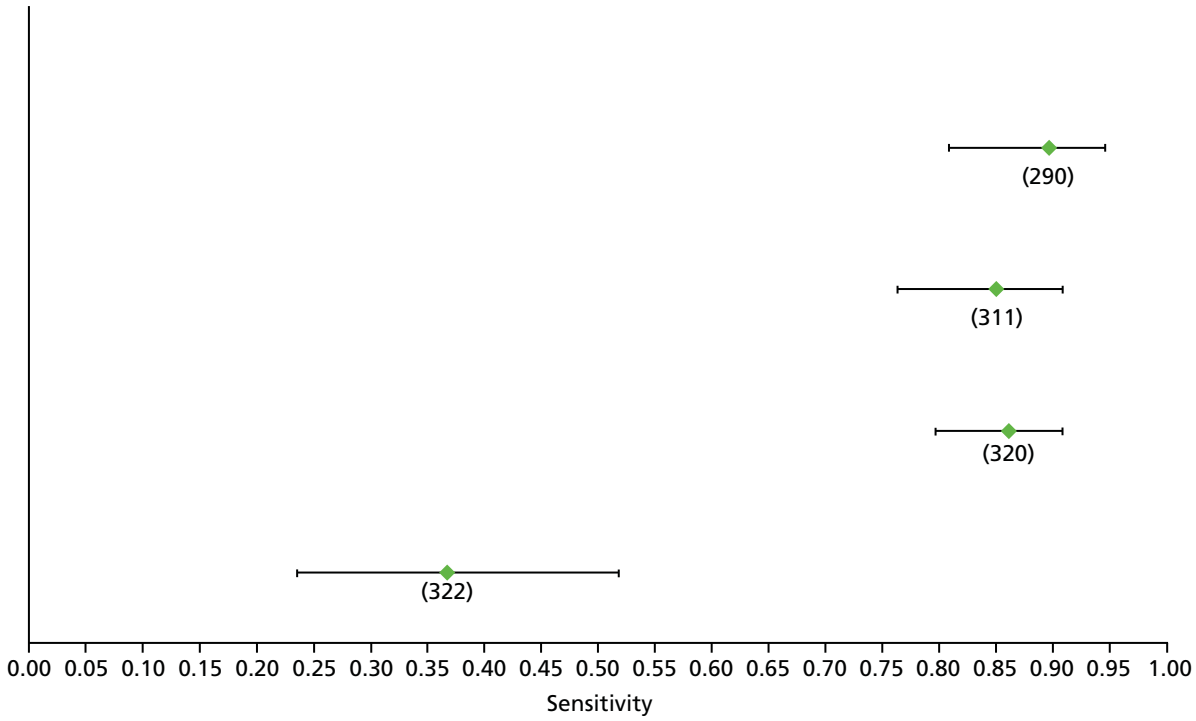

FIGURE 153 Non-alcoholic fatty liver disease: NFS F2 (low cut-off) – sensitivity.

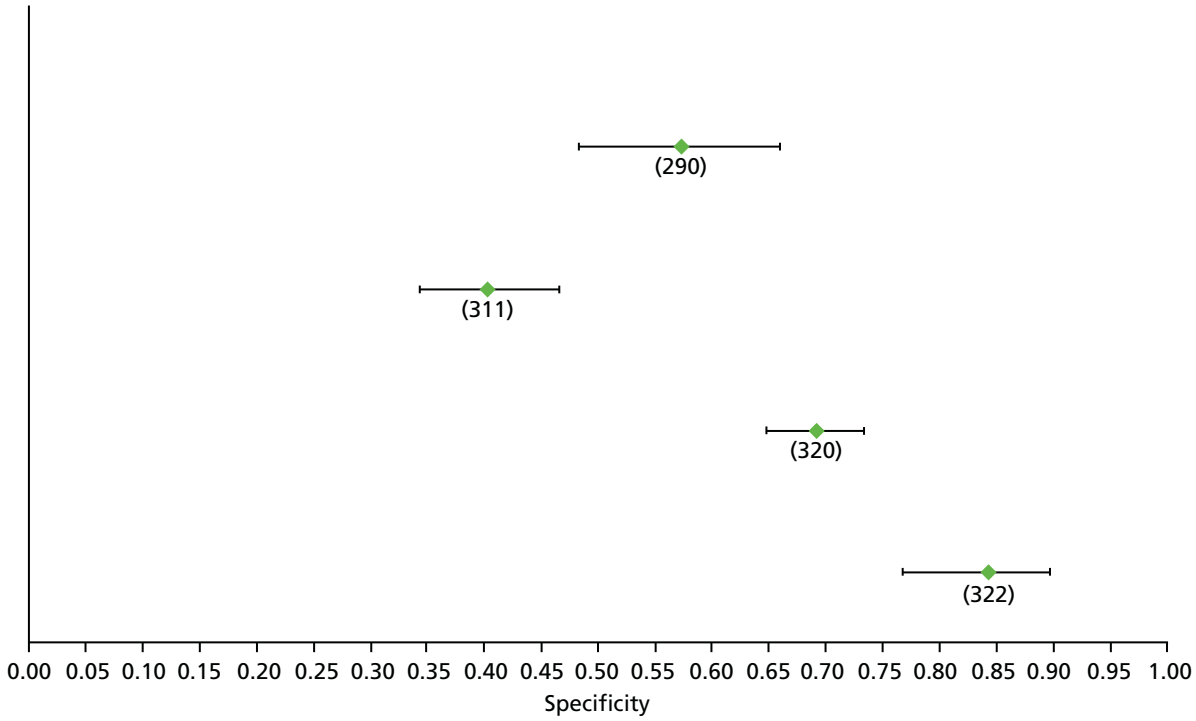

FIGURE 154 Non-alcoholic fatty liver disease: NFS F2 (low cut-off) – specificity.

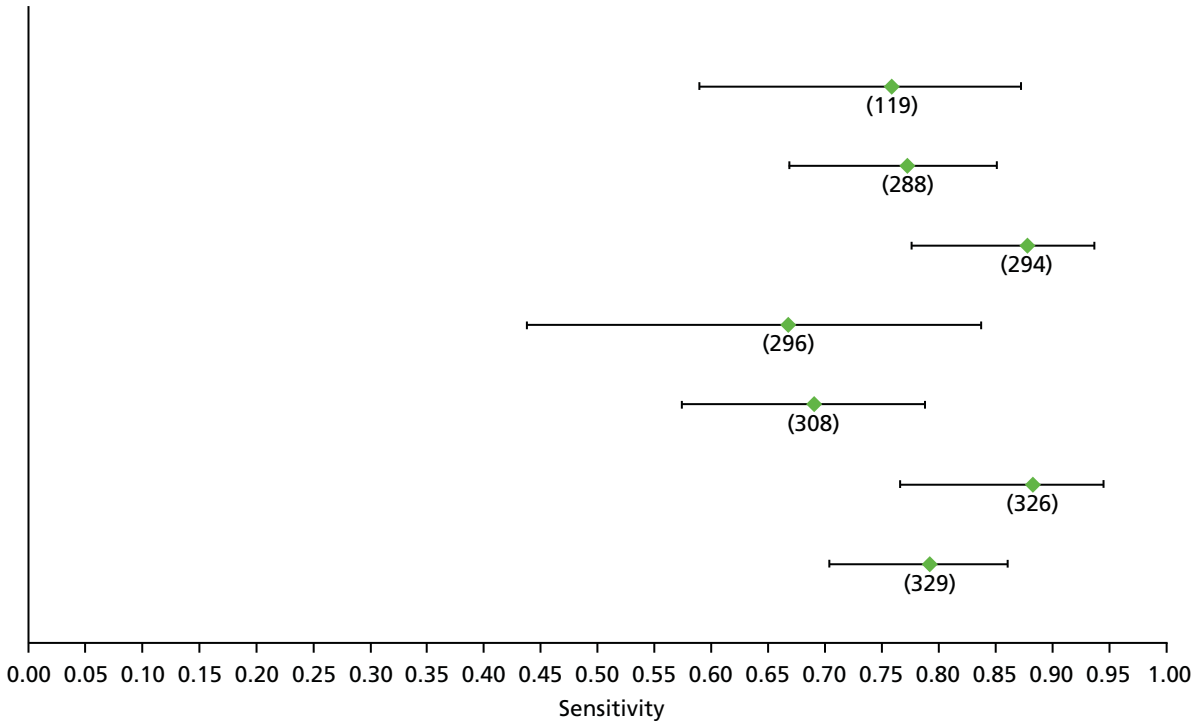

FIGURE 155 Non-alcoholic fatty liver disease: TE F2 – sensitivity.

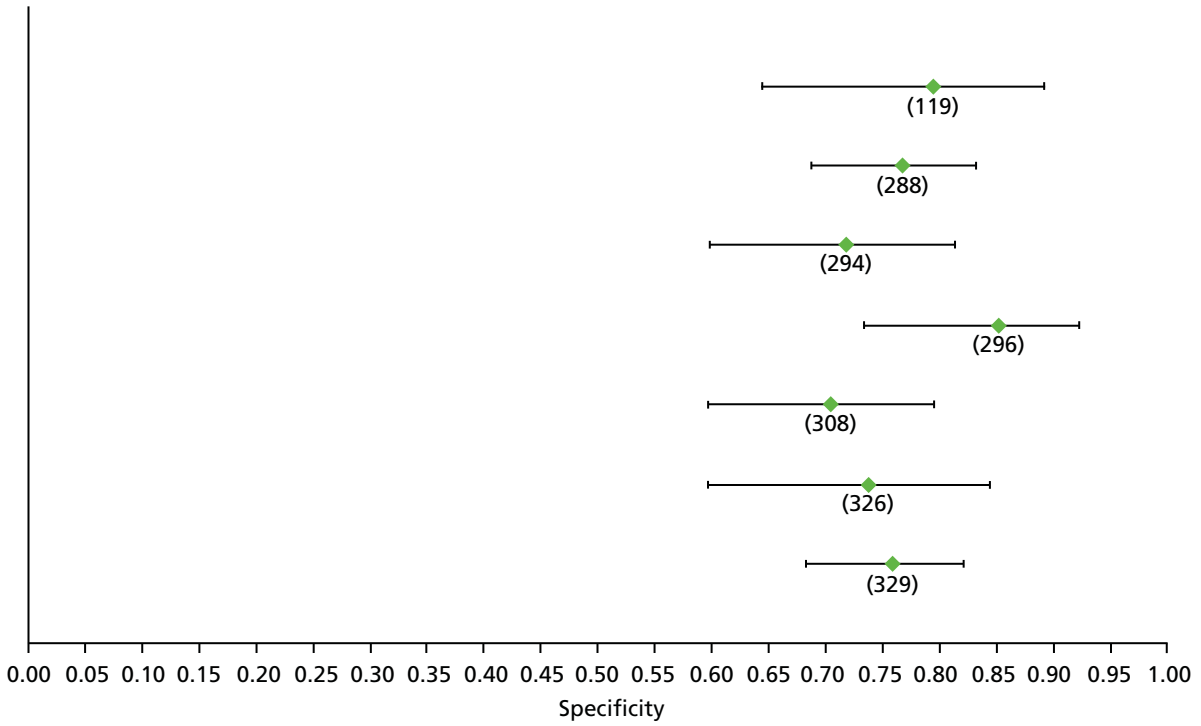

FIGURE 156 Non-alcoholic fatty liver disease: TE F2 – specificity.

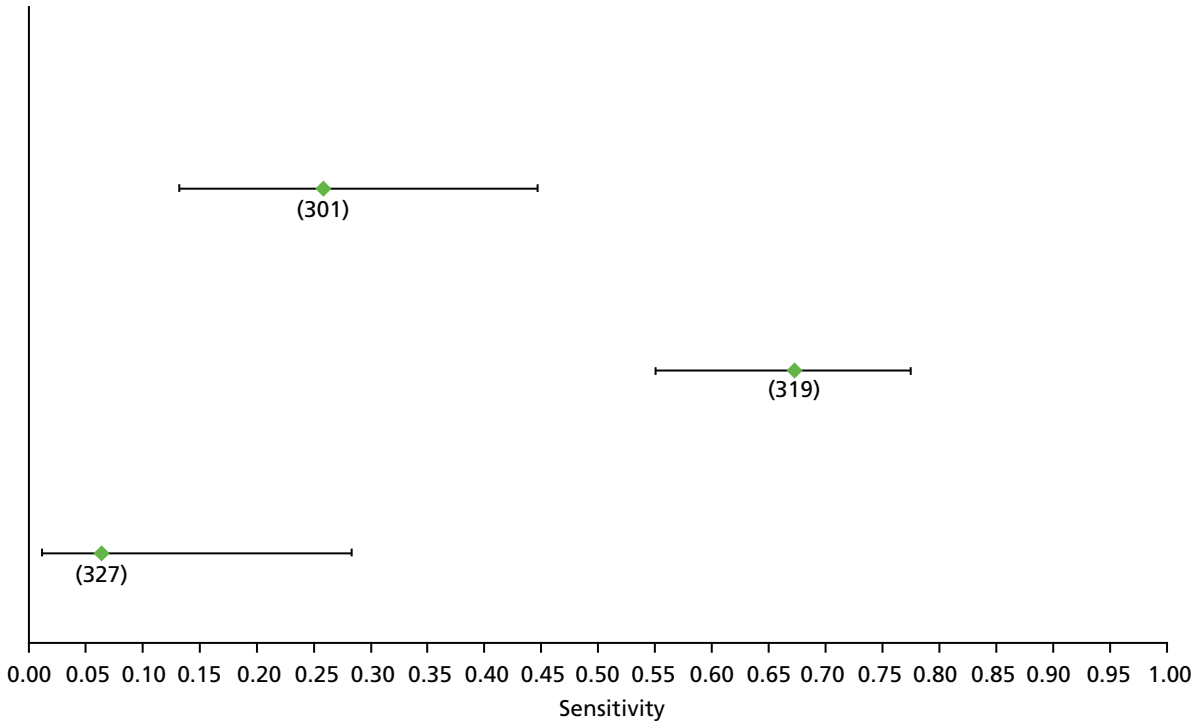

FIGURE 157 Non-alcoholic fatty liver disease: APRI F3 – sensitivity.

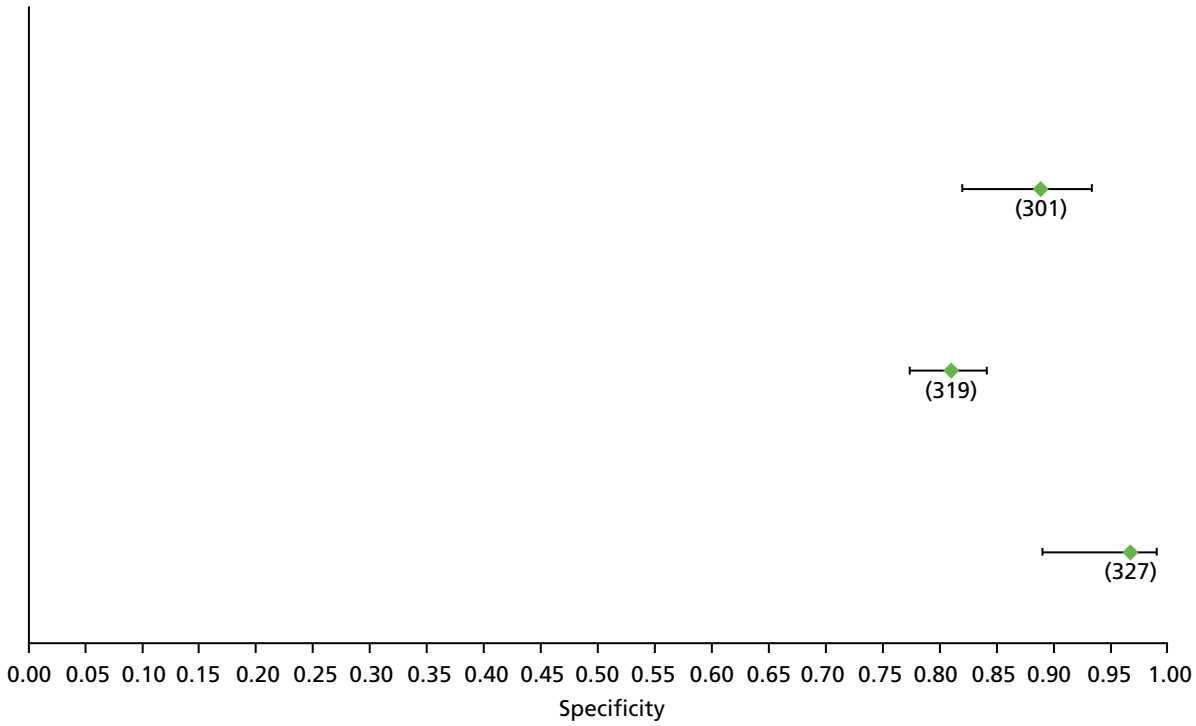

FIGURE 158 Non-alcoholic fatty liver disease: APRI F3 – specificity.

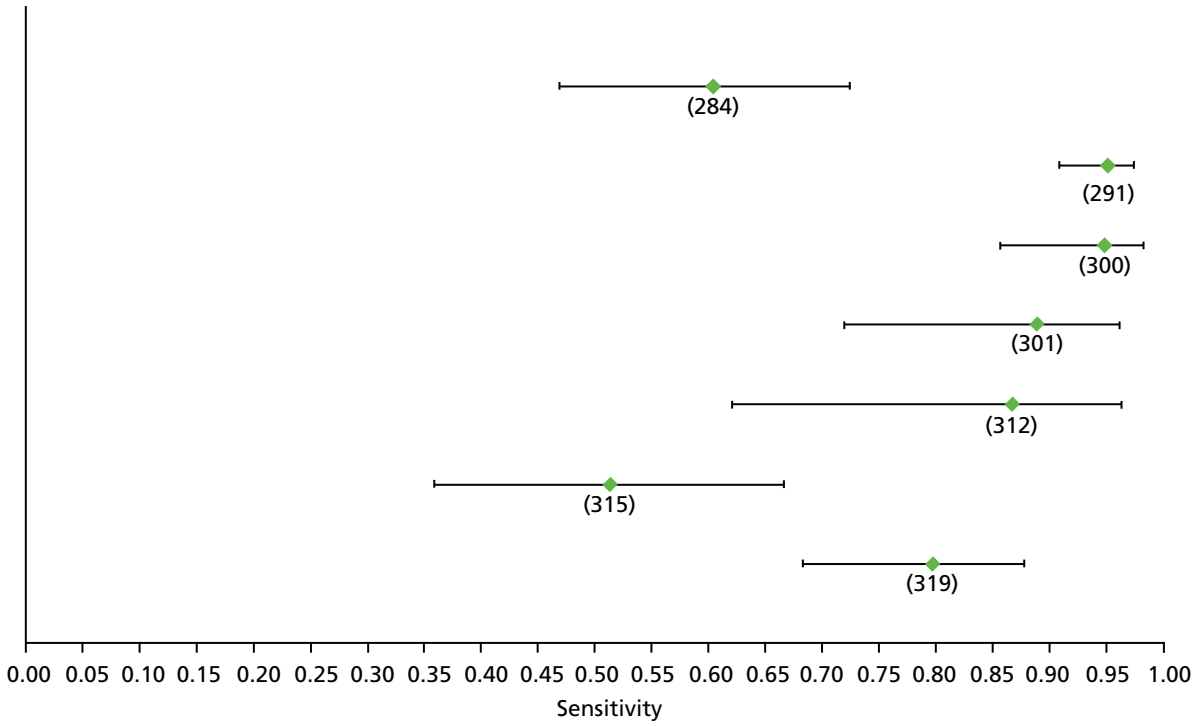

FIGURE 159 Non-alcoholic fatty liver disease: BARD F3 – sensitivity.

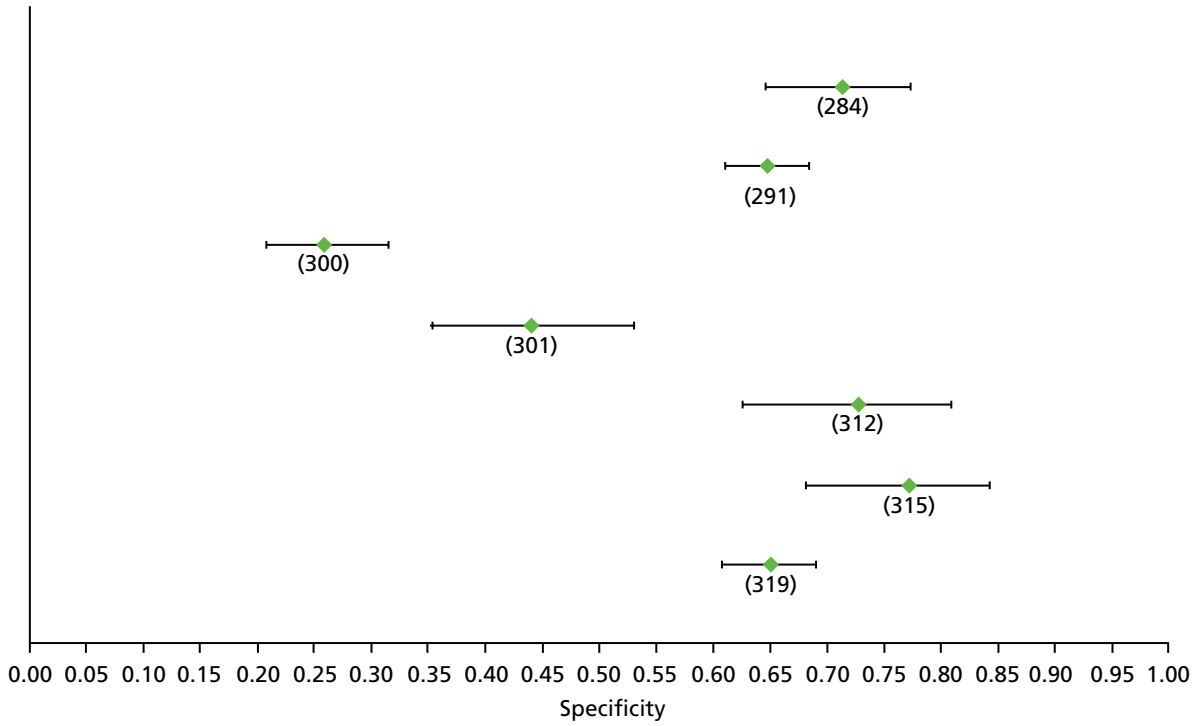

FIGURE 160 Non-alcoholic fatty liver disease: BARD F3 – specificity.

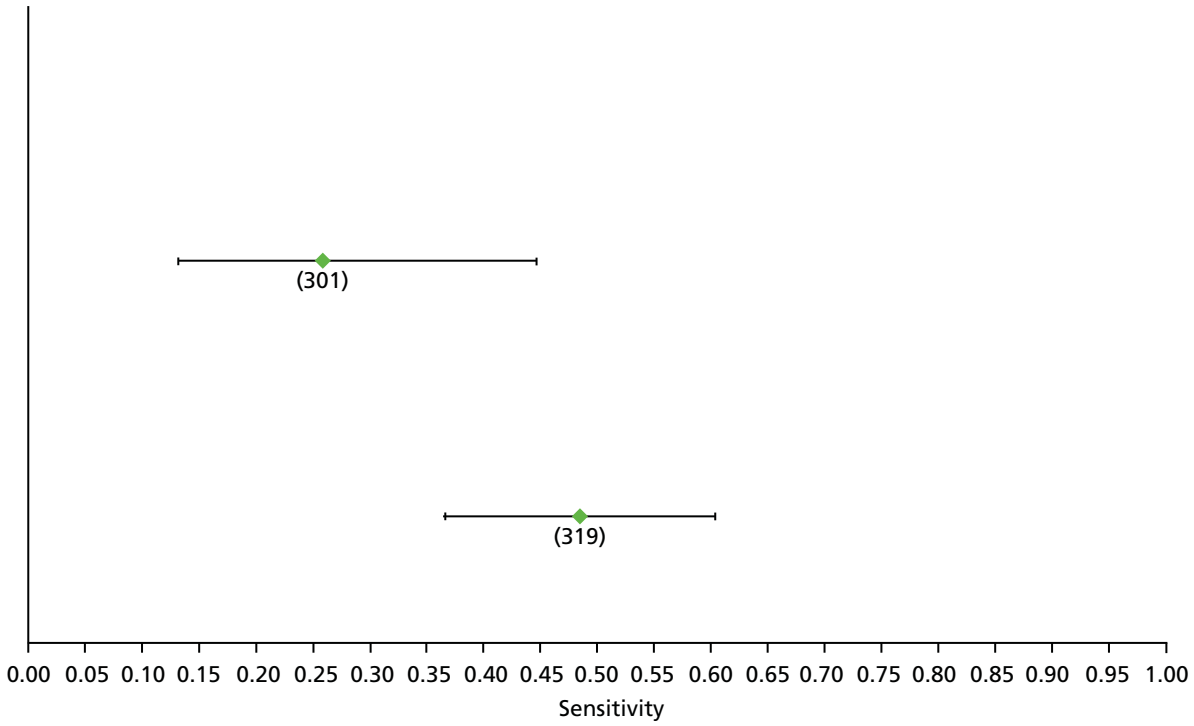

FIGURE 161 Non-alcoholic fatty liver disease: FIB-4 F3 (high cut-off) – sensitivity.

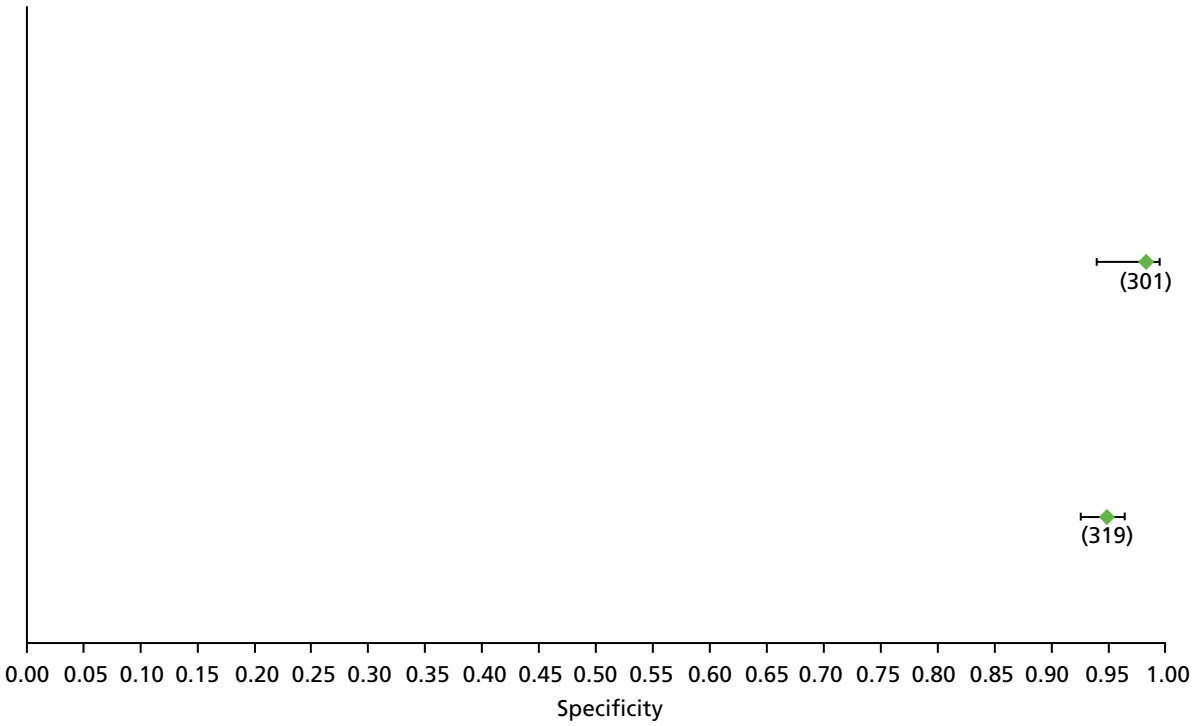

FIGURE 162 Non-alcoholic fatty liver disease: FIB-4 F3 (high cut-off) – specificity.

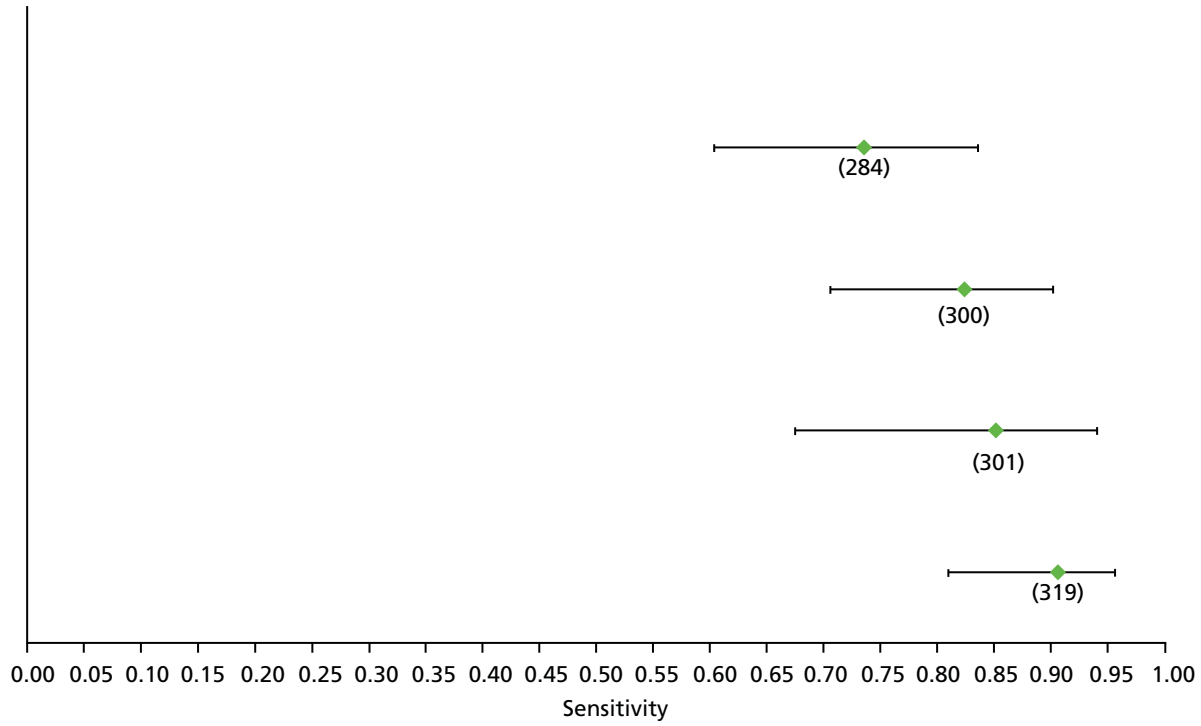

FIGURE 163 Non-alcoholic fatty liver disease: FIB-4 F3 (low cut-off) – sensitivity.

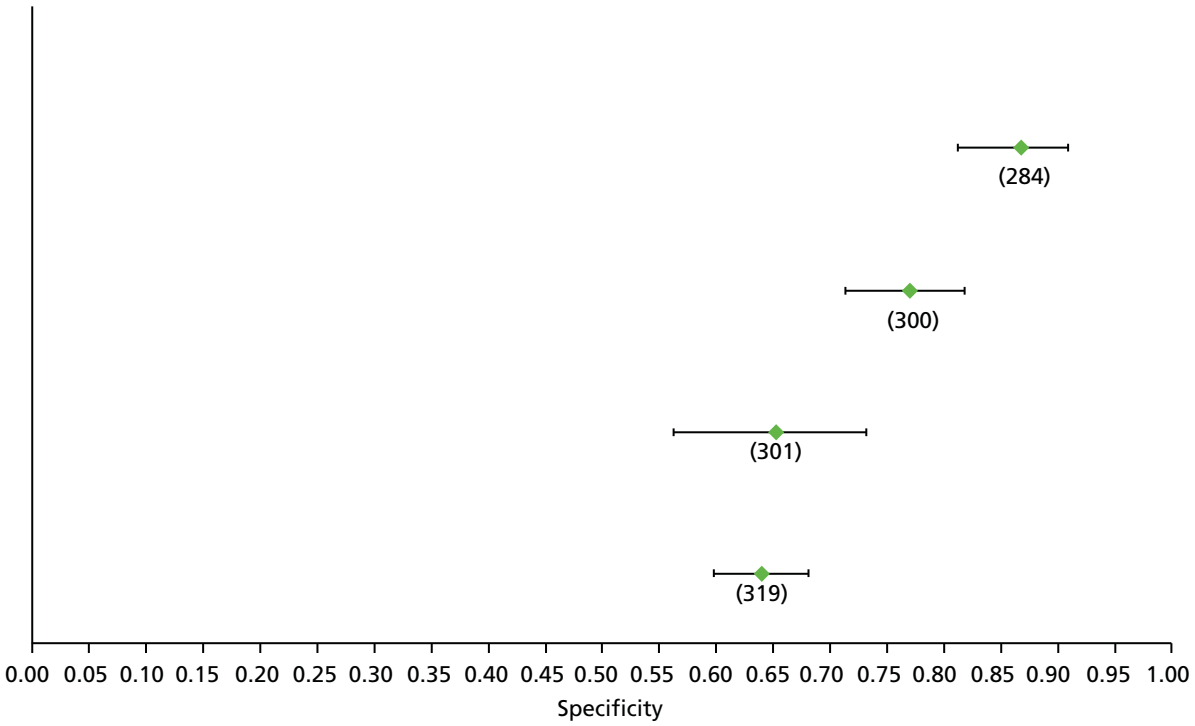

FIGURE 164 Non-alcoholic fatty liver disease: FIB-4 F3 (low cut-off) – specificity.

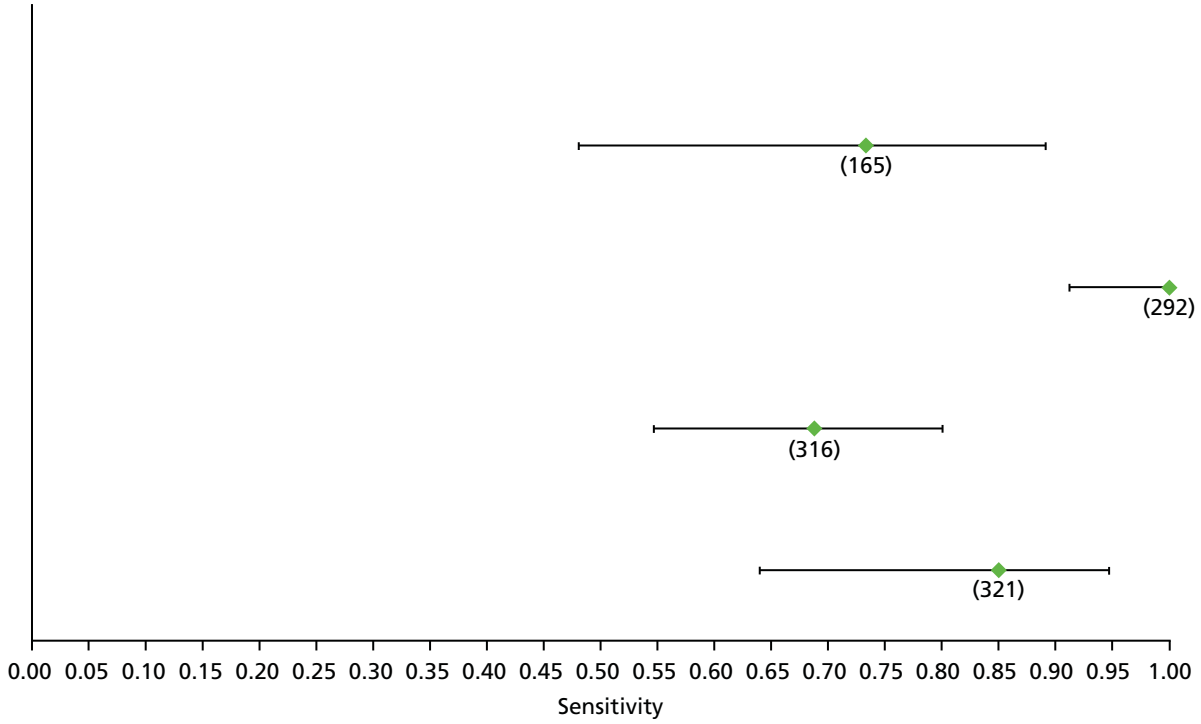

FIGURE 165 Non-alcoholic fatty liver disease: hyaluronic acid F3 – sensitivity.

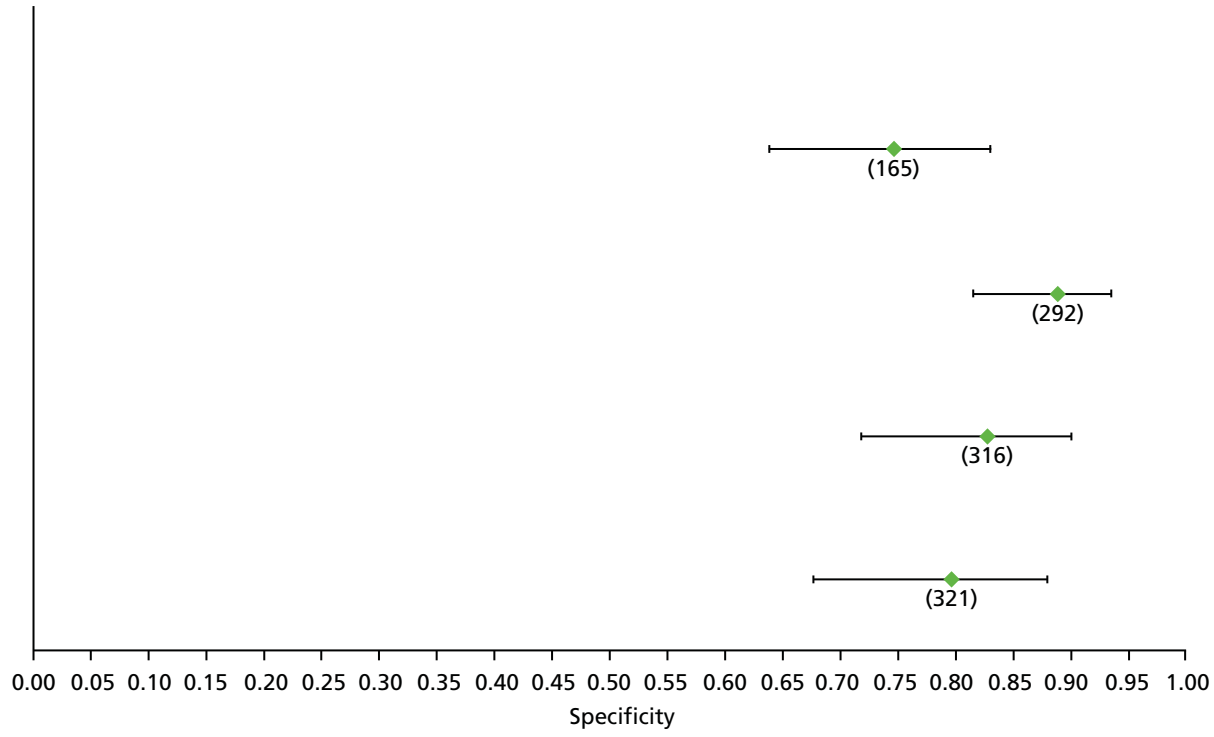

FIGURE 166 Non-alcoholic fatty liver disease: hyaluronic acid F3 – specificity.

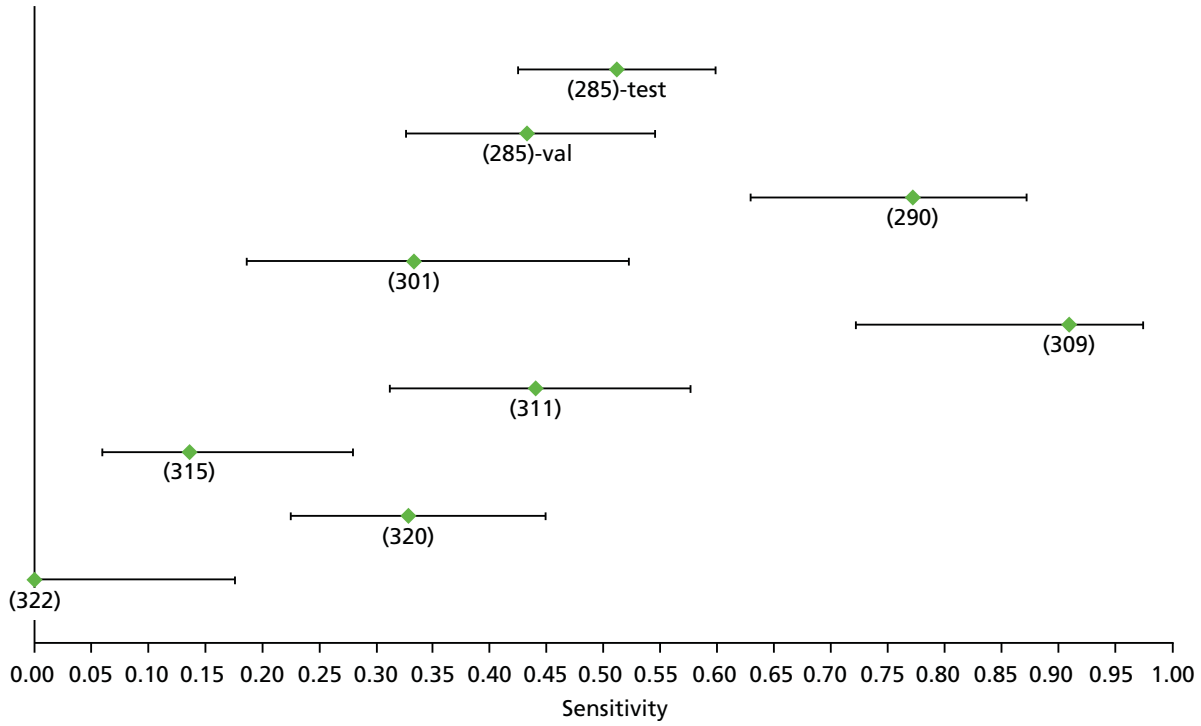

FIGURE 167 Non-alcoholic fatty liver disease: NFS F3 (high cut-off) – sensitivity. Val, validation cohort.

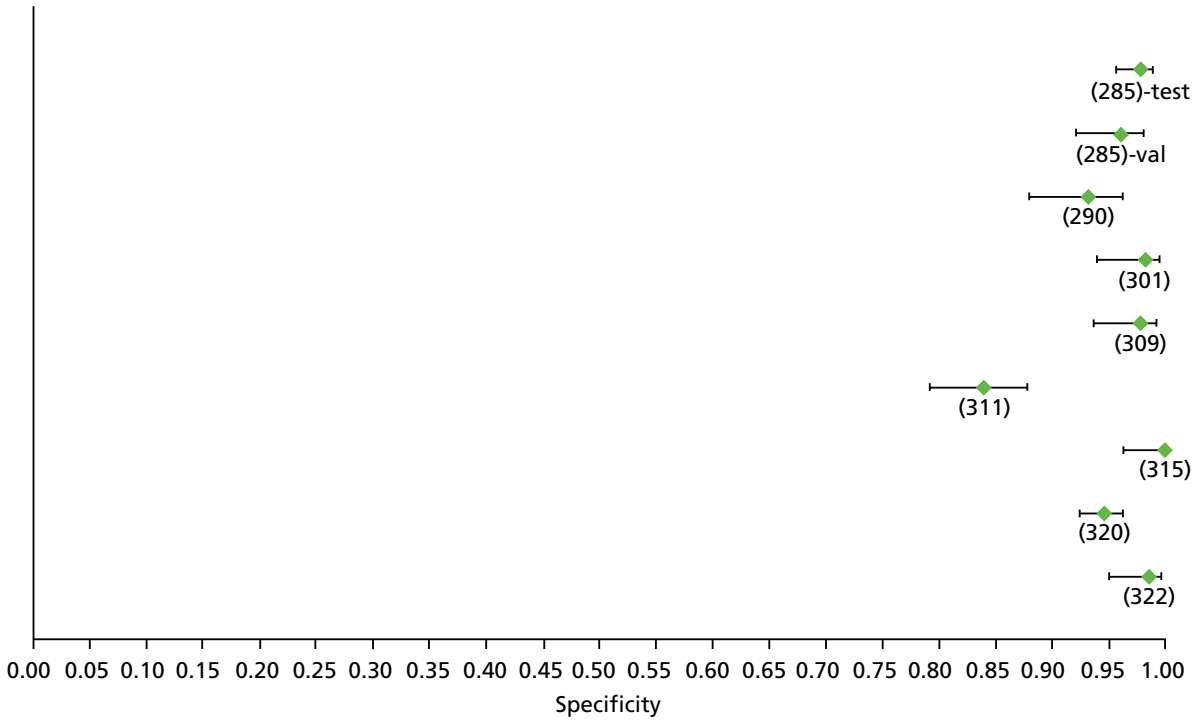

FIGURE 168 Non-alcoholic fatty liver disease: NFS F3 (high cut-off) – specificity. Val, validation cohort.

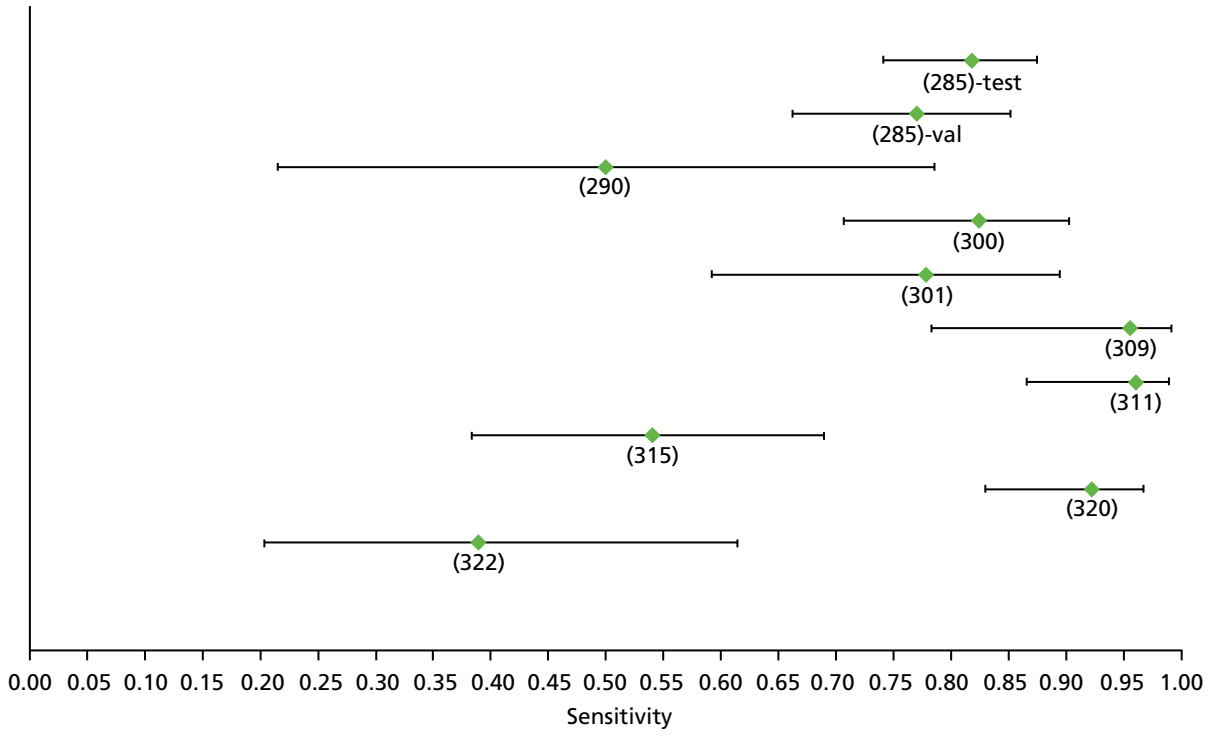

FIGURE 169 Non-alcoholic fatty liver disease: NFS F3 (low cut-off) – sensitivity. Val, validation cohort.

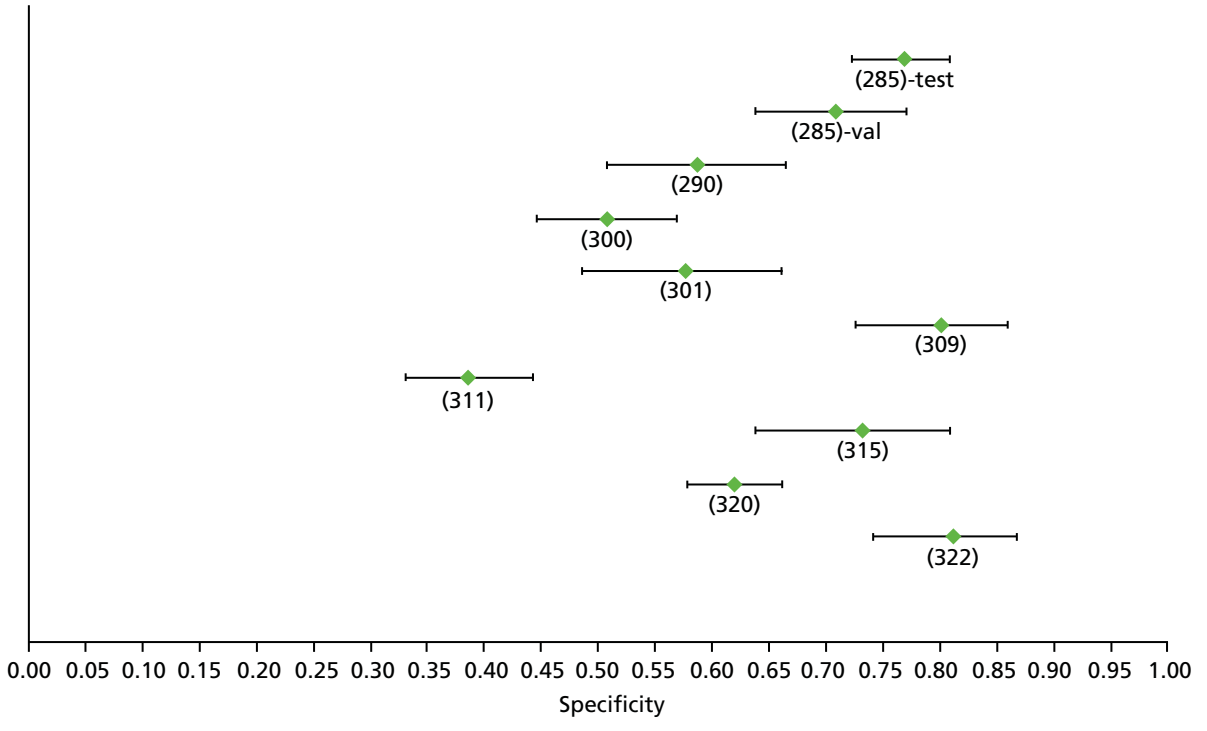

FIGURE 170 Non-alcoholic fatty liver disease: NFS F3 (low cut-off) – specificity. Val, validation cohort.

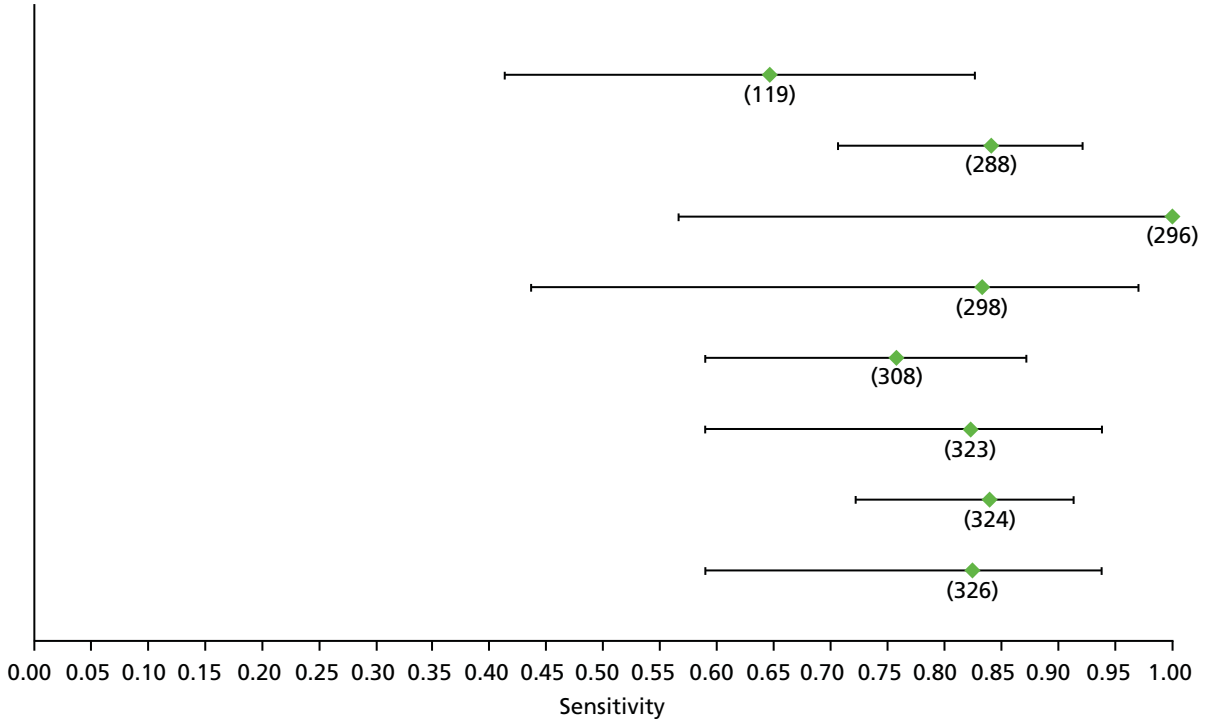

FIGURE 171 Non-alcoholic fatty liver disease: TE F3 – sensitivity.

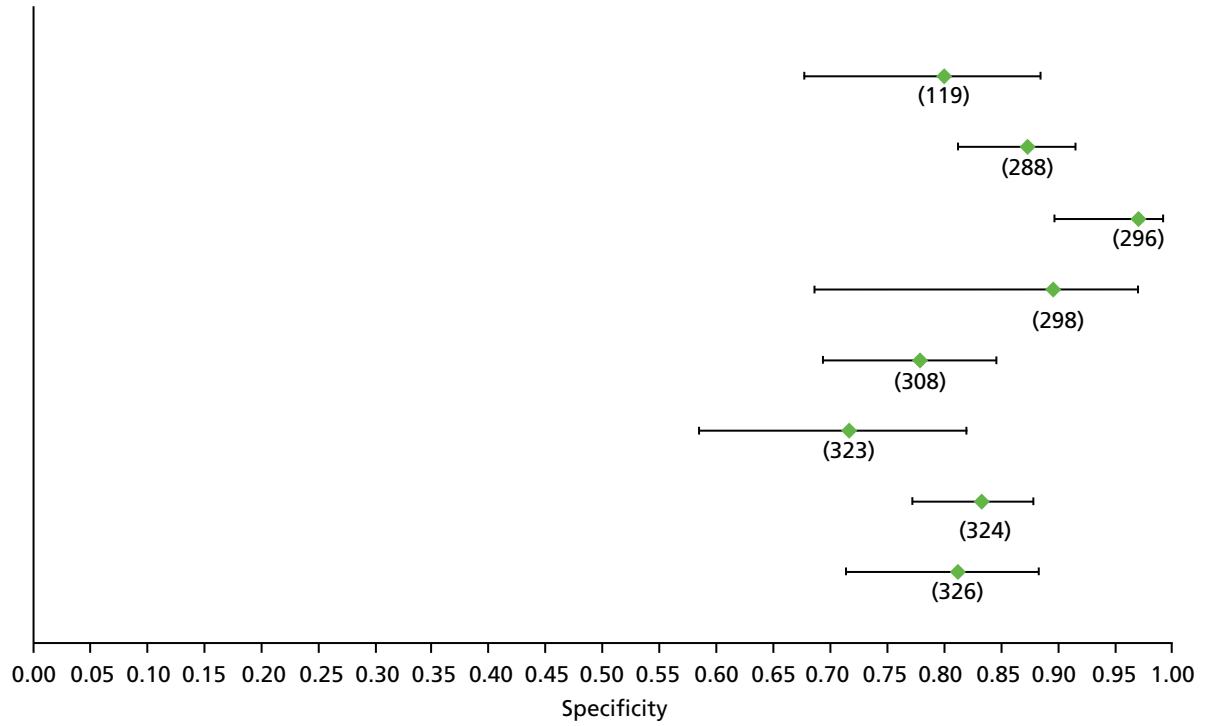

FIGURE 172 Non-alcoholic fatty liver disease: TE F3 – specificity.

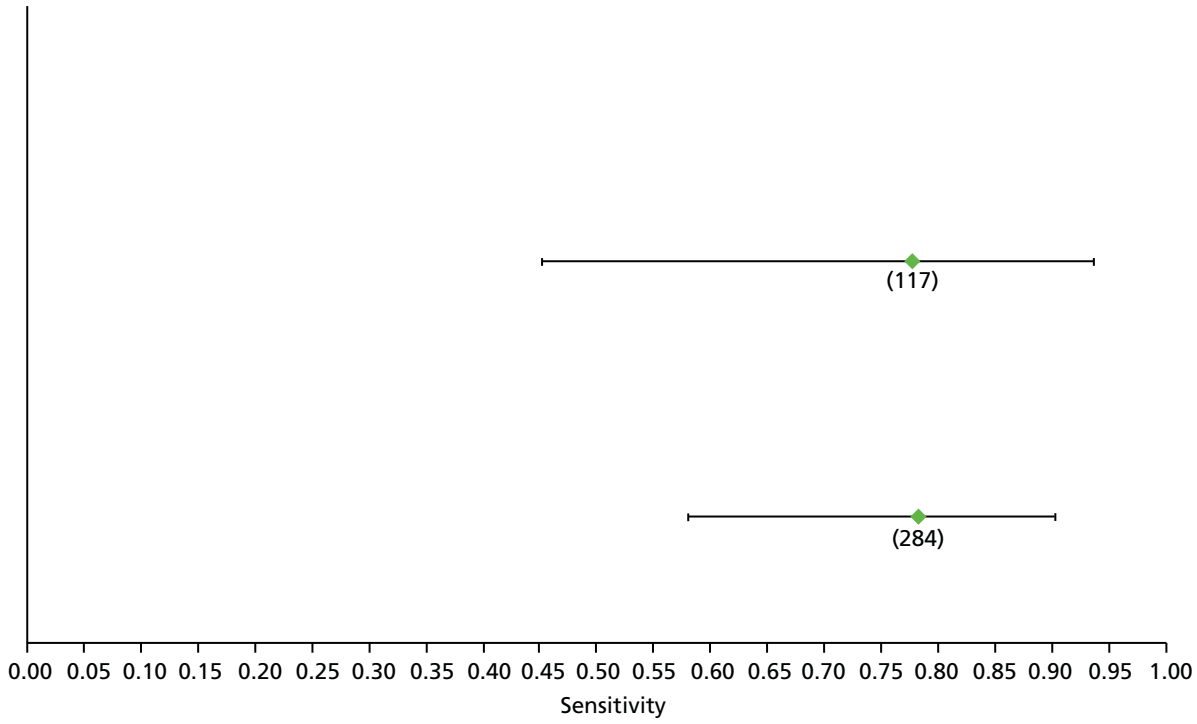

FIGURE 173 Non-alcoholic fatty liver disease: APRI F4 – sensitivity.

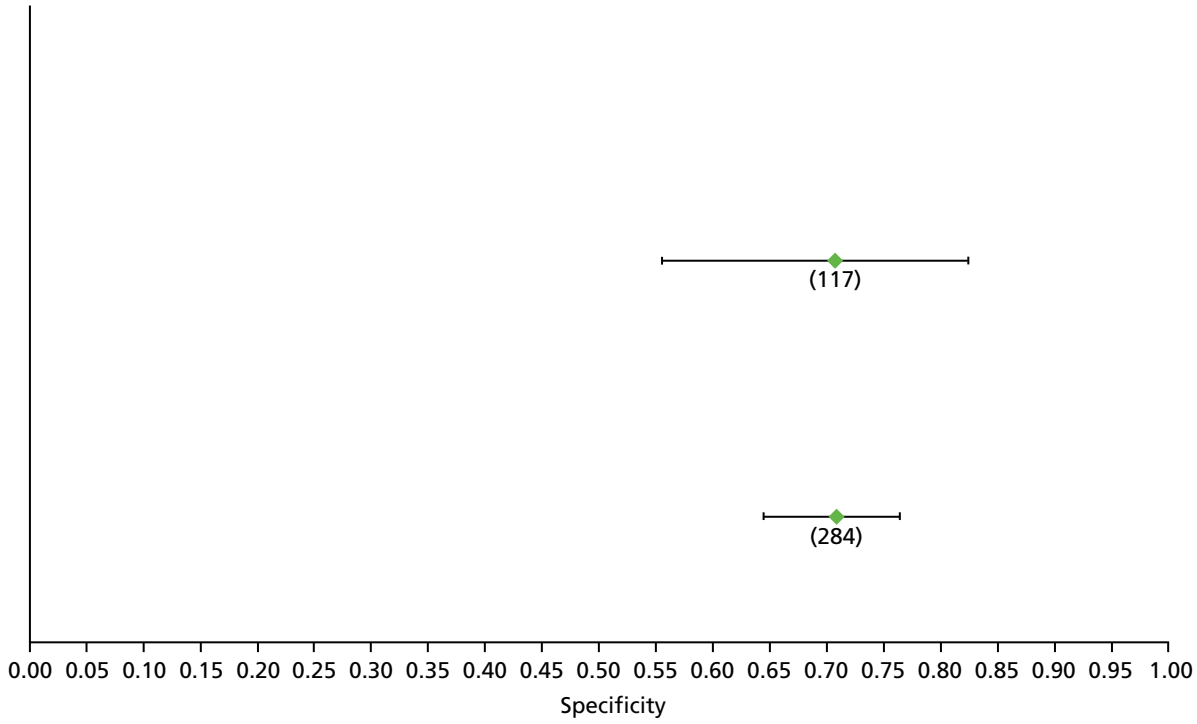

FIGURE 174 Non-alcoholic fatty liver disease: APRI F4 – specificity.

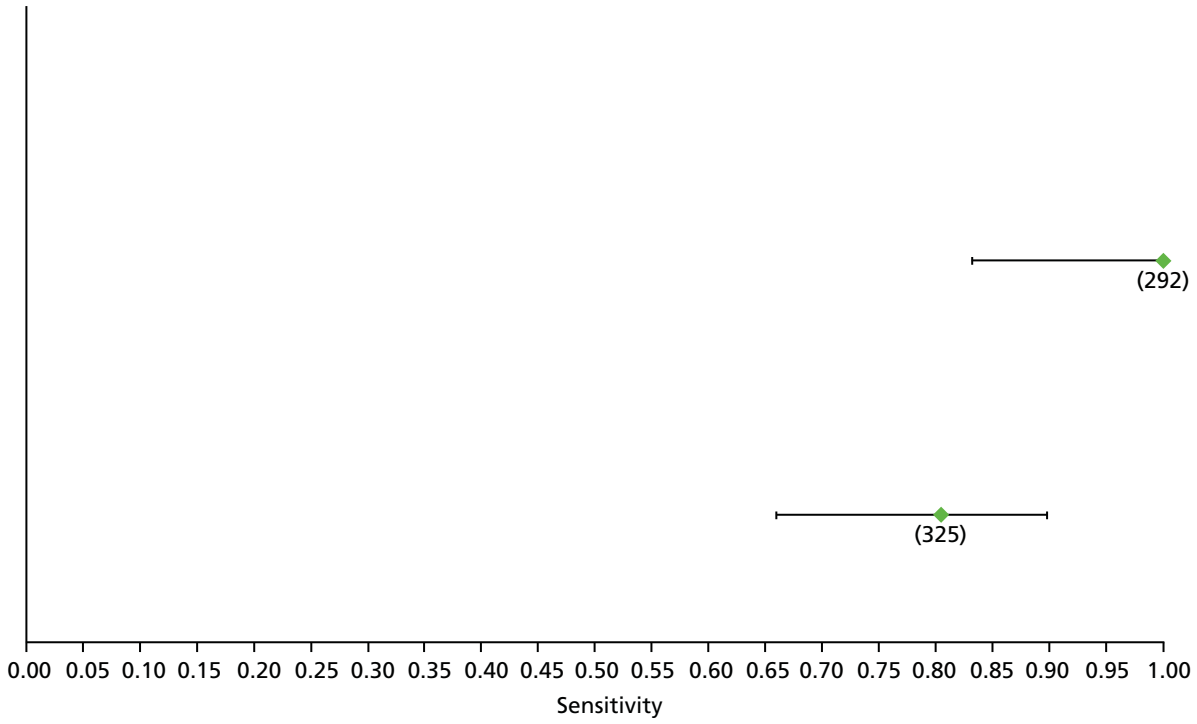

FIGURE 175 Non-alcoholic fatty liver disease: platelet F4 – sensitivity.

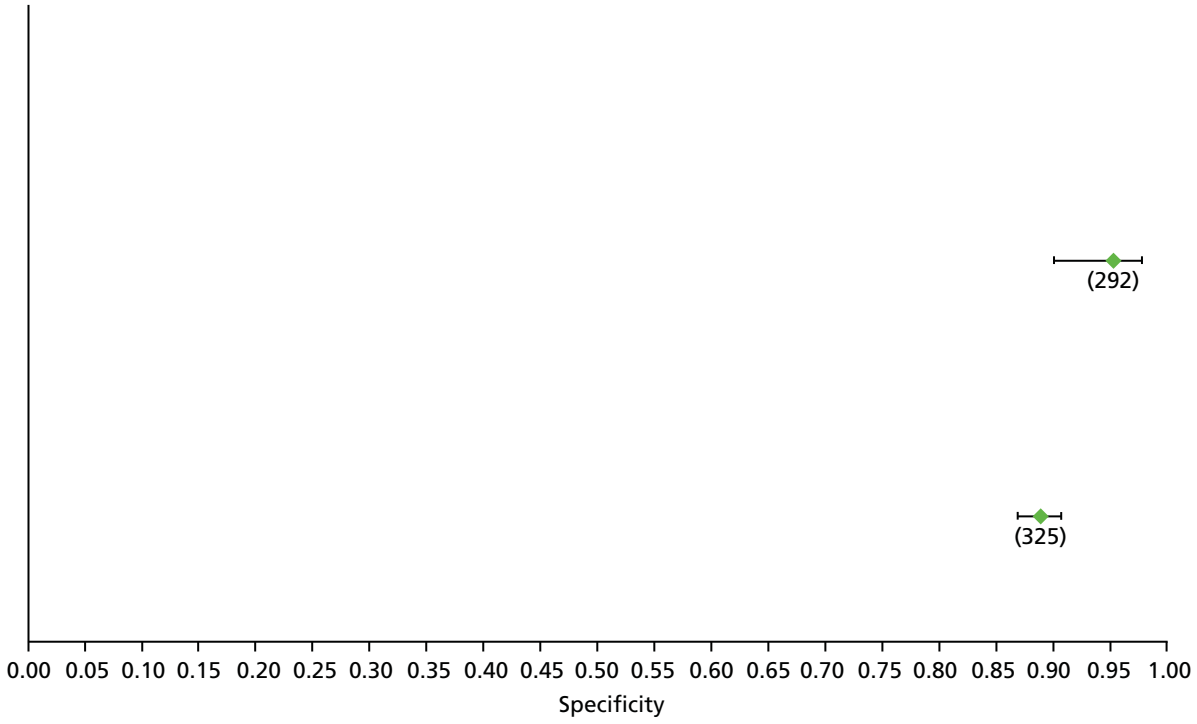

FIGURE 176 Non-alcoholic fatty liver disease: platelet F4 – specificity.

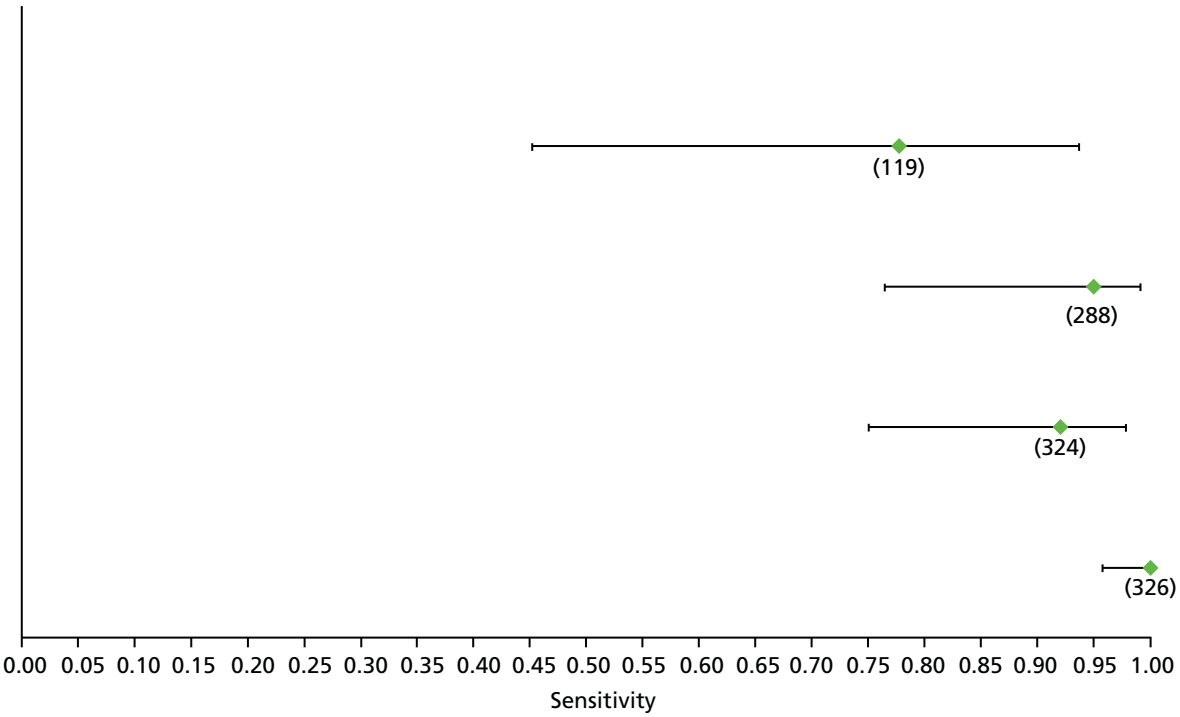

FIGURE 177 Non-alcoholic fatty liver disease: TE F4 – sensitivity.

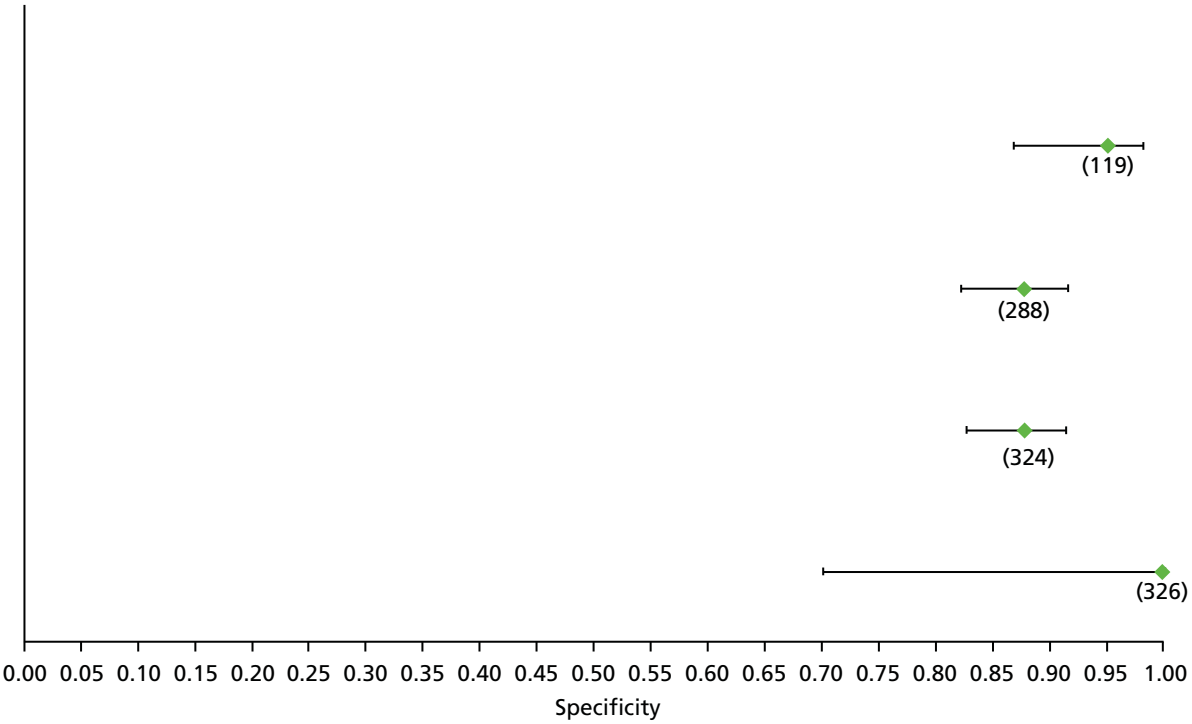

FIGURE 178 Non-alcoholic fatty liver disease: TE F4 – specificity.

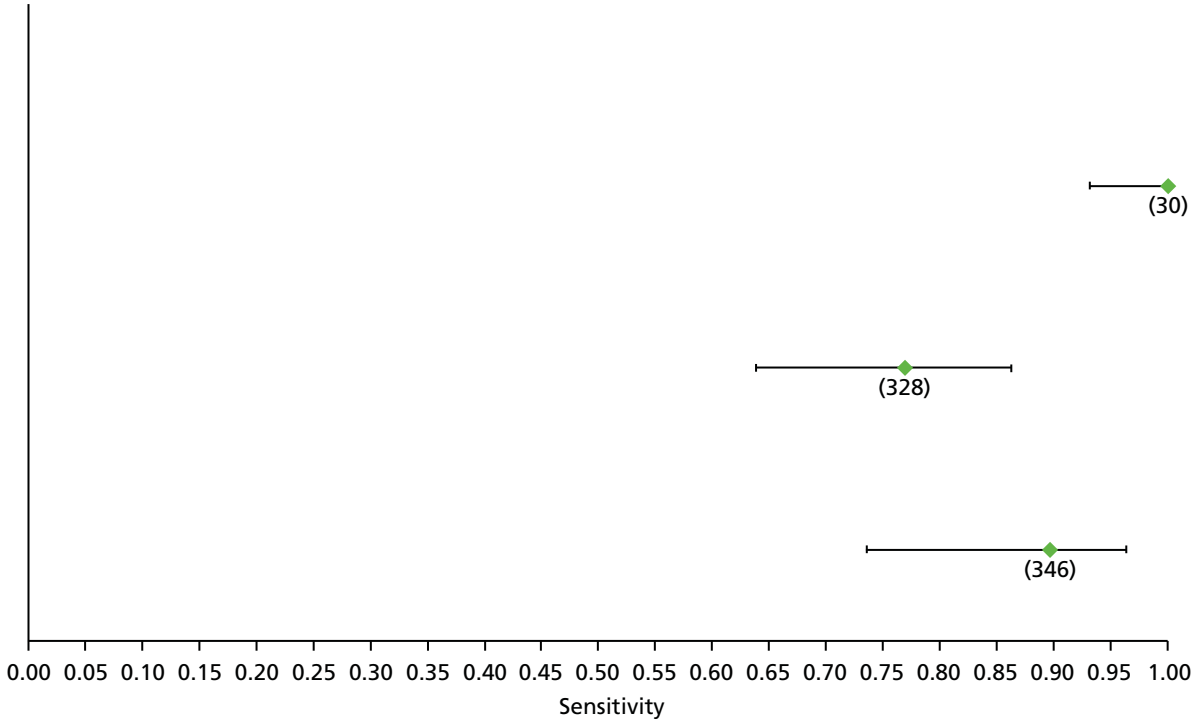

FIGURE 179 Radiology: MR elastography F2 – sensitivity.

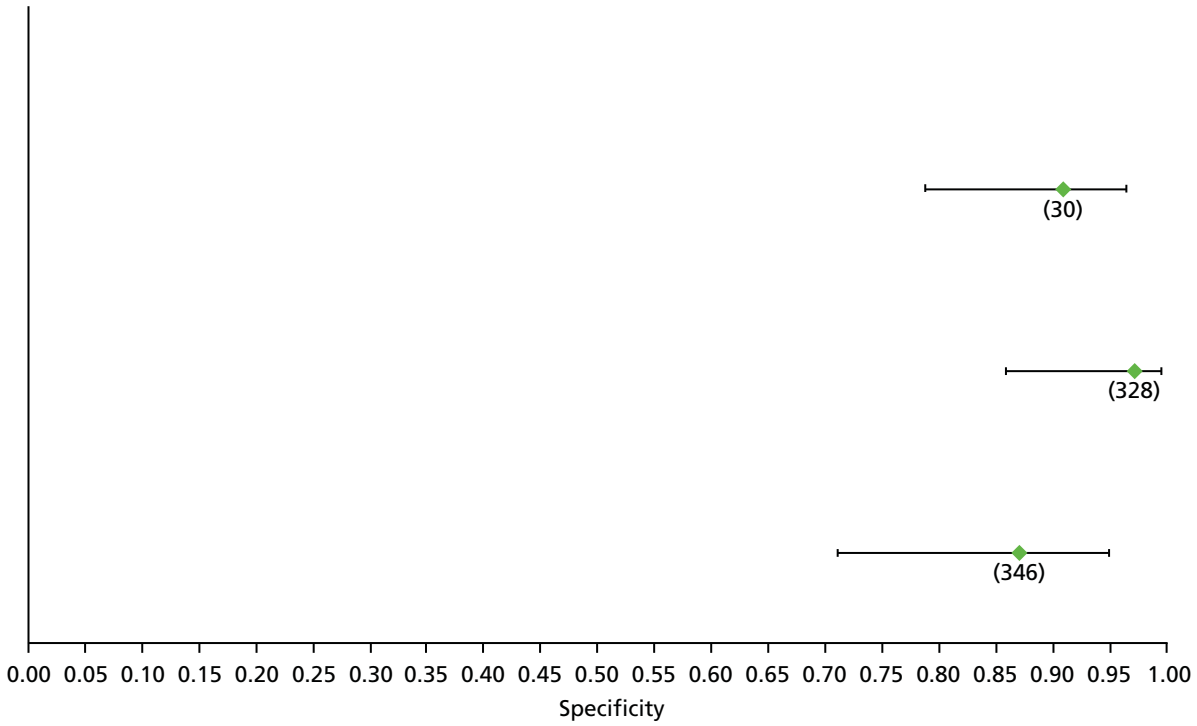

FIGURE 180 Radiology: MR elastography F2 – specificity.

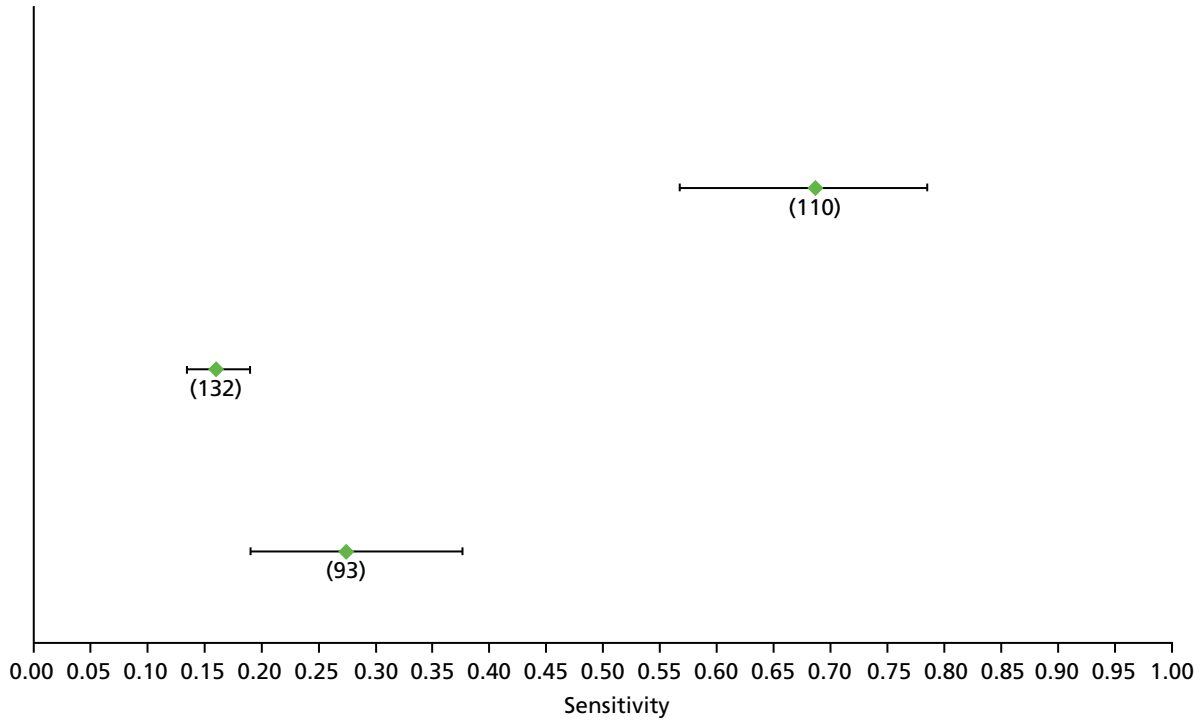

FIGURE 181 Radiology: ultrasound F2 – sensitivity.

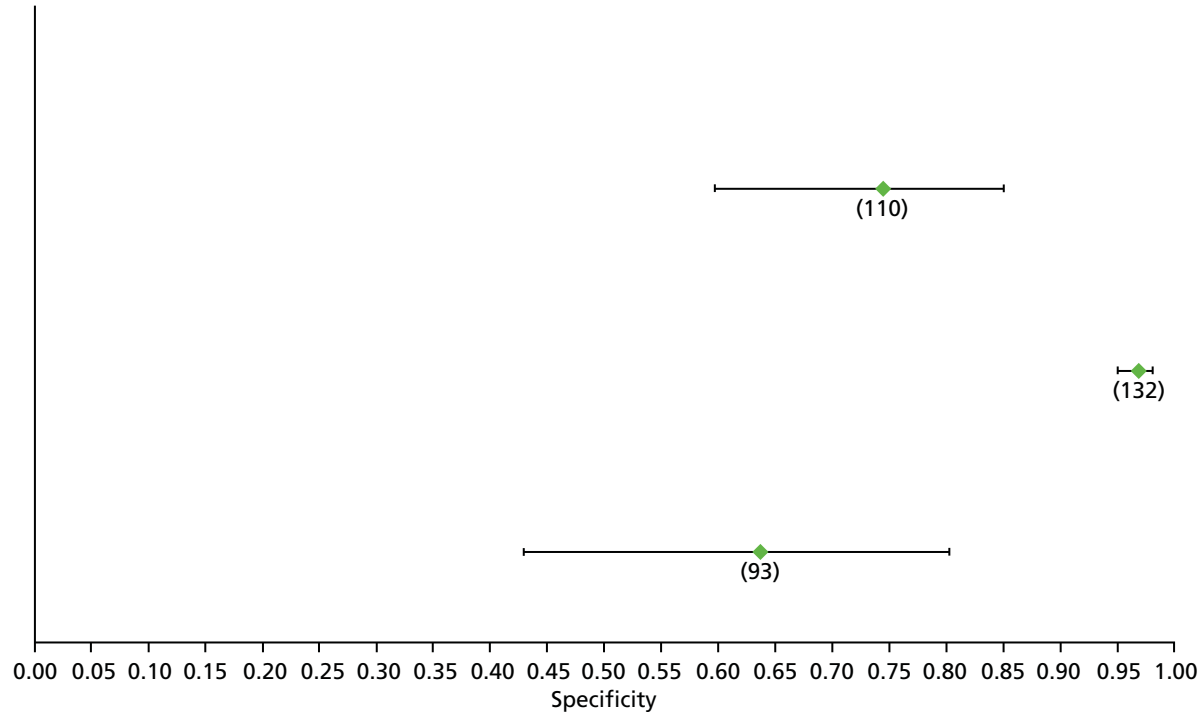

FIGURE 182 Radiology: ultrasound F2 – specificity.

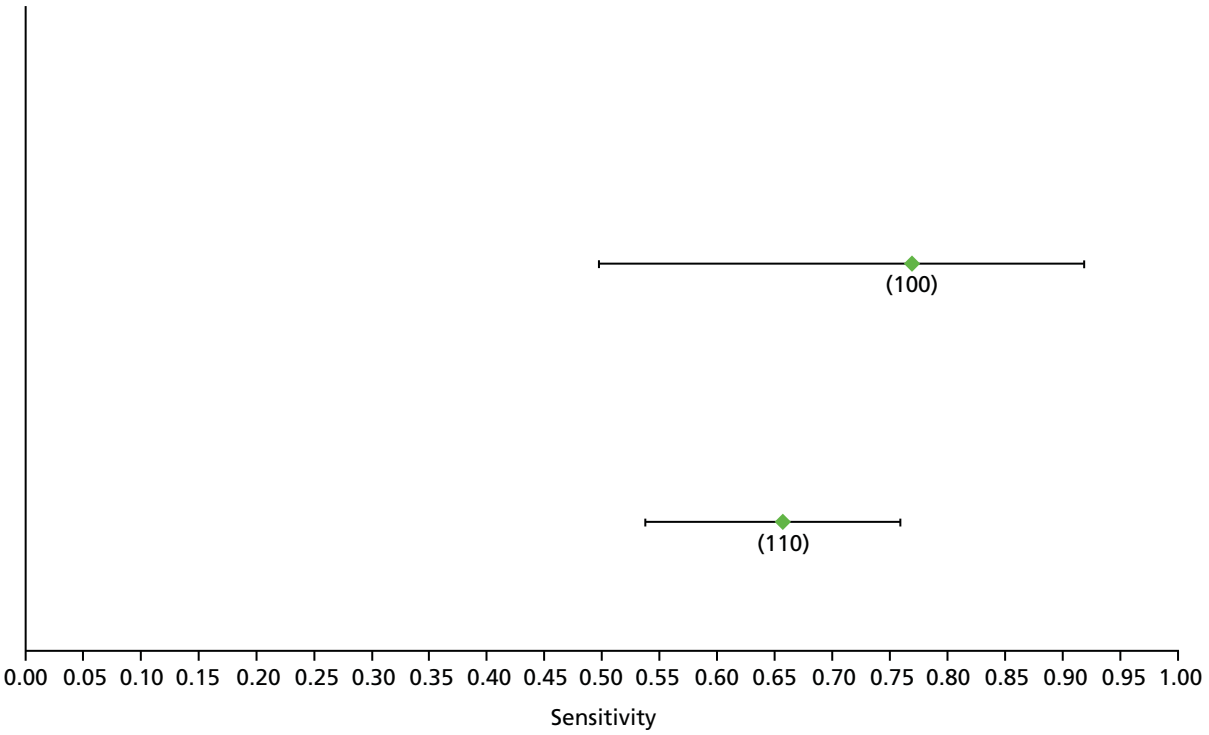

FIGURE 183 Radiology: ultrasound SAPI F2 – sensitivity.

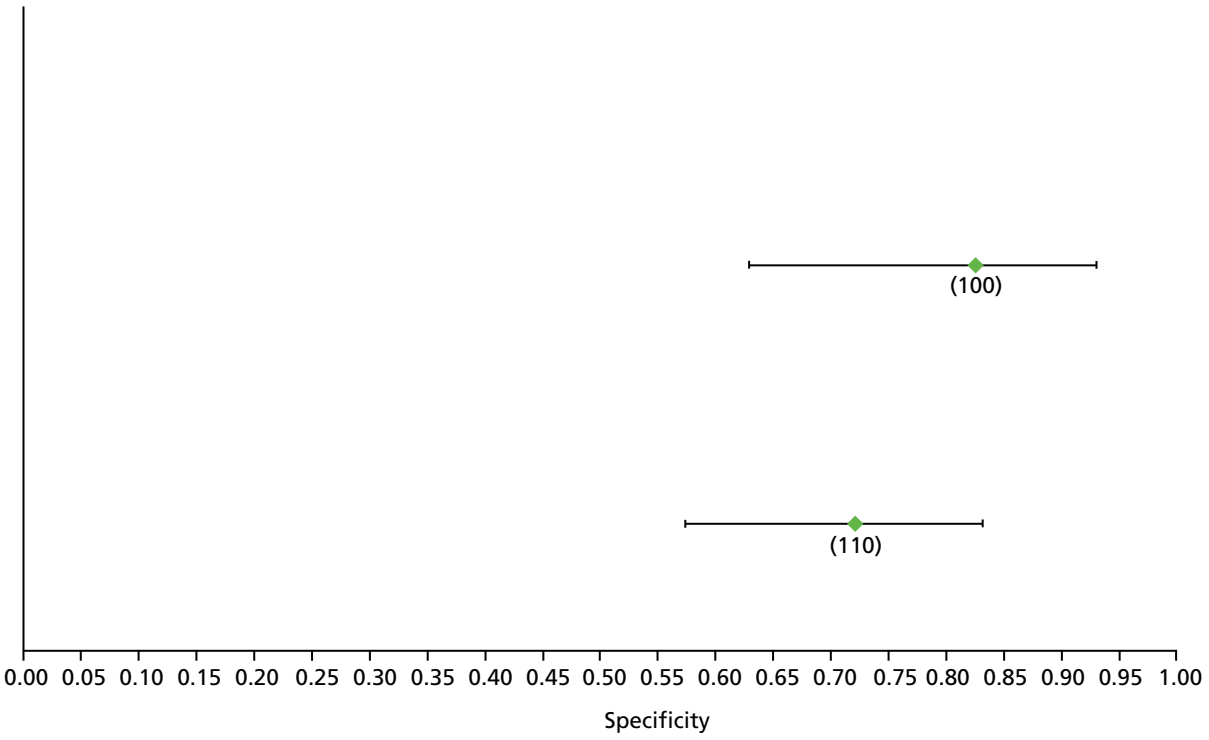

FIGURE 184 Radiology: ultrasound SAPI F2 – specificity.

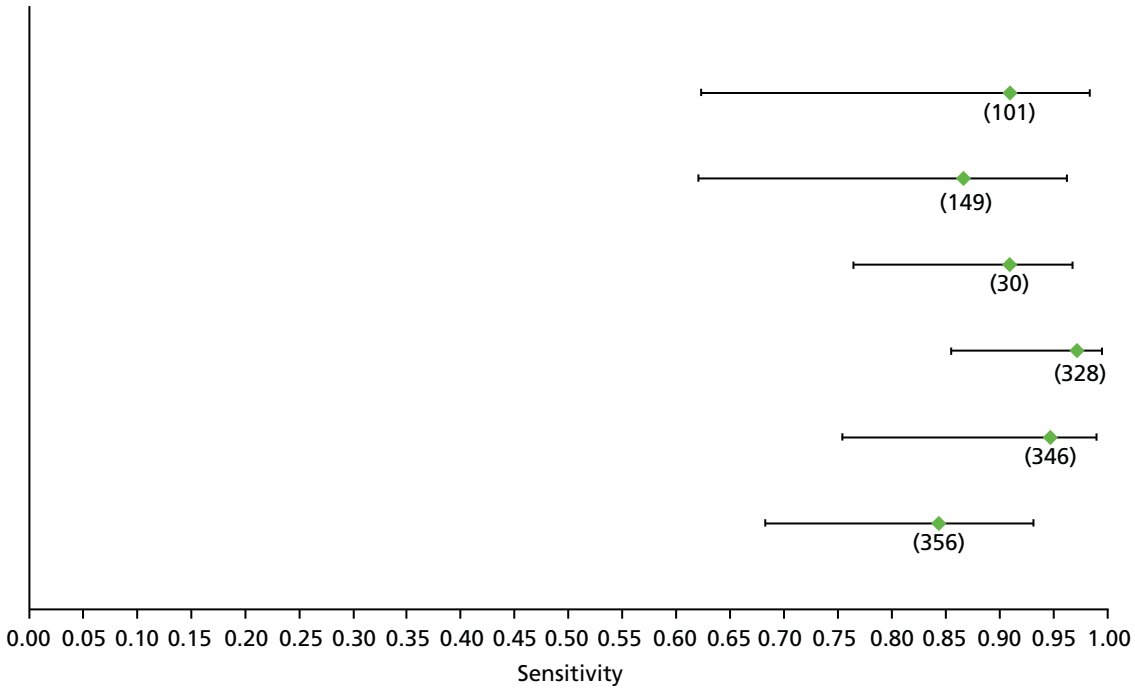

FIGURE 185 Radiology: MR elastography F3 – sensitivity.

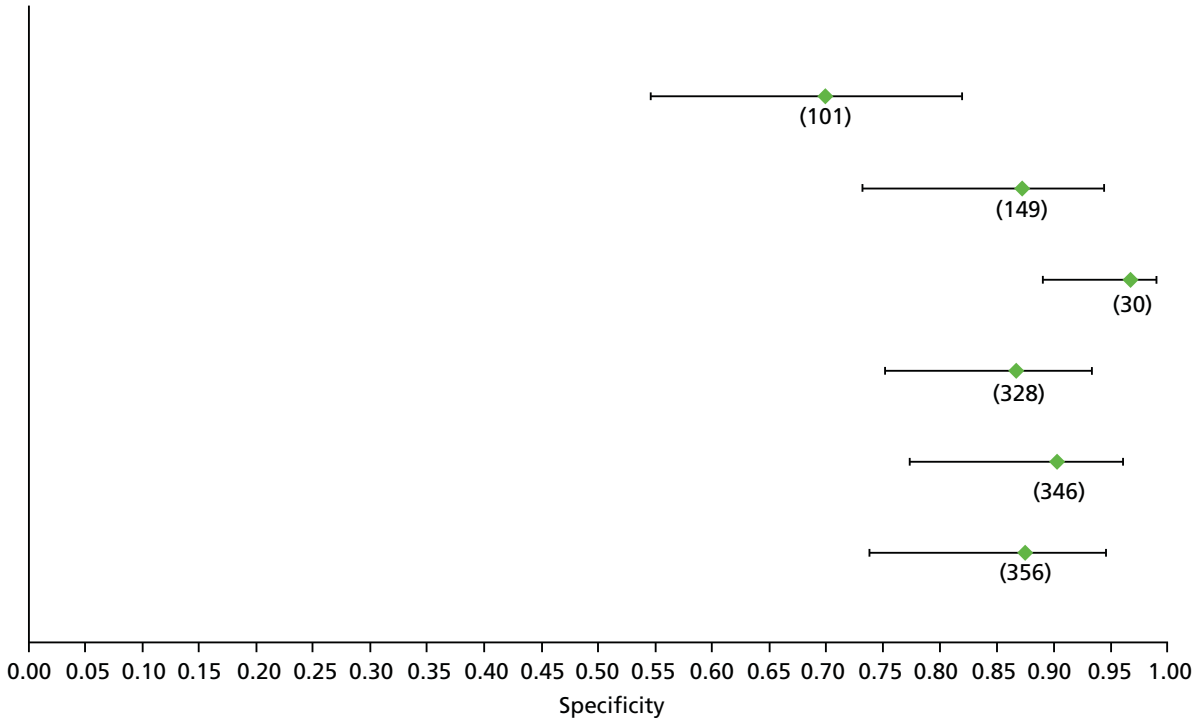

FIGURE 186 Radiology: MR elastography F3 – specificity.

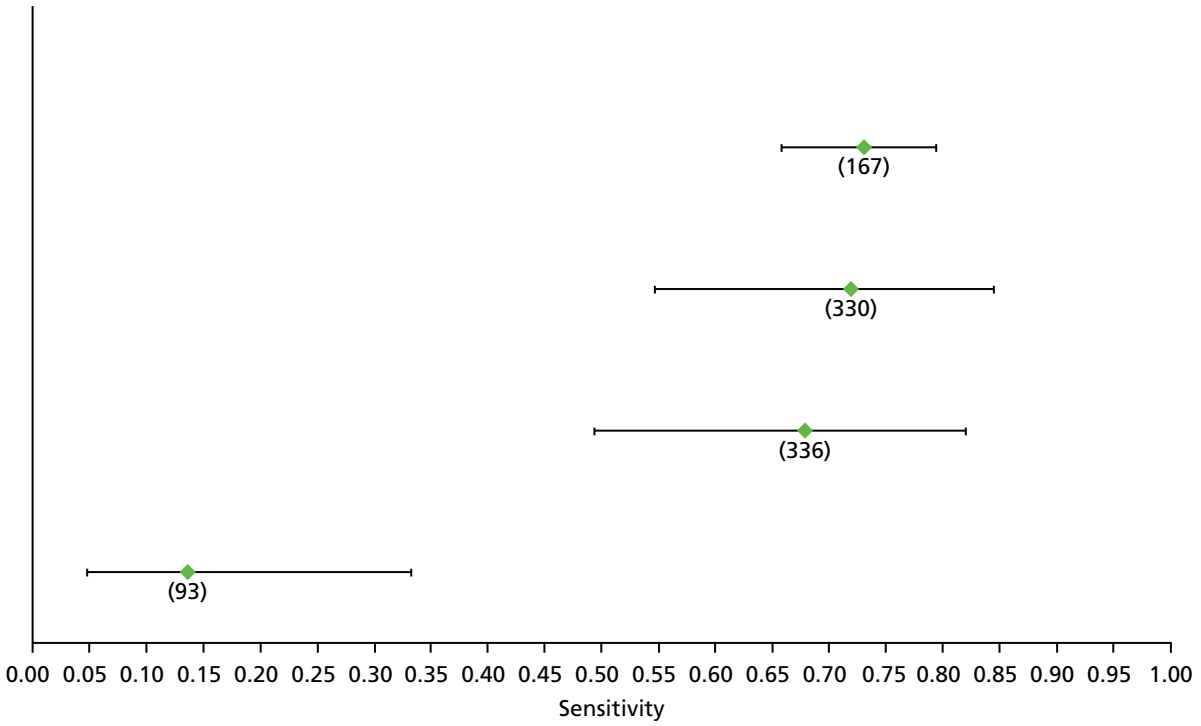

FIGURE 187 Radiology: ultrasound F3 – sensitivity.

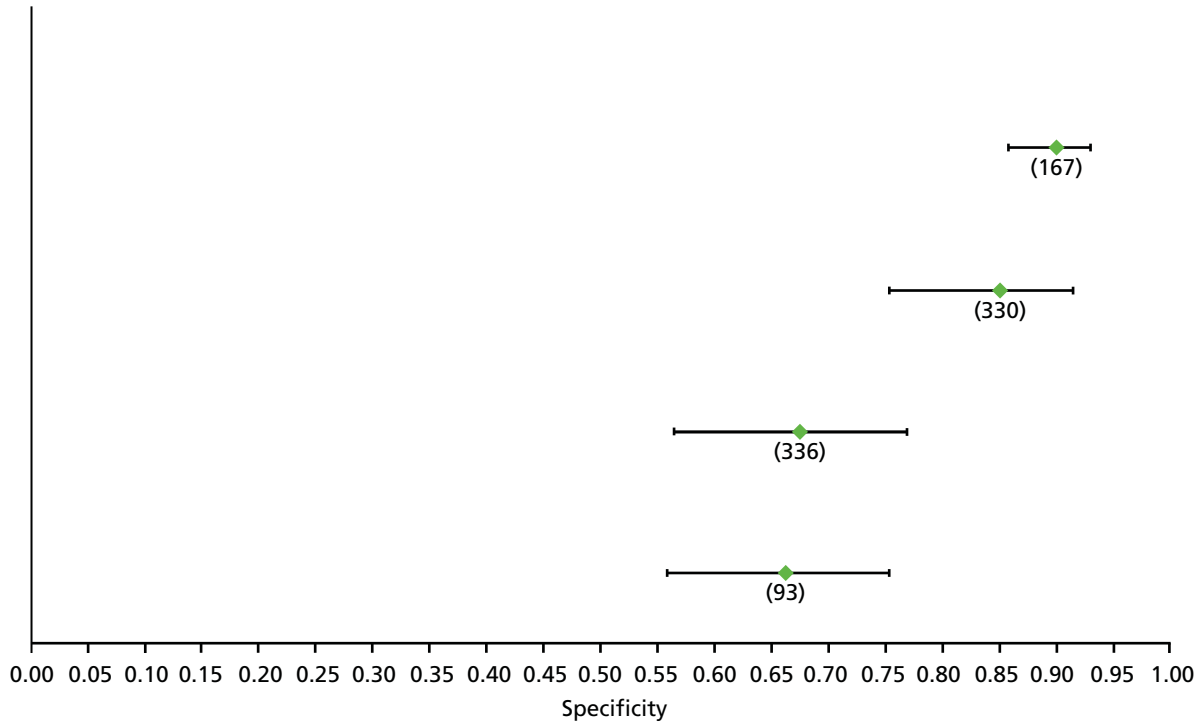

FIGURE 188 Radiology: ultrasound F3 – specificity.

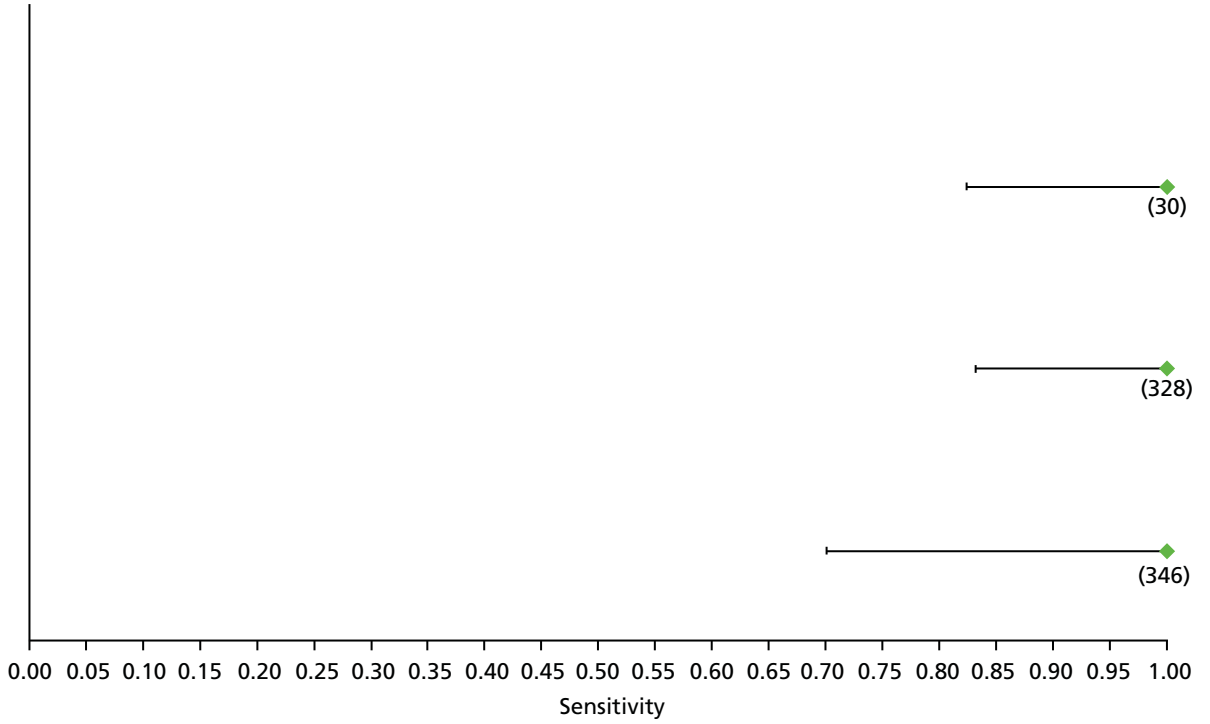

FIGURE 189 Radiology: MR elastography F4 – sensitivity.

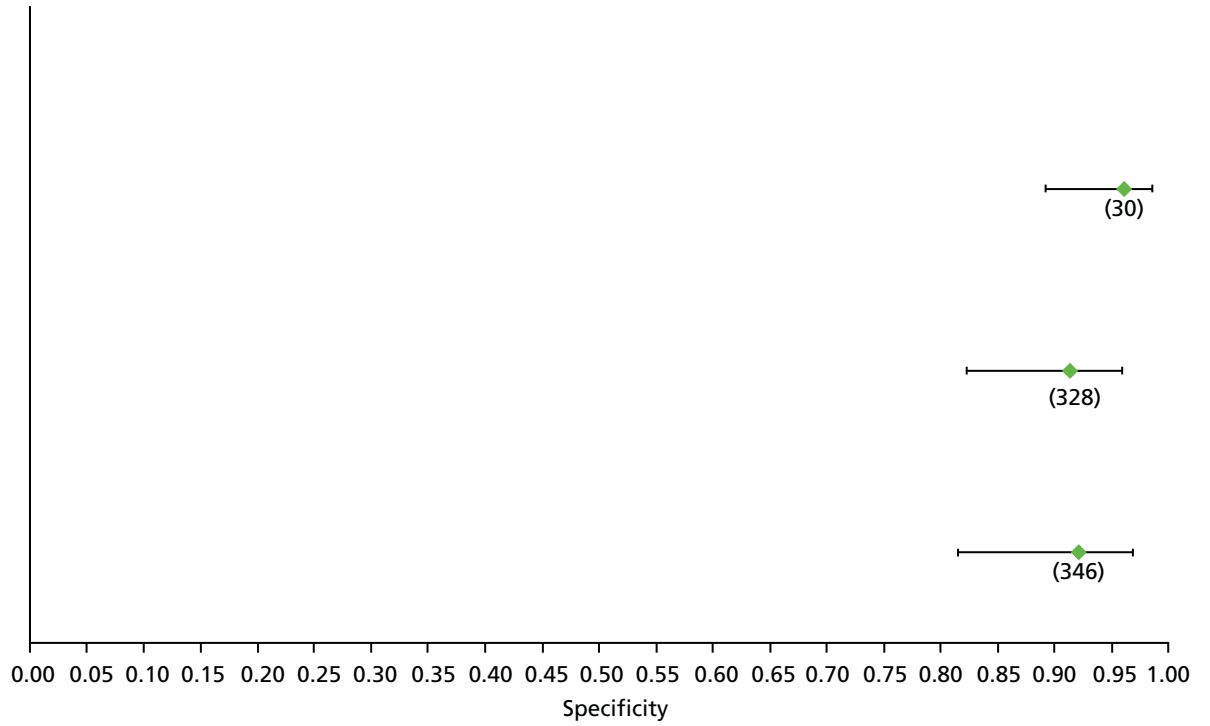

FIGURE 190 Radiology: MR elastography F4 – specificity.

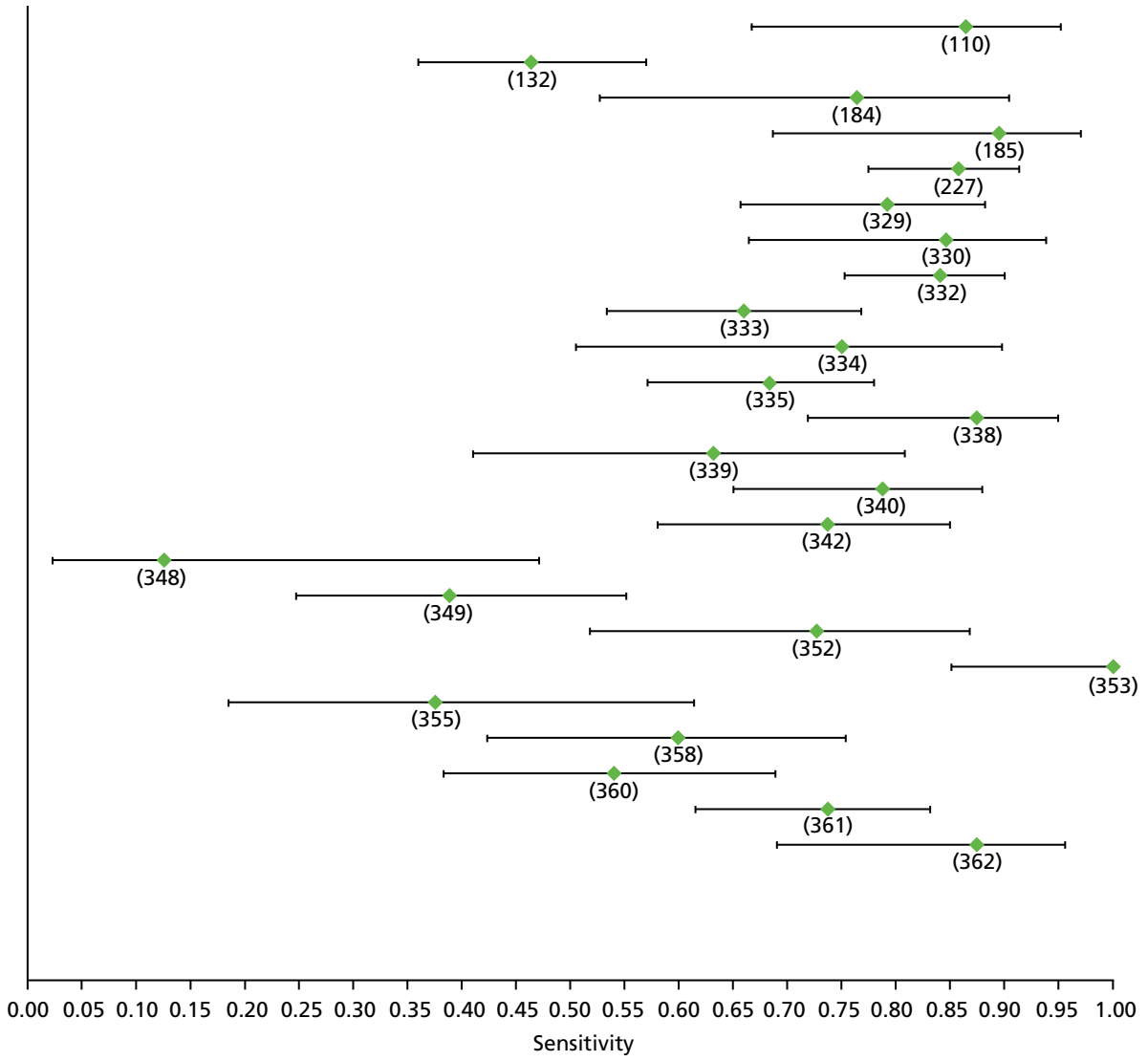

FIGURE 191 Radiology: ultrasound F4 – sensitivity.

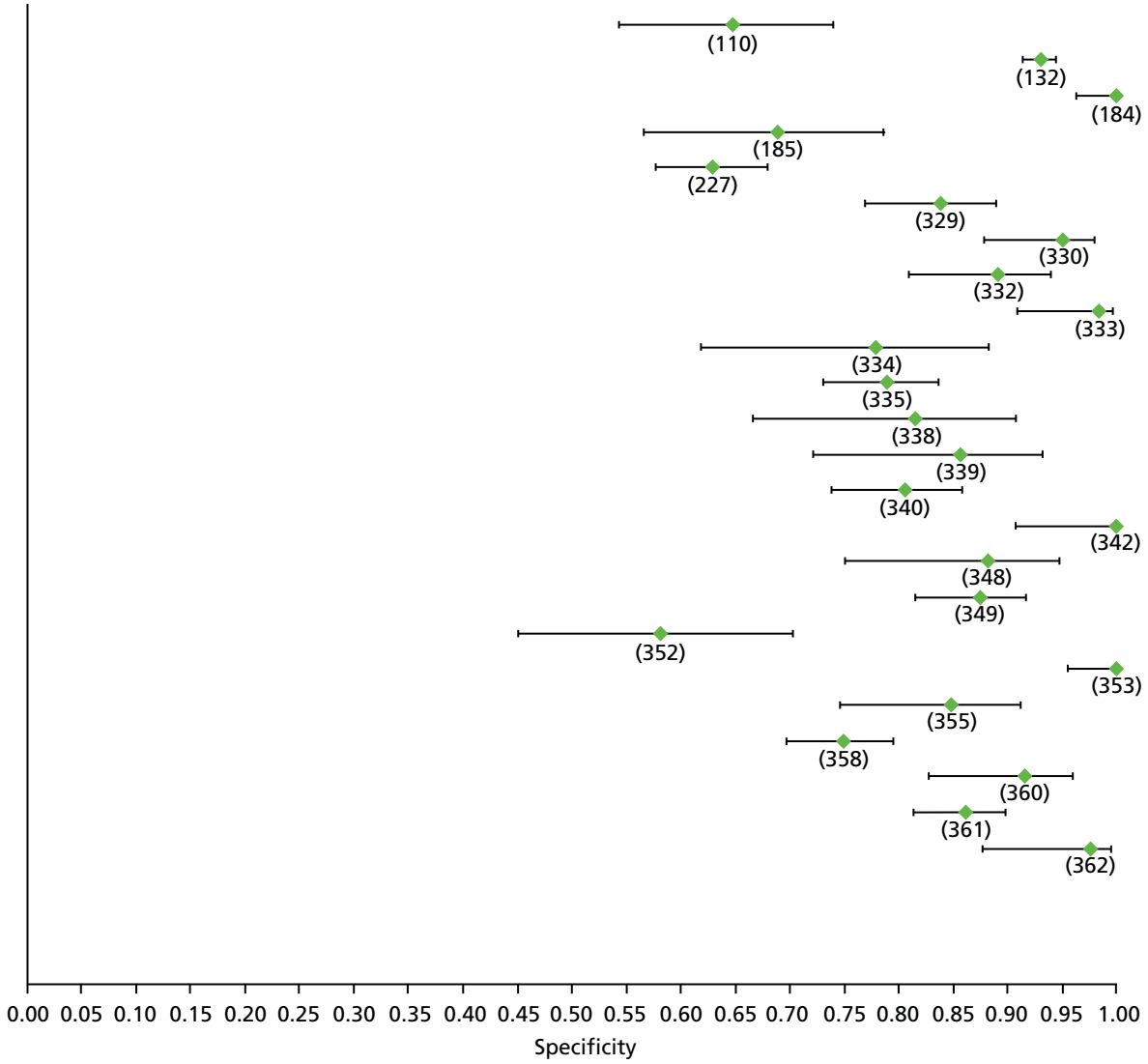

FIGURE 192 Radiology: ultrasound F4 – specificity.

Appendix 6 Summary receiver operating characteristic curves

This appendix presents SROC curves of non-invasive tests across different disease aetiologies and fibrosis stages. We included only non-invasive tests that had available results from at least four studies with convergence by METADAS. The dot represents the relative sensitivity or specificity and the line represents the 95% CIs.

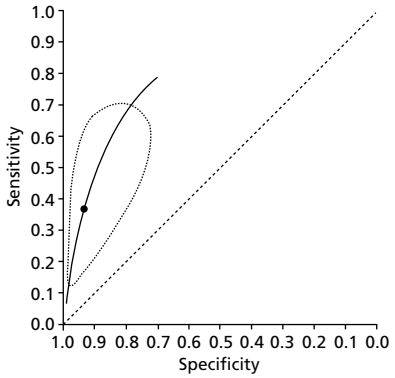

FIGURE 193 APRI (high cut-off) in diagnosing ≥ F2 in patients with chronic HBV.

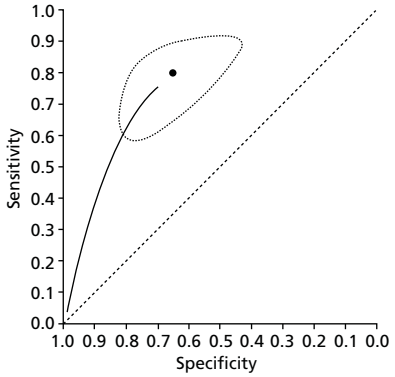

FIGURE 194 APRI (low cut-off) in diagnosing ≥ F2 in patients with chronic HBV.

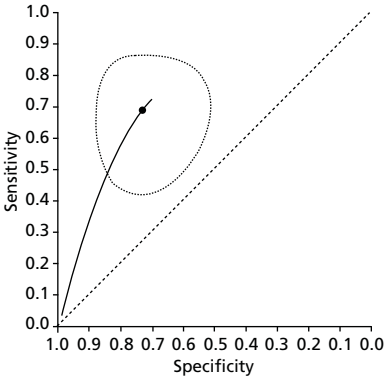

FIGURE 195 FIB-4 (low cut-off) in diagnosing ≥ F2 in patients with chronic HBV.

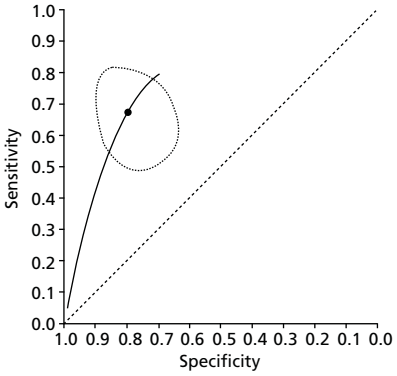

FIGURE 196 Fibrotest in diagnosing ≥ F2 in patients with chronic HBV.

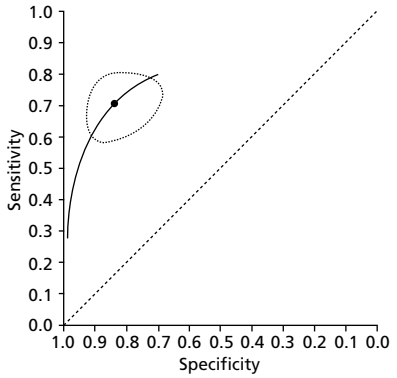

FIGURE 197 Transient elastography (Fibroscan) in diagnosing ≥ F2 in patients with chronic HBV.

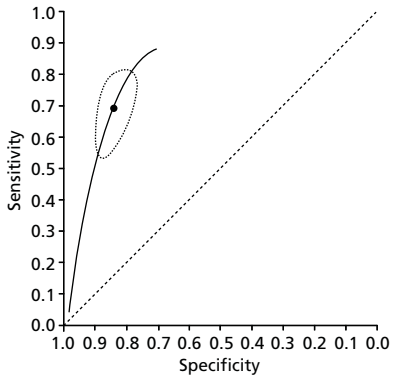

FIGURE 198 Transient elastography (Fibroscan) in diagnosing ≥ F3 in patients with chronic HBV.

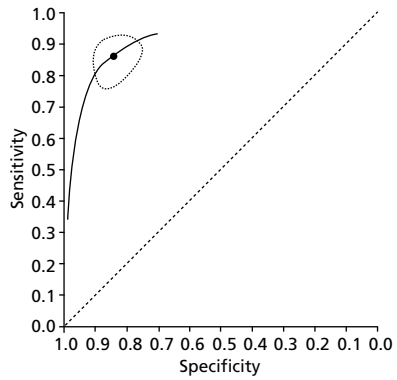

FIGURE 199 Transient elastography (Fibroscan) in diagnosing F4 in patients with chronic HBV.

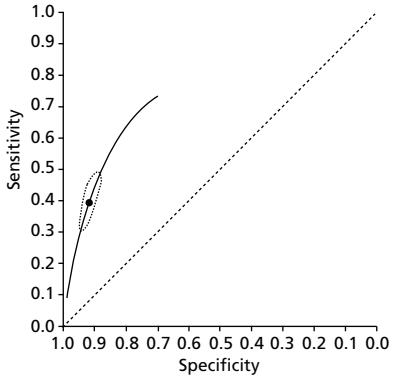

FIGURE 200 APRI (high cut-off) in diagnosing ≥ F2 in patients with chronic HCV.

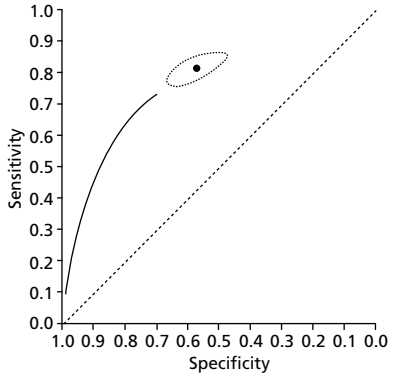

FIGURE 201 APRI (low cut-off) in diagnosing ≥ F2 in patients with chronic HCV.

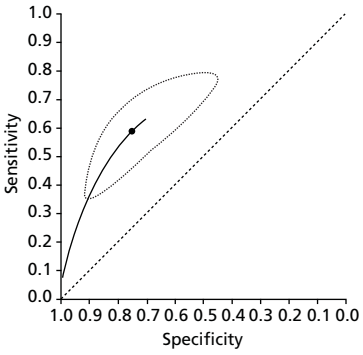

FIGURE 202 FIB-4 (high cut-off) in diagnosing ≥ F2 in patients with chronic HCV.

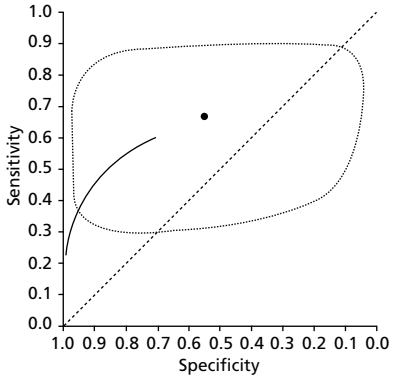

FIGURE 203 Lok's index in diagnosing ≥ F2 in patients with chronic HCV.

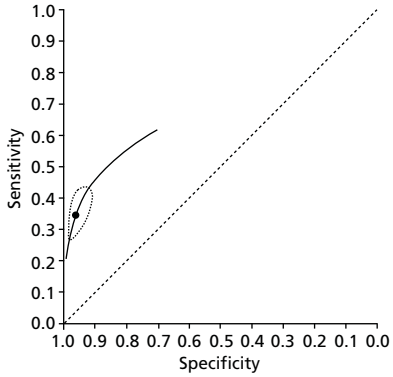

FIGURE 204 Forns index (high cut-off) diagnosing ≥ F2 in patients with chronic HCV.

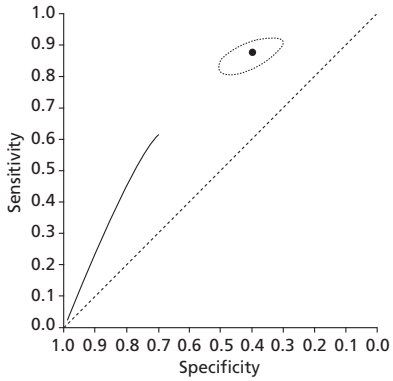

FIGURE 205 Forns index (low cut-off) in diagnosing ≥ F2 in patients with chronic HCV.

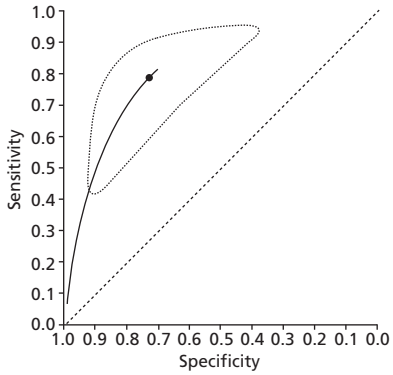

FIGURE 206 Fibrometer in diagnosing ≥ F2 in patients with chronic HCV.

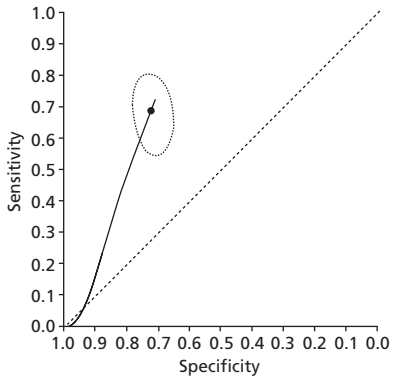

FIGURE 207 Fibrotest in diagnosing ≥ F2 in patients with chronic HCV.

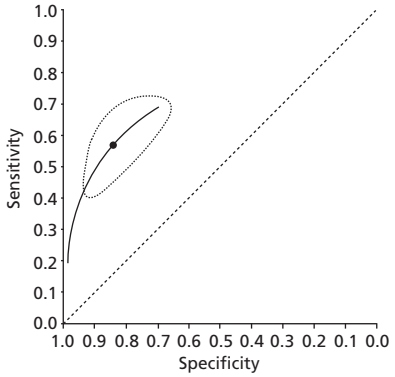

FIGURE 208 Fibrotest (high cut-off) in diagnosing ≥ F2 in patients with chronic HCV.

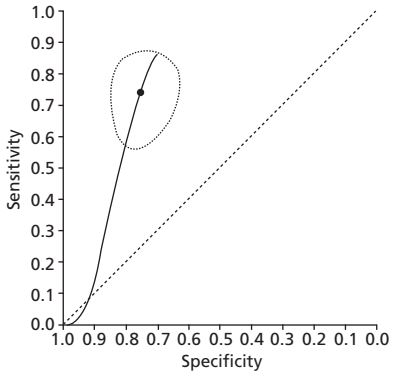

FIGURE 209 Hyaluronic acid in diagnosing ≥ F2 in patients with chronic HCV.

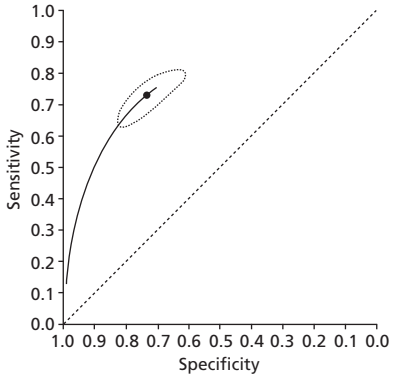

FIGURE 210 Hepascore in diagnosing ≥ F2 in patients with chronic HCV.

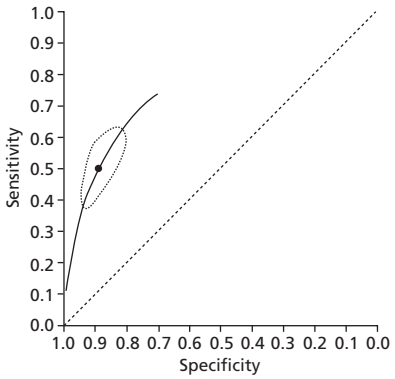

FIGURE 211 Platelet count in diagnosing ≥ F2 in patients with chronic HCV.

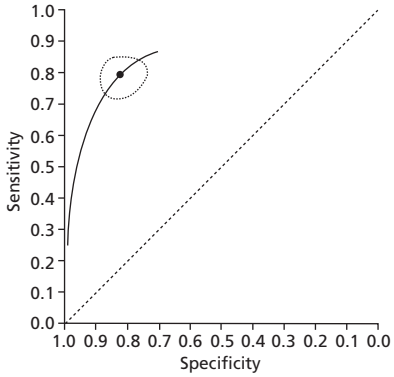

FIGURE 212 Transient elastography (Fibroscan) in diagnosing ≥ F2 in patients with chronic HCV.

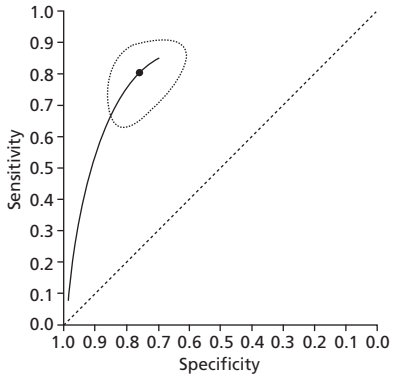

FIGURE 213 Hepascore in diagnosing ≥ F3 in patients with chronic HCV.

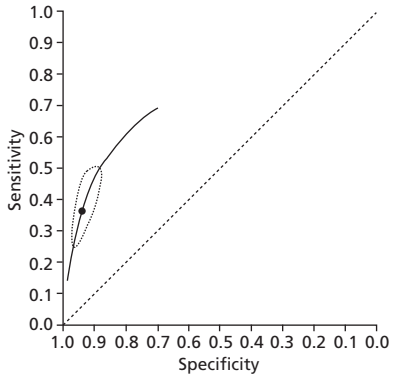

FIGURE 214 FIB-4 (high cut-off) in diagnosing ≥ F3 in patients with chronic HCV.

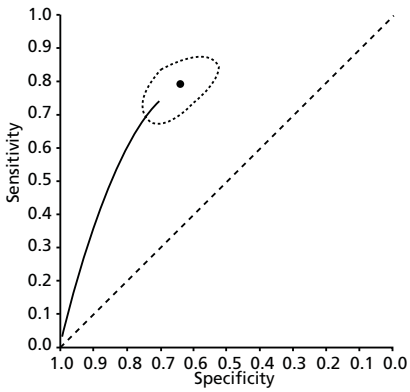

FIGURE 215 FIB-4 (low cut-off) in diagnosing ≥ F3 in patients with chronic HCV.

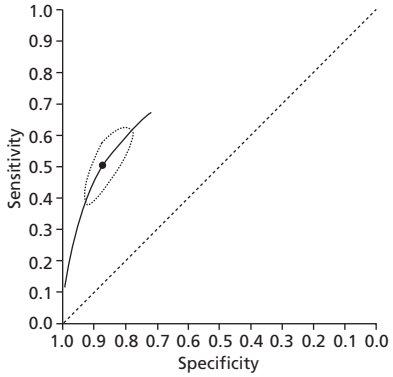

FIGURE 216 APRI (high cut-off) in diagnosing ≥ F3 in patients with chronic HCV.

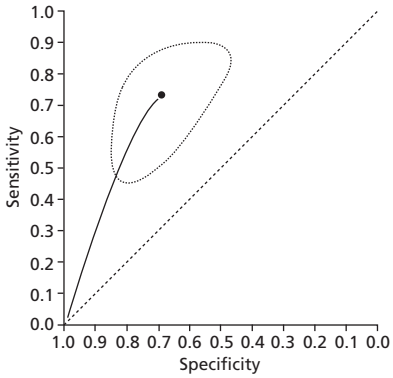

FIGURE 217 Fibrotest in diagnosing ≥ F3 in patients with chronic HCV.

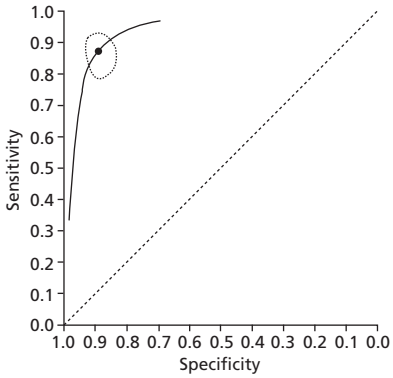

FIGURE 218 Transient elastography (Fibroscan) in diagnosing ≥ F3 in patients with chronic HCV.

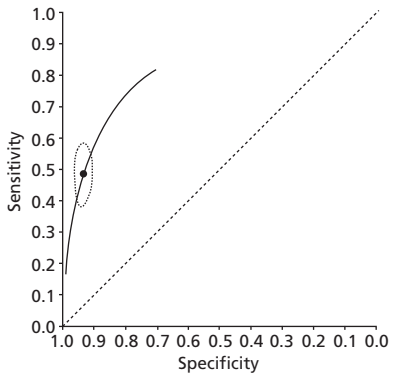

FIGURE 219 APRI (high cut-off) in diagnosing F4 in patients with chronic HCV.

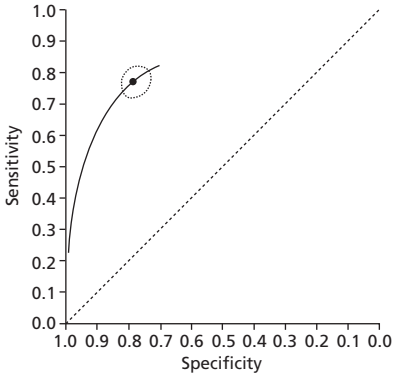

FIGURE 220 APRI (low cut-off) in diagnosing F4 in patients with chronic HCV.

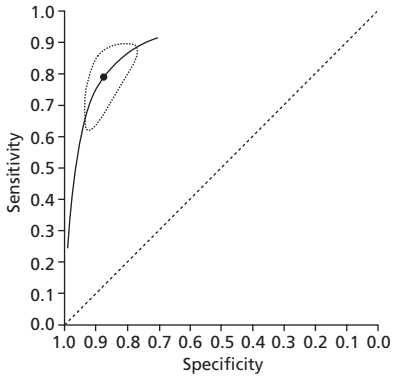

FIGURE 221 AST–ALT ratio in diagnosing F4 in patients with chronic HCV.

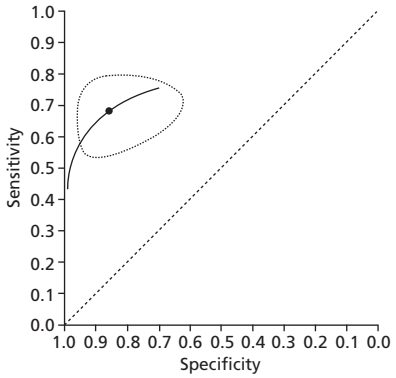

FIGURE 222 Platelet count in diagnosing F4 in patients with chronic HCV.

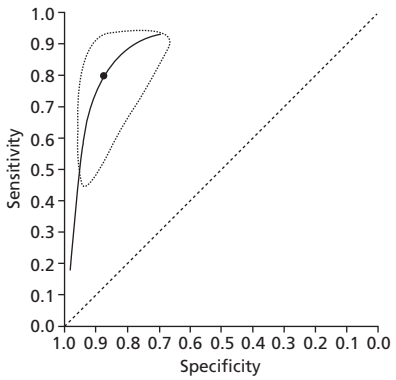

FIGURE 223 Hyaluronic acid in diagnosing F4 in patients with chronic HCV.

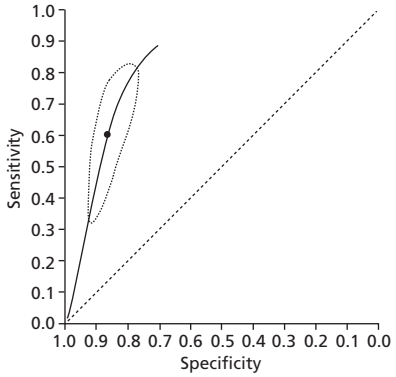

FIGURE 224 Fibrotest in diagnosing F4 in patients with chronic HCV.

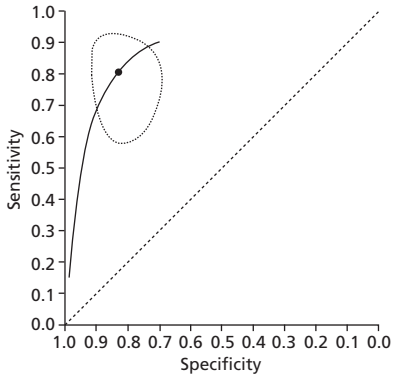

FIGURE 225 Hepascore in diagnosing F4 in patients with chronic HCV.

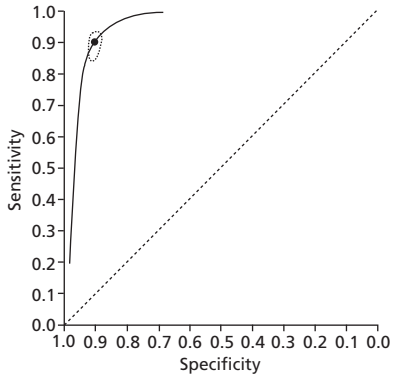

FIGURE 226 Transient elastography (Fibroscan) in diagnosing F4 in patients with chronic HCV.

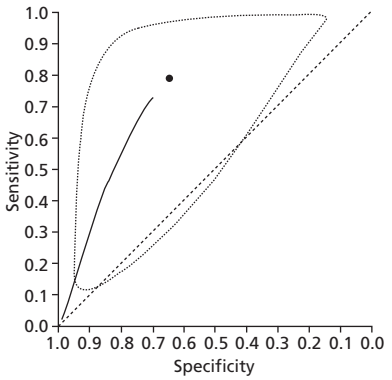

FIGURE 227 Non-alcoholic fatty liver disease fibrosis score (low cut-off) in diagnosing ≥ F2 in patients with NASH.

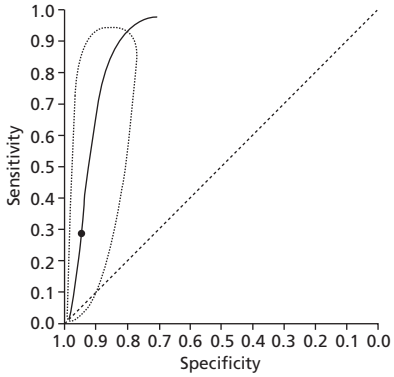

FIGURE 228 Non-alcoholic fatty liver disease fibrosis score (high cut-off) in diagnosing ≥ F2 in patients with NASH.

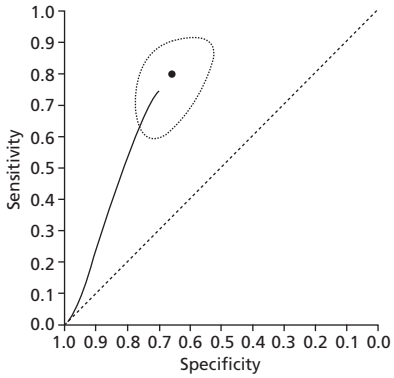

FIGURE 229 Non-alcoholic fatty liver disease fibrosis score (low cut-off) in diagnosing ≥ F3 in patients with NASH.

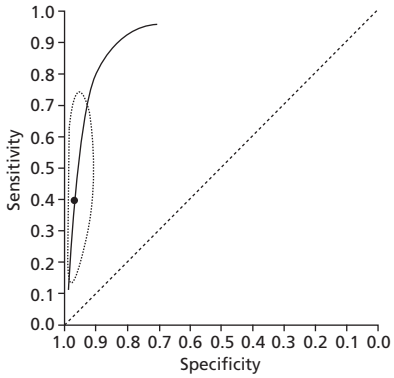

FIGURE 230 Non-alcoholic fatty liver disease fibrosis score (high cut-off) in diagnosing ≥ F3 in patients with NASH.

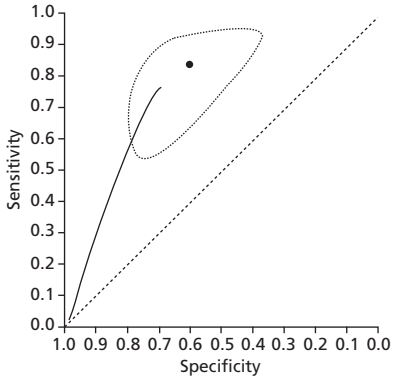

FIGURE 231 BARD score in diagnosing ≥ F3 in patients with NASH.

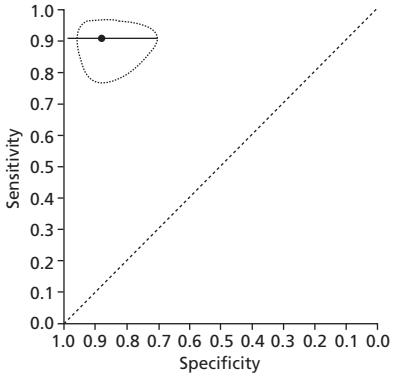

FIGURE 232 Magnetic resonance elastography in diagnosing ≥ F3 in patients irrespective of liver disease aetiology.

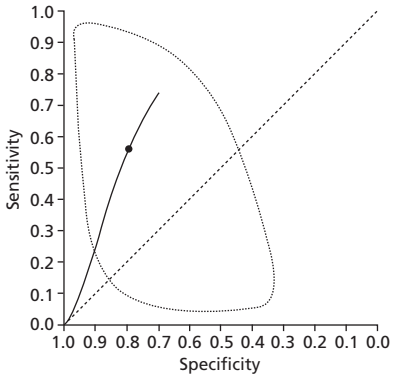

FIGURE 233 Ultrasound in diagnosing ≥ F3 in patients irrespective of liver disease aetiology.

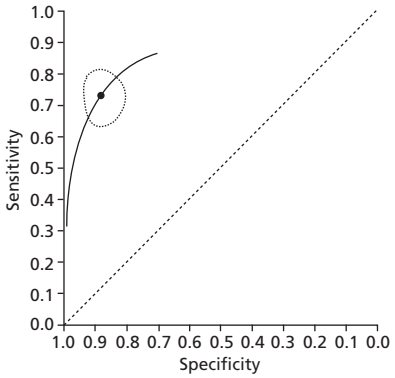

FIGURE 234 Ultrasound in diagnosing F4 in patients irrespective of liver disease aetiology.

Appendix 7 Probability of non-invasive liver tests returning true-positive, false-negative, true-negative or false-positive results

TABLE 72 Hepatitis C: outcome of diagnostic tests (% TP, FN, TN and FP results; average prevalence 53%) (≥ F2 by METAVIR)

Test	TP, %	FN, %	TN, %	FP, %	Summary sensitivity	Summary specificity
Age–Platelet Index	33	20	38	10	0.58	0.70
AST–ALT	23	29	33	14	0.44	0.71
APRI	40	12	38	9	0.77	0.81
APRI (high cut-off)	21	32	44	4	0.39	0.92
APRI (low cut-off)	43	10	27	20	0.82	0.57
ARFI	42	11	42	5	0.79	0.89
CT	37	16	30	17	0.70	0.64
Bordeaux algorithm	46	6	42	5	0.88	0.89
CDS	38	13	23	24	0.66	0.49
ELF	44	9	33	14	0.84	0.70
ELF (high cut-off)	25	28	43	5	0.47	0.90
ELF (low cut-off)	47	5	25	23	0.90	0.52
EOB-MRI	34	19	38	10	0.64	0.79
FIB-4	18	35	41	6	0.34	0.86
FIB-4 (high cut-off)	31	22	35	12	0.59	0.74
FIB-4 (low cut-off)	47	6	20	28	0.89	0.42
Fibrosis Index	38	15	40	7	0.71	0.84
Fibroindex (high cut-off)	13	40	46	1	0.24	0.98
Fibroindex (low cut-off)	44	9	27	20	0.83	0.57
Fibrometer	42	11	34	14	0.79	0.73
Fibropaca algorithm	45	8	43	5	0.85	0.90
FibroQ	41	12	31	16	0.78	0.66
FibroSpect	41	11	33	14	0.78	0.71
Forns index	16	37	18	29	0.30	0.39
Forns index (high cut-off)	18	35	45	2	0.35	0.96
Forns index (low cut-off)	46	7	19	28	0.88	0.40
Fibrosis Probability Index (high cut-off)	22	30	45	2	0.42	0.95
Fibrosis Probability Index (low cut-off)	48	5	21	26	0.91	0.45
Fibrotest	36	17	34	13	0.68	0.72
Fibrotest (high cut-off)	30	23	14	7	0.57	0.85

continued

TABLE 72 Hepatitis C: outcome of diagnostic tests (% TP, FN, TN and FP results; average prevalence 53%) (≥ F2 by METAVIR) (*continued*)

Test	TP, %	FN, %	TN, %	FP, %	Summary sensitivity	Summary specificity
Fibrotest (low cut-off)	48	5	20	28	0.91	0.41
GUCI	34	18	37	10	0.65	0.79
Hyaluronic acid	39	13	36	12	0.75	0.75
Hyaluronic acid (high cut-off)	13	41	43	4	0.23	0.92
Hyaluronic acid (low cut-off)	36	17	29	18	0.67	0.62
Hepascore	37	14	34	13	0.73	0.73
Hepascore (high cut-off)	17	35	44	4	0.33	0.92
King's	44	8	33	14	0.84	0.70
King's (high cut-off)	31	22	37	10	0.58	0.79
King's (low cut-off)	33	20	38	9	0.62	0.81
Lok's index	35	17	26	21	0.67	0.55
MP3	43	9	35	13	0.85	0.73
MR	45	8	43	4	0.85	0.90
PIINP/MMP-1 index	38	18	40	7	0.65	0.85
PIINP	41	12	36	11	0.78	0.76
PLT	26	26	42	5	0.50	0.89
PLT–Spleen	46	6	5	13	0.88	0.73
Pohl Index	3	50	47	1	0.06	0.99
Fibroscan	42	11	39	8	0.79	0.83
Type IV collagen	46	6	35	13	0.88	0.73
YKL-40 (high cut-off)	18	35	38	9	0.33	0.80
YKL-40 (low cut-off)	42	11	16	32	0.80	0.33
US	18	34	40	7	0.35	0.86
US SAPI	39	14	37	10	0.74	0.79
US SAPI (high cut-off)	32	20	45	2	0.61	0.96
US SAPI (low cut-off)	50	3	19	29	0.94	0.39
CEUS	46	6	35	13	0.88	0.73
DW-MRI	41	11	37	10	0.78	0.78
MR elastography	50	3	43	4	0.94	0.92
APRI (combined cut-off)	40	13	41	7	0.75	0.86
ELF (combined cut-off)	43	9	40	8	0.82	0.84
FIB-4 (combined cut-off)	44	9	35	14	0.83	0.73
Fibroindex (combined cut-off)	30	22	45	2	0.58	0.95
Fibrospect (combined cut-off)	53	0	47	0	1.00	1.00
Forns (combined cut-off)	39	14	43	4	0.74	0.91
Fibrotest (combined cut-off)	47	7	35	12	0.87	0.74

TABLE 72 Hepatitis C: outcome of diagnostic tests (% TP, FN, TN and FP results; average prevalence 53%) (≥ F2 by METAVIR) (*continued*)

Test	TP, %	FN, %	TN, %	FP, %	Summary sensitivity	Summary specificity
Hyaluronic acid (combined cut-off)	34	19	34	13	0.64	0.72
Hepascore (combined cut-off)	22	31	43	4	0.42	0.91
YKL-40 (combined cut-off)	33	20	29	18	0.63	0.62
Leroy algorithm	47	5	46	1	0.90	0.98
SAFE algorithm	53	0	38	9	1.00	0.81

CDS, Cirrhosis Discriminant Score; CEUS, contrast-enhanced ultrasound; DW-MRI, diffusion-weighted magnetic resonance imaging; EOB-MRI, (gadolinium-ethoxybenzyl-diethylenetriamine-penta-acetic-acid) enhanced magnetic resonance imaging; FN, false negative; FP, false positive; MMP-1, matrix metalloproteinase-1; MP3, metalloproteinase-3; PLT, platelet; TN, true negative; TP, true positive; US, ultrasound.

TABLE 73 Hepatitis B: outcome of diagnostic tests (% TP, FN, TN and FP results; average prevalence 54%) (≥ F2 by METAVIR)

Test	TP, %	FN, %	TN, %	FP, %	Summary sensitivity	Summary specificity
AAR	31	23	27	19	0.57	0.59
APGA	9	45	8	38	0.17	0.98
Age-Platelet Index	4	51	28	17	0.07	0.62
APRI (combined cut-off)	39	15	41	4	0.73	0.91
APRI (high cut-off)	20	34	43	3	0.37	0.93
APRI (low cut-off)	43	11	30	16	0.80	0.65
ARFI	38	16	31	15	0.71	0.67
FIB-4 (combined cut-off)	9	46	45	1	0.16	0.98
FIB-4 (high cut-off)	5	49	45	1	0.09	0.99
FIB-4 (low cut-off)	37	17	33	12	0.68	0.73
Fibrotest	36	18	37	9	0.66	0.80
Forns (combined cut-off)	14	40	46	0	0.26	1.00
Forns index (high cut-off)	8	46	46	0	0.15	1.00
Forns index (low cut-off)	31	23	35	10	0.58	0.77
GUCI	36	18	44	1	0.67	0.97
Hyaluronic acid	46	9	38	8	0.84	0.83
Hepascore	43	12	34	12	0.79	0.74
Hui index	27	27	42	4	0.50	0.91
PAPAS	39	15	36	10	0.73	0.78
Fibroscan	38	16	38	7	0.71	0.84
US	19	35	39	7	0.35	0.86
CEUS	48	6	33	12	0.88	0.73
DW-MRI	42	12	36	10	0.78	0.78
MR elastography	51	3	42	4	0.94	0.92
US SAPI	40	14	36	10	0.74	0.79
US SAPI (high cut-off)	33	21	44	2	0.61	0.96
US SAPI (low cut-off)	51	3	18	28	0.94	0.39
CT	38	16	29	17	0.70	0.64

AAR, AST–ALT ratio; APGA, AST, platelet count, GGT, α-fetoprotein; CEUS, contrast-enhanced ultrasound; DW-MRI, diffusion-weighted magnetic resonance imaging; FN, false negative; FP, false positive; PAPAS, age, ALP, α-fetoprotein, AST; TN, true negative; TP, true positive; US, ultrasound.

TABLE 74 Alcoholic liver disease: outcome of diagnostic tests (% TP, FN, TN and FP results; average prevalence 37%) (≥ F4 by METAVIR)

Test	TP, %	FN, %	TN, %	FP, %	Summary sensitivity	Summary specificity
APRI (high cut-off)	15	24	22	39	0.40	0.62
Fibrotest (high cut-off)	33	8	3	55	0.91	0.87
Fibrotest (low cut-off)	36	2	0	32	1.00	0.50
	29	7	8	56	0.78	0.89
Fibroscan	32	11	5	53	0.86	0.83

FN, false negative; FP, false positive; TN, true negative; TP, true positive.

TABLE 75 Non-alcoholic fatty liver disease: outcome of diagnostic tests (% TP, FN, TN and FP results; average prevalence 19%) (≥ F3 by Kleiner)

Test	TP, %	FN, %	TN, %	FP, %	Summary sensitivity	Summary specificity
Age–Platelet Index	14	7	61	17	0.66	0.78
APRI	9	13	65	14	0.40	0.82
ARFI	20	2	70	8	0.90	0.90
AST–ALT (high cut off)	10	12	71	7	0.46	0.91
AST–ALT (low cut-off)	17	5	55	23	0.79	0.70
Bard	18	3	48	31	0.84	0.61
Type IV collagen	17	4	63	16	0.79	0.80
ELF	17	4	70	8	0.80	0.90
FIB-4 (high cut-off)	8	13	76	3	0.98	0.97
FIB-4 (low cut-off)	18	3	58	20	0.84	0.74
Fibrotest (high cut-off)	9	13	75	3	0.40	0.96
Fibrotest (low cut-off)	19	3	57	21	0.88	0.73
Fibrotest: Fibroscan	9	13	75	3	0.39	0.96
Hyaluronic acid	19	3	64	13	0.88	0.82
Hepascore	16	5	66	12	0.75	0.84
NAFIC (high cut-off)	18	4	64	14	0.84	0.82
NAFIC (low cut-off)	21	1	53	26	0.96	0.67
NDP: advanced fibrosis	19	3	55	24	0.88	0.70
NFS ELF (high cut-off)	19	3	78	1	0.86	0.99
NFS ELF (low cut-off)	14	7	75	3	0.91	0.96
NFS (high cut-off)	9	13	76	2	0.40	0.97
NFS (low cut-off)	17	4	52	27	0.80	0.66
NFS Fibroscan	2	20	77	1	0.08	0.98
PLT	14	8	60	19	0.63	0.76
Fibroscan (TE)	18	4	66	13	0.82	0.84
MR elastography	20	2	69	9	0.91	0.88
FIB-4 combined cut-off (inconclusive results retested with Fibroscan)	17	5	73	5	0.79	0.93
NFS (combined cut-off) (inconclusive results retested with Fibroscan)	15	6	75	3	0.71	0.96
NAFIC (combined cut-off) (inconclusive results retested with Fibroscan)	21	1	62	16	0.95	0.79
NFS ELF (combined cut-off) (inconclusive results retested with Fibroscan)	20	2	78	1	0.90	0.99
Fibrotest (combined cut-off) (inconclusive results retested with Fibroscan)	18	4	76	3	0.83	0.87

FN, false negative; FP, false positive; NDP, NAFLD diagnostic panel; PLT, platelet; TN, true negative; TP, true positive.

TABLE 76 Cirrhosis: outcome of diagnostic tests (% TP, FN, TN and FP results; average prevalence 20%)

Test	TP, %	FP, %	FN, %	TN, %	Summary sensitivity	Summary specificity
Age–Platelet Index	18	22	2	59	0.88	0.73
APRI	16	24	4	56	0.79	0.70
APRI (combined cut-off)	13	7	7	73	0.64	0.91
APRI (high cut-off)	9	6	11	74	0.45	0.93
APRI (low cut-off)	15	17	5	63	0.75	0.78
ARFI	17	18	3	62	0.84	0.87
AST–ALT ratio	10	10	10	70	0.49	0.87
BARD	10	13	10	67	0.52	0.84
Bordeaux	17	4	3	76	0.87	0.95
CDS	18	26	2	54	0.88	0.67
CDS (high cut-off)	7	1	13	79	0.33	1.00
CDS (low cut-off)	18	8	2	72	0.89	0.90
ELF	19	17	1	63	0.93	0.79
ELF (combined cut-off)	17	13	3	67	0.84	0.84
ELF (high cut-off)	10	8	10	72	0.52	0.90
ELF (low cut-off)	18	38	2	42	0.90	0.53
Fibrosis Index (FI)	8	1	12	79	0.38	1.00
FIB-4	16	18	4	62	0.80	0.78
FIB-4 (combined cut-off)	15	6	5	74	0.75	0.93
FIB-4 (high cut-off)	8	6	12	74	0.42	0.92
FIB-4 (low cut-off)	17	23	3	57	0.84	0.71
Fibroindex	14	7	6	73	0.70	0.91
Fibrometer	14	10	6	70	0.72	0.88
Fibrometer (combined cut-off)	18	2	2	78	0.89	0.97
Fibrometer (high cut-off)	7	2	13	78	0.39	0.98
Fibrometer (low cut-off)	19	23	1	57	0.96	0.71
Fibropaca	15	2	5	78	0.73	0.97
Fontana	16	27	4	53	0.79	0.66
Forns index	20	21	0	59	1.00	0.74
Forns index (combined cut-off)	19	20	1	60	0.97	0.75
Forns index (high cut-off)	13	7	7	73	0.67	0.91
Forns index (low cut-off)	18	50	2	30	0.88	0.37
Fibrotest	12	11	8	70	0.61	0.87
Fibrotest (combined cut-off)	14	4	6	76	0.70	0.95
Fibrotest (high cut-off)	15	5	5	75	0.73	0.94
Fibrotest (low cut-off)	18	28	2	52	0.89	0.65

continued

TABLE 76 Cirrhosis: outcome of diagnostic tests (% TP, FN, TN and FP results; average prevalence 20%) (*continued*)

Test	TP, %	FP, %	FN, %	TN, %	Summary sensitivity	Summary specificity
GUCI	13	11	7	69	0.64	0.86
Hyaluronic acid	16	9	4	71	0.81	0.88
Hepascore	16	13	4	67	0.82	0.84
Hepascore (combined cut-off)	13	1	7	79	0.66	0.99
Hepascore (high cut-off)	8	1	12	79	0.39	0.99
Hepascore (low cut-off)	16	14	4	66	0.80	0.83
King's	15	8	5	72	0.74	0.90
Lok's index (high cut-off)	8	4	12	76	0.40	0.95
Lok's index (low cut-off)	17	27	3	53	0.84	0.66
MR	15	22	5	58	0.75	0.72
PGAA	16	9	4	71	0.78	0.89
PIIINP	14	17	6	63	0.70	0.79
PLT	14	10	6	70	0.72	0.88
PLT–Spleen	17	14	3	66	0.85	0.82
SAFE	15	6	5	74	0.74	0.93
Fibroscan	18	9	2	71	0.89	0.89
Type IV collagen	14	20	6	61	0.71	0.76
CEUS	17	10	3	70	0.84	0.88
DW-MRI	18	22	2	58	0.88	0.73
MR elastography	20	5	0	75	1.00	0.93
MRI	15	16	5	64	0.75	0.80
US	15	9	5	70	0.73	0.88
US SAPI	15	26	6	54	0.73	0.67

CDS, Cirrhosis Discriminant Score; CEUS, contrast-enhanced ultrasound; DW-MRI, diffusion-weighted magnetic resonance imaging; FN, false negative; FP, false positive; PLT, platelet; TN, true negative; TP, true positive; US, ultrasound.

Appendix 8 Probabilistic sensitivity analysis parameters

TABLE 77 Probabilistic sensitivity analysis parameters

Parameter	Base-case value	Standard error	Distribution
HBV health-state costs (£ 2012)			
Mild fibrosis	185	36	Gamma
Moderate fibrosis	959	102	
Compensated cirrhosis	1521	309	
Decompensated cirrhosis	36,194	9967	
HCC	36,194	9967.190	
Liver transplant	64,122	5886	
Post liver transplant	16,321	7933	
HBV utilities			
Mild fibrosis	0.77	0.035	Gamma
Moderate fibrosis	0.66	0.018	
Compensated cirrhosis	0.55	0.032	
Decompensated cirrhosis	0.57	0.076	
HCC	0.57	0.076	
Liver transplant	0.73	0.016	
Post liver transplant	0.78	0.064	
HCV health-state costs (£ 2012)			
Mild fibrosis	185	36	Gamma
Moderate fibrosis	959	102	
Compensated cirrhosis	1521	309	
Decompensated cirrhosis	38,871	9410	
HCC	38,871	9410	
Liver transplant	69,174	7055	
Post liver transplant	4356	862	
HCV utilities			
Mild fibrosis	0.77	0.035	Beta
Moderate fibrosis	0.66	0.018	
Compensated cirrhosis	0.55	0.032	
Decompensated cirrhosis	0.49	0.056	
HCC	0.49	0.056	
Liver transplant	0.51	0.053	
Post liver transplant	0.52	0.061	

continued

TABLE 77 Probabilistic sensitivity analysis parameters (*continued*)

Parameter	Base-case value	Standard error	Distribution
During treatment: mild fibrosis	0.65	0.03	
During treatment: moderate fibrosis	0.55	0.018	
During treatment: cirrhosis	0.44	0.032	
Following successful treatment: mild treatment	0.82	0.04	
Following successful treatment: moderate treatment	0.71	0.05	
Following successful treatment: compensated cirrhosis	0.60	0.04	
ALD			
Probability cirrhosis if continue to drink	20%	0.02	Beta
Abstinence rate after diagnosis of no cirrhosis with liver biopsy	41%	0.041	
Abstinence rate after diagnosis of cirrhosis with liver biopsy	62%	0.0612	
Abstinence rate after diagnosis of no cirrhosis with a NILT	31%	0.031	
Abstinence rate after diagnosis of cirrhosis with a NILT	52%	0.052	
Probability adverse event after liver biopsy	0.72%	0.00072	
Mortality risk after liver biopsy	0.09%	0.00009	
Cirrhosis			
Reduction in mortality with HCC screening	37%	0.037	Beta
HBV and HCV			
Transition probabilities and all-cause mortality rate			Dirichlet
All aetiologies			
Summary sensitivity and specificity estimates			Beta

Appendix 9 Unit costs of non-invasive liver tests and liver biopsy

TABLE 78 Unit costs of NILTs and liver biopsy (2012–13)

Test	Unit cost, £	Source
AAR	0.90	Royal Free, 12 December 2012, personal communication
AP (age–PLT ratio) (API?)	3.50	Royal Free, 12 December 2012, personal communication
APGA	4.95	Royal Free, 22 January 2013, personal communication
APRI	4.05	Royal Free, 12 December 2012, personal communication
ARFI	51.00	As per personal communication Royal Free: costed at same cost as Fibroscan
AST–ALT ratio (AAR)	0.90	Royal Free, 12 December 2012, personal communication
BARD	0.90	Royal Free, 29 May 2013, personal communication
Bordeaux	94.60	Costed as combination strategy (Fibrotest and Fibroscan)
CDS	7.19	Royal Free, 30 January 2013, personal communication
CEUS	113.70	Department of Health reference costs 2011–12[427] (US > 20 minutes) plus contrast (SonoVue) cost: £48.70 (Royal Free personal communication)
CT	105.00	Department of Health reference costs 2011–12[427] (CT with contrast pre and post scan): Diagnostic Imaging Outpatients
DW-MRI	199.00	Cost as per MRI with contrast (as per personal communication, Royal Free, 4 December 2012)
ELF	108.00	Wiktoria Jonasson, Royal Free, 9 May 2012, personal communication: cost of ELF is £90 + VAT
EOB-MRI	199.00	Cost as per MRI with contrast (as per Royal Free)
FIB-4	4.40	Royal Free, 12 December 2012, personal communication
Fibroindex	48.00	Royal Free, 14 January 2013, personal communication
Fibrometer	44.00	Anne Laure Gilles, BioLiveScale, 22 May 2012, personal communication: quoted approximate price €50 (converted to UK cost using OECD indices)
Fibropaca-algorithm	509.89	Tests: APRI, Forns index, liver biopsy, Fibrotest: proportion calculated using Sebastiani *et al*.[31]
FibroQ	7.19	Royal Free, 30 January 2013, personal communication
Fibroscan (TE)	51.00	Department of Health Reference Costs 2011–12:[427] US < 20 minutes. From Diagnostic Imaging, Outpatients (DIAGIM-OP) code RA23Z. As per advice from Royal Free
Fibrosis Index (FI)	4.40	Royal Free, 12 December 2012, personal communication
Fibrospect	35.34	Royal Free, 14 January 2013, personal communication
Fibrotest	43.60	Jean Marie Castille, Directeur General (Biopredictive), 31 May 2012, personal communication: converted to GBP Sterling (OECD PPP and exchange rates data: rate 0.871929)
Fontana F4	31.50	Royal Free, 24 July 2013, personal communication
Forns index	4.26	Royal Free, 22 January 2013, personal communication
FPI high	8.58	Royal Free, 22 January 2013, personal communication

continued

TABLE 78 Unit costs of NILTs and liver biopsy (2012–13) (*continued*)

Test	Unit cost, £	Source
FPI low	8.58	Royal Free, 22 January 2013, personal communication
Fibrotest Fibroscan	94.60	Costed as Fibrotest and Fibroscan
GUCI	6.84	Royal Free, 22 January 2013, personal communication
Hyaluronic acid	8.00	Royal Free, 22 January 2013, personal communication
Hepascore	16.24	Royal Free, 22 January 2013, personal communication
Hui index	4.60	Royal Free, 22 January 2013, personal communication
King's	6.84	Royal Free, 22 January 2013, personal communication
Leroy algorithm	724.74	Tests: APRI, liver biopsy, Fibrotest: proportion calculated using Sebastiani *et al.*[31]
Liver Biopsy	956.61	Stevenson *et al.*[428]
Lok's index (HALT-C)	7.19	Royal Free, 30 January 2013, personal communication
MP3	20.00	Royal Free, 30 January 2013, personal communication
MR elastography	199.00	Department of Health Reference Costs 2011–12:[427] Diagnostic Imaging Outpatients, MRI one area pre and post contrast (code RA23Z)
MRI	199.00	Cost as per MRI with contrast (as per Royal Free)
NAFIC	28.17	Royal Free, 29 May 2013, personal communication
NDP	21.18	Royal Free, 29 May 2013, personal communication
NFS high	4.95	Royal Free, 29 May 2013, personal communication
NFS ELF	112.95	Royal Free, 29 May 2013, personal communication
NFS TE	55.95	Sum of NFS and Fibroscan (TE)
PAPAS	5.15	Royal Free, 22 January 2013, personal communication
PGAA	9.07	Royal Free, 24 July 2013, personal communication
PIIINP/MMP-1 index	48.00	Royal Free, 14 January 2013, personal communication
PIINP	28.00	Royal Free, 12 December 2012, personal communication
PLT	3.50	Royal Free, 29 May 2013, personal communication
PLT–Spleen (SPRI)	54.50	Royal Free, 4 December 2012, personal communication
Pohl Index	4.40	Royal Free, 30 January 2013, personal communication
SAFE	743.22	Tests: APRI, liver biopsy, Fibrotest: proportion calculated using Sebastiani *et al.*[31,187]
TE	51.00	Department of Health Reference Costs 2011–12:[427] US < 20 minutes [from Diagnostic Imaging, Outpatients (DIAGIM-OP) code RA23Z as advised by Royal Free]
Type IV collagen	20.00	Royal Free, 30 January 2013, personal communication
US	51.00	Department of Health Reference Costs 2011–12:[427] US < 20 minutes [from Diagnostic Imaging, Outpatients (DIAGIM-OP) code RA23Z]
US SAPI	65.00	Department of Health Reference Costs 2011–12:[427] US > 20 minutes
YKL-40	20.00	Royal Free, 8 February 2013, personal communication

AAR, AST–ALT ratio; APGA, AST, platelet count, GGT, α-fetoprotein; API, Age-Platelet Index; CDS, Cirrhosis Discriminant Score; CEUS, contrast-enhanced ultrasound; DW-MRI, diffusion-weighted magnetic resonance imaging; EOB-MRI, (gadolinium-ethoxybenzyl-diethylenetriamine-penta-acetic-acid) enhanced magnetic resonance imaging; FPI, Fibrosis Probability Index; HALT-C, Hepatitis C Antiviral Long-Term Treatment Against Cirrhosis; MMP-1, matrix metalloproteinase-1; MP3, metalloproteinase-3; NDP, NAFLD diagnostic panel; OECD, Organisation for Economic Co-operation and Development; PAPAS, age, ALP, α-fetoprotein, AST; PLT, platelet; US, ultrasound; VAT, value-added tax.
The costs of non-invasive tests were sourced through our clinical collaborators (authors on the report) who sourced these from the biochemistry department and the finance department at the Royal Free Hospital.

Appendix 10 Cost-effectiveness acceptability curves

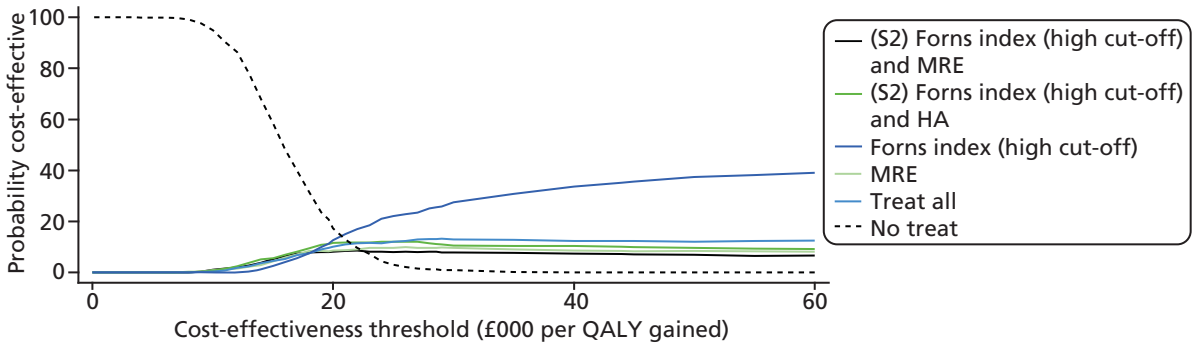

FIGURE 235 Cost-effectiveness acceptability curve for HBeAg-positive. HA, hyaluronic acid; MRE, MR elastography.

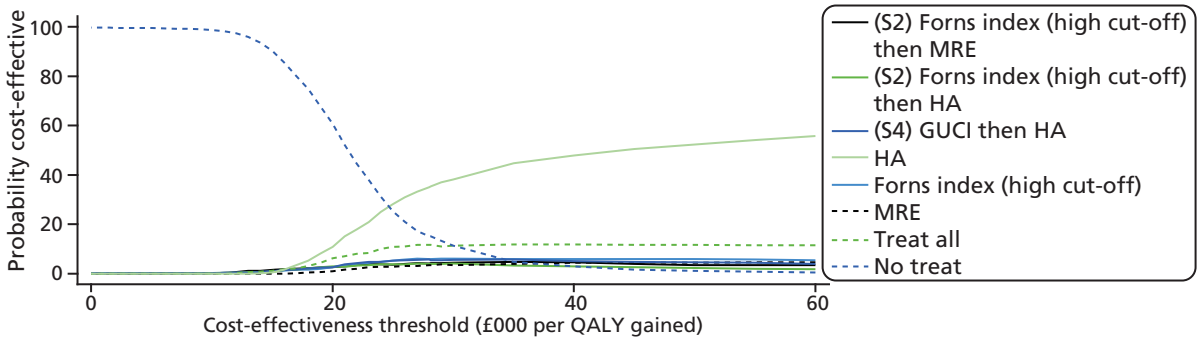

FIGURE 236 Cost-effectiveness acceptability curve for HBeAg-negative. HA, hyaluronic acid; MRE, MR elastography.

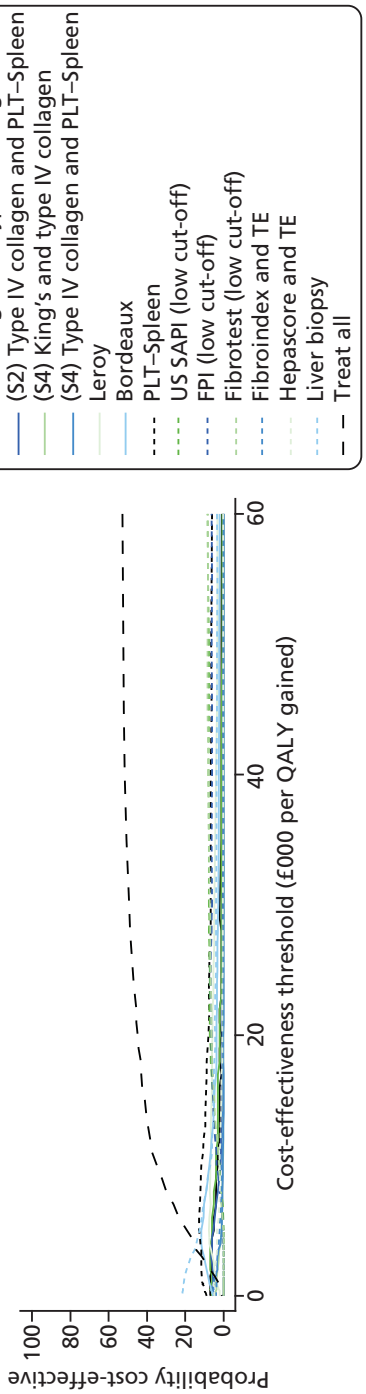

FIGURE 237 Cost-effectiveness acceptability curve for HCV. FPI, Fibrosis Probability Index; US, ultrasound.

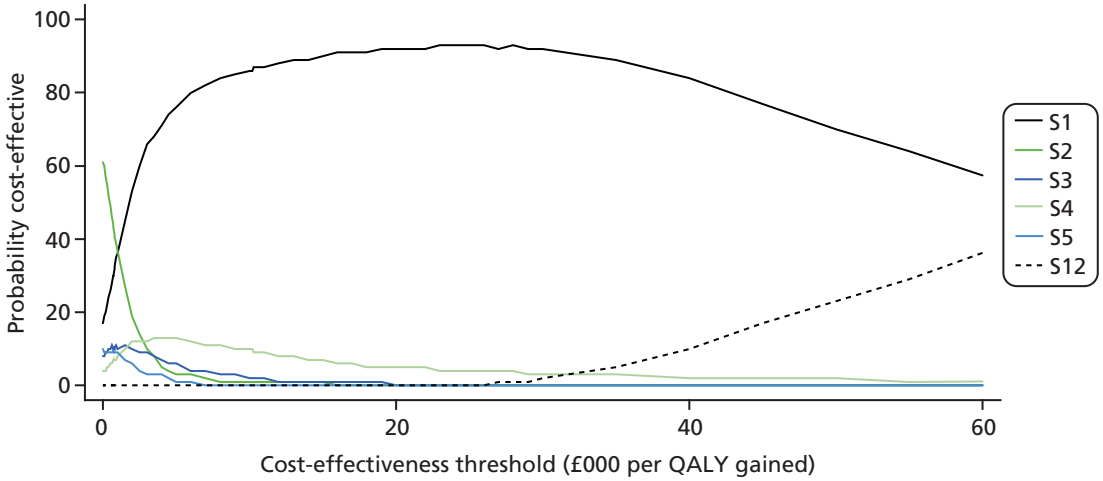

FIGURE 238 Cost-effectiveness acceptability curve for ALD.

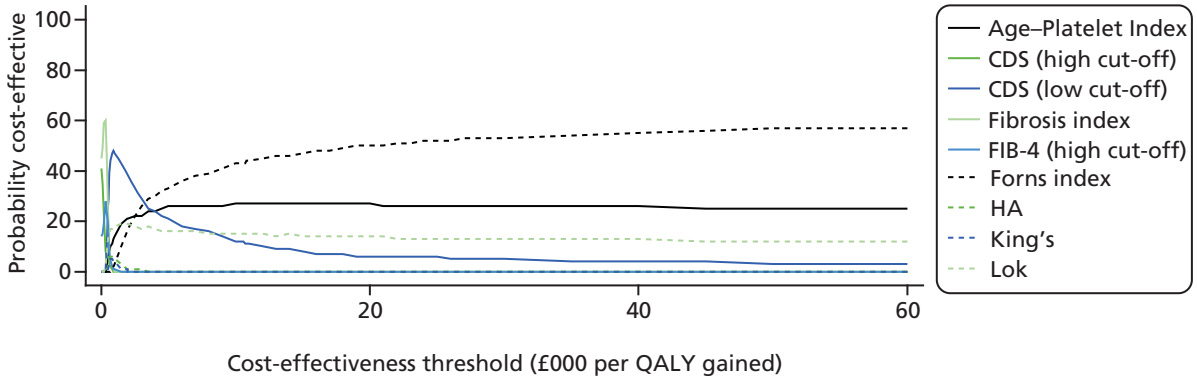

FIGURE 239 Cost-effectiveness acceptability curve for cirrhosis. HA, hyaluronic acid.

Appendix 11 Sensitivity analysis

Results of sensitivity analyses: for clarity and presentation, all testing strategies which were 'dominated' or 'extendedly dominated' are excluded from the tables.

Sensitivity analyses tables for hepatitis B

HBeAg-positive

TABLE 79 HBeAg-positive: sensitivity analysis of tests where bivariate model converged (only)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,831	9.64	–	–	–
APRI (high cut-off)	75,210	11.45	37,380	1.81	20,673
Fibroscan	79,000	11.61	3790	0.16	23,345
Treat all	101,484	12.18	22,484	0.57	39,747

TABLE 80 HBeAg-positive: sensitivity analysis using minimum disease prevalence

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	29,410	12.23	–	–	–
GUCI	57,054	13.73	27,644	1.50	18,486
MR elastography	60,233	13.81	3179	0.09	37,348
Treat all	99,263	14.76	39,030	0.95	41,177

TABLE 81 HBeAg-positive: sensitivity analysis using maximum disease prevalence

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	49,392	6.30	–	–	–
FIB-4 (high cut-off)	108,296	9.59	58,904	3.30	17,871

TABLE 82 HBeAg-positive: sensitivity analysis using 25th quartile value

Test strategy	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
No treat	34,436	10.709	–	–	–
GUCI	67,387	12.432	32,951	1.72	19,121
MR elastography	70,256	12.520	2869	0.09	32,618
Hyaluronic acid	72,298	12.570	2042	0.05	40,722
Treat all	100,896	13.264	28,598	0.69	41,229

TABLE 83 HBeAg-positive: sensitivity analysis using 75th quartile value

Test strategy	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
No treat	41,192	8.68	–	–	–
Forns index (high cut-off)	79,060	10.45	37,868	1.77	21,387
Treat all	103,038	11.35	23,978	0.90	26,718

TABLE 84 HBeAg-positive: sensitivity analysis assuming no treatment benefit for patients who are incorrectly treated while in a mild health state

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
No treat	37,831	9.64	–	–	–
GUCI	74,924	11.50	37,093	1.86	19,934
MR elastography	77,610	11.59	2686	0.08	32,200
Treat all	102,064	11.63	24,454	0.04	550,668

TABLE 85 HBeAg-positive: sensitivity analysis using alternative utility values

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	38,109	9.58	–	–	–
GUCI	75,108	11.48	36,999	1.90	19,476
MR elastography	77,641	11.57	2533	0.10	26,589
US SAPI (low cut-off)	91,418	11.91	13,777	0.34	41,083
Treat all	101,954	12.15	10,535	0.25	42,996

US, ultrasound.

TABLE 86 HBeAg-positive: sensitivity analysis no disutility associated with liver biopsy

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,831	9.64	–	–	–
GUCI	74,925	11.52	37,095	1.88	19,733
MR elastography	77,585	11.64	2660	0.11	23,449
Treat all	101,484	12.18	23,899	0.54	44,019

TABLE 87 HBeAg-positive: sensitivity analysis assuming larger disutility associated with liver biopsy

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,831	9.64	–	–	–
GUCI	74,918	11.52	37,087	1.88	19,727
MR elastography	77,594	11.64	2676	0.11	23,846
Treat all	101,484	12.18	23,890	0.54	43,922

TABLE 88 HBeAg-positive: sensitivity analysis on health state costs (decompensated cirrhosis and HCC)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	19,248	9.64	–	–	–
GUCI	56,277	11.52	37,029	1.88	19,694
MR elastography	59,473	11.66	3196	0.14	22,918
Treat all	84,468	12.24	24,994	0.58	43,095

TABLE 89 HBeAg-positive: sensitivity analysis – retest cost set to APRI (all costs set to APRI)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
No treat	37,831	9.64	–	–	–
GUCI	74,915	11.52	37,084	1.88	19,725
MR elastography	77,013	11.64	2098	0.12	17,810
Treat all	101,484	12.18	24,471	0.54	45,456

TABLE 90 HBeAg-positive: sensitivity analysis on sensitivity and specificity of retest [set to APRI (low cut-off)]

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,831	9.64	–	–	–
MR elastography	95,175	12.02	57,345	2.38	24,077
Treat all	101,484	12.18	6309	0.15	40,836

TABLE 91 HBeAg-positive: sensitivity analysis on sensitivity and specificity of retest (set to Fibrotest)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,831	9.64	–	–	–
MR elastography	90,312	11.94	52,481	2.29	22,868
Treat all	101,484	12.18	11,172	0.24	46,317

TABLE 92 HBeAg-positive: sensitivity analysis on sensitivity and specificity of retest (set to Fibroscan)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,831	9.64	–	–	–
MR elastography	91,679	11.97	53,848	2.32	23,182
Treat all	101,484	12.18	9805	0.21	45,952

TABLE 93 HBeAg-positive: sensitivity analysis on selection of tests for second stage of analysis (sequential testing)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,831	9.64	–	–	–
(S2) Hyaluronic acid and US SAPI (high cut-off)	75,246	11.52	37,414	1.88	19,903
(S2) Hyaluronic acid and US SAPI (low cut-off)	77,560	11.62	2315	0.09	24,883
Hyaluronic acid	79,005	11.66	1444	0.04	34,084
Treat all	101,484	12.18	22,480	0.52	43,150

S2, strategy 2; US, ultrasound.

TABLE 94 HBeAg-positive: sensitivity analysis on length of successful response rate to treatment with peginterferon (reduced to 15 years – and reintroduces risk of progression to more severe health states)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,831	9.64	–	–	–
GUCI	75,271	11.52	37,440	1.88	19,928
MR elastography	77,930	11.64	2659	0.12	22,601
Treat all	101,484	12.18	23,554	0.54	43,636

HBeAg-negative

TABLE 95 HBeAg-negative: sensitivity analysis of tests where bivariate model converged (only)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,579	8.83	–	–	–
Treat all	96,525	10.92	58,947	2.09	28,137

TABLE 96 HBeAg-negative: sensitivity analysis using minimum disease prevalence

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	28,696	11.23	–	–	–
Treat all	120,532	15.24	91,836	4.02	22,871

TABLE 97 HBeAg-negative: sensitivity analysis using the maximum disease prevalence

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	49,584	5.68	–	–	–
MR elastography	96,726	7.87	47,142	2.18	21,581
US SAPI (low cut-off)	99,174	7.93	2448	0.07	36,897

TABLE 98 HBeAg-negative: sensitivity analysis using 25th quartile value

Test strategy	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
No treat	33,514	9.837	–	–	–
Treat all	94,495	11.842	60,981	2.01	30,413

TABLE 99 HBeAg-negative: sensitivity analysis using 75th quartile value

Test strategy	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
No treat	40,908	7.927	–	–	–
Treat all	97,007	9.971	56,099	2.04	27,447

TABLE 100 HBeAg-negative: sensitivity analysis – all retest costs set to cost of APRI

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
No treat	37,579	8.83	–	–	–
Treat all	96,525	10.92	58,947	2.09	28,137

TABLE 101 HBeAg-negative: sensitivity analysis assuming no treatment benefit for patients who are incorrectly treated while in a mild health state

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,579	8.83	–	–	–
MR elastography	71,699	9.87	34,120	1.04	32,694
Treat all	95,989	10.32	24,290	0.45	53,660

TABLE 102 HBeAg-negative: sensitivity analysis to utility values

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,589	8.76	–	–	–
Treat all	96,314	10.82	58,724	2.05	28,603

TABLE 103 HBeAg-negative: sensitivity analysis assuming no disutility associated with liver biopsy

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no none	37,579	8.83	–	–	–
Treat all	96,525	10.92	58,947	2.09	28,137

TABLE 104 HBeAg-negative: sensitivity analysis of greater disutility associated with liver biopsy (increased to 0.3)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,579	8.83	–	–	–
Treat all	96,525	10.92	58,947	2.09	28,137

TABLE 105 HBeAg-negative: sensitivity analysis on health state costs (decompensated cirrhosis and HCC)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	18,894	8.83	–	–	–
Treat all	77,894	10.90	59,000	2.07	28,456

TABLE 106 HBeAg-negative: sensitivity analysis to amend sensitivity and specificity of retest [set to APRI (low cut-off)]

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,579	8.83	–	–	–
MR elastography	89,722	10.72	52,144	1.90	27,476
Treat all	96,525	10.92	6803	0.20	34,501

TABLE 107 HBeAg-negative: sensitivity analysis to amend sensitivity and specificity of retest (set to Fibrotest)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,579	8.83	–	–	–
MR elastography	86,285	10.64	48,706	1.82	26,831
Treat all	96,525	10.92	10,241	0.28	36,615

TABLE 108 HBeAg-negative: sensitivity analysis to amend sensitivity and specificity of retest (set to Fibroscan)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	37,579	8.83	–	–	–
MR elastography	85,041	10.63	47,462	1.81	26,260
Treat all	96,525	10.92	11,484	0.29	39,934

TABLE 109 HBeAg-negative: sensitivity analysis to amend selection of tests for second stage of analysis (sequential testing)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
No treat	37,579	8.83	–	–	–
Treat all	96,525	10.92	58,947	2.09	28,137

TABLE 110 HBeAg-negative: sensitivity analysis to reduce length of successful response rate to treatment with peginterferon (reduced to 15 years – and reintroduces risk of progression to more severe health states)

Test	Cost £	QALY	Incremental cost, £	Incremental QALY	ICER £
Treat no one	37,579	8.83	–	–	–
Treat all	96,525	10.92	58,947	2.09	28,137

TABLE 111 HBeAg-negative: sensitivity analysis to amend sex distribution and starting age changed to reflect HBeAg-positive model

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	38,737	9.26	–	–	–
GUCI	73,824	10.78	35,088	1.52	23,065
MR elastography	76,631	10.89	2807	0.11	25,547
US SAPI (low cut of)	90,200	11.17	13,569	0.28	48,775
Treat all	99,905	11.36	9705	0.20	49,720

US, ultrasound.

Sensitivity analysis tables for hepatitis C

TABLE 112 Hepatitis C: sensitivity analysis of tests where bivariate model converged (only)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Forns index (high cut-off)	47,426	14.119	–	–	–
Fibroscan	47,448	14.278	22	0.16	141
Treat all	51,241	14.732	3793	0.45	8370

TABLE 113 Hepatitis C: sensitivity analysis using minimum disease prevalence

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Pohl	29,537	15.26	–	–	–
US SAPI (high cut off)	29,638	15.33	100	0.07	1424
MR elastography	29,974	15.39	337	0.06	5939
CEUS	30,269	15.42	294	0.03	9149
Treat all	38,159	16.18	7890	0.75	10,457

CEUS, contrast-enhanced ultrasound; US, ultrasound.

TABLE 114 Hepatitis C: sensitivity analysis using maximum disease prevalence

Test strategy	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
(S4) Type IV collagen and MR elastography	70,627	12.40	–	–	–
MR elastography	70,710	12.42	84	0.02	3893
Treat all	72,058	12.64	1348	0.22	6155

S4, strategy 4.

TABLE 115 Hepatitis C: sensitivity analysis utility values set equal to Shepherd *et al.*[392]

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	47,212	19.52	–	–	–
Treat all	51,488	19.92	4276	0	10,813

TABLE 116 Hepatitis C: sensitivity analysis increasing utility values by 0.1

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	47,138	16.34	–	–	–
Type IV collagen	47,792	16.41	654	0.08	8615
Treat all	51,369	16.80	3577	0.39	9181

TABLE 117 Hepatitis C: sensitivity analysis allowing for small risk of progression to decompensated cirrhosis and HCC after SVR in compensated cirrhosis health state

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	49,455	14.16	–	–	–
Treat all	53,211	14.57	3757	0.41	9112

TABLE 118 Hepatitis C: sensitivity analysis with specific SVR rates for mild, moderate and cirrhotic health states

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	49,668	14.12	–	–	–
Treat all	52,924	14.62	3256	0.50	6517

TABLE 119 Hepatitis C: sensitivity analysis to change test and retest costs [all set as a NILT (APRI serum marker)]

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	46,603	14.26	–	–	–
Treat all	51,241	14.73	4638	0.47	9938

TABLE 120 Hepatitis C: sensitivity analysis setting sensitivity and specificity of retest to APRI

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat all	51,241	14.73	–	–	–

TABLE 121 Hepatitis C: sensitivity analysis setting sensitivity and specificity of retest to Fibrotest

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat all	51,241	14.73	–	–	–

TABLE 122 Hepatitis C: sensitivity analysis setting sensitivity and specificity of retest to Fibroscan

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat all	51,241	14.73	–	–	–

TABLE 123 Hepatitis C: sensitivity analysis on selection of tests for second-stage analysis (most effective and least costly tests)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
(S4) PIINP/MMP and MR elastography	46,772	14.19	–	–	–
MR elastography	46,891	14.27	119	0.08	1452
Treat all	51,241	14.73	4350	0.46	9516

MM, matrix metalloproteinase; S4, strategy 4.

Hepatitis C: sensitivity analysis of genotype distribution used in analysis

TABLE 124 Hepatitis C: sensitivity analysis of genotype distribution used in analysis

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	43,423	14.33	–	–	–
Treat all	46,182	14.76	1898	0.44	4352

TABLE 125 Hepatitis C: sensitivity analysis assuming no disutility for liver biopsy applied

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	46,851	14.27	–	–	–
Treat all	51,241	14.73	4391	0.46	9468

TABLE 126 Hepatitis C: sensitivity analysis increasing disutility for liver biopsy to 0.3

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	46,875	14.27	–	–	–
Treat all	51,241	14.73	4366	0.457	9546

TABLE 127 Hepatitis C: sensitivity analysis of disutility associated with adverse events from telaprevir and boceprevir

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	47,117	14.26	–	–	–
Type IV collagen	47,752	14.33	635	0.07	8824
Treat all	51,350	14.70	3598	0.37	9704

TABLE 128 Hepatitis C: reduction in effectiveness of treatment (decreasing SVR rate by 80%)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
US SAPI (high cut-off)	47,259.70	14.16	–	–	–
MR elastography	47,281.65	14.22	22	0.07	336

US, ultrasound

TABLE 129 Reduction in effectiveness of treatment (decreasing SVR rate by 70%)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	47,283	14.23	–	–	–

TABLE 130 Hepatitis C: reduction in effectiveness of treatment (decreasing SVR rate by 60%)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	47,142	14.23	–	–	–
Treat all	55,028	14.24	7886	0.011	723,503

TABLE 131 Hepatitis C: reduction in effectiveness of treatment (decreasing SVR rate by 50%)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	47,158	14.24	–	–	–
Treat all	54,153	14.32	6995	0.08	83,697

TABLE 132 Reduction in effectiveness of treatment (measured using SVR rate by 40%)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	47,125	14.25	–	–	–
Treat all	53,800	14.40	6675	0.15	45,877

TABLE 133 Hepatitis C: reduction in effectiveness of treatment (decreasing SVR rate by 30%)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	47,071	14.26	–	–	–
US SAPI (low cut-off)	50,623	14.38	3552	0.12	28,435
Treat all	53,193	14.47	2571	0.09	29,740

US, ultrasound.

TABLE 134 Hepatitis C: reduction in effectiveness of treatment (decreasing SVR rate by 20%)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	46,951	14.26	–	–	–
Treat all	52,545	14.56	5594	0.30	18,830

TABLE 135 Reduction in effectiveness of treatment (decreasing SVR rate by 10%)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	46,902	14.27	–	–	–
Treat all	51,241	14.73	4339	0.462	9384

TABLE 136 Hepatitis C: sensitivity analysis increasing effectiveness of treatment (increase in SVR rate) and price (£20,000)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
MR elastography	43,631	14.71	–	–	–
TE	44,188	14.72	557	0.01	56,413
Treat all	49,207	15.27	5576	0.56	10,009

TABLE 137 Hepatitis C: sensitivity analysis increasing effectiveness of treatment (increase in SVR rate) and price (£40,000)

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Treat no one	54,940	12.46	–	–	–
Pohl	55,973	14.46	1033	2.00	517
US SAPI (high cut-off)	56,931	14.62	314	0.07	6083
MR elastography	57,913	14.73	982	0.11	9189
Treat all	69,108	15.26	11,195	0.53	21,174

US, ultrasound.

Sensitivity analyses tables for alcoholic liver disease

TABLE 138 Alcoholic liver disease sensitivity analysis: probability of progressing to cirrhosis if continue to drink with diagnosis of no cirrhosis (probability set to 10%)

Strategy	Tests	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Strategy 2	APRI (high cut-off) and liver biopsy	15,787	9.06	–	–	–
Strategy 5	PGAA and liver biopsy	15,908	9.35	122	0.29	416
Strategy 3	Fibrotest (high cut-off) and liver biopsy	16,016	9.46	108	0.10	1041
Strategy 4	Fibrotest (low cut-off) and liver biopsy	16,187	9.53	171	0.08	2271
Strategy 1	Liver biopsy	16,321	9.56	134	0.03	5199
Strategy 12	All patients treated as having cirrhosis (receive HCC screening)	30,574	9.65	14,253	0.10	146,491

TABLE 139 Alcoholic liver disease sensitivity analysis: probability of progressing to cirrhosis if continue to drink with diagnosis of no cirrhosis (probability set to 30%)

Strategy	Tests	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Strategy 2	APRI (high cut-off) and liver biopsy	19,199	8.52	–	–	–
Strategy 1	Liver biopsy	19,409	9.08	210	0.56	377
Strategy 12	All patients treated as having cirrhosis (receive HCC screening)	31,377	9.35	11,967	0.27	44,302

TABLE 140 Alcoholic liver disease sensitivity analysis: probability of progressing to cirrhosis if continue to drink with diagnosis of no cirrhosis (probability set to 40%)

Strategy	Tests	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Strategy 1	Liver biopsy	20,971	8.82	–	–	–
Strategy 12	All patients treated as having cirrhosis (receive HCC screening)	31,672	9.20	10,701	0.37	28,747

TABLE 141 Alcoholic liver disease sensitivity analysis: probability of progressing to cirrhosis if continue to drink with diagnosis of no cirrhosis (probability set to 50%)

Strategy	Tests	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Strategy 1	Liver biopsy	22,352	8.58	–	–	–
Strategy 12	All patients treated as having cirrhosis (receive HCC screening)	31,944	9.04	9591	0.46	20,835

TABLE 142 Alcoholic liver disease sensitivity analysis: probability of progressing to cirrhosis if continue to drink with diagnosis of no cirrhosis (probability set to 60%)

Strategy	Tests	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Strategy 1	Liver biopsy	24,107	8.31	–	–	–
Strategy 12	All patients treated as having cirrhosis (receive HCC screening)	32,461	8.85	8355	0.55	15,232

TABLE 143 Alcoholic liver disease sensitivity analysis: adverse events (associated with liver biopsy) decreased to 0.05% from 0.72%

Strategy	Tests	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Strategy 2	APRI (high cut-off) and liver biopsy	17,480	8.80	–	–	–
Strategy 5	PGAA and liver biopsy	17,703	9.08	223	0.28	795
Strategy 1	Liver biopsy	17,913	9.32	209	0.24	854
Strategy 12	All patients treated as having cirrhosis (receive HCC screening)	30,984	9.51	13,071	0.18	71,616

TABLE 144 Alcoholic liver disease sensitivity analysis: adverse events (associated with liver biopsy) increased from 0.72% to 0.90%

Strategy	Tests	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Strategy 2	APRI (high cut-off) and liver biopsy	17,486	8.79	–	–	–
Strategy 5	PGAA and liver biopsy	17,674	9.06	188	0.27	693
Strategy 1	Liver biopsy only	17,882	9.29	208	0.24	884
Strategy 12	All treated as having cirrhosis	30,954	9.48	13,072	0.19	69,697

TABLE 145 Alcoholic liver disease sensitivity analysis: mortality rate (associated with liver biopsy) reduced to 0.05%

Strategy	Tests	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Strategy 2	APRI (high cut-off) and liver biopsy	17,480	8.80	–	–	–
Strategy 5	PGAA and liver biopsy	17,703	9.08	223	0.28	795
Strategy 1	Liver biopsy	17,913	9.32	209	0.24	854
Strategy 12	All patients treated as having cirrhosis (receive HCC screening)	30,984	9.51	13,071	0.18	71,616

TABLE 146 Alcoholic liver disease sensitivity analysis: mortality rate (associated with liver biopsy) increased to 0.20%

Strategy	Tests	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Strategy 2	APRI (high cut-off) and liver biopsy	17,488	8.78	–	–	–
Strategy 1	Liver biopsy	17,840	9.29	169	0.23	721
Strategy 12	All patients treated as having cirrhosis (receive HCC screening)	30,973	9.49	13,133	0.20	65,679

Cirrhosis sensitivity analysis

TABLE 147 Cirrhosis sensitivity analysis: increased benefit for patients diagnosed as false positive who receive screening for HCC

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Fibrosis Index	24,908	2.09	–	–	–
CDS (low cut-off)	24,946	2.18	39	0.09	439
Forns index	24,974	2.22	28	0.04	653
Forns index (low cut-off)	25,032	2.27	58	0.05	1106

CDS, Cirrhosis Discriminant Score.

TABLE 148 Cirrhosis sensitivity analysis: screening cost for α-fetoprotein testing removed from analysis

Test	Cost, £	QALY	Incremental cost, £	Incremental QALY	ICER, £
Fibrosis Index	24,908	2.09	–	–	–
CDS (low cut-off)	24,946	2.16	38	0.07	526
Forns index	24,974	2.18	28	0.01	1884

CDS, Cirrhosis Discriminant Score.

Appendix 12 Scatterplots of cost-effectiveness analysis results

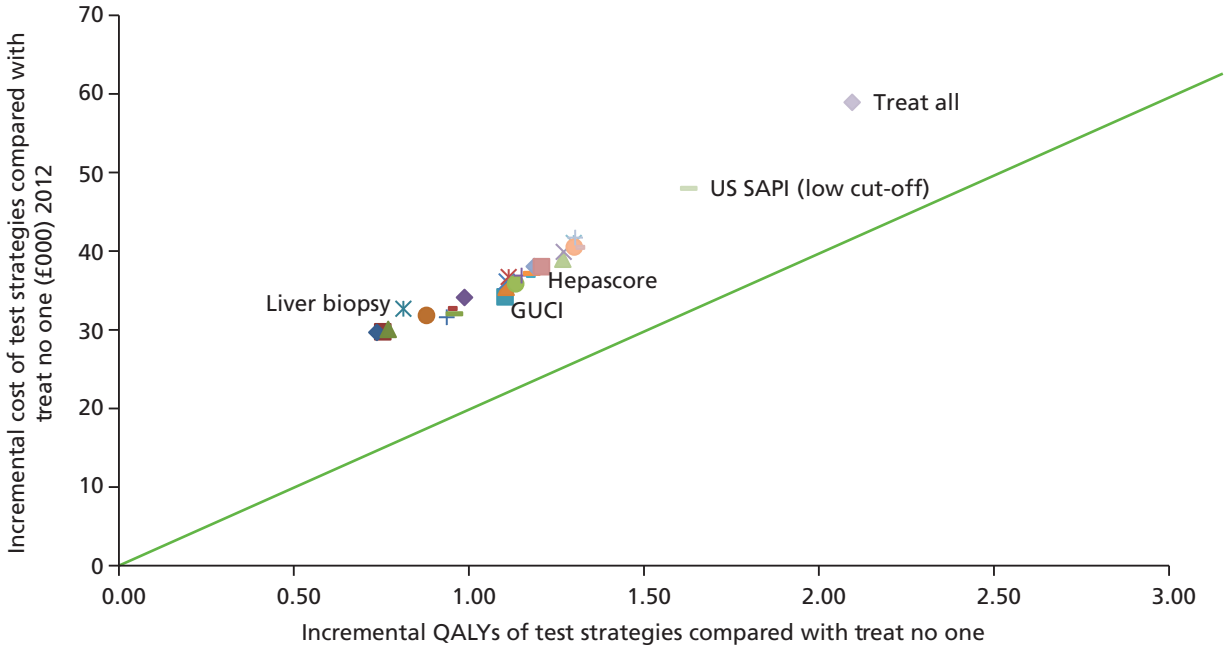

FIGURE 240 Scatterplot of results: HBeAg (negative) – first stage of the analysis (all tests strategies compared with 'treat no one'). US, ultrasound.

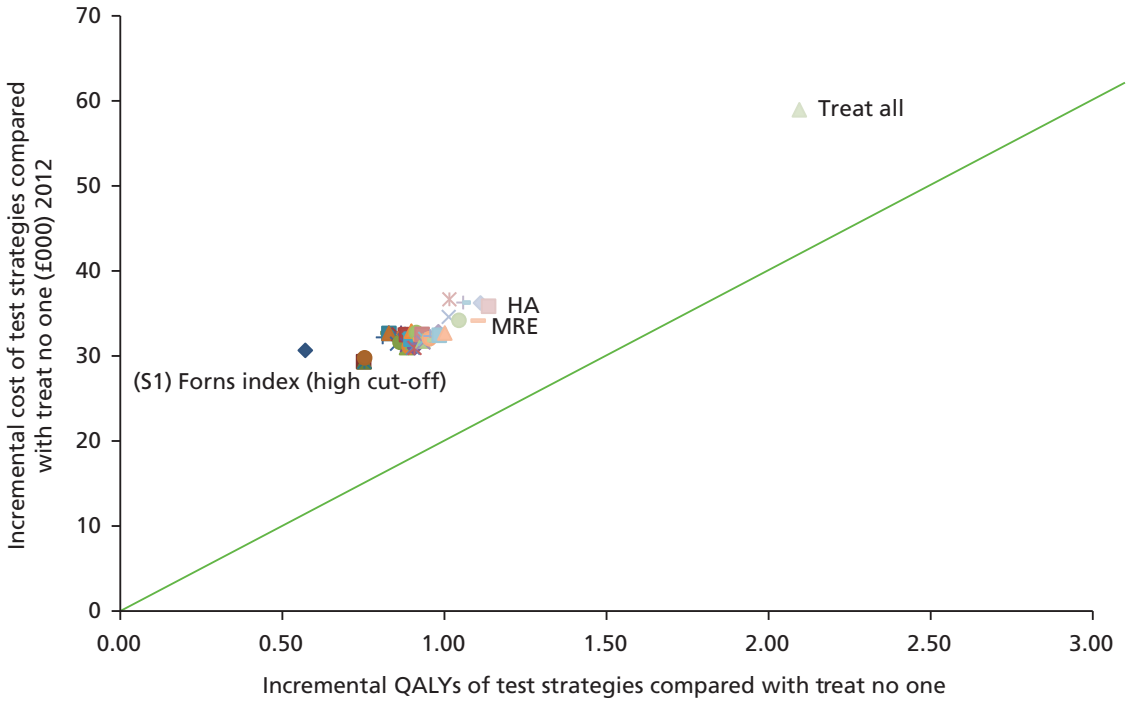

FIGURE 241 Scatterplot of results: HBeAg (negative) – second stage of the analysis (all tests strategies compared with 'treat no one'). HA, hyaluronic acid; MRE, MR elastography; S1, strategy 1.

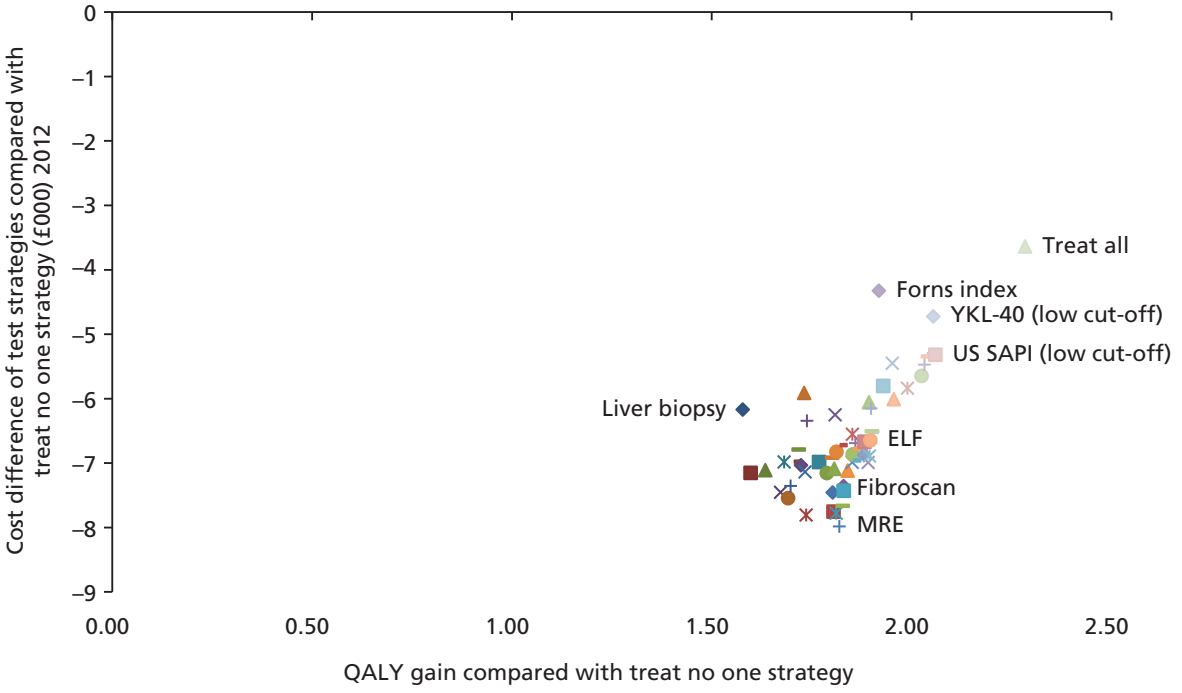

FIGURE 242 Scatterplot of results: HCV – first stage of the analysis (all tests strategies compared with 'treat no one'). MRE, MR elastography; US, ultrasound.

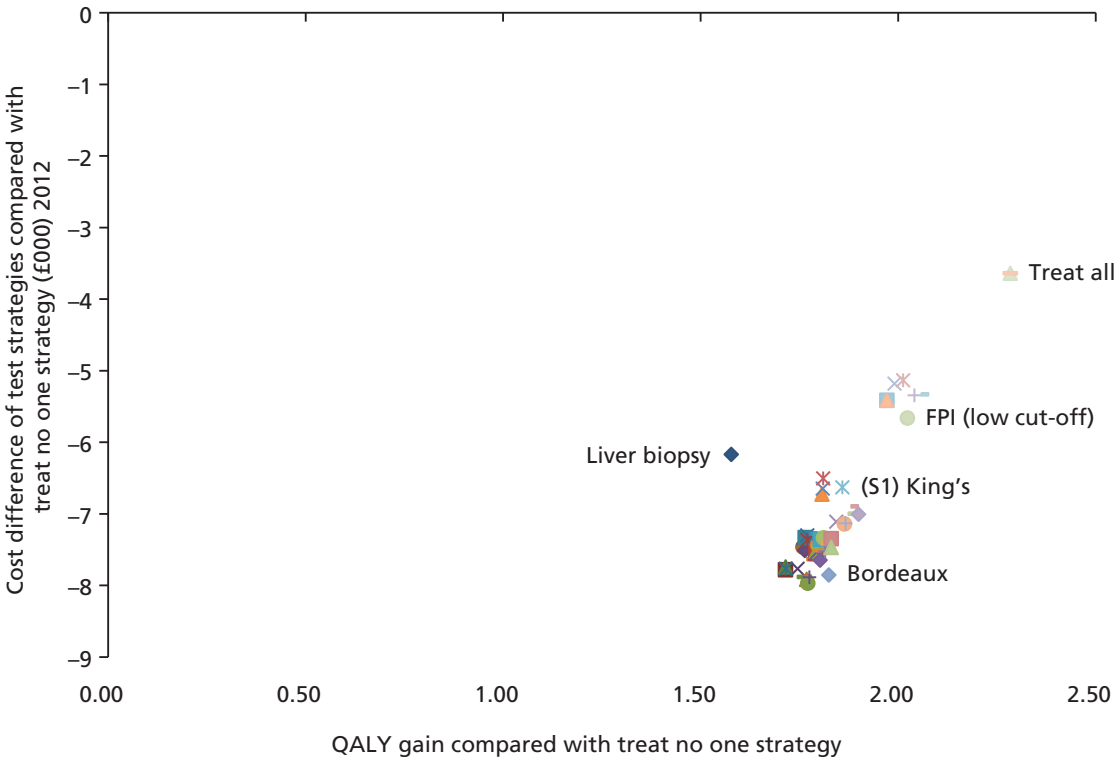

FIGURE 243 Scatterplot of results: HCV – second stage of the analysis (all tests strategies compared with 'treat no one'). HA, hyaluronic acid; MRE, MR elastography; S1, strategy 1.

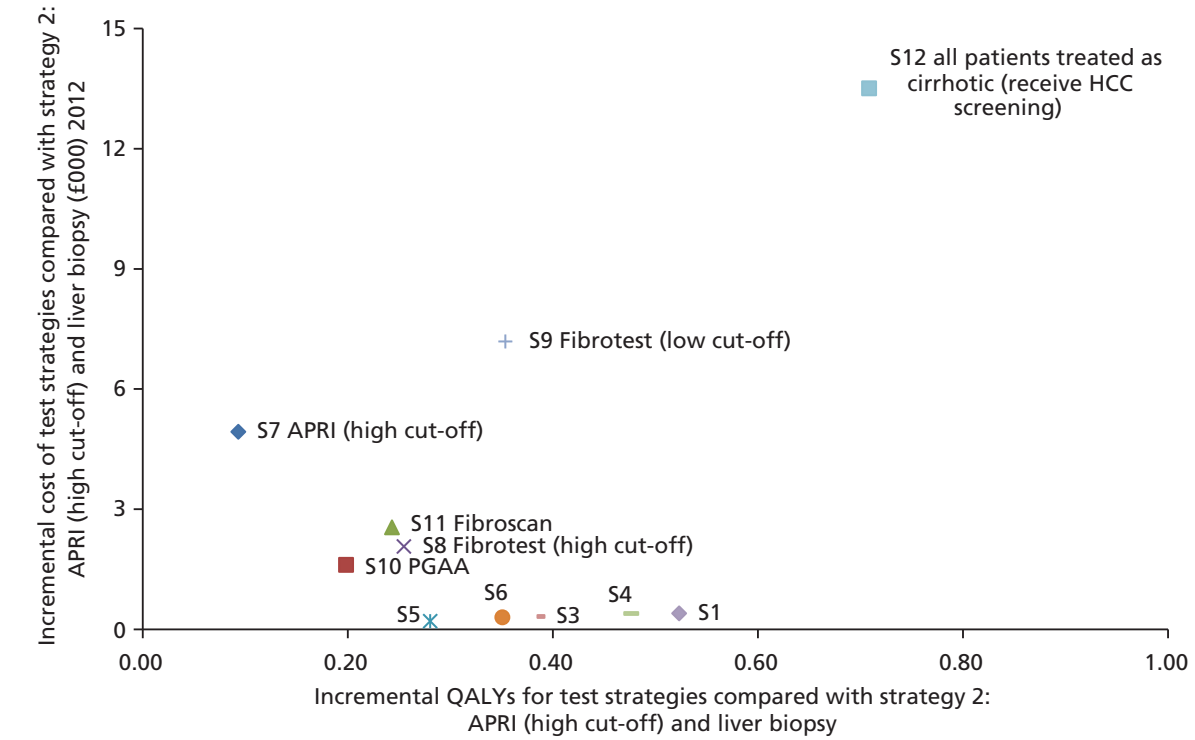

FIGURE 244 Scatterplot of results: ALD (all test strategies compared with least costly test strategy).

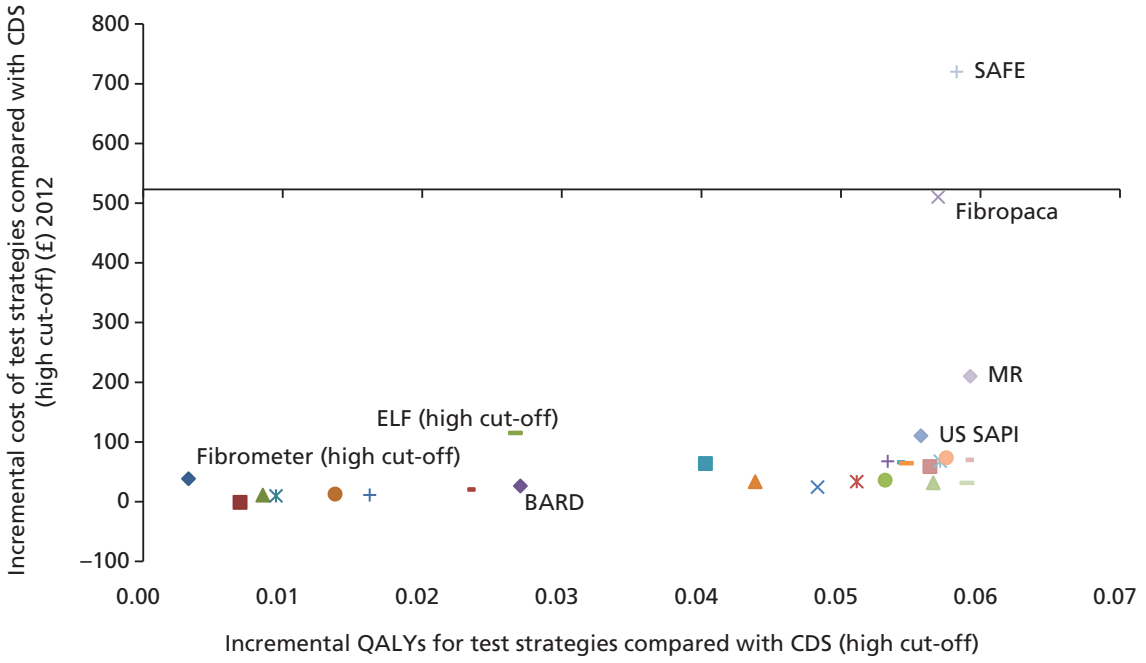

FIGURE 245 Scatterplot of results: cirrhosis (all tests strategies compared with least costly test strategy). CDS, Cirrhosis Discriminant Score; US, ultrasound.

Appendix 13 Reporting patient and public involvement

Patient representatives were not present in the steering committee meetings, as these were dedicated to technical discussions around meta-analysis and disease-modelling data and assumptions.

Upon completion of our study, the British Liver Trust, which is the leading liver charity in the UK, was contacted in order to arrange the presentation and dissemination of the findings of our study in patients and their representatives.

There were no slots available in their latest meeting; however, these will be presented in a future meeting.

We further collaborated with Research Media and produced a flyer, which explains in simplified language our research on non-invasive tests, their effectiveness and their cost-effectiveness. This will be published online in the British Liver Trust website and also in International Innovation, which is considered one of the leading global dissemination resources (www.research-europe.com/index.php/digital_magazine/). Fifty hard copies of this publication will be sent to key stakeholders.

EME
HS&DR
HTA
PGfAR
PHR

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

Published by the NIHR Journals Library