JOURNAL OF VIRAL HEPATITIS
JVH

Journal of Viral Hepatitis, 2015 doi:10.1111/jvh.12469

Cost-effectiveness of noninvasive liver fibrosis tests for treatment decisions in patients with chronic hepatitis B in the UK: systematic review and economic evaluation

C. Crossan,[1,*] E. A. Tsochatzis,[2,*] L. Longworth,[1] K. Gurusamy,[3] V. Papastergiou,[1] E. Thalassinos,[1] K. Mantzoukis,[1] M. Rodriguez-Peralvarez,[1] J. O'Brien,[1] A. Noel-Storr,[4] G. V. Papatheodoridis,[5] B. Davidson[3] and A. K Burroughs[2] [1]*Health Economics Research Group, Brunel University London, London, UK;* [2]*Sheila Sherlock Liver Unit and UCL Institute for Liver and Digestive Health, The Royal Free Hospital and UCL, London, UK;* [3]*Division of Surgery, Royal Free Campus, UCL Medical School, London, UK;* [4]*Cochrane Dementia and Cognitive Improvement Group, Nuffield Department of Medicine, Oxford University, Oxford, UK; and* [5]*Laiko Hospital, University Department of Gastroenterology, Athens, Greece*

Received December 2014; accepted for publication July 2015

SUMMARY. We compared the cost-effectiveness of various noninvasive tests (NITs) in patients with chronic hepatitis B and elevated transaminases and/or viral load who would normally undergo liver biopsy to inform treatment decisions. We searched various databases until April 2012. We conducted a systematic review and meta-analysis to calculate the diagnostic accuracy of various NITs using a bivariate random-effects model. We constructed a probabilistic decision analytical model to estimate health care costs and outcomes quality-adjusted-life-years (QALYs) using data from the meta-analysis, literature, and national UK data. We compared the cost-effectiveness of four decision-making strategies: testing with NITs and treating patients with fibrosis stage ≥F2, testing with liver biopsy and treating patients with ≥F2, treat none (watchful waiting) and treat all irrespective of fibrosis. Treating all patients without prior fibrosis assessment had an incremental cost-effectiveness ratio (ICER) of £28 137 per additional QALY gained for HBeAg-negative patients. For HBeAg-positive patients, using Fibroscan was the most cost-effective option with an ICER of £23 345. The base case results remained robust in the majority of sensitivity analyses, but were sensitive to changes in the ≥F2 prevalence and the benefit of treatment in patients with F0–F1. For HBeAg-negative patients, strategies excluding NITs were the most cost-effective: treating all patients regardless of fibrosis level if the high cost-effectiveness threshold of £30 000 is accepted; watchful waiting if not. For HBeAg-positive patients, using Fibroscan to identify and treat those with ≥F2 was the most cost-effective option.

Keywords: cirrhosis, fibroscan, fibrosis, incremental cost-effectiveness ratio, prognosis, quality-adjusted-life-years.

INTRODUCTION

It is estimated that 350–400 million people worldwide are hepatitis B virus (HBV) carriers [1] and are therefore at risk for progressive liver disease leading to cirrhosis and hepatocellular cancer (HCC) with associated increased morbidity and mortality. To prevent such progression, antiviral treatment is administered to patients who fulfil certain criteria, based on a combined assessment of viral load, transaminase level and/or liver fibrosis. Indeed, EASL guidelines state that patients should be considered for antiviral treatment when they have HBV DNA levels >2000 IU/mL, serum ALT levels above the upper limit of normal and severity of liver disease as assessed by liver biopsy (or noninvasive markers once validated in HBV-infected patients) showing at least moderate necroinflammation and/or fibrosis [1].

The current reference standard for assessing liver fibrosis is the histological evaluation of liver biopsy with scoring systems such as METAVIR or Ishak [2,3]. However, liver biopsy is an invasive and costly procedure, inconvenient to patients and associated with a small risk of significant

Abbreviations: CEAF, cost-effectiveness frontier; CHB, chronic hepatitis B; CI, confidence intervals; FN, false negative; FP, false positive; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; ICER, incremental cost-effectiveness ratio; NITs, noninvasive tests; QUADAS, Quality Assessment of Diagnostic Accuracy Studies; QUALYs, quality-adjusted-life-years; TN, true negative; TP, true positive.

Correspondence: Emmanuel A. Tsochatzis, Sheila Sherlock Liver Unit and UCL Institute for Liver and Digestive Health, The Royal Free Hospital and UCL, Pond Street, NW3 2QG, London, UK.
E-mail: e.tsochatzis@ucl.ac.uk
*These two authors contributed equally to this work and are joint first authors.

bleeding (1.1–1.6%) [4,5]. Alternatives to liver biopsy include noninvasive fibrosis tests (NITs), which can be broadly divided into three categories: simple/indirect serum markers, direct or patented serum markers and imaging modalities. Simple serum NITs, such as APRI and FIB4, consist of readily available indirect markers of fibrosis such as ALT, AST and platelet count, and are associated with lower costs. Other serum tests such as Fibrotest are patented and must be performed in laboratories that meet certain quality standards, and are therefore more expensive and less available. Transient elastography, which is performed with Fibroscan, is the most widely used imaging modality and measures liver stiffness based on ultrasound principles using dedicated equipment. Simple NITS, such as APRI and FIB-4 have two cut-offs for diagnosing specific fibrosis stages, as the use of a single cut-off would result in suboptimal sensitivity and specificity. These are a high cut-off with high specificity or a low cut-off with high sensitivity. A combined cut-off uses the low cut-off to rule out the presence of a particular stage of fibrosis, and the high cut-off to confirm that the patient has a particular stage of fibrosis [16]. However, a number of patients fall in the indeterminate range of test results (i.e. their score is between the low and the high cut-off) and need alternative testing or future re-testing.

Noninvasive tests may offer cost-effective alternatives to liver biopsy. However, currently there are no studies that assess their cost-effectiveness in patients with chronic hepatitis B (CHB) within a health care setting. We assessed the diagnostic accuracy and cost-effectiveness of NITs in patients with CHB, who would normally undergo liver biopsy to inform treatment decisions.

METHODS

Systematic review

We performed a systematic review and meta-analysis to determine the diagnostic accuracy of NITs compared to liver biopsy in adult patients with CHB. This study was part of a larger project funded by the UK National Institute of Health Research (NIHR) Health Technology Assessment Programme [6] that determined the cost-effectiveness of noninvasive tests in patients with CHB, hepatitis C [7], alcoholic liver disease and nonalcoholic fatty liver disease.

Study selection and data extraction

We included full papers and abstracts which provided the data necessary to determine the number of true positive (TP), true negative (TN), false positive (FP) and false negative (FN) results of the NITs for ≥F2 using liver biopsy as the reference standard, irrespective of language or publication status and whether the data were collected prospectively or retrospectively. We excluded studies which reported on fewer than 10 patients and when the maximum interval between liver biopsy and the NITs was >6 months.

The search strategy, data extraction and analysis were performed as previously described [7] and are detailed in the Appendix S1.

The quality of the included studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS)-2 assessment tool [8]. The criteria used for QUADAS-2 assessment are shown in the Appendix S1.

Economic model

Approach to analysis

The systematic review identified a total of 25 NITs; however the quantity and quality of the evidence varied between the NITs (Table 1). The base case cost-effectiveness analysis compared those NITs for which data on sensitivity and specificity were considered robust, defined as whether the bivariate model used in the meta-analysis converged (n = 5): APRI (high cut-off and low cut-off); Fibroscan; Fibrotest; and FIB4 (low cut-off). We compared the cost-effectiveness of NITs based on four different antiviral treatment scenarios: (i) treat all patients without prior fibrosis evaluation, (ii) treat no patients and no fibrosis evaluation to inform this decision (watchful waiting), (iii) biopsy all patients and treat those with a METAVIR fibrosis stage ≥F2, (iv) test with NIT and treat patients with a METAVIR fibrosis stage ≥F2.

Model structure

The analyses were based on a decision tree framework, combined with a Markov model to estimate the long-term costs and outcomes associated with each potential test diagnosis: TP, FP, FN or TN and the 'treat all' and 'treat no one' testing strategies.

The Markov model estimated the health outcomes and costs for a hypothetical cohort of 1000 CHB patients with suspected fibrosis, who would usually present for liver biopsy. Separate Markov models were constructed for the HBeAg- positive and HBeAg-negative patient cohorts to reflect differences in disease progression and patient characteristics. The structural assumptions underlying the state transition models applied to both groups of patients. Figure 1 provides a schematic illustration of the pathway. The cohort of 1000 CHB patients started the model in either the mild, moderate or cirrhotic health states (≤F4), depending on the outcome being modelled (TN, TP, FP, FN, treat all and treat no-one). The distribution of the population among the different health states was based on the prevalence data from the systematic review. The Markov models were evaluated over a lifetime period with a cycle length of 1 year. Health outcomes were expressed as quality-adjusted-life-years (QALYs), which combine data on life expectancy with data reflecting quality of life. All unit costs

Table 1 Diagnostic accuracy of noninvasive tests for detection of ≥F2 in patients with chronic hepatitis B

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	0.10–0.27	0.98	0.95–0.99	Single study
APRI (low cut-off)	8	0.4–0.6	0.80	0.68–0.88	0.65	0.52–0.77	Bivariate random-effects model with correlation between sensitivity and specificity
APRI (high cut-off)	6	1.5	0.37	0.22–0.55	0.93	0.85–0.97	Bivariate random-effects model with correlation between sensitivity and specificity
Age_Platelet index	1	3	0.68	0.61–0.74	0.62	0.57–0.67	Single study
AST_ALT_ratio	1	0.67	0.57	0.51–0.64	0.59	0.54–0.63	Single study
FIB-4 (low cut-off)	4	1.1–1.7	0.68	0.60–0.75	0.73	0.67–0.79	Bivariate random-effects model with correlation between sensitivity and specificity
FIB-4 (high cut-off)	1	3.25	0.58	0.04–0.17	0.99	0.96–1.00	Single study
Forns index (low cut-off)	1	4.2	0.58	0.47–0.68	0.77	0.61–0.88	Single study
Forns index (high cut-off)	1	6.9	0.15	0.08–0.24	1.00	0.90–1.00	Single study
GUCI	1	0.2	0.67	0.55–0.76	0.97	0.85–0.99	Single study
Hui index	1	0.15	0.50	0.39–0.61	0.91	0.78–0.97	Single study
PAPAS	1	1.67	0.73	0.62–0.81	0.78	0.71–0.84	Single study
Direct serum noninvasive serum tests							
Hyaluronic acid	1	185.3	0.84	0.73–0.91	0.83	0.66–0.93	Single study
Hepascore	1	0.5	0.79	0.68–0.86	0.74	0.65–0.81	Single study
Commercial noninvasive serum tests							
Fibrotest	6	0.40–0.48	0.66	0.57–0.75	0.80	0.72–0.86	Bivariate random-effects model with correlation between sensitivity and specificity
Imaging modalities							
ARFI	1	1.33	0.71	0.59–0.80	0.67	0.30–0.90	Single study
Fibroscan	13	6.3–8.9	0.71	0.62–0.78	0.84	0.74–0.91	Bivariate random-effects model with correlation between sensitivity and specificity
CEUS	3	–	0.88	0.07–1	0.73	0.11–0.98	Random-effects model for sensitivity and specificity without correlation
DWMRI	5	–	0.78	0.63–0.88	0.78	0.51–0.93	Random-effects model for sensitivity and specificity without correlation
MRE	3		0.94	0.13–1	0.92	0.72–0.98	Random-effects model for sensitivity and fixed-effect model for specificity without correlation

(continued)

Table 1 *(continued)*

Test	Number of studies	Cut-off	Summary sensitivity	95% CI	Summary specificity	95% CI	Statistics
US	3	–	0.35	0.14–0.63	0.86	0.59–0.96	Bivariate random-effects model with correlation between sensitivity and specificity
US_SAPI	3	–	0.74	0.69–0.79	0.79	0.72–0.85	Fixed-effects model for sensitivity and specificity without correlation
US_SAPI High cut-off	2	–	0.61	0.54–0.68	0.96	0.9–0.98	Fixed-effects model for sensitivity and specificity without correlation
US_SAPI_F2low cut-off	2	–	0.94	0.9–0.97	0.39	0.31–0.49	Fixed-effects model for sensitivity and specificity without correlation

CI, confidence intervals; APGA, AST, Platelet count; GGT, A-fetoprotein; APRI, AST-to-platelet ratio index; FIB-4, Fibrosis-4; GUCI, Goteborg University Cirrhosis Index; PAPAS, Age; ALP, a-fetoprotein; AST, ARFI; acoustic radiation force impulse; CEUS, contrast enhanced ultrasound; DWMRI, diffusion-weight magnetic resonance imaging; MRE, magnetic resonance elastography; US, ultrasound; US_SAPI, splenic artery pulsatility index determination by ultrasound.

reported in the analysis for health states and liver biopsy, are priced for the year 2012 and were considered from the perspective of the UK National Health System (NHS). Costs and QALYs were discounted using a rate of 3.5% [9]. Cost-effectiveness was assessed according to thresholds used by the UK National Institute for Health and Care Excellence, which include a threshold of £20 000, increasing up to around £30 000 when specific additional considerations are important [9].

The decision tree was populated with results from the Markov model, summary sensitivity and specificity estimates and average disease prevalence estimate (METAVIR ≥F2 extracted from the meta-analysis), to estimate the cost-effectiveness of all comparators.

Input parameters

Current recommended treatment options for chronic CHB include indefinite treatment with nucleoside or nucleotide analogues or pegylated interferon alpha for a period of 12 months [1]. Treatment with antiviral agents, entecavir or tenofovir (50–50 split) was initiated in the model for a patient with positive (METAVIR ≥F2) diagnosis. Only patients in the moderate fibrosis, compensated cirrhosis or 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 effectiveness was sourced from a published meta-analysis by Woo *et al.* [10]; treatment efficacy of the drugs (represented by relative risks in the model) are listed in Table 2.

Treatment with peginterferon alfa-2a was given for a 1 year period to 10% of those who tested negative to reflect a proportion of patients who may not have fibrosis but would receive treatment for necroinflammation [1]. We assumed that 30% of these TN patients who received treatment for 1 year would have a successful response to treatment and would no longer progress to more advanced disease stages. The remaining 90% who initially tested negative and those who had an unsuccessful response to peginterferon alfa-2a treatment underwent a 'watchful waiting' process which incorporated a re-test with an NIT every 2 years. If patients in the model had progressed to a ≥F2 health state at the time of re-test, they received immediate treatment with antivirals (entecavir and tenofovir). We tested this assumption in a sensitivity analysis.

The rate of disease progression in the Markov model was sourced from a published cost-effectiveness study by Dakin *et al.* [11]; however, this study did not report separate transition probability data for the precirrhotic health states (mild and moderate fibrosis). In the absence of transition probability data for the mild and moderate health states, we used data from a study of patients with mild chronic hepatitis C [12]. We assumed that the progression of early fibrosis would not differ significantly in patients with HCV and HBV.

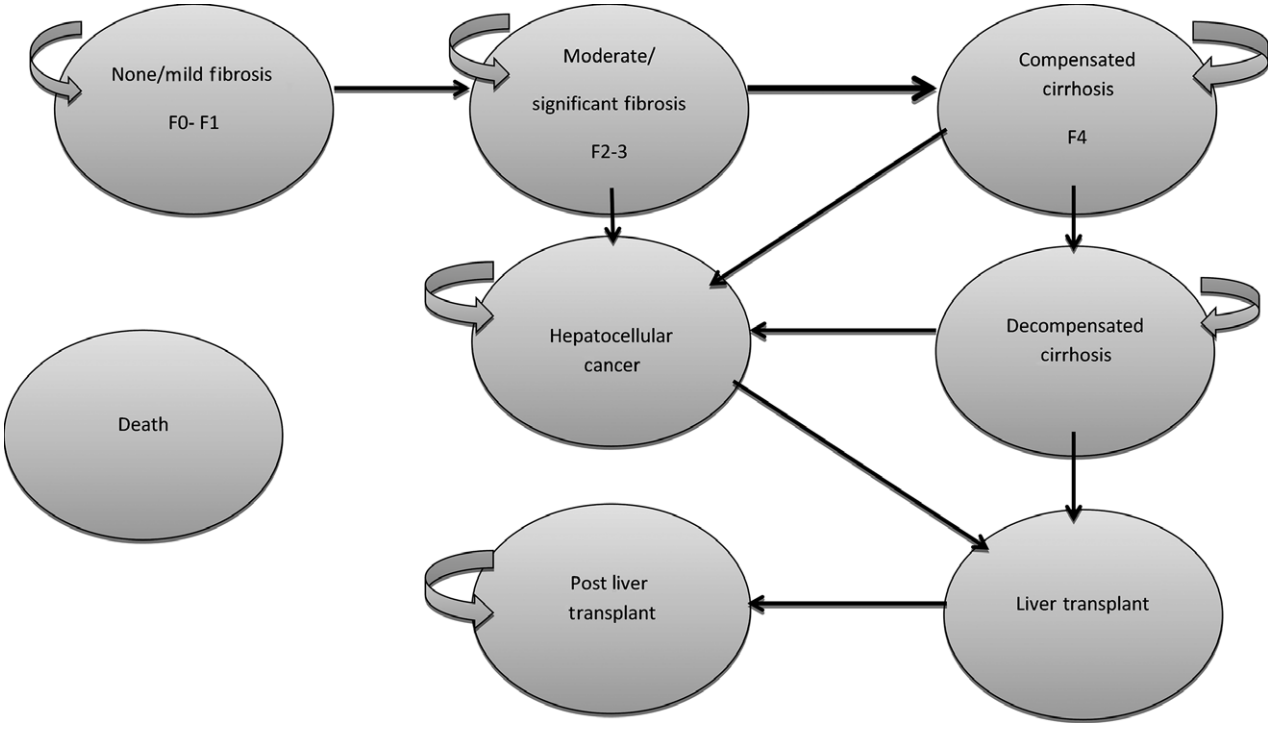

Fig. 1 Illustration of the Markov Model used for economic analysis. The disease stages reflect the METAVIR staging score for liver fibrosis and cirrhosis. The cohort represents people suspected of liver fibrosis who can enter the models in one of three disease stages; mild fibrosis (METAVIR stages F0–F1), moderate fibrosis (METAVIR stages F2–3) and compensated cirrhosis (METAVIR stage F4) with the proportions determined by the prevalence estimated from the results of the systematic review. Within the model, people can remain within any disease stage for longer than one cycle (length of cycle is set as 1 year) except for the liver transplant disease stage where patients can only progress to either a post-liver transplant stage or death.

We sourced health state costs and health-related utility estimates, based on patients' self-reported health status using the EQ-5D questionnaire [13], for the earlier disease stages (≤F4) from the study of patients with mild hepatitis C [12]. We assumed that the resource use identified and collected for this study (inpatient, outpatient and procedures, excluding medication costs) would be similar for patients with HCV and HBV; additionally this was the only study which estimated separate utilities for the mild and moderate health states (using EQ-5D). We sourced costs and utilities for later disease stages (>F4) from a cost-effectiveness study of liver transplantation [14], for which resource-use and EQ-5D data were available for a subgroup of patients with CHB. As we used two different sources of data, we conducted a sensitivity analysis using different utility values and health state costs for the later disease stages.

Cohort characteristics, mortality data, and treatment costs were sourced from published literature and routine national UK sources. Input parameters and sources are listed in Table 2. Costs for the NITs and liver biopsy are listed in the Appendix S1.

Analysis and uncertainty

We conducted an incremental analysis to identify the cost-effective testing strategy. We ruled out tests strategies which were more costly and less effective 'dominated'. We then estimated Incremental Cost-Effectiveness Ratios (ICER), for the remaining NITs, where they were compared to the next best alternative, calculated using the formula:

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

where C_1 = lifetime cost of strategy 1; C_0 = lifetime cost of (the next best) strategy; E_1 = QALYs from strategy 1 and E_0 = QALYs from (the next best) strategy.

Test strategies with an ICER greater than that of a more effective intervention (extendedly dominated) were also ruled out and the remaining tests were then compared to identify the NIT with the highest ICER given a £20 000 per QALY threshold.

A probabilistic sensitivity analysis was run to estimate uncertainty in the mean results. We constructed the cost-effectiveness frontier (CEAF), which plots the uncertainty associated with the optimal option (test with highest expected net benefit), for different values of the

Table 2 Input parameters for model

Model inputs	Parameter value	Distribution	Source
Cohort characteristics			
HBeAg-positive			
Starting age	31 years		
Gender % male	70%		Shepherd *et al.* [25]
HBeAg-negative			
Starting age	40 years		
Gender % male	90%		
Natural history data (Transition probabilities)			
Mild-moderate fibrosis	0.025	Dirichlet	
Moderate fibrosis-compensated Cirrhosis (HBeAg-positive)	0.037		Wright *et al.* [12]
Moderate Fibrosis-compensated cirrhosis (HBeAg-negative)	0.09		
Compensated cirrhosis-decompensated cirrhosis	0.05		
Moderate Fibrosis-Hepatocellular Cancer (HCC)	0.048		Dakin *et al.* [11]
Cirrhosis to HCC	0.024		
Decompensated cirrhosis/HCC- liver transplant	0.016		
Decompensated cirrhosis-death	0.30		
Liver transplant-death	0.21		
Post Liver transplant-death	0.057		
Excess mortality			
Moderate fibrosis	0.035		
Compensated cirrhosis	0.051		
HCC	0.56		
Treatment dosage			
Peginterferon alfa-2a	180 mg (weekly)		
Entecavir	500 mg (daily)		British National Formulary 64
Tenofovir	245 mg (daily)		

Treatment efficacy (Relative risks)

	Mean	CI upper	CI lower	Distribution	Source
HBeAg-positive					
Entecavir	0.56	0.12	0.94	Gamma	
Tenofovir	0.53	0.06	0.95	Gamma	Woo *et al.* [10]
Peginterferon alfa-2a	0.52	0.06	0.95	Lognormal	
HBeAg-negative					
Entecavir	0.64	0.01	1.00	Gamma	
Tenofovir	0.65	0.01	1.00	Gamma	Woo *et al.* [10]
Peginterferon alfa-2a	0.52	0.06	0.95	Log Normal	

Health state costs

	Mean	SE	Distribution	Source
Mild fibrosis	185	36.39	Gamma	Wright *et al.* [12]
Moderate fibrosis	986	101.69		
Compensated cirrhosis	1521	309.05		
Decompensated cirrhosis	36 194	9967.19	Gamma	Longworth *et al.* [14]
Hepatocellular cancer	36 194	9967.19		
Liver transplant	64 122	5584.70		

(continued)

Table 2 *(continued)*

Health state costs				
	Mean	SE	Distribution	Source
Post-liver transplant	16 321	7932.51		
Utilities				
Mild fibrosis	0.77	0.035		Wright *et al.* [12]
Moderate fibrosis	0.66	0.0.18	Gamma	
Compensated cirrhosis	0.55	0.032		
Decompensated cirrhosis	0.57	0.076	Gamma	Longworth *et al.* [14]
Hepatocellular cancer	0.57	0.076		
Liver transplant	0.55	0.016		
Post Liver transplant	0.78	0.064		
Mild Fibrosis (during treatment)	0.65	0.035	Gamma	Wright *et al.* [12]
Moderate fibrosis (during treatment)	0.55	0.018		
Compensated cirrhosis (during treatment)	0.45	0.040		
Disutility following liver biopsy	0.2			Assumed

Treatment costs	Annual Cost (UK £2012)	
Peginterferon alfa-2a	6469	
Entecavir	4420	British National Formulary 64
Tenofovir	2926	

cost-effectiveness threshold (threshold value range varied from £0 to £60 000) [15].

Secondary analyses

We conducted a secondary analysis where we evaluated the sequential use of NITs (see appendix for sequential testing strategies).

We also carried out an analysis where we evaluated all identified NITs (n = 25), regardless of their robustness (robustness was defined as to whether the bivariate model converged for meta-analysis).

Sensitivity analysis

A number of one way sensitivity analyses were undertaken to vary some of the input parameters in the model, including changes to utility values, health state costs, average disease prevalence, assumption of treatment benefit for FP patients (mild health state F0-1) and the re-test assumption of perfect sensitivity and specificity.

RESULTS

Systematic review

The selection flow chart for studies is shown in the Appendix S1. Data on patients with CHB were extracted from 52 studies that evaluated 25 different NITs. NIT cut-offs for the diagnosis of specific histological stages were not always predetermined, and consequently varied. We opted not to perform a separate meta-analysis for each stage-specific cut-off of a NIT, but to group together cut-offs if the range was reasonable. Therefore, when a range of cut-offs is mentioned in the results tables, the reported sensitivities and specificities are probably overestimated.

The summary sensitivities and specificities for each test are presented in Table 1. The average prevalence of METAVIR score ≥F2 in the included studies was 54%. All but one study had a high risk of bias as assessed by the QUADAS-2 tool (Appendix S1) therefore our results should be interpreted with caution. Strikingly, the cut-offs of the NITs were predefined in 11/52 (21%) studies, while liver biopsy was of adequate quality (≥6 portal tracts and ≥15 mm) in 12/52 (23%) studies.

Economic modelling

HBeAg-negative chronic hepatitis B

The most effective strategy to employ is 'treat all without prior diagnostic testing', which had an ICER of £28 137. This would only be considered cost-effective if the £30 000 upper bound of the UK cost-effectiveness threshold range is considered acceptable; if not, a strategy of 'treat no one' would be the most cost-effective. All other testing strategies are dominated by the 'treat all' strategy, as they are more costly and less effective. Table 3 displays results of the base case analysis.

Table 3 HBeAg –negative base case analysis

Testing option	Cost £	QALYs	Incremental cost £	Incremental QALY	ICER £
Treat no one (no diagnostic test)	35 579	8.83	–	–	–
Liver biopsy	70 274	9.64	–	–	Dominated
APRI (high cut off)	69 429	9.71	–	–	Extendedly Dominated
Fibroscan	72 986	9.93	–	–	Extendedly Dominated
Fibrotest	73 857	9.94	–	–	Extendedly Dominated
FIB 4 (low cut off)	75 702	10.01	–	–	Extendedly Dominated
APRI (low cut off)	77 981	10.13	–	–	Extendedly Dominated
Treat all (no diagnostic test)	96 525	10.92	58 947	2.09	28 137

The CEAF (Appendix S2) for the HBeAg-negative model shows that the probability of 'treat all' being on average the most cost-effective testing option given a cost-effectiveness threshold value of £30 000, is 39%.

HBeAg-positive chronic hepatitis B

For HBeAg-positive patients at a cost-effectiveness threshold of £30 000: the cost effective option to use when all tests were compared was Fibroscan, with an ICER of £23 345 (Table 4). However, the CEAF (Figure Appendix) for the HBeAg-positive model shows that the probability of Fibroscan being on average the most cost-effective testing option given a cost-effectiveness threshold value of £30 000 is low at 21%.

A testing strategy 'treat all' without prior diagnostic testing provided a higher health gain (QALY) than Fibroscan, however, this option had an ICER of £39 747 which would not be acceptable given a £20 000–30 000 cost-effectiveness threshold.

Secondary analyses

A secondary analysis where we evaluated the use of more than one NIT found that 'treat all' without prior diagnostic testing remained the most cost-effective option for the HBeAg-negative population with an ICER of £28 138, whereas for the HBeAg-positive population, using APRI (low cut-off) followed by Fibroscan using the second sequential testing strategy (appendix) was the most cost-effective testing strategy with an ICER of £23 901 (see Appendix S1).

An analysis evaluating all of the NITs identified during the systematic review (irrespective of robustness) found similar results for the HBeAg-negative population as the base case analysis (Treat all with an ICER of £28 137). However, for the HBeAg-positive population, Magnetic Resonance Elastography (MRE) was the most cost-effective option with an ICER of £23 468. This analysis found that several NITs had similar outcomes (cost and QALYs) (See Appendix S1 for incremental analysis tables).

Sensitivity analysis results

HBeAg-negative

The base results remained robust to the majority of sensitivity analyses; those they were sensitive to are detailed below.

Using the starting age and gender split used in the HBeAg-positive model changed the results for the HBeAg-negative population so that Fibroscan became the most cost-effective testing option with an ICER of £25 575.

Amending the ≥F2 prevalence used in the model (54%) also changed the base results; when using a prevalence of 43% (lower quartile), the no testing and no treatment option was the most cost-effective testing strategy (treat all testing strategy had an ICER of £30 413). When the

Table 4 HBeAg –positive base case analysis

Testing option	Cost £	QALYs	Incremental cost £	Incremental QALY	ICER £
Treat no one (no diagnostic test)	37 831	9.64	–	–	–
Liver biopsy	75 957	11.41	–	–	Dominated
APRI (high cut off)	75 210	11.45	37 380	1.81	20 673
Fibroscan	79 000	11.61	3 790	0.16	23 345
Fibrotest	79 462	11.62	–	–	Extendedly Dominated
FIB 4 (low cut off)	81 382	11.67	–	–	Extendedly Dominated
APRI (low cut off)	83 788	11.75	–	–	Extendedly Dominated
Treat all (no diagnostic test)	101 484	12.18	22 484	0.57	39 474

maximum prevalence of 92% was used, Fibroscan was the most cost-effective test with an ICER of £21 853.

We also tested the assumption that not all patients were diagnosed correctly at re-test (watchful waiting strategy for patients who initially test negative) by using the sensitivity and specificity estimates of three commonly used tests [APRI (low cut-off), Fibrotest, and Fibroscan]. With this analysis, Fibroscan became the most cost-effective test when the sensitivity and specificity of Fibrotest or Fibroscan were used (ICERs of £27 584 and £27 088 respectively). There was no change to the base case results when the sensitivity and specificity for APRI (low cut-off) was used for the re-test.

In the base case all patients who tested positive (≥F2) received treatment with antiviral agents. When we assumed that patients in a mild health state (F0-1), who were incorrectly diagnosed as having more advanced fibrosis (FP patients) would receive no benefit from treatment, the results changed so that 'treat no one' became the most cost-effective option (the ICER for treat all increased to £35 081).

HBeAg-positive

The results were robust to the majority of sensitivity analyses, apart from the analyses detailed below.

We carried out four analyses where we amended the prevalence of ≥F2 used in the model. Amending the prevalence from 54% to 43% (lower quartile) changed the most cost-effective test to APRI (high cut-off) with an ICER of £19 989. When changing the prevalence to 65% (third quartile), the most cost-effective testing option became 'treat all' irrespective of fibrosis stage with an ICER of £24 615. When using the maximum prevalence of 92%, APRI (high cut-off) became the most cost-effective test with an ICER of £18 186. Using the minimum ≥F2 prevalence (27%) also changed the most cost-effective test to APRI (high cut-off) with an ICER of £19 464.

When we assumed that patients in a mild health state, who were incorrectly diagnosed as having advanced fibrosis (FP patients), would receive no benefit from treatment, the most cost effective test was APRI (high cut-off) with an ICER of £21 122.

The results changed when we tested the assumption that all persons were diagnosed correctly when a re-test was performed (watchful waiting); APRI (low cut-off) became the most cost-effective test when the sensitivity and specificity of APRI (low cut-off) or Fibrotest were used (ICERs of £24 651 and 29 644 respectively). There was no change to the base case results when the sensitivity and specificity for Fibroscan was used for the re-test.

DISCUSSION

We compared five NITs with each other, liver biopsy and a 'treat all' and 'treat no one' approach for informing treatment decisions in chronic HBV patients. The overall results differed for the HBeAg-positive and HBeAg-negative patients.

In the HBeAg-negative population, treating all patients based on viral load and ALT irrespective of the degree of fibrosis offered the largest QALY gain with an ICER of £28 317. There was some uncertainty in these results and the 'treat all strategy' had a 39% probability of having the highest net benefit, given a £30 000 cost-effectiveness threshold. In the HBeAg-positive population the most cost-effective testing option was Fibroscan, however this NIT had a low (21%) probability of being the optimal testing option given a £30 000 cost-effectiveness threshold.

Similar findings for treatment in patients with CHB have been reported in assessments conducted to inform national guidelines. For example, in the UK, entecavir for the treatment of HBeAg-negative patients (assuming lifetime treatment duration) had an ICER of £27 124/QALY gained [17], similar to our base case analysis result. The fact that antiviral treatment for CHB in the UK is cost-effective at a £30 000 but not at a £20 000 cost-effectiveness threshold reflects the high unit cost and lifetime duration of treatment [17,18]. Currently, a third of UK patients with HBV followed in liver centres are on antiviral treatment, however, only 18% of those treated are on recommended first-line treatment [19].

The difference between the HBeAg-positive and negative results can be attributed to differences in baseline characteristics such as an older starting age within the model, higher male population, higher risk of developing cirrhosis, and the effectiveness of treatment in the HBeAg-negative group.

Our data therefore suggest that liver biopsy is not cost-effective for informing treatment decisions in patients with CHB in the UK and most probably in most developed countries, with the caveat regarding local costs, preferences and decision rules. A pragmatic cost-effective approach therefore would be to treat HBeAg(−) patients based on high viral load and deranged ALT, irrespective of their fibrosis level. For HBeAg(+) patients, the most cost-effective strategy is to perform a Fibroscan to inform treatment decisions in patients with high viral load and deranged ALT. This is a deviation from the current EASL guidelines, where liver biopsy is the preferred strategy irrespective of HBeAg status. It also simplifies treatment decisions in HBeAg(−) patients, who could be potentially seen in nurse-led outpatient clinics.

Our meta-analysis of NITs 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 striking finding of our meta-analysis was that the vast majority of studies had high risk of bias and failed in important methodological aspects, such as the absence of predetermined test cut-offs and suboptimal quality of

liver biopsy as the reference standard. Moreover, there was a paucity of sufficiently evaluated and validated NITs in patients with CHB. Indeed, there were adequate data for a reliable meta-analysis in only four NITs, namely APRI, FIB4, Fibroscan and Fibrotest. A recent updated meta-analysis on APRI and FIB-4 showed similar diagnostic accuracy with the data presented in this paper [20]. The increasingly used Fibroscan does not have validated cut-offs for specific fibrosis stages. Published cut-offs are based on post hoc analyses and have not been prospectively validated in independent cohorts, therefore the reported diagnostic accuracy is most likely overestimated [21]. All tests performed significantly better in diagnosing cirrhosis than lesser fibrosis stages. Therefore, NITs in CHB need better quality studies and further validation, particularly for the diagnosis of moderate fibrosis.

Our economic modelling was performed from the perspective of an economy of a developed country and therefore its findings cannot be extrapolated to the developing world. This would require a separate analysis with different utilities and costs. A cost-effectiveness analysis of antiviral treatment in middle-income countries has shown that this is cost-effective, however the use of NITs was not factored in the analysis [22]. However, the greatest burden of HBV infection is encountered in low-income countries in Asia and the sub-Saharan Africa. The World Health Organization has recently launched guidelines for treatment of people with CHB in such countries [23]. Treatment is recommended in patients with deranged ALT and viral load >2000 IU/ml or those with evidence of cirrhosis, based on clinical finding or an APRI score of >2. The choice of APRI was based on widespread availability rather than cost-effectiveness analysis. Clearly, such countries would require subsidised costs of antiviral treatment, similar to the paradigm of HIV infection, for these guidelines to become applicable.

There are limitations to this analysis. Firstly, treatment decisions are far more complex than fibrosis evaluation in CHB, and depend on a global assessment that takes into account HBeAg status, viral load, transaminases, fibrosis and necroinflammation but also family history of HCC and family planning in females. Indeed, in some patients the decision to treat is straightforward without the need of a liver biopsy, and in others even a ≥F2 would not necessarily prompt treatment initiation. Therefore, this analysis is relevant only for patients who would require a liver biopsy to decide on treatment initiation, and not in unselected patients with CHB.

Secondly, the assumption that the re-test carried out during the watchful waiting process correctly identified all patients who had progressed is a potential limitation; a sensitivity analysis incorporating the sensitivity and specificity estimates of three NITs found that the base case results did change for both populations, implying that this assumption may underestimate the ICER of treating all patients.

Thirdly, we did not include a health state specifically describing seroconversion in the model structure. This was for simplicity purposes as our outcome was treatment initiation rather than treatment effectiveness. Although some studies have shown that fibrosis and even cirrhosis can regress with antiviral therapy [24], this is not factored into the model. This may underestimate the cost-effectiveness of treatment.

In conclusion, to our knowledge, this analysis is the first to compare the cost-effectiveness of a number of different NITs for use in patients with CHB. The current reference standard, liver biopsy was more costly and less effective than other tests in both the base case and sensitivity analyses. We identified that treating all patients (if eligible for treatment based on ALT and/or viral load) regardless of fibrosis level is the most cost-effective strategy for HBeAg-negative patients, whereas an NIT, Fibroscan is the most cost-effective diagnostic strategy for HBeAg-positive patients, although there is significant uncertainty around this result. These findings were robust to several sensitivity analyses.

FUNDING

The analysis for Hepatitis B was part of a larger project funded by The National Institute for Health Research Health Technology Assessment (HTA project 09/114/02) and will be published in full in the Health Technology Assessment journal series. Visit the HTA programme website for more details www.hta.ac.uk/link to project page. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Department of Health.

CONFLICTS OF INTEREST

None.

CONTRIBUTORS

Our work was in collaboration between health economists from the Health Economics Research Group (HERG) based at Brunel University London and the Royal Free Hospital. AB, KG, LL, ET and BD developed the project proposal and secured funding for the project. The analysis was undertaken by KG (meta-analysis) and CC (cost-effectiveness analysis) supported by LL, ET, AB, and BD. ANS designed the search strategy for the systematic review on non-invasive tests. VP, EvT, JOB, KM and MR conducted study selection and data extraction. CC wrote the first draft of this paper, which was subsequently edited by all authors who have approved the final version.

REFERENCES

1 EASL clinical practice guidelines: management of chronic hepatitis B virus infection. *J Hepatol* 2012; 57 (1): 167–185.

2 Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. The French METAVIR Cooperative Study Group. *Hepatology*. 1994; 20(1 Pt 1):15–20.

3 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(1): 82–90.

4 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(1): 103–110.

5 Terjung B, Lemnitzer I, Dumoulin FL *et al.* Bleeding complications after percutaneous liver biopsy. *Digestion* 2003; 67(3): 138–145.

6 Crossan C, Tsochatzis EA, Longworth L *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): 1–410.

7 Tsochatzis EA, Crossan C, Longworth L *et al.* Cost-effectiveness of noninvasive liver fibrosis tests for treatment decisions in patients with chronic hepatitis C. *Hepatology* 2014; 60(3): 832–843.

8 Whiting PF, Rutjes AW, Westwood ME *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155(8): 529–536.

9 National Institute for Health and Clinical Excellence. Guide to the Methods of Technology Appraisal. London: NICE, 2013.

10 Woo G, Tomlinson G, Nishikawa Y *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(4): 1218–1229.

11 Dakin H, Bentley A, Dusheiko G. Cost-utility analysis of tenofovir disoproxil fumarate in the treatment of chronic hepatitis B. *Value Health* 2010; 13(8): 922–933.

12 Wright M, Grieve R, Roberts J *et al.* Health benefits of antiviral therapy for mild chronic hepatitis C: randomized controlled trial and economic evaluation. *Health Technol Assess*. 2006; 10(21):iii–93.

13 Brooks R. EuroQol: the current state of play. *Health Policy* 1996; 37(1): 53–72.

14 Longworth L, Young T, Buxton MJ *et al.* Midterm cost-effectiveness of the liver transplantation program of England and Wales for three disease groups. *Liver Transpl* 2003; 9(12): 1295–1307.

15 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(5): 886–897.

16 Papastergiou V, Tsochatzis E, Burroughs AK. Non-invasive assessment of liver fibrosis. *Ann Gastroenterol* 2012; 25(3): 218–231.

17 National Institute for Health and Clinical Excellence. Entecavir for the treatment of chronic hepatitis B. NICE technology appraisal guidance 153. London: NICE, 2009.

18 National Institute for Health and Clinical Excellence. Tenofovir disoproxil for the treatment of chronic hepatitis B. NICE technology appraisal guidance 173. London: NICE, 2009.

19 Tedder RS, Rodger AJ, Fries L *et al.* The diversity and management of chronic hepatitis B virus infections in the United Kingdom: a wake-up call. *Clin Infect Dis* 2013; 56(7): 951–960.

20 Xiao G, Yang J, Yan L. Comparison of diagnostic accuracy of aspartate aminotransferase to platelet ratio index and fibrosis-4 index for detecting liver fibrosis in adult patients with chronic hepatitis B virus infection: a systemic review and meta-analysis. *Hepatology* 2015; 61(1): 292–302.

21 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(4): 650–659.

22 Toy M, Onder FO, Idilman R *et al.* The cost-effectiveness of treating chronic hepatitis B patients in a median endemic and middle income country. *Eur J Health Econ* 2012; 13(5): 663–676.

23 WHO. Prevention, care and treatment of persons with chronic hepatitis B infection. www.whoint/hiv, (accessed 12 March 2015).

24 Tsochatzis EA, Bosch J, Burroughs AK. Liver cirrhosis. *Lancet* 2014; 383(9930): 1749–1761.

25 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): 1–122.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1: Data extraction and analysis, search strategy, utility data, quality assessment of included studies, unit costs for NILTs and liver biopsy, stage 1 incremental analysis for HBeAg(-) and (+) patients.

Appendix S2: Cost-effectiveness frontier (CEAF) for HBeAg(-) and (+) patients.