Surveillance of cirrhosis for hepatocellular carcinoma: systematic review and economic analysis

J Thompson Coon,[1*] G Rogers,[1] P Hewson,[2] D Wright,[2] R Anderson,[1] M Cramp,[3] S Jackson,[4] S Ryder,[5] A Price[6] and K Stein[1]

[1] Peninsula Technology Assessment Group (PenTAG), Peninsula Medical School, Exeter, UK
[2] School of Mathematics and Statistics, University of Plymouth, UK
[3] Hepatology Research Group, Peninsula Medical School, Exeter, UK
[4] Radiology Department, Plymouth NHS Hospitals Trust, UK
[5] Queen's Medical Centre, University of Nottingham, UK
[6] Southampton Health Technology Assessments Group, University of Southampton, UK

* Corresponding author

Executive summary

Health Technology Assessment
NHS R&D HTA Programme
www.hta.ac.uk

◆

Executive summary

Background

Cirrhosis is long-term liver damage from the build-up of scar tissue (fibrosis) which, as it develops, impairs effective blood flow and inhibits the organ's vital functions. There are many causes of cirrhosis, including viral hepatitis [hepatitis B and C virus (HBV and HCV)], excessive alcohol intake, non-alcoholic fatty liver disease, primary biliary cirrhosis and haemochromatosis (iron overload).

Cirrhosis can remain in an asymptomatic (compensated) state for many years. The onset of overt liver failure (decompensation) is characterised by a variety of symptoms including ascites, portosystemic encephalopathy, gastrointestinal bleeding and hepatorenal syndrome, and is often the first indication of previously silent liver disease.

Hepatocellular carcinoma (HCC) is a malignant tumour arising from liver cells (hepatocytes) and occurs mainly in cirrhotic livers. HCC affects around twice as many men as women and is more common in those above the age of 40. There is some evidence for a recent rise in incidence of HCC in England with age-adjusted incidence rising from 1.8 per 100,000 men and 0.6 per 100,000 women in 1995 to 2.8 and 0.8, respectively, in 2002. Curative treatment options include resection and orthoptic liver transplantation. Palliative treatments include percutaneous ethanol injection, radiofrequency ablation and transarterial chemoablation.

Two diagnostic tests are routinely used to detect HCC in clinical practice: serum α-fetoprotein (AFP) and ultrasonography. The sensitivity of AFP as a diagnostic tool is restricted by the existence of non-AFP-secreting tumours. The reliability of ultrasonographic diagnosis depends on a range of factors, including the expertise of the operator, the sophistication of the equipment and the size and nature of the tumour.

Routine periodic surveillance of individuals with cirrhosis is currently recommended by UK, European and American clinical guidelines. A 2002 survey confirmed that approximately three-quarters of UK gastroenterologists undertake a formal programme of surveillance for HCC in cirrhosis, mostly using a combination of AFP and ultrasound. The optimal screening frequency has not been established, although an interval of 6 months is recommended in UK and European guidelines, purportedly on the basis of available evidence on tumour growth rate.

Observational data suggest that HCCs detected during formal surveillance are smaller and more likely to be a single lesion than those that present symptomatically or by chance. Consequently, patients whose disease is detected as a result of surveillance are more likely to receive curative treatment than those whose diagnosis is symptomatic or incidental.

Objective

The objective of this report was to evaluate the effectiveness, cost-effectiveness and cost–utility of surveillance of patients with cirrhosis [alcoholic liver disease (ALD)-, HBV- and HCV-related], using periodic serum AFP testing and/or liver ultrasound examination, to detect HCC, followed by treatment with liver transplantation or resection, where appropriate.

Methods

Systematic review
Electronic databases were searched for randomised clinical trials of surveillance (with AFP and ultrasound) of people with cirrhosis of known underlying cause (ALD, HBV, HCV) for HCC. Updated searches were performed in March 2006.

Economic analysis
A computerised decision-analytic model was developed to compare various surveillance strategies. Comparisons were made between:

- no surveillance
- annual surveillance using AFP as the initial screening test
- annual surveillance using ultrasound alone
- annual surveillance using AFP and ultrasound
- 6-monthly surveillance using AFP as the initial screening test

▶

- 6-monthly surveillance using ultrasound alone
- 6-monthly surveillance using AFP and ultrasound.

The study modelled a population with a diagnosis of compensated cirrhosis who are also eligible to enter a surveillance programme. Those deemed eligible were aged 70 years or less, with no pre-existing medical conditions that would preclude treatment with liver transplantation or hepatic resection (including current alcohol or intravenous drug abuse).

Previously published research and a panel of clinical experts helped to inform the structure of the model. A Markov model was used to capture both natural disease progression and diagnostic and treatment pathways reflective of current best clinical practice in the UK NHS. HCC tumours are detected as a result of regular surveillance and symptomatic or incidental diagnosis. Surgical treatment of small or medium-sized tumours is predominantly by liver transplantation; liver resection is also possible, particularly in people with small tumours. Surgical treatment of people with large tumours is not possible within the confines of the model. People with decompensated cirrhosis or a surgically treatable tumour enter the transplant waiting list and have an equal chance of receiving a liver transplant. People deemed to have surgically untreatable HCC enter a range of simpler model states which simulate the costs and effectiveness of palliative care and best supportive care.

Parameter estimates were obtained from comprehensive literature reviews. No methodological restrictions were applied, but searches were limited to papers published or available in English.

The technical performance of the alternative testing strategies was modelled using decision trees. Test sensitivity for AFP and ultrasound was varied according to tumour size. Expected costs and utilities for each surveillance strategy were calculated using both a cohort and a Monte Carlo simulation approach. The model runs for the lifetime of the cohort. Costs (base year 2004) and benefits (QALYs) were discounted at 3.5%.

The model was developed to allow separate analysis of each of the three main cirrhosis aetiologies (ALD, HBV and HCV).

Uncertainty in the model was explored using extensive one-way sensitivity analyses, selected scenario analyses and probabilistic sensitivity analysis. A value of information analysis was conducted to determine the maximum possible value of further research.

Results

Systematic review
The searches returned 214 separate references. From screening of abstracts, 207 of these were excluded, leaving seven potentially relevant studies to be reviewed in full. All seven were excluded at this stage [no results for patients with cirrhosis ($n = 3$), modelling study ($n = 1$), narrative review ($n = 2$), uncontrolled cohort study ($n = 1$)].

Economic analysis
Effectiveness of surveillance
Based on the assumptions used in the model, the most effective surveillance strategy uses a combination of AFP testing and ultrasound at 6-monthly intervals. Compared with no surveillance, this strategy is estimated to more than triple the number of people with operable HCC tumours at time of diagnosis, and almost halve the number who die from HCC. This is a result of the identification of over ten times as many small HCC tumours (less than 2 cm in diameter) and over twice as many medium-sized tumours (between 2 and 5 cm in diameter). Consequently, more tumours are suitable for surgical intervention. Under the conditions of the model, this surveillance strategy would lead to an increase in the percentage of liver transplantations performed for known HCC (as opposed to decompensated cirrhosis) from 8% to 28%, compared with no surveillance.

On all effectiveness measures and at both testing frequencies, AFP- and ultrasound-led surveillance strategies are very similar. This may be because test sensitivity was varied according to tumour size, which means that AFP testing is capable of identifying many more small tumours than ultrasound. The best available evidence suggests that AFP tests will detect approximately six times as many small tumours as ultrasound. Increasing the frequency of either test to 6-monthly intervals is more effective than performing combined testing on an annual basis.

Cost of surveillance
The undiscounted lifetime cost of the surveillance strategies, including all care and treatment costs, ranges from £40,300 (annual AFP triage) to £42,900 (6-monthly AFP and ultrasound). The equivalent discounted costs are £28,400 and

£30,400. Only a small proportion (<4% of undiscounted costs) of these total costs results from the cost of the screening tests. However, screening test costs, and the cost of liver transplants and caring for people post-transplant, accounted for most of the incremental cost differences between alternative surveillance strategies.

Cost-effectiveness of surveillance

The results suggest that different surveillance strategies may provide the best value for money in patient groups of different cirrhosis aetiologies. The surveillance of people with HBV-related cirrhosis for HCC provides the best value for money, while surveillance in people with ALD-related cirrhosis provides the poorest value for money.

In people with HBV-related cirrhosis, at an assumed maximum willingness to pay (WTP) for a quality-adjusted life-year (QALY) of £30,000, both the deterministic and probabilistic cost–utility analyses suggest the optimal surveillance strategy would be 6-monthly surveillance with the combination of AFP testing and ultrasound.

In contrast, for those with ALD-related cirrhosis, annual screening with AFP as a triage test is the only surveillance strategy that is likely to be considered cost-effective at this WTP. NHS investment in more effective surveillance strategies probably represents an unacceptable cost for the extra benefits gained (e.g. £35,000 per QALY gained by moving to a 6-monthly AFP triage strategy). In addition, there is a high degree of uncertainty in the ALD model. The probabilistic analysis implies that the estimated benefits of a 6-monthly AFP triage strategy will only be worth the cost when society's WTP for a QALY exceeds around £40,000.

For people with HCV-related cirrhosis, and again applying this WTP threshold, the model suggests that the most cost-effective surveillance strategy would be surveillance with a 6-monthly AFP triage strategy.

It may not be considered practical to have separate screening strategies for people with cirrhosis of different aetiology. Results for an artificially produced mixed cohort containing people with HBV, ALD and HCV suggest that, if one surveillance policy had to be applied across cirrhosis cohorts of all three aetiologies, 6-monthly AFP with ultrasound is always the most effective option and, at commonly accepted levels of WTP

for QALYs, 6-monthly AFP as a triage test is the most cost-effective strategy.

These results should be viewed with caution for a number of reasons. Considerable uncertainty still surrounds some of the underlying parameters that influence the cost-effectiveness estimates. In addition, some of the differences between aetiologies may be more attributable to the mean age at diagnosis than to any inherent differences in the nature of disease progression. Lastly, the discounting of costs and benefits has a major impact on the cost-effectiveness results.

Conclusions

In a mixed aetiology cohort, the most effective surveillance strategy is to screen each patient with AFP assay and ultrasound imaging on a 6-monthly basis. However, when costs are taken into account it is doubtful whether ultrasound should be routinely offered to those with blood AFP of less than 20 ng/ml, unless policy makers are prepared to pay a very high price (over £60,000 per QALY) for the extra benefits achieved. Furthermore, the cost-effectiveness of surveillance for HCC varies considerably depending on the aetiology of cirrhosis; it is much more likely to be cost-effective in those with HBV-related cirrhosis, and much less likely to be cost-effective in those with ALD-related cirrhosis. This may be largely due to the younger age at diagnosis of cirrhosis in patients with HBV. This raises the possibility that there may be further subgroups of patients with ALD and HCV, diagnosed with cirrhosis at a younger age, in whom more intensive surveillance might provide value for money.

Implications for policy

The results show that surveillance strategies for HCC are effective, and can often be considered cost-effective in patients with cirrhosis. We believe that the implementation of formal surveillance programmes should be considered where they do not currently exist.

The results also suggest that different surveillance strategies in patient groups with different underlying causes of cirrhosis may provide the best value for money, if appropriate recall systems could be implemented, and also if this was judged to be ethically acceptable.

A surveillance strategy in which AFP testing is used as a triage step probably represents the best value for money.

These results also suggest a possible shift in the clinical settings where cirrhosis surveillance is conducted; as AFP triage appears to be a highly cost-effective strategy, either annually or 6-monthly, it may be more appropriate to perform the initial screening test in the primary care setting.

If effective surveillance programmes were to become widespread across the UK against a background of limited organ supply, the waiting list for liver transplants would undoubtedly increase. Detailed exploration of this was beyond the scope of this project, but preliminary findings suggest that this might be an important issue.

Recommendations for further research

Model development
- Extensive value of information analysis should be used to identify which parameters or groups of parameters contribute most to the uncertainty in the cost–utility results, and therefore suggest priorities for further primary research.
- Alternative modelling methods should be used to account for heterogeneity in the patient population, so that the impact of factors such as tumour growth rate, tumour characteristics and variability in individual patients' serial test results may be accurately assessed. Such methods could also be used to investigate the optimal surveillance strategy, optimal surveillance interval and the effects of surveillance on waiting lists for liver transplantation.
- Further investigation is needed into the accepted cut-off levels for AFP tests and how different cut-off levels impact on the effectiveness and cost-effectiveness of surveillance for HCC.
- Further modelling studies should investigate innovative surveillance strategies not currently undertaken in clinical practice.
- Further modelling studies should investigate the impact of alternative treatment modalities (e.g. more resection of small tumours, radiofrequency ablation as a 'curative' treatment of small tumours), because identifying more operable HCC tumours will probably lead to longer transplant waiting lists.
- Further modelling is needed of the impact of age at diagnosis of cirrhosis on the cost-effectiveness of surveillance strategies.

- Anecdotal reports suggest that non-alcohol fatty liver disease is increasing in incidence and will soon represent the second largest cause of cirrhosis in the UK. Further modelling studies are needed to assess the effectiveness and cost-effectiveness of surveillance in this patient group.

Other research
- The effectiveness and cost-effectiveness of microbubble ultrasound technology to detect HCC tumours should be evaluated, and the test performance in various stages of cirrhosis/aetiologies compared with explant pathology.
- Epidemiological research in the UK is needed to assess the incidence and rate of tumour growth of HCC in different cirrhosis aetiologies.
- The association between the level of AFP secreted and tumour size in different cirrhosis aetiologies needs to be assessed.
- Detailed observational research is needed on the epidemiology and natural history of ALD-related cirrhosis. Despite existing evidence that ALD accounts for the majority of the UK's disease burden of cirrhosis, and emerging evidence that alcohol consumption is rising, ALD-related cirrhosis remains particularly poorly described in the literature.
- Observational studies could be conducted which collect AFP measurements on the same population of people with cirrhosis over time, and investigate the relationship between the emergence or presence of HCC tumours and patterns of change in AFP levels over time (as opposed to the predictive ability of particular absolute AFP thresholds).
- Quality of life studies should assess the utility of all stages of disease, during assessment for treatment, during and post-treatment in all cirrhosis aetiologies in a UK population.

Publication

NIHR Health Technology Assessment Programme

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

The research findings from the HTA Programme directly influence decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC). HTA findings also help to improve the quality of clinical practice in the NHS indirectly in that they form a key component of the 'National Knowledge Service'.

The HTA Programme is needs-led in that it fills gaps in the evidence needed by the NHS. There are three routes to the start of projects.

First is the commissioned route. Suggestions for research are actively sought from people working in the NHS, the public and consumer groups and professional bodies such as royal colleges and NHS trusts. These suggestions are carefully prioritised by panels of independent experts (including NHS service users). The HTA Programme then commissions the research by competitive tender.

Secondly, the HTA Programme provides grants for clinical trials for researchers who identify research questions. These are assessed for importance to patients and the NHS, and scientific rigour.

Thirdly, through its Technology Assessment Report (TAR) call-off contract, the HTA Programme commissions bespoke reports, principally for NICE, but also for other policy-makers. TARs bring together evidence on the value of specific technologies.

Some HTA research projects, including TARs, may take only months, others need several years. They can cost from as little as £40,000 to over £1 million, and may involve synthesising existing evidence, undertaking a trial, or other research collecting new data to answer a research problem.

The final reports from HTA projects are peer-reviewed by a number of independent expert referees before publication in the widely read monograph series *Health Technology Assessment*.

Criteria for inclusion in the HTA monograph series
Reports are published in the HTA monograph series 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 referees 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.