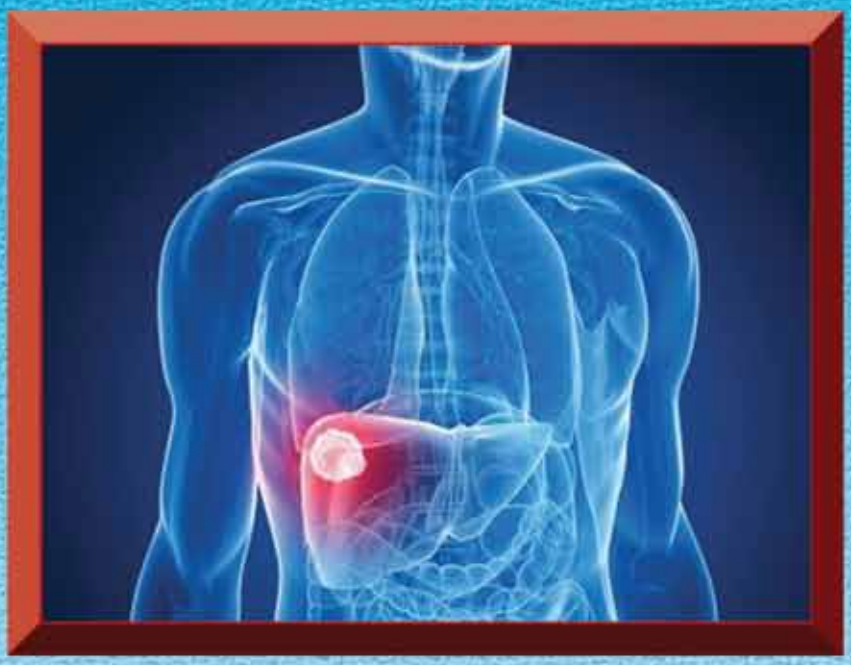
2012

HEALTH TECHNOLOGY ASSESSMENT REPORT

MOH/P/PAK/256.12(TR)

MINISTRY OF HEALTH MALAYSIA



SERUM ALPHA-FETOPROTEIN (AFP) AND/OR ULTRASOUND (US) FOR HEPATOCELLULAR CARCINOMA (HCC) SCREENING

MaHTAS

Malaysian Health Technology Assessment Section

MEDICAL DEVELOPMENT DIVISION
MINISTRY OF HEALTH MALAYSIA

Published by
Malaysian Health Technology Assessment Section, (MaHTAS)
Medical Development Division, Ministry of Health Malaysia,
Level 4, Block E1, Complex E, Precinct 1,
Federal Government Administrative Centre
62590, Putrajaya, Malaysia

ISBN: 978-967-0399-36-2

MINISTRY OF HEALTH MALAYSIA

Health Technology Assessment Report

SERUM ALPHA-FETOPROTEIN (AFP) AND/OR ULTRASOUND (US) FOR HEPATOCELLULAR CARCINOMA (HCC) SCREENING

DISCLAIMER

This Health Technology Assessment has been developed from analysis, interpretation and synthesis of scientific research and/or technology assessment conducted by other organizations. It also incorporates, where available, Malaysian data, and information provided by experts to the Ministry of Health Malaysia. While effort has been made to do so, this document may not fully reflect all scientific research available. Additionally, other relevant scientific findings may have been reported since completion of the review.

Please contact : www.htamalaysia@moh.gov.my, if you like further information.

For further information please contact:
Health Technology Assessment Section (MaHTAS),
Medical Development Division
Ministry of Health Malaysia
Level 4, Block E1, Precinct 1
Government Office Complex
62590 Putrajaya
Tel: 603 8883 1246
Fax: 603 8883 1230

Available at the following website:
http://www.moh.gov.my

AUTHORS

Mr. Syful Azlie bin Md Fuzi
Senior Assistant Director
Health Technology Assessment Section (MaHTAS)
Medical Development Division
Ministry of Health Malaysia

Dr. Roza binti Sarimin
Senior Principal Assistant Director
Health Technology Assessment Section (MaHTAS)
Medical Development Division
Ministry of Health Malaysia

EXPERT COMMITTEE

Dr. Nor Saleha binti Ibrahim Tamin
Senior Principal Assistant Director
Disease Control Division
Ministry of Health Malaysia

Dr. Syed Mohd Redha bin Syed Nasir
Consultant Hepatologist
Department of Hepatology
Hospital Selayang, Selangor

Prof. Dr. Rosmawati binti Mohamed
Consultant Hepatologist
Faculty of Medicine
University Malaya, Kuala Lumpur

Dr. Noraini binti Abdul Rahim
Consultant Radiologist & Head of Department
Department of Radiology (Diagnostic Imaging)
Hospital Serdang, Selangor

Dr. Tengku Norita binti Tengku Yazid
Consultant Chemical Pathologist
Department of Patology
Hospital Selayang, Selangor

Datin Dr. Rugayah binti Bakri
Head, Health Technology Assessment Section (MaHTAS),
Deputy Director
Medical Development Division
Ministry of Health Malaysia

EXTERNAL REVIEWERS

Assoc. Prof. Dr. Razman bin Jarmin
Consultant Hepatobiliary Surgeon & Head of Department
Department of Surgery
Faculty of Medicine
National University of Malaysia

Dr. Tan Soek Siam
Senior Consultant Clinical Hepatologist & Head of Department
Department of Hepatology
Hospital Selayang, Selangor
(Head, Hepatology Service, Ministry of Health Malaysia)

Datin Dr. Zaharah binti Musa
Senior Consultant Radiologist & Head of Department
Department of Radiology (Diagnostic Imaging)
Hospital Selayang, Selangor
(Head, Radiology Service, Ministry of Health Malaysia)

Dr. Sia Koon Ket
Consultant Gastroenterologist
Hospital Ampang, Selangor

Dr Muhammad Arif bin Mohd Hashim
Consultant Chemical Pathologist & Head of Department
Department of Pathology
Hospital Kuala Lumpur

ACKNOWLEDGEMENT

The authors for this Health Technology Assessment Report would like to express their gratitude and appreciation to the following for their contribution and assistance:

- Health Technology Assessment and Clinical Practice Guidelines Council.
- Technical Advisory Committee for Health Technology Assessment.
- Madam Sin Lian Thye and Mr. Fadhlurahman bin Kamarudin from MaHTAS for their contribution in retrieval of the evidence.

DISCLOSURE

The authors of this report have no competing interest in this subject and the preparation of this report is totally funded by the Ministry of Health, Malaysia.

EXECUTIVE SUMMARY

Background

Most primary liver cancers are classified as hepatocellular carcinoma (HCC). According to the World Health Organization (WHO) and GLOBOCAN 2008, liver cancer is the seventh most common form of cancer worldwide and the third leading cause of cancer-related death globally. While more recent data is yet to be published, going by the numbers provided in the National Cancer Registry (NCR) in 2006, liver cancer was then already ranked the sixth most frequent cancer, fifth among males and ninth among females in Malaysia. Chronic hepatitis B virus (HBV) affects around a million patients in Malaysia (2004) which accounts for majority of the diagnosed HCC (> 80.0%). However, there is currently no formal/ structured national liver cancer screening programme being implemented. Screening for HCC in the high-risk group using test method such as serum alpha-fetoprotein (AFP) and imaging method such as ultrasound (US) on the other hand, has become widely applied despite lack of published evidence of benefit on these various methods used. With the significant burden of liver cancer globally and locally, one of the strategies for early detection of cancer in the Malaysian National Cancer Management Blueprint 2008-2015 is to provide service on liver cancer screening. Therefore, this Health Technology Assessment (HTA) is requested to review evidence on the efficacy/effectiveness and cost-effectiveness of using serum AFP and/or US for HCC detection in the high-risk group in MOH facilities. The high-risk group for developing HCC include chronic liver infection due to hepatitis and liver cirrhosis. With regards to hepatitis, the common infection is HBV and infrequently common is the HCV infection.

Technical features

Two commonly used methods for liver cancer screening are serum alpha-fetoprotein (AFP) and ultrasound (US) examination of the liver. However, there are limitations to the usefulness of these methods. It is uncertain which method is better or whether the two should be combined.

Policy Question

i. Should a screening programme for HCC in the high-risk group be established as part of the Malaysian National Cancer Control Programme?

ii. Which methods namely using serum AFP alone or US alone or combined is most suitable to be used for HCC detection in the high-risk group?

Objective

i. To determine the benefits of HCC screening programme in the high-risk group using serum AFP and/or US compared with no screening, with regards to patient outcomes such as detection rate, mortality rate, survival rate, quality of life (QOL) and quality adjusted life years (QALY) gained.

ii. To determine the diagnostic accuracy of serum AFP and/or US for HCC detection in the high-risk group.

iii. To determine the cost-effectiveness of using serum AFP and/or US for HCC detection in the high-risk group

Methods

The following electronic databases were searched: MEDLINE (1950-Week 3 March 2012), EBM Reviews-Cochrane Database of Systematic Reviews (2005 to May 2012), EBM Reviews-Cochrane Central Register of Controlled Trials (1st Quarter 2012), EBM Reviews-HTA Databases (1st Quarter 2012), EBM Reviews-Cochrane Methodology Register (1st Quarter 2012), EBM Reviews-ACP Journal Club (1991 to May 2012), EBM Reviews-NHS Economic Evaluation Database (1st Quarter 2012) via OVID, PubMed, INAHTA database, HTA database and US FDA database. The last search was run on 10 May 2012. No limits were applied to the search. Detailed search strategy is as in **Appendix 4**. Additional articles were identified from reviewing the bibliographies of retrieved articles and hand-searching of journals. General search engine was used to get additional web-based information.

Results and conclusion

Effectiveness of HCC screening using serum AFP and/or US

The available evidences on HCC screening for mortality and survival from three large randomised controlled trials and one cohort study were conflicting. However, most studies showed that using serum AFP and/or US were more effective than no screening. Zhang BH *et al.* trial in Shanghai reported a significantly lower HCC mortality rate in the screened group than in controls, being 83.2 per 100,000 and 131.5 per 100,000 respectively with a mortality rate ratio of 0.63 (95% CI: 0.41 to 0.98). In contrast, two other trials, Yang *et al.* (Shanghai) and Chen *et al.* (Qidong) did not differ significantly regarding HCC mortality. Yang *et al.* and Zhang BH *et al.* reported the overall survival rate at five-year was better in screened group than in controls (52.7% and 46.4%, respectively), whereas there was no difference in five-year survival between the screened group and the control group in Chen *et al.* trial. A prospective 16-year population-based cohort study showed a significant survival benefits at five and 10 years in screened hepatitis B surface antigen positive (HBsAg+ve) Alaskan native carriers compared with historical controls (42.0% and 30.0%, respectively).

Diagnostic accuracy of serum AFP and/or US

There was good level of evidence to suggest that the sensitivity and specificity of serum AFP varies with the cut-off value or positivity threshold. At cut-off value between 20 ng/mL and 200 ng/mL, the sensitivity of AFP alone ranged from 41.0% to 80.0% and 20.0% to 45.0%, whereas specificity ranged from 80.0% to 95.0% and 99.0% to 100.0%, respectively. For US alone, the sensitivity varies from 60.0% to 94.0% and specificity from 94.0% to 97.1% in detecting HCC tumour nodules, varied by characteristics of liver diseases in screened patients, screening frequency, operator expertise as well as on the type of US equipment available. The sensitivity and specificity improved when a combination of AFP and US were used sequentially, at 92.2% and 95.0% respectively, particularly for HCC related with chronic liver infection (HBV). However for HCC related to cirrhosis, the combination of AFP and US gave the sensitivity of 69.0%. The overall positive predictive value (PPV) ranged from 3.0% to 6.6% whereas false positive rate ranged from 2.9% to 7.5%.

Economic evaluation

There was good level of evidence to suggest that:

i. The use of US alone at 6-month intervals in detecting HCC was not only more effective but cheaper than AFP test alone. However, the combined tests not only increased the efficacy of HCC detection but also increased the false positive rate and the cost

ii. The use of US at 12-month intervals and AFP at 6-month intervals was a reasonable strategy, offering the greatest gain in life-expectancy while still maintaining an ICER < USD$50,000 (RM150,000) per QALY

iii. At willingness to pay threshold of £30,000 (RM150,000) per QALY, the most cost-effective strategy is 6-monthly AFP-triage with ICER of £27,600 (RM138,000) per QALY gained

However, economic evaluation review was subjected to several limitations. It must be emphasized that the cost-effectiveness of HCC screening were assessed by retrospective analysis or by using decision models. Although retrospective studies suffered from selection bias, decision-analysis models were based on simulation of costs and health outcomes and therefore, their results may vary greatly according to different assumptions, such as the incidence of HCC in the screening population, the screening interval, the modality of diagnosis, the type of treatment after diagnosis, the doubling time of tumours, and so forth. In the Malaysian context, the fees charged by MOH hospital for serum AFP is approximately [redacted] per test, while US varied from [redacted] to [redacted] per imaging. Ultrasound machines cost about [redacted] and range up to [redacted]. The price depends largely on the level of complexity of the machine.

Recommendation

Based on this review, good level of evidence on effectiveness (with respect to mortality and survival rate) showed that there was benefits in screening for HCC using serum AFP and/or US in the high-risk group and hence, can be established as part of the Malaysian National Cancer Control Programme. The decision to enter a patient into a *screening programme is determined by the level of risk for HCC and hence, **surveillance is recommended for the following groups of patients:

i. Hepatitis B carriers:
   - Asian males ≥ 40 years
   - Asian females ≥ 50 years

ii. All cirrhotic hepatitis B regardless of age

iii. Family history of HCC

iv. Liver cirrhosis
   - Hepatitis C
   - Alcoholic cirrhosis
   - Genetic hemochromatosis
   - Primary biliary cirrhosis

There was also good level of evidence to show that the combination of serum AFP and US is the most suitable method to be used for HCC detection, particularly for HCC related with chronic liver infection due to HBV. In addition, the recommended cut-off level of serum AFP was ≥ 20.0 ng/mL, as evidence showed that there was optimal balance between sensitivity and specificity in detecting HCC at this cut-off level.

From the cost-effectiveness perspective, most of the studies in the review indicated that 12-months screening interval using serum AFP plus US was as cost-effective as the 6-months interval using serum AFP alone. Hence, the screening interval of 6 to 12 months was a reasonable cost-effective strategy for surveillance of HCC.

However, before commencing the screening programme for HCC detection, it should be noted that currently in Malaysia, serum AFP test are conducted at laboratory hospitals with immunoassay facilities which covers MOH state hospitals and hospitals with specialist amounting to 36 MOH hospitals (personal communication with Head of Chemical Pathology Activities, MOH). Meanwhile, US examination of the liver is only conducted in 39 MOH hospitals with radiologist (personal communication with Head of Radiology Service, MOH).

Footnote:

***Screening –** application of diagnostic tests in patients at risk for HCC, but in whom there is no a priori reason to suspect that HCC is present.

****Surveillance –** the repeated application of screening tests.

TABLE OF CONTENTS

ABBREVIATIONS

HCC	Hepatocellular carcinoma
WHO	World Health Organization
NCR	National Cancer Registry
ASR	Age-standardised incidence rate
HBV	Hepatitis B virus
HCV	Hepatitis C virus
HBsAg	Hepatitis B surface antigen
HBeAg	Hepatitis B e-antigen
PLC	Primary liver cancer
OLT	Orthotopic liver transplantation
ALD	Alcohol liver disease
AFP	Alpha-fetoprotein
US	Ultrasound
USPSTF	United States Preventive Services Task Force
AASLD	American Association for the Study of Liver Disease
EASL	European Association for the Study of Liver
HTA	Health Technology Assessment
CT	Computed tomography
MRI	Magnetic resonance imaging
TACE	Transarterial chemoembolization
PEI	Percutaneous ethanol injection
QOL	Quality of life
QALY	Quality adjusted life year
LY	Life year
RCT	Randomised controlled trial
PPV	Positive predictive value
NPV	Negative predictive value
CASP	Critical Appraisal Skills Programme
CI	Confident interval
OR	Odds ratio
SROC	Summary receiver operator characteristics curve
AUC	Area under the curve
LR	Likelihood ratio
cm	centimetre
ng/mL	nano gram per millilitre
µg/L	micro gram per litre
ICER	Incremental cost-effectiveness ratio

HEALTH TECHNOLOGY ASSESSMENT (HTA) SERUM ALPHA-FETOPROTEIN (AFP) AND/OR ULTRASOUND (US) FOR HEPATOCELLULAR CARCINOMA (HCC) SCREENING

1.0 BACKGROUND

Most primary liver cancers are classified as hepatocellular carcinoma (HCC). According to World Health Organization (WHO) and GLOBOCAN 2008, liver cancer is the seventh most common form of cancer worldwide and the third leading cause of cancer-related death globally. The occurrence of HCC varies widely depending on geographic location. The highest liver cancer rates are found in Eastern and South Eastern Asia, Middle and Western Africa, whereas rates are low in Central and Eastern Europe, as well as in Northern Europe and South-Central Asia (**Figure 1**).[1, 2, 3] While more recent data is yet to be published, going by the numbers provided in the National Cancer Registry (NCR) in 2006, liver cancer was then already ranked the sixth most frequent cancer, fifth among males and ninth among females. A total of 793 cases were registered with NCR in 2006 and which comprises of 568 males and 225 females. The incidence of liver cancer increased with age with the overall age-standardised incidence rate (ASR) of 4.9 per 100,000 populations. It was higher among males (ASR of 7.2 per 100,000) compared to females (ASR of 2.7 per 100,000).[4]

Hepatocellular carcinoma (HCC) is strongly associated with chronic liver infection or hepatitis, especially hepatitis B (HBV) and hepatitis C (HCV) virus infection. In Asia, the high incidence of chronic liver infection is the main cause of HCC. Other important risk factors include liver cirrhosis from excessive alcohol consumption as well as ingestion of aflatoxin, a substance which is found in mouldy nuts and grain. Hepatocellular carcinoma (HCC), however, is not hereditary and therefore do not run in the families in absence of the above risk factors.[5, 6]

The liver is the largest solid organ in the human body. It is shaped like a pyramid and is divided into right and left lobes (**Figure 2**). It has a rich blood supply coming from both arterial and venous systems, namely the hepatic artery and portal vein, making it a common site of spread for cancers from other organs, such as the colon and breast. Primary liver cancer or HCC on the other hand, arises from the liver cell itself. The liver does important work to keep human body healthy: [5, 6]

- It breaks down and stores many of the nutrients absorbed from the intestine
- It makes some of the clotting factors needed to stop bleeding from a cut or injury
- It makes bile that goes into the intestine to help absorb nutrients
- It filters out and breaks down toxic wastes in the blood, which are then removed from the body

Liver cells (hepatocytes) can become cancerous when there is a breakdown of normal cell processes. The cancerous liver cells grow out of control and developed into a tumour, called primary liver cancer or HCC. The treatment of HCC is often complicated because many patients also have chronic liver disease. Secondary liver cancer, or metastases of other cancers from other organs, also grows in the liver, but it is made up of cells that have travelled (metastasized) from another part of the body (such as colon, stomach, lung, breast, lung, etc.). Once liver cancer has spread to other parts of the body, advanced cancer treatments may be used to try to slow the progress of the disease. At this stage, palliative care is also offered to all patients to reduce their pain and control other symptoms with prolongation of life as a secondary objective.[7]

Screening for HCC offers the best hope for early detection of preclinical or early symptomatic phase which can improve survival. Many physicians screen patients in high-risk groups with either serum alpha-fetoprotein (AFP) or ultrasound (US) of the liver or both. These include people with known cirrhosis, especially if it is so bad that the patient is waiting to get a liver transplant. Otherwise a cancer may start during the wait and become so advanced that it cannot be cured. Some people with chronic hepatitis B or hepatitis C infections should also be screened, like those with liver cancer in the family. For other people at higher risk, the benefits of screening may not be as clear.[7]

There are guidelines developed by various medical entities for HCC screening. However, their recommendations are conflicting because the true benefit of screening remains uncertain. Several major organizations, including the United States Preventive Services Task Force (USPSTF), a group of experts convened by the U.S. Public Health Service, National Comprehensive Cancer Network, and the American Cancer Society do not have any specific guidelines for screening patients for HCC. The United States National Cancer Institute is against routine screening for lack of survival benefit. More recently, American Association for the Study of Liver Diseases (AASLD) and European Association for the Study of the Liver (EASL) recommended ultrasound of the liver every six months for high-risk patients.[8] On the other hand, the Hong Kong Guidelines on Cancer Prevention, Early Detection & Screening for Liver Cancer stated that high-risk group patients such as chronic hepatitis carrier and known liver cirrhosis patients should receive blood tests for liver function and AFP together with US of liver every sixth to ninth months.[9]

The diagnosis of HCC raises many questions and there is a need for clear, understandable answer. By definition, screening is the one-time application of a test that allows the detection of a disease at a stage when intervention may significantly improve the natural course and outcome. In contrast, surveillance is the repeated application of such tests over time. The objective of both is to reduce disease specific mortality. Although there is clearly a need for early diagnosis of HCC, there is considerable controversy about the role of screening and surveillance in its management. It appears that some form of screening and surveillance for HCC is widely practiced in patients with chronic liver disease, but the optimal methods for surveillance on patient survival remain uncertain.[10]

With the significant burden of liver cancer globally and locally, one of the strategies for early detection of cancer in the Malaysian National Cancer Management Blueprint 2008-2015 is to provide service on liver cancer screening. Therefore, a Health Technology Assessment (HTA) is required to review evidence on the efficacy/effectiveness and cost-effectiveness of using serum AFP and/or US for HCC detection in the high-risk group in MOH facilities. This HTA was requested by the Senior Principal Assistant Director of Cancer Unit, Disease Control Division, Ministry of Health Malaysia.

Figure 1: Age-standardised liver cancer incidence rate by sex and world area. Source: GLOBOCON 2008

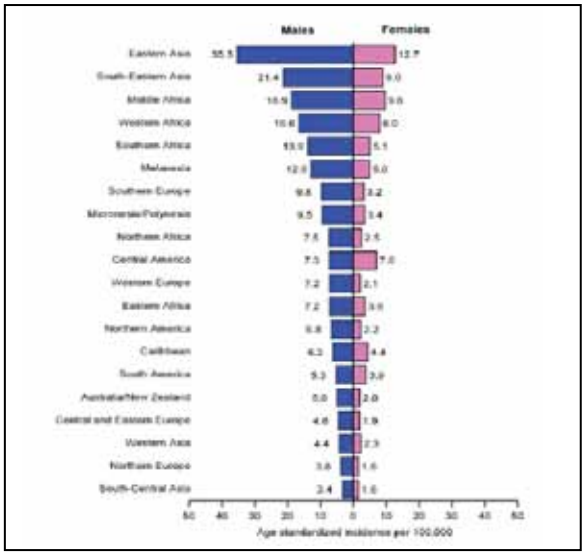
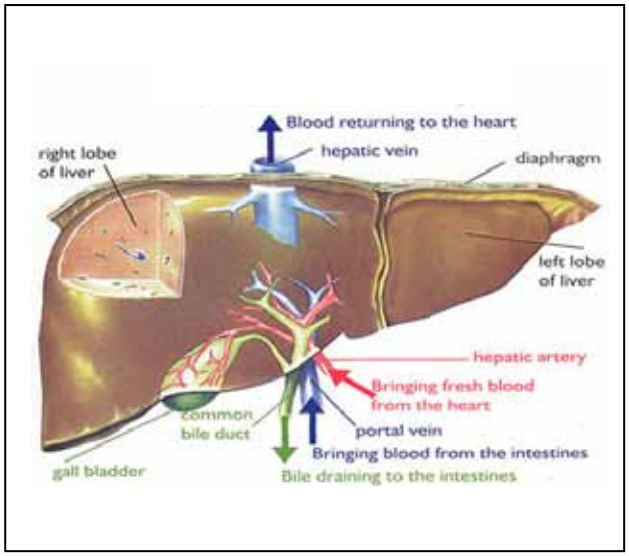


Figure 2: Liver anatomy



2.0 TECHNICAL FEATURES

2.1 Screening methods for hepatocellular carcinoma (HCC)

Because of the liver's size and location, it is impossible to detect liver tumours upon physical examination. The most commonly used screening methods are serum alpha-fetoprotein (AFP) and ultrasound (US) (**Figure 3**).

Figure 3: Serum alpha-fetoprotein (AFP) test and Ultrasound (US) machine

2.1.1 Serum alpha-fetoprotein (AFP)

Serum AFP is a protein produced primarily by fetal liver and the portion of a developing embryo that is similar to the yolk cavity in bird eggs (yolk sac tissues). Its concentration is typically elevated when a baby is born and decline rapidly. In healthy children and non-pregnant adults, AFP is normally only detectable at very low levels. Liver damage and certain cancers such as testicular and ovarian cancers can increase AFP concentrations significantly. It is also produced whenever liver cells are regenerating such as with chronic liver diseases and tumours. Generally, normal range of AFP is < 10 ng/mL. Moderate levels of AFP (> 500 ng/ml) can be seen in patients with chronic hepatitis. Moreover, many patients with various types of acute and chronic liver diseases without documentable liver cancer can have mild or even moderate elevations of AFP. Serum AFP is used in the screening and diagnosis of liver cancer because AFP is above the normal range in 60.0% to 70.0% of primary liver cancer cases. Unfortunately, AFP levels are normal in 30.0% to 40.0% of all liver cancer, in which case cancer can only be detected by finding a mass on ultrasound or computed tomography (CT) scan.[11, 12]

2.1.2 Ultrasound (US)

Limitations in the sensitivity and specificity of AFP in surveillance of high-risk populations led to the use of US as an additional method for detection of HCC. Ultrasound (US) uses reflected sound waves to produce images of organs and other structures in the body. It is commonly used as an initial test for detecting solid lesions in the liver. If a tumour exists, the US may produce a different echo pattern than it does with healthy tissue. The machine will record the images of the liver and a report will be made by the radiologist. The report will only show the surface and the shape of the liver but will not show a detailed view of any abnormalities that may exist in the liver. Ultrasound (US) is very operator-dependent. This means that unlike regular x-rays, the ability of the person doing the US is important. The quality of US examination is very variable. Furthermore, in the presence of cirrhosis, ultrasound becomes even less reliable, increasing the probability of missing a significant number of cancers.[6, 7]

2.2 Confirmatory tests for hepatocellular carcinoma (HCC)

Previously, abnormal screening results may lead to liver biopsy for confirmation of diagnosis. However, the use of biopsy to confirm HCC is controversial. It can be difficult to distinguish large cirrhotic nodules from well-differentiated HCC or low grade dysplastic nodules from HCC in either needle or wedge biopsies. Liver biopsy also carries a small risk of tumour spread along the needle track, and it can be painful. New technology can reduce the risk of pain and complications from liver biopsy. A variety of radiological investigations have been used to confirm ultrasound findings in patients with cirrhosis and chronic hepatitis with an isolated raised AFP. These include computed tomography (CT), magnetic resonance imaging (MRI), lipiodol-CT, and hepatic angiography.[13]

2.3 Treatment alternative for hepatocellular carcinoma (HCC)

Treatment choices depend on the type and stage of the cancer, how well the liver is functioning, and the overall health of the individual. For patients with localised disease, surgery is the treatment of choice and generally offers the only chance of cure. This involves removing the whole tumour or tumours with a margin of normal liver tissue. Sometimes a whole lobe of liver may need to be removed but as the human body requires only about 25 per cent of a normal liver to function, this can be done safely. For locally advanced disease which cannot be removed, local therapies may be attempted to reduce the size of the tumour in the liver and relieve the patient's symptoms. These include arterial infusion of chemotherapy or lipiodol into the tumour, ultrasonic radiofrequency ablation and direct ethanol injections. Intravenous chemotherapy may also be given. However, the effectiveness of these therapies is limited and the tumours do not generally disappear with treatment. For patients with advanced disease, treatment is generally aimed at relieving symptoms with prolongation of life as a secondary objective.[5]

3.0 POLICY QUESTION

3.1 Should a screening programme for HCC in the high-risk group be established as part of the Malaysian National Cancer Control Programme?

3.2 Which methods namely using serum AFP alone or US alone or combined is most suitable to be used for HCC detection in the high-risk group?

4.0 OBJECTIVE

4.1 To determine the benefits of HCC screening programme in the high-risk group using serum AFP and/or US compared with no screening, with regards to patient outcomes such as detection rate, mortality rate, survival rate, quality of life (QOL) and quality adjusted life years (QALY) gained

4.2 To determine the diagnostic accuracy of serum AFP and/or US for HCC detection in the high-risk group

4.3 To determine the cost-effectiveness using AFP and/or US for HCC detection in the high-risk group

5.0 METHODS

5.1 Literature search strategy

Literature search was done by two Information Specialists who searched for published articles pertaining to the use of serum AFP and/or US for HCC screening. The following electronic databases were searched: MEDLINE (1950-Week 3 March 2012), EBM Reviews-Cochrane Database of Systematic Reviews (2005 to May 2012), EBM Reviews-Cochrane Central Register of Controlled Trials (1st Quarter 2012), EBM Reviews-HTA Databases (1st Quarter 2012), EBM Reviews-Cochrane Methodology Register (1st Quarter 2012), EBM Reviews-ACP Journal Club (1991 to May 2012), EBM Reviews-NHS Economic Evaluation Database (1st Quarter 2012) via OVID, PubMed, INAHTA database, HTA database and US FDA database. The last search was run on 10 May 2012. No limits were applied to the search. Detailed search strategy is as in **Appendix 4**. Additional articles were identified from reviewing the bibliographies of retrieved articles and hand-searching of journals. General search engine was used to get additional web-based information.

5.2 Study selection

Based on the policy question the following inclusion and exclusion criteria were used:-

5.2.1 Inclusion criteria

a.	Study design	HTA reports, systematic review, randomised controlled trial (RCT), diagnostic accuracy studies, cross-sectional, cohort, case-control, and economic evaluation studies.
b.	Population	High-risk group: i. Chronic hepatitis B carriers ii. Chronic hepatitis C carriers iii. Known liver cirrhosis patients
c.	Intervention	i. AFP ii. US iii. Combination of AFP and US
d.	Comparator	i. No screening ii. Usual care
e.	Outcomes	i. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). ii. Detection rate, mortality rate, survival rate, quality of life (QOL), and quality adjusted life years gained (QALY) gained. iii. Cost, cost-benefit, cost-effectiveness, cost utility, and economic evaluation
f.	Publication	Full text articles published in English

5.2.2 Exclusion criteria

Based on the policy question the following inclusion and exclusion criteria were used:-

i. Animal study
ii. Narrative review
iii. Experimental study
iv. Non English full text articles

Based on the above inclusion and exclusion criteria, study selection were carried out independently by two reviewers. The titles and abstracts of all studies were assessed for the above eligibility criteria. If it is absolutely clear from the title and/or the abstract that the study was not relevant, it was excluded. If it was unclear from the abstract and/or the title, the full text article was retrieved. Two reviewers assessed the content of the full text articles. Disagreement was resolved by discussion.

5.3 Quality assessment strategy

The methodological quality of all the relevant full text articles retrieved was assessed using the Critical Appraisal Skills Programme (CASP) tool depending on the type of study design.[14] Quality assessment was conducted by two reviewers. Disagreements were resolved by discussion. All full text articles were graded based on guidelines from the U.S./Canadian Preventive Services Task Force (**Appendix 1**) or NHS Centre for Reviews and Dissemination (CRD) University of York, Report Number 4(2nd Edition), March 2001 for test accuracy studies (**Appendix 2**).[15, 16]

5.4 Data extraction strategy

Data were extracted from included studies by a reviewer using a pre-designed data extraction form (evidence table as shown in **Appendix 6**) and checked by another reviewer. Disagreements were resolved by discussion and the extracted data was presented and discussed with the expert committee. The data extracted was as follows: (1) Details of methods including study design, (2) Study population characteristics including age, trial inclusion and exclusion criteria, (3) Details of intervention and comparator, and (4) Types of outcome measures including diagnostic accuracy (sensitivity, specificity, PPV, and NPV), effectiveness of HCC screening (detection rate, mortality rate, survival rate, QOL, and QALY gained), cost, cost-benefit, cost-effectiveness, cost utility, and economic evaluation of using serum AFP and/or US for HCC detection.

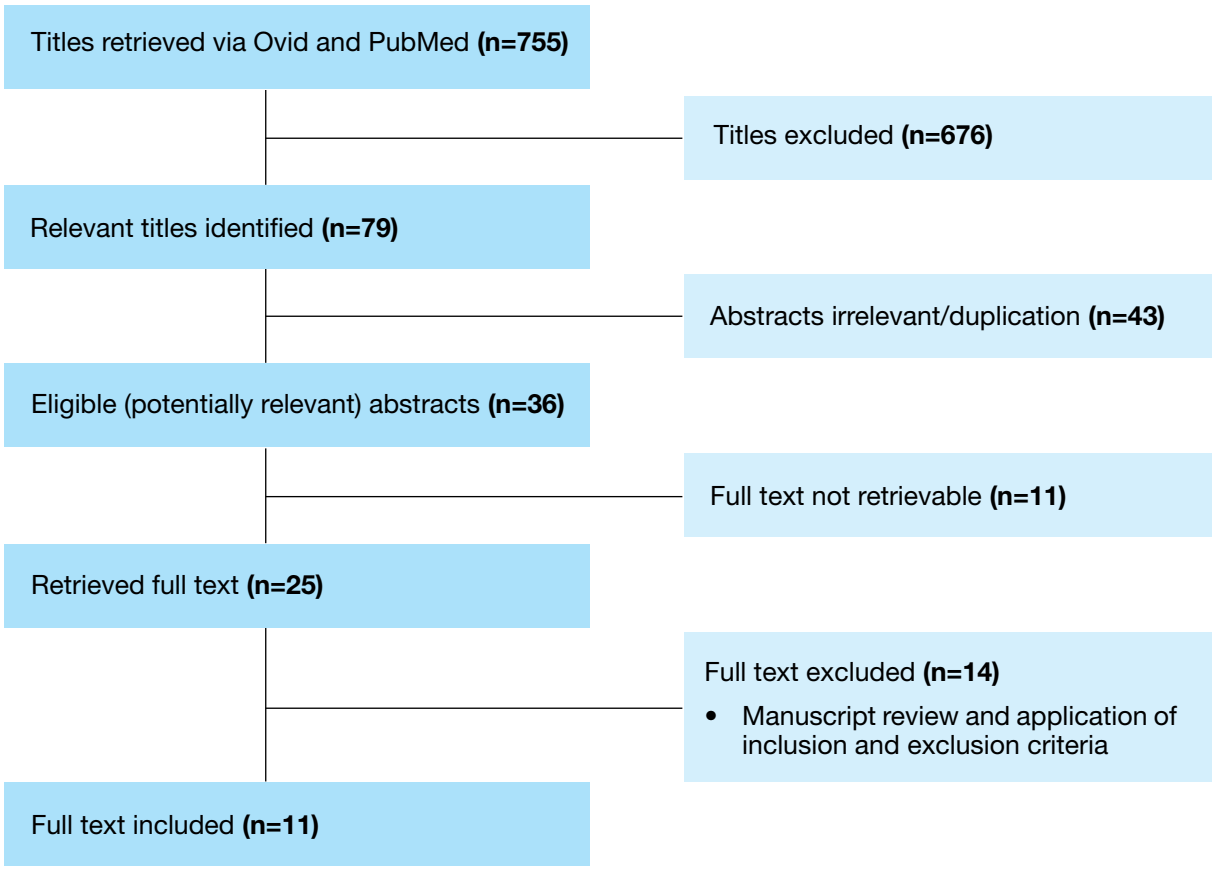
5.5 Methods of data synthesis

Data on the diagnostic accuracy, effectiveness, and cost-effectiveness of AFP and/or US for HCC screening were presented in tabulated format with narrative summaries. No meta-analysis was conducted for this review.

6.0 RESULTS

Search strategies yielded many published articles related to serum AFP and/or US for HCC screening. A total of **755** titles were retrieved and **79** relevant titles were identified. Of these, **43** abstracts were found to be irrelevant or duplication. **Thirty-six** potentially relevant abstracts were attempted for retrieval. Of these, **25** potentially relevant abstracts were retrieved in full text, and full text for **11** abstracts cannot be retrieved. After reading and appraising the full text articles, **11** full text articles were included, as shown in **Figure 4**. **Fourteen** full text articles were excluded based on inclusion and exclusion criteria, and quality of the studies are listed in **Appendix 7**. The articles comprised three systematic reviews, two systematic review with meta-analysis, three RCTs, one cohort study, and two economic evaluation papers. We also included one article by the World Health Organization (WHO).

Figure 4: Flow chart of retrieval of articles used in the results



6.1 Diagnostic accuracy of serum afp and/or us

Cochrane systematic review on alpha-fetoprotein (AFP) and/or liver ultrasonography (US) for liver cancer screening in patients with chronic hepatitis B in 2003 and edited in 2009 included two randomised controlled trials. The included trials were from Shanghai, China (Yang *et al.* 1999) and Toronto, Canada (Sherman *et al.* 1995). The Yang *et al.* trial compared biannual AFP plus US screening with no screening in 18,816 urban residents aged 35 to 55 years for five years. They randomised 9,373 people into the screened group and 9,443 people into the control group; all aged 35 to 55 years. The diagnosis of HCC was made by CT scan and/or MRI. In contrast, the Sherman *et al.* trial compared biannual AFP plus US screening with AFP in 1,069 urban residents aged over 18 years (mean 39 ± 12 years) for five years. They were randomised into two arms: one (n=538) screened with AFP plus US and another (n=531) with AFP. The diagnosis of HCC was made by histology or CT scan. The review showed that in the Yang *et al*. trial, the sensitivity was 68.6% for AFP alone, 84.3% for US alone, and 92.2% for AFP plus US. They calculated a specificity of 95.0% for AFP alone, 97.1% for US alone, and 95.0% for AFP plus US. The PPV was 3.3%, 6.6%, and 3.0% for AFP, US, and AFP plus US, respectively whereas the NPV was 99.0% for AFP and US, and 100.0% for AFP plus US. Unfortunately, the results from Sherman *et al.* trial indicated that sample size was not adequate to compare the two screening methods, and the outcome of screening was difficult to know exactly because the data and the figures were not clearly given. The authors concluded that there were not enough quality trials to support or refute screening of HBsAg-positive patients for HCC. It is possible that screening may be effective, but also that harm caused by screening treatment may outweigh any gain.[17 level 3]

A systematic review with meta-analysis by Singal *et al.* using the MEDLINE and SCOPUS databases through 1st July 2007 yielded 19 studies that evaluated the accuracy (performance characteristics) of US for HCC screening. There was total of 6,555 cirrhotic patients from the included studies. The first aim of the study was to determine the sensitivity and specificity of surveillance using US to detect HCC at any stage. The second aim was to determine the sensitivity of surveillance using US to detect early HCC, and if there was any additional benefit of concurrently checking AFP. From the review, six studies were found on US for detecting HCC at any stage with a pooled sensitivity of 94.0% (95% confidence intervals, CI: 83.0% to 98.0%), a pooled specificity of 94.0% (95% CI: 89.0% to 97.0%), and a pooled diagnostic odds ratio (OR) of 232.7 (95% CI: 105.9 to 511.2). A summary receiver operator characteristics curve (SROC) analysis demonstrated an area under the curve (AUC) of 0.98 (95% CI: 0.96 to 0.99) suggesting a high diagnostic accuracy (**Figure 5**). Meanwhile, 13 studies were used in the second part of the analysis in which the sensitivity of US with and without AFP to detect early HCC was assessed. The pooled sensitivity was reported 63.0% (95% CI: 49.0% to 76.0%). However, AFP improved the performance of US with the pooled sensitivity increased to 69.0% (95% CI: 53.0% to 81.0%; P=0.65). The forest plot of the sensitivity of US and AFP for detecting early HCC is shown in **Figure 6**. A wide range of AFP cut-offs (15 ng/mL to 400 ng/mL) were used to diagnose HCC in the included studies, although the cut-off level did not appear to affect the utility of AFP (P=0.95). Furthermore, studies with surveillance intervals of < 6 months had a pooled sensitivity of 70.1% (95% CI: 55.6% to 84.6%), while the studies with surveillance intervals between 6 and 12 months had a pooled sensitivity of 50.1% (95% CI: 40.0% to 59.2%). The authors concluded that surveillance with US demonstrated limited sensitivity for early HCC, although this may be improved by testing at 6-month intervals. Serum AFP provided no additional benefit to US. Currently available evidence evaluating surveillance using US had significant limitations (such as verification bias) and future studies are necessary to determine optimal surveillance methods for early HCC.[18 level 3]

Figure 5: Performance characteristics of US for the detection of HCC at any stage

(a) forest plot for the sensitivity of ultrasound to detect HCC at any stage; (b) forest plot for the specificity of ultrasound to detect HCC at any stage; (c) forest plot for the odds ratio of ultrasound to detect HCC at any stage; (d) summary receiver operator curve plot for the detection of HCC at any stage by ultrasound. *Q*, chi-squared test of heterogeneity; I^2, inconsistency index; *AUC*, area under the curve.

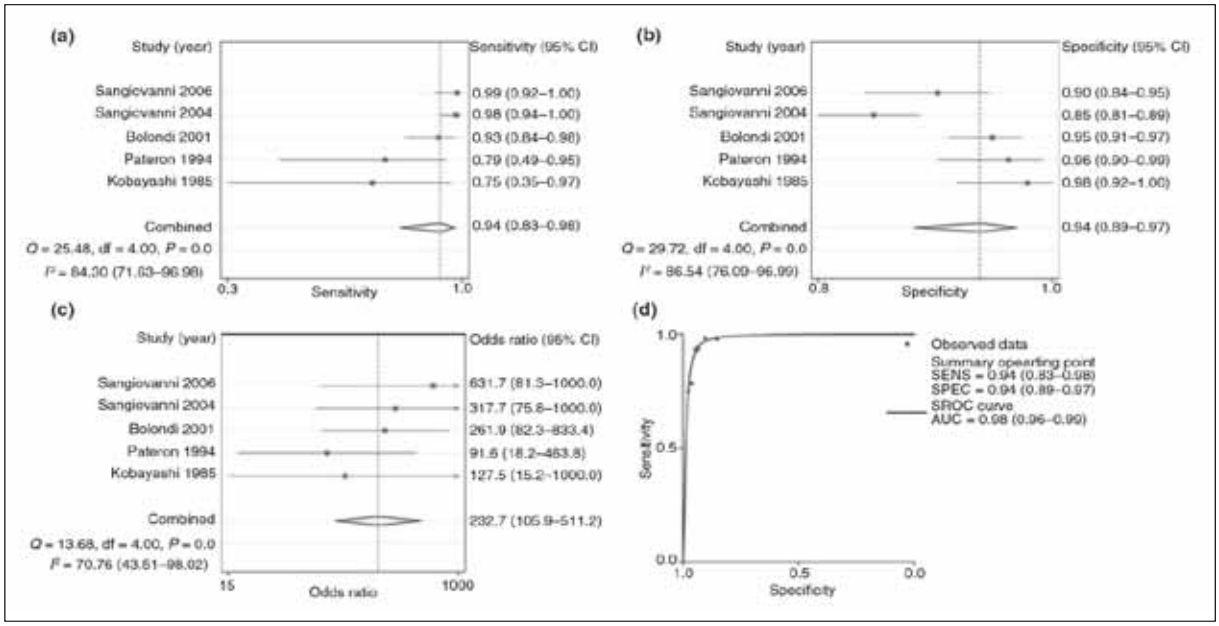
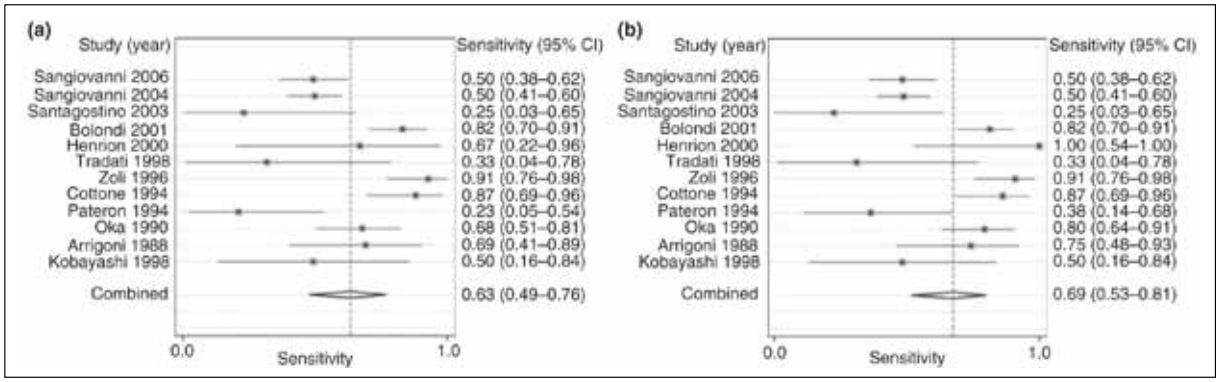


Figure 6: Sensitivity of US with and without AFP for the detection of early-stage HCC

(a) forest plot for the sensitivity of ultrasound to detect early HCC; (b) forest plot for the sensitivity of ultrasound with AFP to detect early HCC. *Q*, chi-squared test of heterogeneity; I^2, inconsistency index.



Another systematic review with meta-analysis was conducted by Colli *et al.* on the accuracy of AFP, US, spiral CT, and MRI in diagnosing HCC. The review only examined studies including patients with chronic liver disease, using the pathological findings of the explanted or resected liver, or the history of focal liver lesion(s) as the reference standard. Pooled sensitivity, specificity, and likelihood ratios (LR) were calculated using the random effect model. From the nine AFP studies, the sensitivity and specificity varied widely, and this could not be entirely attributed to the threshold effect of the different cut-off levels used (**Table 1**). The pooled estimates of 14 US studies were 60.0% (95% CI: 44.0% to 76.0%) for sensitivity, 97.0% (95% CI: 95.0% to 98.0%) for specificity, 18.0 (95% CI: 8.0 to 37.0) for LR+, and 0.5 (95% CI: 0.4 to 0.6) for LR- (**Table 2**). For the 10 CT studies, pooled sensitivity was 68.0% (95% CI: 55.0% to 80.0%), specificity 93.0% (95% CI: 89.0% to 96.0%), LR+ 6.0 (95% CI: 3.0 to 12.0), and LR- 0.4 (95% CI: 0.3 to 0.6). For the nine MRI studies, pooled sensitivity was 81.0% (95% CI: 70.0% to 91.0%), specificity 85.0% (95% CI: 77.0% to 93.0%), positive Likelihood Ratio 3.9 (95% CI: 2.0 to 7.0), and negative Likelihood Ratio 0.3 (95% CI: 0.2 to 0.5). The accuracy of CT and MRI are shown in **Table 3**. The authors concluded that US was highly specific but insufficiently sensitive to detect HCC in many cirrhotic or to support an effective surveillance program. The role of CT and MRI remains to be defined, particularly as additional tests aimed at confirming positive US results. High-quality prospective studies are needed to define the actual diagnostic role of AFP.[19 level 3]

Gebo *et al.* conducted a systematic review on screening test for HCC in patients with chronic HCV. The search strategy involved searching MEDLINE and other electronic databases between January 1985 and March 2002 on two key questions: *(1) What is the efficacy of using screening test for HCC to improve clinical outcomes, and (2) What are the sensitivity and specificity of screening tests for HCC? They found one surveillance study that answered Key Question no. (1)*. In this prospective cohort surveillance study, a total of 360 patients who were followed in a clinic for Hepatitis received AFP and US screening twice a year. The control group was 2,170 patients who received usual care in other hepatology clinics. During a mean follow-up of 56 months, focal lesions that prove to be HCC were found in 24/360 (6.7%) of the patients in the screening group and 129/2,170 (5.9%) in the control group. Of the 24 malignancies noted in the screening group, 75.0% were unifocal and < 3 cm, compared with 16.0% in the control group, showing a statistically significant difference. In the screening group, at the time of diagnosis, serum AFP was normal (< 20 ng/mL) in 11 patients, between 20 ng/mL and 200 ng/mL in nine patients, and > 200 ng/mL in four patients. At these thresholds, sensitivities for detecting HCC were 46.0%, 38.0%, and 17.0%, respectively. Overall, this study indicated that HCC could be detected earlier and was more often resectable when the screening group was compared with patients who received usual care. In contrast, they identified 24 studies to answer the Key Question no. (2), which included patients with chronic HCV or HBV or both, to address the sensitivities and specificities of the screening tests.

The sensitivity of AFP for detecting HCC was moderately high at 60.0% to 80.0%, with a specificity of 70.0% to 90.0%, for threshold values that decreased from 400.0 ng/mL to 10.0 ng/mL. However, the sensitivity and specificity of US were limited in that some were designed to assess the incidence of HCC and not to assess the performance characteristics of US. Therefore, studies are needed to determine whether screening improves clinical outcome.[20 level 3]

The test characteristic of serum AFP for detecting HCC in HCV- infected patients with or without cirrhosis was evaluated in another systematic review by Gupta S, Bent S, and Kohlwes which included five studies. Two studies were prospective cohort design and three were case-control designs. Computed tomography (CT), MRI, histopathology, and disease-free time greater than two years were considered adequate gold standards. By using the most commonly reported cut-off value of a positive test result for HCC (AFP level > 20 µg/L), the sensitivity ranged from 41.0% to 65.0%, while specificity ranged from 80.0% to 94.0%. Positive likelihood ratio (LR+) ranged from 3.1 to 6.8 and LR- ranged from 0.4 to 0.6 (**Table 4**). Four of the five studies reported sensitivity and specificity for an AFP cut-off value of > 200 µg/L, a value that is frequently reported to be specific for the diagnosis of HCC. The range of specificities was very high at this cut-off value (99.0% to 100.0%), but the sensitivity was very low (20.0% to 45.0%) as shown in **Table 5**. The authors concluded that current studies had substantial methodology limitations, making it difficult to define clear estimates of sensitivity and specificity for this test. A prospective study done with careful attention to limitation of bias is clearly needed to define whether any screening strategy can provide clinically important benefits.[21 level 3]

Table 1: Accuracy of AFP in diagnosing HCC

Author	Yr	References	Cut-off (ng/mL)	Sensitivity (%)	Specificity (%)	Likelihood Ratio	
						Positive	Negative
Piantino	89	20	>50*	67	87	5.2	0.38
"			>50†	76	67	2.3	0.35
"			>100†	62	97	20.6	0.39
Lopez	96	22	>200	59	97	22.5	0.42
Bayati	98	25	>10	93	67	2.8	0.1
"			>17.8	35	99	35	0.65
Chalasani	99	26	>20	63	86	4.8	0.42
Gambarin	00	27	>20	58	91	6.4	0.46
"			>50	47	96	11.8	0.6
Trevisani	01	33	>20	60	91	6.6	0.43
Tong	01	41	>11	86	89	7.8	0.15
"			>21	41	94	6.8	0.62
Nguyen	02	44	>10	78	61	2.0	0.36
"			>20	63	80	3.1	0.46
"			>50	51	89	4.6	0.55
"			>100	41	97	13.6	0.61
"			>200	32	100	320	0.68
Marrero	02	45	>11	77	71	2.6	0.32
"			>20	68	86	4.8	0.37
"			>100	47	98	23.5	0.54

*Cross-sectional study.
†Longitudinal study.

Table 2: Accuracy of US in diagnosing HCC

Author	Yr	References	Sensitivity (%)	Specificity (%)	Likelihood Ratio	
					Postive	Negative
Okazaki	84	16	86	99	66	0.14
Maringhini	84	17	92	86	6.5	0.09
Kobayashi	85	18	75	98	32.6	0.26
Tanaka	86	19	47	100	589	0.41
Dodd	92	21	43	98	21.5	0.58
Saada	97	23	33	100	333	0.67
Chalasani	99	26	59	92	8.4	0.45
Gambarin	00	27	58	94	9.6	0.44
Rode	01	30	46	95	9.2	0.57
Kim	01	31	38	92	4.7	0.67
Bennett	01	35	30	97	7.4	0.72
Teefey	03	40	89	73	3.3	0.15
Tong	01	41	100	98	50	0.0
Libbrecht	03	42	40	100	400	0.6
Pooled estimates (95% CI)			60.5 (44–76)	96.9 (95–98)	17.7 (8.5–36.9)	0.5 (0.4–0.6)

Table 3: Accuracy of spiral CT (upper panel) and MRI (lower panel) in diagnosing HCC

Author	Yr	References	Sensitivity (%)	Specificity (%)	Likelihood Ratio	
					Positive	Negative
Chalasani	99	26	93	96	22.7	0.09
Gambarin	00	27	53	94	8.8	0.5
Lim	00	28	80	96	20	0.21
Peterson	00	29	44	92	5.1	0.72
Rode	01	30	54	93	7.7	0.49
Mortelè	01	32	82	100	82	0.18
De Ledinghen	02	36	81	85	5.2	0.22
Zacherl	01	37	65	56	1.49	0.62
Teefey	03	40	67	72	2.4	0.46
Libbrecht	03	42	50	79	2.4	0.63
Pooled estimates (95% CI)			67.5 (55–80)	92.5 (89–96)	6.1 (3.1–12)	0.4 (0.3–0.6)
Born	98	24	69	92	8.25	0.34
Rode	01	30	77	57	1.79	0.4
Krinsky	01	34	54	86	3.8	0.53
de Ledinghen	02	36	90	100	905	0.09
Mori	02	39	85	74	3.2	0.2
Teefey	03	40	56	72	2.1	0.61
Bhartia	02	38	93	100	900	0.07
Libbrecht	03	42	70	82	0.71	0.37
Burrell	03	43	100	95	20.0	0.00
Pooled estimates (95% CI)			80.6 (70–91)	84.8 (77–93)	3.9 (2.4–6.5)	0.3 (0.2–0.5)

Table 4: Test characteristic of AFP levels > 20 µg/L for detecting HCC

Study, Year (Reference)	Sensitivity of AFP Level > 20 µg/L (95% CI), %	Specificity of AFP Level > 20 µg/L (95% CI), %	Positive Likelihood Ratio (95% CI)	Negative Likelihood Ratio (95% CI)
Peng et al., 1999 (20)	65 (58–71)	87 (79–93)	4.9 (3.0–8.0)	0.5 (0.3–0.5)
Cedrone et al., 2000 (18)†‡	55	88	4.6	0.5
Tong et al., 2001 (15)‡	41	94	6.8	0.6
Trevisani et al., 2001 (21)‡	60	91	6.7	0.4
Nguyen et al., 2002 (19)	63 (56–70)	80 (73–86)	3.1‡	0.5‡

* AFP = α-fetoprotein.
† Data for patients with hepatitis C virus and hepatitis B virus analyzed together.
‡ Data for CIs are not available or calculable.

Table 5: Test characteristic of AFP levels > 200 µg/L for detecting HCC

Study, Year (Reference)	Sensitivity of AFP Level > 200 µg/L (95% CI), %	Specificity of AFP Level > 200 µg/L (95% CI), %	Positive Likelihood Ratio (95% CI)	Negative Likelihood Ratio (95% CI)
Peng et al., 1999 (20)	45 (38–52)	100 (97–100)	∞†‡	0.6 (0.5–0.6)
Cedrone et al., 2000 (18)‡	20	99	29	0.8
Tong et al., 2001 (15)‡	NR	NR		
Trevisani et al., 2001 (21)‡	22	99	37	0.8
Nguyen et al., 2002 (19)	32 (25–39)	100 (100–100)	∞†‡	0.7‡

* AFP = α-fetoprotein; NR = not reported.
† When the reported specificity is 100%, the likelihood ratio is theoretically infinite.
‡ Data for CIs are not available or calculable.

A randomised controlled trial by Chen *et al.* studied 5,581 men aged 30 to 69 years who were chronic carriers of HBV between 1989 and 1995 in Qidong Country, China. Of these men, 3,712 were randomly assigned to a screening group and 1,869 to a control group. Screening entailed 6-monthly AFP assays, with follow-up of patients having an abnormal test result (≥ 20 µg/L). All patients were followed up for liver cancer and/or death. Confirmation of the diagnosis of the majority of the primary liver cancer (PLC) patients was by US. A few (4.3%) cases were diagnosed with US plus CT examination. Some patients, especially those considered to be candidates for surgery, received a liver biopsy or fine needle aspiration, so that a pathological confirmation was available for 10.4%. In addition, the diagnosis was confirmed in some subjects by autopsy, or aspiration of liver tissue post mortem (12.0%). The overall (on one or more occasion) sensitivity and specificity of the screening programme for liver cancer using AFP was 55.3% and 86.5%, respectively. These figures include cases of liver cancer among subjects who did not attend for screening at some or all of the scheduled examinations. In subjects who complied with all scheduled screening tests, the sensitivity was 80.0% and specificity 80.9% (**Table 6**).[22 level 3]

Table 6: Cases of liver cancer, and AFP results among subjects in screening group

(a) All subjects

	PLC		
AFP	Yes	No	Total
Positive	142	466	608
Negative	115	2989	3104
Total	257	3455	3712

Sensitivity = 142/257 = 55.3%; Specificity = 2989/3455 = 86.5%.

(b) Subjects completing all scheduled tests

	PLC		
AFP	Yes	No	Total
Positive	32	197	229
Negative	8	833	841
Total	40	1030	1070

Sensitivity = 32/40 = 80.0%; Specificity = 833/1030 = 80.9%.

Zhang B and Yang B assessed the validity and cost of screening for PLC using AFP or US, and combination of AFP and US. A total of 18,816 subjects aged 35 to 59 years with HBV or a history of chronic hepatitis were recruited and randomly assigned to one of two groups in urban Shanghai, China. The screening group, comprising 9,373 subjects, was offered AFP testing and US examination every six months; the control group, comprising 9,443 subjects, was offered no screening. A screening test was defined as positive when AFP > 20 μg/L, and when US demonstrated a new solid lesion in the liver. The final diagnosis was reached by CT, MRI, and biopsy. Primary liver cancer (PLC) was detected in 51 subjects, 36 of whom had small PLC. They found that when AFP and US were used in parallel the detection rate, false-positive rate, and PPV were 92.0%, 7.5%, and 3.0% respectively. When AFP was used alone the detection rate, false-positive rate, and PPV were 69.0%, 5.0%, and 3.3% respectively. When US was used alone the detection rate, false-positive rate, and PPV were 84.0%, 2.9%, and 6.6% respectively (**Table 7**). These findings indicated that US had a higher detection rate, a lower false-positive rate, and higher PPV than AFP. The combined test not only increased the efficacy of screening, but also increased the false-positive rate (7.5% versus 2.9%). Therefore, combined AFP and US as a screening test was found to be suitable for relatively developed areas of China, but otherwise, US alone is the method of choice.[23 level 1]

Table 7: Detection rate, false positive rate, PPV of screening tests and their 95% CI

	Detection rate (%)	*False positive rate (%)*	*Positive predictive value (%)*
AFP and/or US positive	92 (80 to 97)	7.5 (7.1 to 7.9)	3.0 (2.2 to 3.8)
AFP alone positive	69 (54 to 80)	5.0 (4.7 to 5.3)	3.3 (2.2 to 4.4)
US alone positive	84 (73 to 93)	2.9 (2.7 to 3.1)	6.6 (4.7 to 8.5)

AFP = α fetoprotein; US = ultrasonography.

6.2 Effectiveness of HCC screening

Cochrane systematic review on AFP and/or US for liver cancer screening in patients with chronic hepatitis B in 2003 and edited in 2009 included two randomised controlled trials. Yang *et al.* trial (n=18,816) compared biannual AFP plus US screening with no screening for five years. A total of 86 HCC were detected in the screened group and 51 HCC in the control group after five years (OR=1.69, 95% CI: 1.20 to 2.36). The two groups did not differ significantly regarding HCC mortality (OR=0.81, 95% CI: 0.54 to 1.22). The survival rate of patients with resected HCC in the screened group reached 52.7% after three to five years, but was 0.0% in the control group. Another trial, Sherman *et al.* (n=1,069) compared AFP plus US versus AFP screening, but could not decide which approach was superior due to small sample size. The authors concluded that there were not enough quality trials to support or refute screening of HBsAg-positive patients for HCC. It is possible that screening may be effective, but also that harm caused by screening treatment may outweigh any gain.[17 level I]

Chen *et al.* conducted a randomised control trial to investigate the effectiveness of screening for liver cancer in reducing mortality from the disease in a high-risk population in Qidong, China during the period 1989 to 1995. A total of 5,581 men aged 30 to 69 years (HBsAg positive) were identified by population screening and randomly assigned to a screening group (group A= 3,712), and controls (group B= 1,869). Screening was planned to be 6-monthly AFP assays, with follow-up of subjects having an abnormal test (≥ 20 µg/L). All subjects were followed up for liver cancer and/or death until 31st December 1995. The final diagnosis was reached by US, US plus CT, pathological, autopsy, or aspiration of liver tissue post mortem. Overall, 374/5,581 (6.7%) liver cancer were diagnosed. The incidence was higher (1,341.7 per 100,000 person-years) in the screened group A, than in the group B controls (1,195.6 per 100,000), but the difference was not significant (p=0.30). The mortality rate from liver cancer was 1,138.1 per 100,000 person-years in the group A and 1,113.9 per 100,000 in the group B, but the difference was not statistically significant (p=0.86). The one-, three, and five-year relative survival rates were 23.7%, 7.0%, and 4.0% in group A, and 9.7%, 4.0%, and 4.1% in group B, respectively, with no difference in five-year survival between the groups. The authors concluded that screening with AFP resulted in earlier diagnosis of liver cancer, but the gain in lead time did not result in any overall reduction in mortality, because therapy for the patients found by screening was ineffective. Further studies using improved methods of screening, diagnosis and treatment were indicated.[22 level I]

The effect of screening on HCC mortality in people at increased risk was evaluated by Zhang BH, Yang BH, and Tang ZY. The study included 18,816 people, aged 35 to 59 years with hepatitis B virus infection or a history of chronic hepatitis in urban Shanghai, China. Participants were randomly allocated to a screening (n=9,373) or control (n=9,443) group. Screening group participants were invited to have an AFP test and US examination every six months. Controls received no screening and continued to use health-care facilities. Screening was stopped in December 1997 and all participants were followed up until December 1998. The screened group completed 58.2% of the screened offered. When the screened group was compared to the control group, the number of HCC was 86 versus 67; subclinical HCC being 52/86 (60.5%) versus 0; small HCC was 39/86 (45.3%) versus 0; resection achieved was 40/86 (46.5%) versus 5/67 (7.5%); the one-, three, and five-year survival rates were 65.9%, 52.6%, 46.4% versus 31.2%, 7.2%, 0.0%, respectively. Thirty-two people died from HCC in the screened group versus 54 in the control group. They reported a significantly lower HCC mortality rate in the screened group than in controls, being 83.2 per 100,000 and 131.5 per 100,000 respectively with a mortality rate ratio of 0.63 (95% CI: 0.41 to 0.98), as shown in **Table 8**. They concluded that biannual screening with a combination of AFP and US reduced HCC mortality after five-year follow-up by 37.0%. [24 level I]

Table 8: Outcome of screening

	Screening group	Control group
HCC occurrence		
No. of cases	86	67
Total incidence(per 100,000)	223.7	163.1
Rate ratio (95% CI)	1.37(0.99, 1.89)	
Deaths from HCC		
No. of death	32	54
Total mortality(per 100,000)	83.2	131.5
Rate ratio (95% CI)	0.63(0.41, 0.98)	

McMahon *et al.* conducted a prospective 16-years, population based cohort study to determine the impact of screening for HCC in 1,487 HBsAg+ve Alaska native carriers with AFP determinations every six months. Men and non-pregnant women with an elevated AFP level were evaluated for the presence of HCC by US examination. The long-term survival rate for patients whose HCC was detected by the screening programme was compared with a historical control group of Alaska native patients with HCC from the same population who were clinically diagnosed with HCC between 1969 and October 1982, through a National Cancer Institute-sponsored Cancer Registry. Between 1982 and 1993, an AFP level > 25 ng/mL was used as a cut-off level. After 1993, the cut-off level was lowered to 15 ng/mL because a carrier with an AFP of 15 ng/mL had been found to have a large nonresectable tumour. The final diagnosis was reached by CT scan if the US examination was unsatisfactory or suggested a lesion. Between October 1982 and December 1998, 26,752 HBsAg carriers in which AFP determinations were performed, and one or more AFP elevations were found in 61 men and 39 non-pregnant woman. During follow-up, HCC was diagnosed in 32 patients (24 men and eight women). Hepatocellular carcinoma (HCC) tumours < 6 cm were found in 23 patients; 22 patients had resections, and one patient refused a resection. In comparison with 12 patients with HBV-related HCC (diagnosed from 1969 to October 1982) before this programme, the five- and 10-year survival rate for the 32 patients with HCC were 42.0% (P=0.008) and 30.0% (P=0.07), respectively. Five- and 10-year tumour-free survival rates for carriers who had a normal AFP level on initial screening and subsequently developed HCC were 29.0% (P=0.004) and 24.0% (P=0.024), respectively. This population based screening programme using AFP determination revealed that in most HBsAg+ve carriers, HCC can be detected at an early, potentially resectable stage and showed significant survival benefits at five and 10 years in screened carriers compared with historical controls.[25 level II-2]

6.3 Economic evaluation

Zhang B and Yang B assessed the cost of screening for primary liver cancer (PLC) using AFP or US, and combination of AFP and US. A total of 18,816 subjects aged 35 to 59 years with HBV or a history of chronic hepatitis were recruited and randomly assigned to one of two groups in urban Shanghai, China. The screening group, comprising 9,373 subjects, was offered AFP testing and US examination every six months; the control group, comprising 9,443 subjects, was offered no screening. A screening test was defined as positive when AFP > 20 µg/L, and when US demonstrated a new solid lesion in the liver. The final diagnosis was reached by CT, MRI, and biopsy. The cost was 5 RMB (USD$0.60) for an AFP test, 10 RMB (USD$1.20) for an US examination, and about 500 RMB (USD$60) for investigating a screen positive patient (USD$1=8.3 RMB). Detection of a small PLC (diameter < 5 cm) was used as a measure of the effectiveness in the cost analysis.From the study, PLC was detected in 51 subjects, 36 of whom had small PLC (**Table 9**). **Table 10** on the other hand, shows the costs of screening, calculated from the data in **Table 9**. The cost for each case detected by US was lower than for each case detected by AFP (16,451 RMB versus 25,139 RMB). The combined test had the highest costs for each small PLC detected (30,206 RMB). The incremental cost of using both tests rather than AFP alone (11 extra cases detected) was 41,722 RMB (USD$5,027) for each additional case detected; for both tests rather than US alone (four additional cases detected) was 140,242 RMB (USD$16,897).

The results of this study showed that the use of US is not only more effective, but cheaper than AFP testing alone. The combined test not only increased the efficacy of screening, but also increased the false positive rate and the cost. Therefore, combined AFP and US as a screening test is suitable for relatively developed areas of China, but otherwise, US alone is the method of choice.[23]

Table 9: Results of screening

		Number with PLC	*Number with no PLC*
AFP+	US+	31 (21)*	104
AFP+	US−	4 (4)	915
AFP−	US+	12 (11)	500
AFP−	US−	4 (0)	18 724
Total		51 (36)	20 243

* Number out of the total that were small (≤5 cm diameter) when first detected.
AFP = α fetoprotein; US = ultrasonography; PLC = primary liver cancer.

Table 10: Cost analysis for screening using AFP and US combined and separately

	AFP and US	*AFP*	*US*
Small PLC (n)	36	25	32
Total cost	1 087 410	628 470	526 440
Cost for each small PLC detected	30 206	25 139	16 451

*All costs in the table are given as RMB (1 RMB = $0.1205).
AFP = α fetoprotein; US = ultrasonography; PLC = primary liver cancer.

Lin O *et al.* conducted cost-effectiveness analysis using Markov decision model. Several screening strategies with abdominal US or CT and AFP at six to 12 month intervals in 40-year-old patients with chronic hepatitis C and compensated cirrhosis were simulated from a societal perspective. Three plausible strategies were modelled: AFP and US every six months (biannual AFP/US), AFP and US every 12 months (annual AFP/US), and AFP every six months with US every 12 months (biannual AFP/annual US). Only direct medical costs were considered. Discounting was implemented at 3.0% per year for QALYs. Results were expressed in terms of discounted costs, life years (LYs), QALYs, and incremental cost-effectiveness ratios (ICERs) for QALYs gained. Incremental comparisons were performed by rank ordering the alternatives by increasing efficacy after eliminating those that were more costly and less efficacious than an alternative (i.e. dominated). The ICER was defined as the additional cost per additional gain in QALY for any particular screening strategy versus the next less efficacious (in terms of QALYs gained) screening strategy. The least efficacious strategy was compared against the no screening strategy. The analysis demonstrated that for the least efficacious strategy, annual AFP/US, the ICER (versus no screening) was USD$23,043 per QALY. For the strategy most commonly used in United States (biannual AFP/annual US), the ICER was USD$33,083 per QALY (versus annual AFP/US). The most efficacious strategy (biannual AFP/US) entailed a higher ICER of USD$73,789 per QALY (versus biannual AFP/annual US). Screening using CT and AFP resulted in better survival compared with the corresponding strategy using US and AFP, with ICER ranging from approximately USD$23,000 to USD$96,000 per QALY (see **Table 11**). Based on analysis, the authors concluded that the best screening protocol had yet to be formally defined, but based on the results, US at 12-month intervals and AFP at 6-month intervals was a reasonable strategy, offering the greatest gain in life-expectancy while still maintaining an ICER < USD$50,000 per QALY. More frequent screening with US provided some additional benefit, but was more expensive. Screening with CT instead of US was more efficacious and appeared to be cost-effective, and deserved a further study. However, this study was subjected to important limitations. True costs were used instead of charges, and life-years were adjusted for quality. Liver transplantation, as well as palliative therapy for HCC such as transarterial chemoembolization (TACE), percutaneous ethanol injection (PEI) or thermoabalation for unresectable cases was modelled.[26]

Table 11: Cost-effectiveness of strategies for HCC screening in patient with compensated cirrhosis

Strategy	Lifetime cost ($)	Additional cost ($)	Expected QALY	QALY gained	Expected LY	LY gained	Incremental CE ratio ($/QALY)	Incremental CE ratio ($/LY)
Base case								
No screen	46 232	0	6.269	0	8.634	0	–	–
US12AFP12	53 145	6913	6.569	0.300	8.965	0.331	23 043	20 885
US12AFP6	54 733	1588	6.617	0.048	9.021	0.056	33 083	28 357
US6AFP6	57 168	2435	6.650	0.033	9.080	0.036	73 789	67 639
After inclusion of CT screening strategies[2]								
No screen	46 232	0	6.269	0	8.634	0	–	–
US12AFP12	53 145	6913	6.569	0.300	8.965	0.331	23 043	20 885
CT12AFP12	53 655	510	6.583	0.014	8.983	0.018	36 429	28 333
US12AFP6	54 733	1078	6.617	0.034	9.021	0.038	31 706	28 368
CT12AFP6	55 147	414	6.625	0.008	9.031	0.010	51 750	41 400
US6AFP6	57 168	2021	6.650	0.025	9.080	0.026	80 840	77 731
CT6AFP6	58 232	1064	6.610	0.011	9.093	0.013	96 727	81 846

[1] Incremental CE ratios are calculated by dividing incremental cost ($) over incremental outcome (QALYs). Each incremental value is determined by subtracting the value of the strategy of the next less effective strategy (as measured by QALYs gained) from that of the strategy under consideration. The incremental CE ratio of the least efficacious screening strategy (US and AFP every 12 months) is calculated against the no screening strategy.
[2] CT screening: screening strategies with triphasic abdominal CT and AFP and compared against each other as well as against strategies using US. US12AFP6, US at 12-month intervals and AFP levels at 6-month intervals; US6AFP6, US and AFP levels at 6-month intervals; US12AFP12, US and AFP levels at 12-month intervals; QALY, quality-adjusted life-year; US, abdominal ultrasonography; AFP, serum alpha-foetoprotein level.

A cost-utility analysis by Thompson CJ *et al.* evaluated effectiveness and cost-effectiveness of surveillance strategies for HCC in individuals with cirrhosis using Markov decision-analytic model. The population of interest was people with compensated cirrhosis aged ≤ 70 years with no pre-existing medical conditions that might preclude treatment with liver transplantation (OLT) or hepatic resection (including current alcohol or intravenous drug use). The model considered three cirrhosis aetiologies due to alcoholic liver disease (ALD), HBV, and HCV. Results were also combined to approximate a mixed aetiology population. Comparisons were made between varieties of surveillance algorithms using AFP assay and/or US at 6- and 12-monthly intervals. Parameter estimates were obtained from comprehensive literature reviews. Uncertainty was explored using one-way and probabilistic sensitivity analysis. Cost and utility values were attached to each state, and the differences between the aggregated costs and health outcomes in each simulation were used to estimate the cost-effectiveness of surveillance, expressed as incremental cost per QALY. Cost and QALYs were discounted at 3.5% per year.[27]

For the effectiveness of surveillance, the analysis showed that the 6-monthly AFP plus US was most effective across all outcomes, more than tripling the number of HCC diagnosed while operable (16.9%), and almost halving the number dying from HCC (10.8%) when compared with no surveillance. Number needed to be under surveillance (NNS) to prevent one death from HCC was 11. However, the cheapest strategy, annual AFP-triage, still achieved substantial gain: for example, more than doubling the number of operable HCC found (11.9%), and increasing the number of small tumours found more than six-fold (**Table 12**). Cost-utility results on the other hand, are shown in **Table 13**. In an incremental analysis, neither of the US-only strategies would be considered (since they are both slightly less effective and more costly than surveillance at the same frequency with AFP-triage). Therefore, in the mixed aetiology cohort, the cheapest surveillance strategy was annual AFP-triage, with ICER of £20,700 per QALY. The addition of US to this strategy increased the ICER to £60,100 per QALY gained.[27]

Table 12: Lifetime effectiveness of surveillance

	No surveillance	Annual AFP-triage	Annual US	Annual AFP+US	6-monthly AFP-triage	6-monthly US	6-monthly AFP+US
% with operable HCC	5.1%	11.9%	11.7%	13.5%	15.3%	15.0%	16.9%
% HCC_S at diagnosis	0.3%	1.9%	1.5%	2.3%	3.1%	2.6%	3.7%
% HCC_M at diagnosis	2.1%	3.8%	4.1%	4.2%	4.2%	4.6%	4.4%
% getting OLTs	17.1%	19.1%	19.1%	19.2%	20.1%	20.0%	20.3%
% OLTs for known HCC	8.3%	20.3%	20.0%	23.2%	25.3%	24.9%	27.9%
% dying of HCC	19.9%	14.7%	14.9%	13.5%	12.0%	12.3%	10.8%
NNS to prevent 1 death[a]	—	19	20	15	13	13	11
% dead by age 75 years	69.3%	68.4%	68.5%	68.2%	68.0%	68.0%	67.8%
NNS to prevent 1 death[b]	—	114	117	93	78	79	68

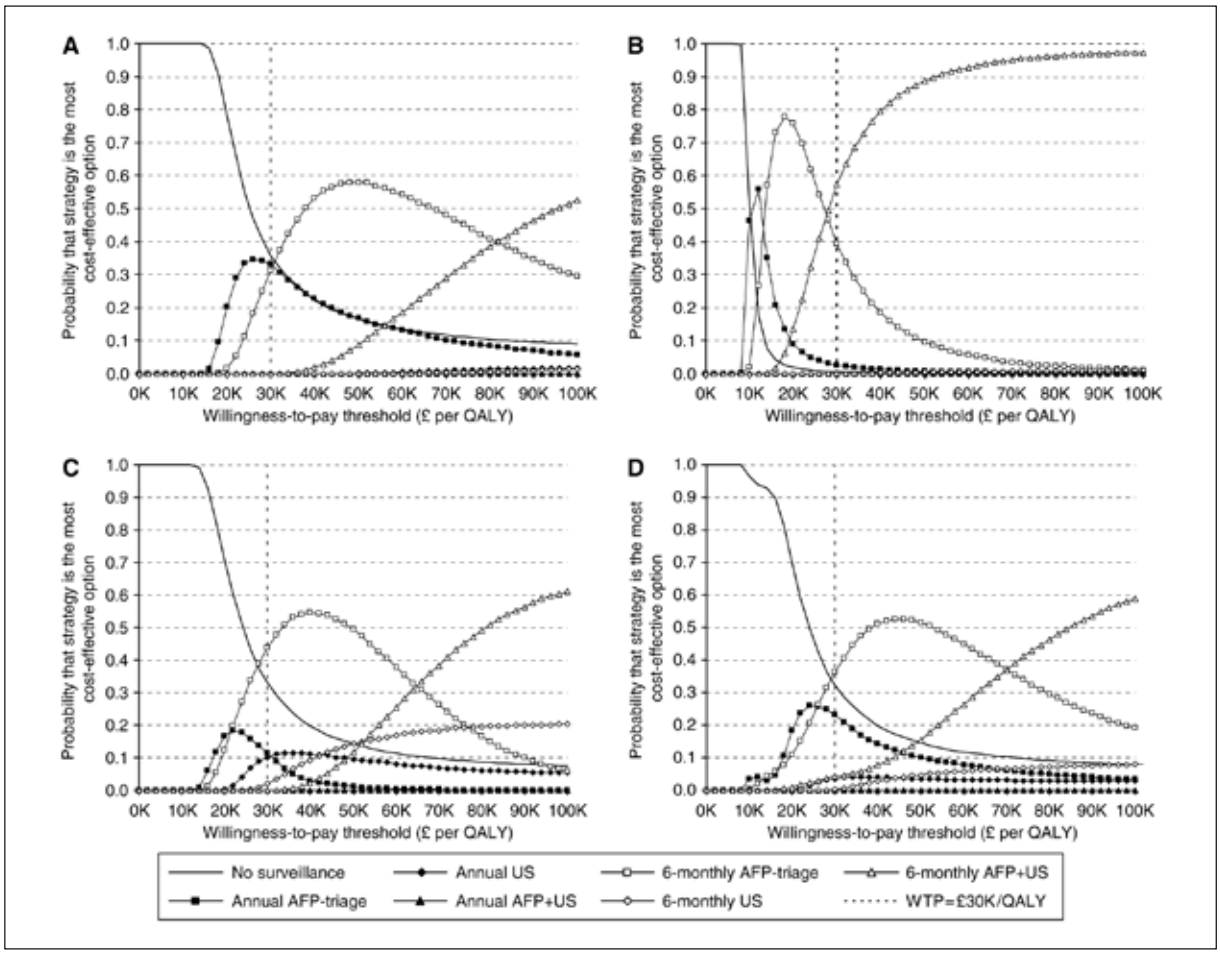
HCC = hepatocellular carcinoma; HCC_L = large hepatocellular carcinoma; HCC_M = medium hepatocellular carcinoma; HCC_S = small hepatocellular carcinoma; NNS = number needed to be under surveillance; OLT = orthotopic liver transplantation. [a]NNS to prevent one death from HCC. [b]NNS to prevent one 'premature' death (age <75 years).

Table 13: Cost-utility analysis

			Incremental analysis		
Strategy	**Cost (£)**	**Utility (QALYs)**	**Cost (£)**	**Utility (QALYs)**	**£/QALY (ICER)**
ALD					
No surveillance	£26 100	9.359			
Annual AFP-triage	£27 400	9.410	£1300	0.051	£24 800
Annual US	£27 700	9.410		Extendedly dominated	
Annual AFP+US	£28 100	9.422		Extendedly dominated	
6-monthly AFP-triage	£28 200	9.433	£800	0.024	£35 500
6-monthly US	£28 800	9.434		Extendedly dominated	
6-monthly AFP+US	£29 200	9.445	£1000	0.011	£88 000
HBV					
No surveillance	£29 600	10.858			
Annual AFP-triage	£31 700	11.069	£2100	0.211	£10 200
Annual US	£32 100	11.066		Dominated	
Annual AFP+US	£32 700	11.119		Extendedly dominated	
6-monthly AFP-triage	£33 000	11.168	£1300	0.099	£12 700
6-monthly US	£33 600	11.164		Dominated	
6-monthly AFP+US	£34 200	11.216	£1300	0.048	£26 800
HCV					
No surveillance	£27 600	8.087			
Annual AFP-triage	£29 500	8.172	£1900	0.085	£22 200
Annual US	£29 700	8.172		Extendedly dominated	
Annual AFP+US	£30 300	8.193		Extendedly dominated	
6-monthly AFP-triage	£30 600	8.212	£1100	0.040	£27 600
6-monthly US	£31 000	8.213		Extendedly dominated	
6-monthly AFP+US	£31 600	8.232	£1000	0.020	£50 400
Mixed aetiology					
No surveillance	£26 900	9.021			
Annual AFP-triage	£28 400	9.096	£1500	0.075	£20 700
Annual US	£28 800	9.096		Dominated	
Annual AFP+US	£29 200	9.114		Extendedly dominated	
6-monthly AFP-triage	£29 400	9.131	£1000	0.035	£27 600
6-monthly US	£29 900	9.131		Dominated	
6-monthly AFP+US	£30 400	9.148	£1000	0.017	£60 100

AFP = α-foetoprotein; ALD = alcoholic liver disease; HBV; hepatitis B virus; HCV = hepatitis B virus; ICER = incremental cost-effectiveness ratio; QALY = quality-adjusted life-year; US = ultrasound. Discount rate of 3.5% per annum applied to all costs and benefits.

Figure 7: Cost-effectiveness acceptability curves



Cost-effectiveness acceptability curves, showing relative probability of maximal cost-effectiveness among surveillance strategies. (**A**) ALD-related cirrhosis; (**B**) HBV-related cirrhosis; (**C**) HCV-related cirrhosis; (**D**) mixed aetiology cohort (weighting: 57.6% ALD; 7.3% HBV; 35.1% HCV). Maximal cost-effectiveness reflects the proportion of Monte Carlo simulations (10 000 per aetiology) in which each strategy generated the highest net monetary benefit.

Figure 7 shows cost-effectiveness acceptability curves for ALD, HBV, HCV, and the mixed aetiology cohort. These graphs showed the probability that each strategy would be considered the most cost-effective (in terms of highest net monetary benefit) at different levels of willingness to pay for a QALY. At willingness to pay threshold of £30,000 per QALY, the most intensive surveillance protocol simulated (6-monthly AFP plus US) is only likely to be considered cost-effective in individuals with HBV-related cirrhosis. In those with HCV-related cirrhosis, 6-monthly AFP-triage is more likely to be considered cost-effective; indeed, willingness to pay would have to rise to around £65,000 per QALY before 6-monthly AFP plus US becomes most likely to be considered cost-effective in this population. In individual with ALD-related cirrhosis, there was uncertainty about which strategy would be most cost-effective, with no surveillance, annual AFP-triage, and 6-monthly AFP-triage having approximately equal likelihood of maximal cost-effectiveness. In the mixed aetiology cohort that approximated the decision framework if a single strategy was to be adopted across all aetiologies, surveillance of any kind can only be recommended if willingness to pay approaches £30,000 per QALY. At this level, 6-monthly AFP-triage appeared to be the most cost-effective surveillance protocol, and remains the foremost option until willingness to pay reached very high levels. This analysis suggested that society would have to be prepared to spend nearly £70,000 per QALY gained before the most effective strategy, 6-monthly AFP plus US, could confidently be assumed to provide best value for money.[27]

The authors concluded that in patients with cirrhosis, surveillance strategies for HCC were effective, and can often be considered cost-effective. The most effective strategy for a mixed aetiology cohort of individuals with cirrhosis is AFP assay combined with US imaging on a 6-monthly basis. However, when costs were taken into account, the use of AFP as a triage may be preferable. Surveillance was much more likely to be cost-effective in those with HBV-related cirrhosis, while surveillance of people with ALD-related cirrhosis appears least economically efficient. Besides, the authors strongly stated as its limitation that there was very little published evidence on which to base many of the parameter estimates for the model, and few data originated within the United Kingdom. This was particularly apparent for defining US performance. Second, as the primary focus of this evaluation was the effectiveness and cost-effectiveness of surveillance, they had used a simplified approach to modelling treatment in which OLT and resection were the only curative options available. Third, they had also assumed that the three cirrhosis aetiologies were mutually exclusive and acknowledged that many people developed cirrhosis as a result of multiple causes.[27]

In the Malaysian context, the fees charged by MOH hospital for serum AFP is approximately [redacted] per test, while US varied from [redacted] to [redacted] per imaging.[28] Ultrasound machines cost about [redacted] and range up to [redacted]. The price depends largely on the level of complexity of the machine.[29]

6.4 Other considerations

Organizational

Chronic hepatitis B virus (HBV) affects around a million patients in Malaysia which accounts for majority of the diagnosed HCC (> 80.0%). However, there is currently no formal/structured national liver cancer screening programme being implemented. In fact, screening carried out in Malaysia in 1997/1998 revealed an HBsAg+ve chronic HBV rate of 5.2%. There was a local study attempts to examine the demographic and clinical features of chronic HBV patients in Malaysia from 2005 to 2006, conducted by SS Tan *et al.* from Department of Hepatology, Selayang Hospital. A total of 3,275 patients with a primary diagnosis of HBV, were admitted into this study. They were seen from 2005 to 2006, in outpatient clinics of 11 Malaysian hospitals (eight public hospitals and three university hospitals), representing geographic locations throughout the country. Patient demographic and clinical data were abstracted from medical charts to a pre-designed case report form by study nurses. In this study, cirrhosis was defined by any of the followings criteria: histology, nodular or shrunken liver seen on radiology, or ≥ 2 combinations of splenomegaly, ascites on imaging, varices on endoscopy or platelet count < 100,000/μL. Majority of chronic HBV patients were males (64.6%) with the mean age of 41.5 ± 14.5 years old; while Chinese patients tend to be older (45.6 ± 14.5 years) compared to other ethnic groups. Among the identified chronic HBV patients, the Chinese population forms the highest proportion (53.1%), followed by Malays (27.3%), indigenous Sabahans (14.0%), indigenous Sarawakians (2.9%), Indians (1.8%) and others (0.8%). Cirrhosis was identified in 10.9% of the chronic HBV patients with a mean age of 51.8 years old. The cumulative prevalence of cirrhosis is higher in males (n=266) compared with females (n=91) (12.58% versus 7.85%). The study found 23.4% of the cirrhotic patients were HBeAg+ve. This study also found indigenous Sarawakian patients had the highest rate of cirrhosis. As conclusion, chronic HBV remains a public health issue and significantly afflicts males in the productive age groups and of Chinese ethnicity in Malaysia. Local treatment guidelines and related health policy need to be established to address the disease severity among various ethnic groups. The authors strongly stated as its limitation that this is a cross-sectional study that cannot establish the cause and effect relationship. However, this study does provide a direction for future investigations on this subject.[30]

Continued public awareness of modifiable risk factors and preventive measures is important. Safe sexual practices, sterilization and disposal of sharps, guidelines for alcohol consumption, and awareness on smoking risks should be continually promoted. Proper training of staff involved in the screening programme is essential since US is very operator-dependent and screening is expensive and labour intensive; which required experience as well as continuous training.

Screening for liver cancer also requires resources. These included not only that the screening tests be available, but that an adequate radiology facility exists to undertake the necessary follow-up x-rays, and that an established treatment facility exists. Such a treatment facility should include hepatologists, expert surgeons, and interventional radiologists. A liver transplant programme should also be accessible. In the absence of such facilities, screening would be able to find early lesions, however it will not evade the risk of HCC related mortality. Currently in Malaysia, serum AFP test are conducted at laboratory hospitals with immunoassay facilities which covers MOH state hospitals and hospital with specialist amounting to 36 MOH hospitals (personal communication with Head of Chemical Pathology Activities, MOH). Meanwhile, US examination of the liver is only conducted in 39 MOH hospitals with radiologist (personal communication with Head of Radiology Service, MOH).

Ethical and legal consideration

Studies have shown that the more sensitive the test, the greater the detection rate and the greater the number of false results. This means that some people with cancer or a pre-malignant condition will be wrongly reassured and others will be subjected to unnecessary procedures and stress. When cancers and tumours have been missed, a negative result will give false reassurance, with the increased possibility that there will be a delay in diagnosis and treatment. The advantages of increased sensitivity have to be weighed against the significant increase in false positive, the consequential number of CT or MRI carried out, and the possible reduction in the specificity of the AFP and/or US. False positive results expose healthy people to unnecessary intervention and alarm, as well as generating considerable additional costs. The credibility of a screening programme can easily be undermined if the screening tests are considered unreliable. In this review, false positive rates for AFP and/or US ranged from 2.9% to 7.5% for HCC.[23 level I]

Not everyone with hepatitis and cirrhosis needs to undergo screening. If, for example the patient is too old or too infirm to undergo an attempt at curative therapy for liver cancer there is no point in screening. Patients who do not have cirrhosis have a negligible risk of developing liver cancer and do not need screening. However, it may be difficult to make the diagnosis of cirrhosis. Moderately advanced cirrhosis can be diagnosed by blood tests and ultrasonography. Early cirrhosis is completely silent, and can only be detected by biopsy. However, patients cannot undergo frequent biopsy simply to find out whether cirrhosis is present or not. Thus, if a biopsy or ultrasound or blood test shows cirrhosis the decision whether to screen or not is easy. The decision is more difficult in patients who have had a biopsy that shows less severe disease than cirrhosis. These patients will, if not treated, progress to cirrhosis over time. The early stages of cirrhosis will be completely undetectable, yet the cancer risk will increase. To deal with this problem some have suggested that even patients with stage III fibrosis (so-called bridging fibrosis, a degree of scarring that is advanced, but not yet cirrhotic) should also undergo screening. However, this makes screening economically unattractive because of the large number of patients who will be screened unnecessarily.[31]

It can be argued that screening without evidence of efficacy is unethical, because surveillance involves not only the inconvenience of regular blood tests, ultrasounds, and extensive secondary radiological imaging, but also results in the diagnosis, albeit early, of tumours that is still untreatable. However, if only small HCC are amenable to therapy, then the approach may be to use the best surveillance tools (currently AFP and US) to find small HCCs and to study the optimal treatment of these lesions through randomized, controlled therapy trials. Given the low resectability rate and survival after surveillance in most Western centers, this is the only way that continued surveillance can be justified. In other words, if treatment trials are not available in a given area for patients with small HCCs, surveillance is inappropriate.[13]

Prorok has enunciated the criteria by which any screening/surveillance programme can be judged: [10]

i. The disease must be common and have a substantial mortality and morbidity.

ii. The target population must be easily identifiable, and there must be an expectation that the physicians caring for the population will accept that screening is necessary and that the population will answer the call for screening.

iii. The screening test must have a low morbidity and a high sensitivity and specificity.

iv. There must be standardized recall procedures.

v. The screening test must be acceptable to the target population.

vi. Finally, and most importantly, there must be an acceptable and effective therapy.

In 1968, Wilson and Jungner authored a WHO document entitled “Principles and Practice of Screening for disease (Public Health Papers, No. 34)”, which has defined ten criteria to be met by mass screening programmes for it to be medically and ethically acceptable. This criterion has been reviewed in 2003 as in **Appendix 5**. Ethical analysis in this context weighs the probable or expected value of mass screening in the population concerned against the assumed or probable risks of adverse physical or psychological effects for those affected if mass screening is or is not done.

7.0 DISCUSSION

When HCC presents with clinical symptoms, the tumour is typically very far advanced and the patient has few therapeutic options. Thus, screening and surveillance for HCC would appear to very appropriate. However, there is no definitive evidence that surveillance improves patient outcomes and current techniques lack of sensitivity and specificity. Nonetheless, serial measurement of AFP levels in serum and hepatic US have become routine practice, despite a lack of evidence of their overall benefit.[32]

7.1 The high-risk group for HCC screening/surveillance

The decision to enter a patient into a screening/surveillance programme is determined by the level of risk for HCC. This, in turn, is related to the incidence of HCC, and it is incidence that most people use to assess risk. However, there are no experimental data to indicate what level of risk or what incidence of HCC should trigger screening/surveillance.[8]

Beasley RP stated that the annual incidence of HCC in male hepatitis B carriers from South East Asia only starts to exceed 0.2% at about age 40 irrespective of presence of cirrhosis or disease activity. In contrast, in Caucasians the risk is related to inflammatory activity and the presence of cirrhosis. Therefore, Asian men should undergo surveillance from age 40 onwards. Hepatocellular carcinoma (HCC) will occur in younger patients, but the efficacy of providing surveillance to all carriers younger than age 40 is likely to be low. The incidence of HCC in women is lower than in men, although age-specific incidence rates are hard to come by. Nonetheless, it seems appropriate to start surveillance at about age 50 in Asian women.[33] Yu MW *et al.* revealed that all hepatitis B carriers with cirrhosis, regardless of age should be screened for HCC. In the presence of a history of a first degree relative with HCC surveillance should start at a younger age than 40, although what that age should be is hard to define.[34]

Hepatitis C infected individuals who do not have cirrhosis have a much lower risk of developing HCC. However, the transition from bridging fibrosis to cirrhosis cannot be determined clinically so that the clinician cannot easily determine when these patients start to develop a significant increase in risk of HCC. For this reason the EASL conference suggested that surveillance may be offered to patients with hepatitis C and cirrhosis or with bridging fibrosis or transition to cirrhosis.[35, 36]

Based on literature, screening/surveillance is recommended for the following high-risk of group:

i. Hepatitis B carriers:
   - Asian males ≥ 40 years
   - Asian females ≥ 50 years

ii. All cirrhotic hepatitis B regardless of age

iii. Family history of HCC

iv. Liver cirrhosis
   - Hepatitis C
   - Alcoholic cirrhosis
   - Genetic hemochromatosis
   - Primary biliary cirrhosis

7.2 Alpha-fetoprotein (AFP) alone

Studies assessing AFP value as a screening tool varied widely in their design and in the characteristics of the patients (type of viral infection, type and severity of liver disease, and so forth). Overall, the studies that assessed the diagnostic accuracy or performance of AFP alone showed that at cut-off level between 20 ng/mL and 200 ng/mL, the sensitivity of AFP ranged from 41.0% to 80.0% and 20.0% to 45.0%, whereas specificity ranged from 80.0% to 95.0% and 99.0% to 100.0% respectively. Positive predictive value (PPV) was reported 3.3% whereas false positive rate was 5.0%.[17, 20-22 level 3, and 23 level 1]

Because AFP test is quantitative, a concern about the optimal cut-off level or positivity threshold was raised. The sensitivity (true-positive rate) and specificity (true-negative rate) depend on the prevalence of HCC in the screened population as well as on the AFP cut-off level chosen for the diagnosis. Most of studies in the review indicated that from the perspective of diagnostic validity, effectiveness, and cost-effectiveness, the recommended cut-off levels was ≥ 20 ng/mL, provides the optimal balance between sensitivity and specificity. The sensitivity of AFP increased as the cut-off level decreased. The specificity of AFP increased as the cut-off level increased.[17, 20-22 level 3, 23, 24 level I, 25 level II-2]

7.3 Ultrasound (US) alone

Similar to AFP, HCC screening studies using US varied by screening frequency, experience of the examiner, and characteristics of liver diseases in screened patients. The reported sensitivity and specificity of US alone in detecting HCC tumour nodules has been quite variable, ranging from 60.0% to 94.0% and 94.0% to 97.1% respectively, depending on the expertise of the operator as well as on the US equipment available (more sophisticated machines produce better quality image and detect smaller tumours). Positive predictive value (PPV) was reported 6.6% whereas false positive rate was 2.9%.[17-20 level 3, 23 level 3] In contrast, two systematic review using US as screening tools in cirrhotic patients were moderately effective; one for detecting early HCC with a sensitivity of 63.0%, while the other only showed 60.0% pooled sensitivity for the 14 US studies.[18-19 level 3]

7.4 Combination of AFP and US

Yang *et al.* found that the sensitivity and specificity for AFP plus US was 92.2% and 95.0%, respectively. Positive predictive value (PPV) was reported 3.0% whereas false positive rate was 7.5%. Sherman *et al.* on the other hand, compared the combination AFP and US with AFP alone in chronic HBV patients. Unfortunately, the results indicated that the sample size was not adequate to compare the two screening methods.[17 level 3]

In a more recent study on the additional benefit of AFP to US during surveillance, it does not substantially improve the sensitivity of surveillance for early HCC, independent of the cut-off level used. Although the pooled sensitivity for early HCC minimally increased from 63.0% to 69.0%, this was not statistically significant (P=0.65). This finding is consistent with the AASLD practice guidelines, which suggested that AFP was not an adequate screening test, but had a role in the diagnosis of HCC with reading of > 200 ng/mL for a mass on imaging.[18 level 3]

7.5 Screening intervals

Based on the diagnostic performance characteristic results, as well as the survival and mortality rate, the interval for screening in high-risk group for HCC has been suggested at six months, [17, 22 level 3 and level I, 23, 24 level I, 25 level II-2] but shorter or longer intervals[18, 19 level 3] have been used. Although there was no prospective comparison of different schedules, in one cost-effectiveness study on chronic HCV and compensated cirrhosis patients, [26] and a Markov decision-analytic model applied to people with compensated cirrhosis[27] suggested that a longer screening interval (12 months) was as cost-effective as the six months interval.

7.6 Detection, mortality and survival rate

Screening for HCC has long been a controversial issue. This review identified three RCTs and one prospective 16-years population-based cohort study on HCC detection, mortality and survival rate.

Yang *et al.* trial reported 86 HCC were detected in the screened group and 51 in controls after five years. However, the two groups did not differ significantly regarding HCC mortality. The survival rate of patients with resected HCC in the screened group reached 52.7% after three to five years, but was 0.0% in the control group.[17 level I]

Chen *et al.* revealed that the mortality rate in the screening group (1,138 per 100,000) was not significantly different from controls (1,114 per 100,000), although AFP screening resulted in earlier diagnosis of liver cancer. There was also no difference in five-year survival between the groups.[22 level

Results from Zhang BH, Yang BH, Tang ZY trial indicated that a higher number of HCCs were diagnosed in the screened than in the control (86 versus 67) at an earlier stage; in addition, overall survival rate at one-, three, and five-year survival rates were better in screened group than in control groups (65.9%, 52.6%, 46.4% versus 31.2%, 7.2%, 0.0%, respectively). A significantly lower HCC mortality rate in the screened group than in controls, being 83.2 per 100,000 and 131.5 per 100,000 respectively with a mortality rate ratio of 0.63 (95% CI: 0.41 to 0.98) were also reported. The results looked promising but there were many factors that contributed to the validity of the results to be considered such as the CI was near to 1.0, intention-to-treat analysis was not used, assessment of outcome was not blinded, and generalizability to other population was uncertain.[24 level I]

The population-based HCC screening programme in Alaska using AFP determination revealed that in most HBsAg+ve carriers, HCC can be detected at an early, potentially resectable stage and showed significant survival benefits at five and 10 years in screened carriers compared with historical controls. However, this finding can be affected by the well-known lead-time and length-time bias of the retrospective screening studies.[25 level II-2]

7.7 Economic evaluation

Given effective screening/surveillance, a programme still faces challenges to its eventual success. These challenges related to cost and the acceptability of screening/surveillance to the physician.

The results of cost-analysis by Zhang B and Yang B demonstrated that the use of US alone is not only more effective, but cheaper than AFP testing alone. Using biannual US screening would detect 32 small PLC at a cost of USD$1,982 each. The combined test not only increased the efficacy of screening, but also increased the false positive rate and the cost.[23]

Lin O *et al.* conducted cost-effectiveness analysis using Markov decision model and based on analysis, they concluded that the best screening protocol has yet to be formally defined. However, based on the results, US at 12-month intervals and AFP at 6-month intervals was a reasonable strategy, offering the greatest gain in life-expectancy while still maintaining an ICER < USD$50,000 per QALY.[26]

From the cost-utility analysis, Thompson CJ *et al.* concluded that the economic efficiency of different surveillance strategies was predicted to vary markedly according to cirrhosis aetiology. In the mixed aetiology cohort, the cheapest surveillance strategy was annual AFP-triage, with ICER of £20,700 per QALY. The addition of US to this strategy increased the ICER to £60,100 per QALY gained.[27]

It must be emphasized that the cost-effectiveness of HCC screening was assessed by retrospective analysis or by using decision models. Although retrospective studies suffered from selection bias, decision-analysis models were based on simulation of costs and health outcomes and, therefore, their results may vary greatly according to different assumptions, such as the incidence of HCC in the screening population, the screening interval, the modality of diagnosis, the type of treatment after diagnosis, the doubling time of tumours, and so forth.

In the Malaysian context, the fees charged by MOH hospital for serum AFP is approximately [redacted] per test, while US varied from [redacted] to [redacted] per imaging.[28] Ultrasound machines start at about [redacted] and range up to [redacted]. The price depends largely on the level of complexity of the machine.[29]

7.8 Limitations

This review has several limitations. Although there was no restriction in language during the search, only English full text articles were included in the report. Although every effort had been made to retrieve full text articles, there were 11 abstracts which the authors failed to retrieve full text. Most of the articles meeting inclusion criteria for this review were systematic reviews and RCT evaluating the diagnostic performance and effectiveness. Most of the diagnostic accuracy studies on AFP and/or US may have introduced bias and limited the conclusions. These limitations included possible verification bias since not all patients undergoing screening test were subjected to CT or MRI or biopsies.

8.0 CONCLUSION

8.1 Diagnostic accuracy of serum AFP and/or US

There was good level of evidence to suggest that:-

a. The sensitivity and specificity of serum AFP alone varies with the cut-off value or positivity threshold. At cut-off value between 20 ng/mL and 200 ng/mL, the sensitivity of AFP ranged from 41.0% to 80.0% and 20.0% to 45.0%, whereas specificity ranged from 80.0% to 95.0% and 99.0% to 100.0% respectively. Positive predictive value (PPV) reported was 3.3% whereas false positive rate was 5.0%. Hence, the recommended cut-off level was ≥ 20 ng/mL.

b. The sensitivity and specificity of US in detecting HCC tumour nodules has been quite variable, ranging from 60.0% to 94.0% and 94.0% to 97.1% respectively, depending on the expertise of the operator as well as on the US equipment available. Positive predictive value (PPV) was reported 6.6% whereas false positive rate was 2.9%.

c. The sensitivity and specificity for a combination of AFP and US was 92.2% and 95.0% respectively, particularly for HCC related with chronic liver infection (HBV). However for HCC related to cirrhosis, the combination AFP and US gave the sensitivity of 69.0%. Positive predictive value (PPV) reported was 3.0% whereas false positive rate was 7.5%.

d. The screening interval of 6 to 12 months was a reasonable strategy for HCC detection

8.2 Effectiveness of HCC screening

There are conflicting evidences from the preliminary findings of three large randomised controlled trials (RCTs) and one cohort study on HCC mortality and survival:-

a. Zhang BH *et al.* trial conducted in Shanghai reported a significantly lower HCC mortality rate in the screened group than in controls, being 83.2 per 100,000 and 131.5 per 100,000 respectively with a mortality rate ratio of 0.63 (95% CI: 0.41 to 0.98). In contrast, two other trials, Yang *et al.* (Shanghai) and Chen *et al.* (Qidong) did not differ significantly concerning HCC mortality.

b. Yang *et al.* and Zhang BH *et al.* reported the overall survival rate at five-year was better in the screened group than in controls (52.7% and 46.4%, respectively). In contrast, there was no difference found in the five-year survival between the screened group and the control groups in Chen *et al.*

c. A prospective 16-year population-based cohort study showed a significant survival benefits at five and 10 years in screened HBsAg+ve Alaskan native carriers compared with historical controls (42.0% and 30.0%, respectively).

8.3 Economic evaluation

There was good level of evidence to suggest that:-

a. The use of US alone at 6-month intervals was not only more effective but cheaper than AFP testing alone. However, the combined tests not only increased the efficacy of HCC screening but also increased the false positive rate and the cost (Zhang B and Yang B in Shanghai).

b. The use of US at 12-month intervals and AFP at 6-month intervals was a reasonable strategy, offering the greatest gain in life-expectancy while still maintaining an ICER < USD$50,000 per QALY (Lin O *et al.* in United States).

c. In cost-utility analysis conducted by Thompson CJ *et al.* (United Kingdom), the cheapest surveillance strategy is annual AFP-triage, with ICER of £20,700 per QALY for the mixed cirrhosis aetiology cohort (alcoholic liver disease, HBV and HCV). At willingness to pay threshold of £30,000 per QALY, the most cost-effective strategy is biannual AFP-triage with ICER of £27,600 per QALY gained.

8.4 Other considerations:

i. The barriers for screening may be different in different countries because of the different health-care system structure and cultural acceptance.

ii. Proper training of staff involved in the screening programme is essential since US is very operator-dependent. Furthermore, screening is expensive and labour intensive thus experience as well as continuous training are definitely needed.

iii. For a mass screening programme to be medically and ethically acceptable, the WHO criteria for mass screening programmes as shown in **Appendix 5** have to be met.

9.0 RECOMMENDATION

Based on this review, good level of evidence on effectiveness (with respect to mortality and survival rate) showed that there was benefits in screening for HCC using serum AFP and/or US in the high-risk group and hence, can be established as part of the Malaysian National Cancer Control Programme. The decision to enter a patient into a *screening programme is determined by the level of risk for HCC and hence, **surveillance is recommended for the follow groups of patients: [8, 33, 34, 35, 36]

i. Hepatitis B carriers:
   - Asian males ≥ 40 years
   - Asian females ≥ 50 years

ii. All cirrhotic hepatitis B regardless of age

iii. Family history of HCC

iv. Liver cirrhosis
   - Hepatitis C
   - Alcoholic cirrhosis
   - Genetic hemochromatosis
   - Primary biliary cirrhosis

There was also good level of evidence to show that the combination of serum AFP and US is the most suitable method to be used for HCC detection, particularly for HCC related with chronic liver infection due to HBV. In addition, the recommended cut-off level of serum AFP was ≥ 20.0 ng/mL, as evidence showed that there was optimal balance between sensitivity and specificity in detecting HCC at this cut-off level.

From the cost-effectiveness perspective, most of the studies in the review indicated that 12-months screening interval using serum AFP and US was as cost-effective as the 6-months interval using serum AFP alone. Hence, the screening interval of 6 to 12 months was a reasonable cost-effective strategy for surveillance of HCC.

However, before commencing the screening programme for HCC detection, it should be noted that currently in Malaysia, serum AFP test are conducted at laboratory hospitals with immunoassay facilities which covers MOH state hospitals and hospitals with specialist amounting to 36 MOH hospitals (personal communication with Head of Chemical Pathologist Activities, MOH). Meanwhile, US examination of the liver is only conducted in 39 MOH hospitals with radiologist (personal communication with Head of Radiology Service, MOH).

Footnote:

***Screening –** application of diagnostic tests in patients at risk for HCC, but in whom there is no a priori reason to suspect that HCC is present.

****Surveillance –** the repeated application of screening tests.

10.0 REFERENCES

1. World Health Organization (WHO). Cancer fact sheets. Available at http://www.who.int/mediacentre/factsheets/fs297/en/. Accessed on 19.01.2012

2. GLOBOCON 2008. Liver cancer incidence and mortality worldwide in 2008 summary. Available at http://globocan.iarc.fr/factsheets/cancers/liver.asp. Accessed on 19.01.2012

3. Jemal A, Bray F, Center M *et al.* Global cancer statistics. CA Cancer J Clin. 2011; 61: 69-90

4. O. Zainal, MA Zainuddin and IT. Nor Saleha. Malaysian Cancer Statistic – Data and Figure Peninsular Malaysia 2006. National Cancer Registry, Ministry of Health Malaysia.

5. Liver Cancer. Malaysian Oncology Society. Available at http://www.malaysiaoncology.org/article.php?aid=35. Accessed on 19.01.2012

6. American Cancer Society. Available at http://www.cancer.org/acs/groups/cid/documents/webcontent/003058-pdf. Accessed on 16.01.2012

7. Bruix J, Sherman M. Practice Guidelines Committee, American Association for the Study of Liver Diseases. Management of hepatocellular carcinoma. Hepatology. 2005; 42: 1208-1236

8. Liver Cancer Overview. Available at http://www.livingwithlivercancer-international.com/scripts/pages/en/about_liver_cancer/ov/index.php. Accessed on 19.01.2012

9. Ying CH. Guidelines on Cancer Prevention, Early Detection & Screening for Liver Cancer. The Hong Kong Anti-Cancer Society. 2009. Available at http://www.hkacs.org.hk/uploadimages/download/00139/Guideline%20on%20liver%20cancer%20prevention,%20detection%20&%20screening%20(full%20recommendation).pdf. Accessed on 19.01.2012

10. Prorok PC. Epidemiology approach for cancer screening. Problems in design and analysis of trials. Am J Pediatr Haematol Oncol. 1992; 14: 117-128

11. AFP Tumor Markes. Lab Tests Online. Available at http://labtestsonline.org/understanding/analytes/afp-tumor/tab/test. Accessed on 19.01.2012

12. Hepatitis B and Liver Cancer Patients. Education Asian Liver Centre. Stanford University. Available at http://liver.stanford.edu?education/hepbpatients.html. Accessed on 19.01.2012

13. Screening for hepatocellular carcinoma. Hepatology Clinical Challenge. 1998; 27(1): 273-278

14. Critical Appraisal Skills Programme (CASP 2004), CASP available at http://www.phru.nhs.uk?Pages/PHD/CASP.htm

15. Harris RP, Helfand M, Woolf SH *et al.* Current Methods of the U.S. Preventive Services Task Force: A review of the Process. Am J Prev Med. 2001;20(suppl 3):21-35

16. Khan KS, Gerben ter Riet, Glanville J *et al.* Undertaking systematic reviews of research on effectiveness. CRD's Guidance for those carrying out or commissioning reviews. *CRD Report Number 4 (2nd Edition) May 2001*. Available at http://www.york.ac.uk/inst/crd/report4.htm

17. Wun YT, and Dickinson JA. Alpha-fetoprotein and/or liver ultrasonography for liver cancer screening in patients with chronic hepatitis B. Cochrane Database of Systematic Review 2003, Issue 2. Edited (no change to conclusion), published in issue 1, 2009.

18. Singal A, Volk M. L, Waljee A *et al.* Meta-analysis: surveillance with ultrasound for early-stage hepatocellular carcinoma in patients with cirrhosis. Alimentary Pharmacology & Therapeutics. 2009; 30: 37-47

19. Colli A, Fraquelli M, Casazza G *et al.* Accuracy of ultrasonography, spiral CT, magnetic resonance, and alpha-fetoprotein in diagnosing hepatocellular carcinoma: a systematic review. American Journal of Gastroenterology. 2006; 101: 513-523

20. Gebo K, Chander G, Jenckes M *et al.* Screening test for hepatocellular carcinoma in patients with chronic hepatitis C: a systematic review. Hepatology. 2002; 36: S84-S92

21. Gupta S, Bent S, and Kohlwes J. Test characteristics of-fetoprotein for detecting hepatocellular carcinoma in patients with hepatitis C. Annals of Internal Medicine. 2003; 139(1): 46-51

22. Chen JG, Parkin DM, Chen QG *et al.* Screening for liver cancer: results of a randomised controlled trial in Qidong, China. Journal of Medical Screening. 2003; 10(4): 204-209

23. Zhang B and Yang B. Combined fetoprotein and ultrasonography as a screening test for primary liver cancer. J Med Screen. 1999; 6: 108-110

24. Zhang BH, Yang BH, and Tang ZY. Randomized controlled trial of screening for hepatocellular carcinoma. J Cancer Res Clin Oncol. 2004; 130: 417-422

25. McMahon B, Bulkow L, Harpster A *et al.* Screening for hepatocellular carcinoma in Alaska natives infected with chronic hepatitis B: a 16-year population-based study. Hepatology. 2000; 32(4)

26. Lin O, Keeffe E, Sanders G *et al.* Cost-effectiveness of screening for hepatocellular carcinoma in patients with cirrhosis due to chronic hepatitis C. Aliment Pharmacol Ther. 2004; 19: 1159-1172

27. Thompson Coon J, Rogers G, Hewson P *et al.* Surveillance of cirrhosis for hepatocellular carcinoma: a cost-utility analysis. British Journal of Cancer. 2008; 98: 1166-1175

28. Akta Fi 1951. Perintah Fi (Perubatan) 2003.

29. How Much Does an Ultrasound Machine Cost. Available at http://www.costowl.com/healthcare/healthcare-ultrasound-machine-costs.html#quotes. Accessed on 29.06.2012

30. SS. Tan, M. Rosmawati, J Menon *et al.* Study of Chronic Hepatitis B (CHB) Patients in Malaysia: Demography and Clinical Features, 19th Conference of the Asian Pacific Association of the Liver. Hepatology International. 2009; 3(1)

31. Sherman M. Screening for hepatocellular carcinoma. HCV advocate. Available at http://www.hcvadvocate.org/hcsp/articles/Sherman-1.html Accessed on 25.05.2012

32. Bisceglie A. Issues in screening and surveillance for hepatocellular carcinoma. Gastroenterology. 2004; 127: 104-107

33. Beasley RP. Hepatitis B virus as the etiologic agent in hepatocellular carcinoma. Hepatology. 1982; 2: 215-265

34. Yu MW, Chang HC, Liaw YF *et al.* Familial risk of hepatocellular carcinoma among chronic hepatitis B carriers and their relatives. J Natl Cancer Inst. 2000; 92: 1159-1164

35. Yoshida H, Shiratori Y, Moriyama M *et al.* Interferon therapy reduces the risk for hepatocellular carcinoma: national surveillance program of cirrhotic and noncirrhotic patients with chronic hepatitis C in Japan. IHIT Study Group. Ann Interm Med. 1999; 131: 174-181

36. Bruix J, Sherman M, Llovet JM *et al.* Clinical management of hepatocellular carcinoma: conclusion of the Barcelona-2000 EASL Conference. J Hepatol 2001; 35: 421-430

11.0 APPENDICIES

APPENDIX 1

HIERARCHY OF EVIDENCE FOR EFFECTIVENESS STUDIES

DESIGNATION OF LEVELS OF EVIDENCE

I Evidence obtained from at least one properly designed randomized controlled trial.

II-I Evidence obtained from well-designed controlled trials without randomization.

II-2 Evidence obtained from well-designed cohort or case-control analytic studies, preferably from more than one centre or research group.

II-3 Evidence obtained from multiple time series with or without the intervention. Dramatic results in uncontrolled experiments (such as the results of the introduction of penicillin treatment in the 1940s) could also be regarded as this type of evidence.

III Opinions or respected authorities, based on clinical experience; descriptive studies and case reports; or reports of expert committees.

SOURCE: US/CANADIAN PREVENTIVE SERVICES TASK FORCE (Harris 2001)

APPENDIX 2

HIERARCHY OF EVIDENCE FOR TEST ACCURACY STUDIES

Level	Description	
1.	A blind comparison with reference standard among an appropriate sample of consecutive patients	
2.	Any one of the following	Narrow population spectrum
3.	Any two of the following	Differential use of reference standard
4.	Any three or more of the following	Reference standard not blind Case control study
5.	Expert opinion with no explicit critical appraisal, based on physiology, bench research or first principles.	

SOURCE: NHS Centre for Reviews and Dissemination (CRD) University of York, Report Number 4 (2nd Edition)

APPENDIX 3

HEALTH TECHNOLOGY ASSESSMENT (HTA) PROTOCOL SERUM ALPHA-FETOPROTEIN (AFP) AND/OR ULTRASOUND (US) FOR HEPATOCELLULAR CARCINOMA (HCC) SCREENING

1.0 BACKGROUND INFORMATION

Most primary liver cancers are classified as hepatocellular carcinoma (HCC). According to World Health Organization (WHO) and GLOBOCON 2008, liver cancer is the seventh most common form of cancer worldwide and the third leading cause of cancer-related death globally. The occurrence of HCC varies widely depending on geographic location. The regions of high incidence are Eastern and South Eastern Asia, Middle and Western Africa, while the lowest in Central and Eastern Europe, Northern Europe, and South-Central Asia.[1, 2, 3] The latest report of the National Cancer Registry (NCR) in Peninsular Malaysia 2006 stated that liver cancer was the sixth most frequent cancer, fifth among male and ninth among female. A total of 793 cases were registered with NCR in 2006 and which comprise of 568 males and 225 females. The incidence of liver cancer increased with age with the overall age-standardised incidence rate (ASR) of 4.9 per 100,000 populations. The incidence was higher among males (ASR of 7.2 per 100,000) compared to females (ASR of 2.7 per 100,000).[4]

The major clinical risk factor for the development of HCC is cirrhosis of the liver. Several different factors can lead to cirrhosis; most common are hepatitis B (HBV) and hepatitis C (HCV) chronic viral infections, excess alcohol consumption, mould toxins (aflatoxins), and non-alcoholic fatty liver disease. Hepatitis B (HBV) is more prevalent in the Asia-Pacific region (excluding Japan) and Africa, whereas HCV is one of the main risk factors for the development of HCC in the United States, Europe, and Japan.[5, 6, 7]

Early detection of liver cancer can be difficult. Often there are no symptoms of liver cancer or liver disease until the disease is in an advanced stage. When symptoms do occur, they may include one or more of the signs such as weight loss (for no apparent reason), ongoing lack of appetite, feeling very full after a small meal, a hard lump on the right side just below the rib cage, pain or discomfort between the stomach and rib cage (possibly around the right shoulder blade), jaundice (yellow-green colour to the skin and whites of the eyes), ascites (abdominal swelling from fluid build up in the abdomen), itching of the skin, unusual tiredness, and nausea.[5, 6, 7]

Based on Malaysian Oncological Society, treatment choices depend on the type and stage of the cancer, how well liver is functioning and the overall health of the individual. Liver cancer generally falls into two categories: [7, 8]

i. *Localized resectable cancer* means that the tumour(s) is located within the liver and has not spread to nearby lymph nodes or other parts of the body; and because the liver is still working well, surgery can typically be performed to remove the tumour(s) or a liver transplant may be considered.

ii. *Localized unresectable cancer* means that even though the cancer has not spread to nearby lymph nodes or other parts of the body, surgery is not possible due to cirrhosis, the location of the tumour within the liver, or other health problems; typically treated with therapies such as radiofrequency ablation, percutaneous ethanol injection therapy or chemoembolization.

Once liver cancer has spread to other parts of the body, advanced cancer treatments may be used to try to slow the progress of the disease. At this stage, palliative care is also offered to all patients to reduce their pain and control other symptoms with prolongation of life as a secondary objective.[7, 8]

Screening for HCC offers the best hope for early detection of preclinical or early symptomatic and improved survival. Many physicians screen patients in high-risk groups with either serum alpha-fetoprotein (AFP) or ultrasound (US) of the liver or both. Two recent randomised controlled trials completed in China demonstrated a significant reduction in HCC-related mortality in patients who underwent screening.[9, 10] Ultrasound (US) of the liver is the preferred screening test because it has a sensitivity of 84% and specificity of more than 90%.[11] A combination of AFP and US has been reported to increase the sensitivity by 5% to 10% over US alone, but it also increases costs and false positive rates.[12, 13] The United States Preventive Services Task Force (USPSTF), National Comprehensive Cancer Network, and American Cancer Society do not have any specific guidelines for screening patients for HCC. The United States National Cancer Institute recommended against routine screening for lack of a survival benefit. More recently, American Association for the Study of Liver Diseases (AASLD) and European Association for the Study of the Liver (EASL) recommended ultrasound of the liver every six months for high-risk patients.[14] The Hong Kong Guidelines on Cancer Prevention, Early Detection & Screening for Liver Cancer stated that high-risk group patients such as chronic hepatitis carrier and known liver cirrhosis patients should received blood test for liver function and AFP together with US of liver every sixth to nine months.[15]

Technical Features

Because of the liver's size and location, it is impossible to detect liver tumours upon physical examination. Instead, an alpha-fetoprotein (AFP) blood test and ultrasound (US) are used for initial screening. An AFP is a protein produced primarily by fetal liver and the portion of a developing embryo that similar to the yolk cavity in bird eggs (yolk sac tissues). Its concentration is typically elevated when a baby is born and decline rapidly. In healthy children and non-pregnant adults, AFP is normally only detectable at very low levels. Liver damage and certain cancers (testicles and ovaries) can increase AFP concentrations significantly. It is also produced whenever liver cells are regenerating such as with chronic liver diseases and tumours. Generally, normal range of AFP is less than 10 ng/ml. Moderate levels of AFP (even almost up to 500 ng/ml) can be seen in patients with chronic hepatitis. Moreover, many patients with various types of acute and chronic liver diseases without documentable liver cancer can have mild or even moderate elevations of AFP. An AFP is used in the screening and diagnosis of liver cancer because AFP is above the normal range in 60% to 70% of primary liver cancer cases. Unfortunately, AFP levels are normal in 30% to 40% of all liver cancer, in which case cancer can only be detected by finding a mass on ultrasound or CT scan.[16, 17]

Limitations in the sensitivity and specificity of AFP in surveillance of high-risk populations led to the use of US as an additional method for detection of HCC. Ultrasound (US) uses reflected sound waves to produce images of organs and other structures in the body. It is commonly used as an initial test for detecting solid lesions in the liver. If a tumour exists, the US may produce a different echo pattern than it does with healthy tissue. The machine will record the images of the liver and a report will be made by the radiologist. Ultrasound (US) is very operator-dependent. This means that unlike regular x-rays, the ability of the person doing the US is important. The quality of US examination is very variable. Furthermore, in the presence of cirrhosis, ultrasound becomes even less reliable, missing a significant number of cancers. Nonetheless, these two tests are widely used. Other alternative screening tools for detection of HCC are computed tomography (CT) or sometimes referred to as computerized axial tomography or CAT scans, and magnetic resonance imaging (MRI). However, both CT and MRI are even more expensive.[3, 7]

With the significant burden of liver cancer all over the world and in Malaysia, one of the strategies for early detection of cancer in the Malaysian National Cancer Management Blueprint 2008-2015 is to provide liver cancer screening service. Therefore, a Health Technology Assessment (HTA) is required to assess the safety, effectiveness and cost-effectiveness of the appropriate test/tests used for HCC screening among the high-risk group in health clinics or hospitals. This HTA was requested by the Senior Principal Assistant Director of Cancer Unit, Disease Control Division, Ministry of Health Malaysia.

2.0 POLICY QUESTION

2.1 Should screening for HCC in the high-risk group be established as part of the Malaysian National Cancer Control Programme?

2.2 Which test method namely using serum AFP alone or US alone or combined is most suitable to be used for HCC detection in the high-risk group?

3.0 OBJECTIVE

3.1 To determine the benefits of HCC screening programme in the high-risk group using AFP and/or US compared with no screening, with regards to patient outcomes such as detection rate, mortality rate, survival rate, quality of life and quality adjusted life years (QALY) gained

3.2 To determine the diagnostic accuracy of AFP and/or US for HCC screening in the high-risk group

3.3 To determine the cost-effectiveness of using serum AFP and/or US for HCC detection in the high-risk group

4.0 METHODOLOGY

4.1 Search Strategy

Electronic database will be searched for published literatures pertaining to AFP and/or US for HCC screening.

4.1.1 Databases as follows; MEDLINE, PubMed, EBM Reviews-Cochrane Database of Systematic Review, EBM-Reviews-Cochrane Central Register of Controlled Trials, EBM Reviews-Health Technology Assessment, EBM Reviews-Cochrane Methodology Register, EBM Reviews-NHS Economic Evaluation Database, Database of Abstracts of Reviews of Effects (DARE), Horizon Scanning, INAHTA Database, HTA database and FDA database.

4.1.2 Additional literatures will be identified from the bibliographies of the related articles.

4.1.3 General search engine will be used to get additional web-based information if there is no retrievable evidence from the scientific databases.

4.1.4 There will be no limitation applied in the search such as year and language.

4.1.5 The search strategy will be included in the appendix.

4.2 Inclusion and exclusion criteria

4.2.1 Inclusion criteria

a. Study design : HTA reports, systematic review, randomised controlled trial (RCT), diagnostic accuracy studies, cross-sectional, cohort, case-control, and economic evaluation studies.

b. Population : High-risk group
   i. Chronic hepatitis B carriers
   ii. Chronic hepatitis C carriers
   iii. Known liver cirrhosis patients

c. Intervention : i. AFP
   ii. US
   iii. Combination of AFP and US

d. Comparators : No screening or usual care

e. Outcome : i. Detection rate, mortality rate, survival rate, quality of life, and quality adjusted life years (QALY) gained.

ii. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of AFP and/or US.

iii. Cost, cost-benefit, cost-effectiveness, and cost utility of using AFP and/or US for HCC screening.

f. Publication : Full text articles published in English

4.2.2 Exclusion criteria

i. Animal study
ii. Narrative review
iii. Experimental study
iv. Non English full text articles

Based on the above inclusion and exclusion criteria, study selection will be carried out independently by two reviewers. Disagreement will be resolved by discussion.

4.3 Data extraction strategy

The following data will be extracted:

4.3.1 Details of methods and study population characteristics.

4.3.2 Detail of intervention and comparators.

4.3.3 Details of individual outcomes for effectiveness, safety and cost associated with AFP and/or US for HCC screening.

4.3.4 Details on diagnostic accuracy (sensitivity, specificity, PPV, NPV) of screening test/tests used in HCC screening.

Data will be extracted from selected studies by a reviewer using a pre-designed data extraction form and checked by another reviewer. Disagreements will be resolved by discussion.

4.4 Quality assessment strategy

The methodology quality of all retrieved literatures will be assessed using the relevant checklist of Critical Appraisal Skill Programme (CASP).

4.5 Methods of analysis/synthesis

Data on the diagnostic accuracy, effectiveness, safety and cost-effectiveness of AFP and/or US for HCC screening will be presented in tabulated format with narrative summaries. No meta-analysis will be conducted for this Health Technology Assessment.

5.0 REPORT WRITING

APPENDIX 4

SEARCH STRATEGY

MEDLINE (OVID) 1950 to August Week 2 2010

1. Carcinoma, Hepatocellular/
2. (hepatocellular adj1 cancer$).tw.
3. (hepatocellular adj1 carcinoma$).tw.
4. hepatoma$.tw.
5. liver cancer$.tw.
6. 1 or 2 or 3 or 4 or 5
7. alpha-Fetoproteins/
8. alpha fetoprotein$.tw.
9. 7 or 8
10. Ultrasonography/
11. Ultrasonography.tw.
12. (Ultrasonic adj1 imaging$).tw.
13. (medical adj1 sonography).tw.
14. 10 or 11 or 12 or 13
15. Magnetic Resonance Imaging/
16. magnetic resonance imaging.tw.
17. (mri adj1 scan$).tw.
18. 15 or 16 or 17
19. Tomography, X-Ray Computed/
20. x ray ct.tw.
21. tomography x-ray computer assisted.tw.
22. computerized tomography x ray.tw.
23. 19 or 20 or 21 or 22
24. 6 or 9 or 14
25. 6 and 18 and 23 and 24
26. limit 25 to (english language and humans)

EBM Reviews: Cochrane Database of Systematic Reviews (2005 to May 2012), Cochrane Central Register of Controlled Trials (1st Quarter 2012), HTA Databases (1st Quarter 2012), Cochrane Methodology Register (1st Quarter 2012), ACP Journal Club (1991 to May 2012), NHS Economic Evaluation Database (1st Quarter 2012) via OVID

1. CCarcinoma, Hepatocellular/
2. (hepatocellular adj1 cancer$).tw.
3. (hepatocellular adj1 carcinoma$).tw.
4. hepatoma$.tw.
5. liver cancer$.tw.
6. 1 or 2 or 3 or 4 or 5
7. alpha-Fetoproteins/
8. alpha fetoprotein$.tw.

9. 7 or 8
10. Ultrasonography/
11. Ultrasonography.tw.
12. (Ultrasonic adj1 imaging$).tw.
13. (medical adj1 sonography).tw.
14. 10 or 11 or 12 or 13
15. Magnetic Resonance Imaging/
16. magnetic resonance imaging.tw.
17. (mri adj1 scan$).tw.
18. 15 or 16 or 17
19. Tomography, X-Ray Computed/
20. x ray ct.tw.
21. tomography x-ray computer assisted.tw.
22. computerized tomography x ray.tw.
23. 19 or 20 or 21 or 22
24. 6 or 9 or 14
25. 6 and 18 and 23 and 24
26. limit 25 to (english language and humans)

PubMed

1. Carcinoma, Hepatocellular (Mesh)
2. hepatocellular cancer*
3. hepatocellular carcinoma*
4. hepatoma*
5. liver cancer*
6. 1 or 2 or 3 or 4 or 5
7. alpha-Fetoproteins(mesh)
8. alpha fetoprotein*
9. 7 or 8
10. Ultrasonography/
11. Ultrasonography
12. Ultrasonic imaging*
13. medical sonography
14. 10 or 11 or 12 or 13
15. Magnetic Resonance Imaging(mesh)
16. magnetic resonance imaging
17. mri scan*
18. 15 or 16 or 17
19. Tomography, X-Ray Computed(Mesh)
20. x ray ct
21. tomography x-ray computer assisted
22. computerized tomography x ray
23. 19 or 20 or 21 or 22
24. 6 or 9 or 14
25. 6 and 18 and 23 and 24

APPENDIX 5

SCREENING CRITERIA

The Wilson-Jungner criteria for appraising the validity of a screening programme

1. The condition being screened for should be an important health problem
2. The natural history of the condition should be well understood
3. There should be a detectable early stage
4. Treatment at an early stage should be of more benefit than at a later stage
5. A suitable test should be devised for the early stage
6. The test should be acceptable
7. Intervals for repeating the test should be determined
8. Adequate health service provision should be made for the extra clinical workload resulting from screening
9. The risks, both physical and psychological, should be less than the benefits
10. The costs should be balanced against the benefits

World Health Organisation 1968

Criteria for appraising the viability, effectiveness and appropriateness of a screening programme 2003

The condition

1. The condition should be an important health problem.
2. The epidemiology and natural history of the condition, including development from latent to declared disease, should be adequately understood and there should be a detectable risk factor, disease marker, latent period or early symptomatic stage.
3. All the cost-effective primary prevention interventions should have been implemented as far as practicable.
4. If the carriers of a mutation are identified as a result of screening the natural history of people with this status should be understood, including the psychological implications.

The test

5. There should be a simple, safe, precise and validated screening test.
6. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed.
7. The test should be acceptable to the population.
8. There should be an agreed policy on the further diagnostic investigation of individuals with a positive test result and on the choices available to those individuals.
9. If the test is for mutations the criteria used to select the subset of mutations to be covered by screening, if all possible mutations are not being tested for, should be clearly set out.

The treatment

10. There should be an effective treatment or intervention for patients identified through early detection, with evidence of early treatment leading to better outcomes than late treatment.
11. There should be agreed evidence-based policies covering which individuals should be offered treatment and the appropriate treatment to be offered.
12. Clinical management of the condition and patient outcomes should be optimised in all healthcare providers prior to participation in a screening programme.

The screening programme

13. There should be evidence from high-quality randomised controlled trials that the screening programme is effective in reducing mortality or morbidity. Where screening is aimed solely at providing information to allow the person being screened to make an 'informed choice' (for example, Down's syndrome and cystic fibrosis carrier screening), there must be evidence from high-quality trials that the test accurately measures risk. The information that is provided about the test and its outcome must be of value and readily understood by the individual being screened.
14. There should be evidence that the complete screening programme (test, diagnostic procedures, treatment/intervention) is clinically, socially. and ethically acceptable to health professionals and the public.
15. The benefit from the screening programme should outweigh the physical and psychological harm (caused by the test, diagnostic procedures and treatment).
16. The opportunity cost of the screening programme (including testing, diagnosis and treatment, administration, training and quality assurance) should be economically balanced in relation to expenditure on medical care as a whole (ie value for money).
17. There should be a plan for managing and monitoring the screening programme and an agreed set of quality assurance standards.
18. Adequate staffing and facilities for testing, diagnosis, treatment, and programme management should be available prior to the commencement of the screening programme.
19. All other options for managing the condition should have been considered (for example, improving treatment and providing other services), to ensure that no more cost-effective intervention could be introduced or current interventions increased within the resources available.
20. Evidence-based information, explaining the consequences of testing, investigation, and treatment, should be made available to potential participants to assist them in making an informed choice.
21. Public pressure for widening the eligibility criteria for reducing the screening interval, and for increasing the sensitivity of the testing process, should be anticipated. Decisions about these parameters should be scientifically justifiable to the public.
22. If screening is for a mutation, the programme should be acceptable to people identified as carriers and to other family members.

http://www.gp-training.net/training/tutorials/management/audit/screen.htm

APPENDIX 6

Evidence Table : **Diagnostic accuracy**
Question : **What is the diagnostic accuracy of AFP and/or US in detecting HCC?**

Bibliographic citation	1. Wun YT, and Dickinson JA. Alpha-fetoprotein and/or liver ultrasonography for liver cancer screening in patients with chronic hepatitis B. Cochrane Database of Systematic Review 2003, Issue 2. Edited (no change to conclusion), published in issue 1, 2009.
Study Type / Methods	Systematic review Included two RCTs: Yang *et al.* 1999 • Shanghai, China • Cluster randomised trial • 300 factories, enterprises, or school were listed: a random sample from these units were drawn and allocated into a screening and control group • No information on generation of the allocation sequence and allocation concealment • No blinding • Used intention to treat analysis Sherman *et al.* 1995 • Toronto, Canada • Randomised clinical feasibility trial • No information on generation of the allocation sequence and allocation concealment • No blinding • Used intention to treat analysis Both randomised trials use the threshold of 20 µg/L for AFP.
LE	3
Number of patients & Patient characteristics	Yang *et al.* 1999 18,816 Shanghai urban residents aged 35 to 55 years with **HBsAg+ve** or **chronic hepatitis** in either screened (n=9,373) or control group (n=9,443) Sherman *et al.* 1995 1,069 confirmed **chronic hepatitis B** aged over 18 years (mean 39 ± 12 years). They were randomised in two arms: AFP arm (n=531) and AFP + US (n=538)
Intervention	Yang *et al.* AFP + US every six months Sherman *et al.* AFP + US every six months
Comparison	Yang *et al.* No screening Sherman *et al.* AFP
Length of follow up (if applicable)	Yang *et al.* 5 years Sherman *et al.* 26 months
Outcome measures/ Effect size	Sensitivity • The Yang *et al.* trial, though screening with AFP + US, evaluated the tests separately, calculating a sensitivity of 68.6% for AFP alone, 84.3% for US alone, and 92.2% for AFP + US. • In the Sherman *et al.* trial, the sensitivity was 64.3% for AFP and 71.4% for US. Specificity • The Yang *et al.* trial calculated a specificity of 95.0% for AFP alone, 97.1% for US alone, and 95.0% for AFP + US. • The Sherman *et al.* trial calculated specificities of 91.4% and 93.8% for AFP and US respectively. PPV • In Yang *et al.*, the PPV was 3.3% for AFP, 6.6% for US, and 3.0% for AFP + US. • In Sherman *et al.*, the PPV for AFP was 9.0% and the PPV for US was 13.2%. NPV • In Yang *et al.*, the NPV was 99.9% for AFP, 99.9% for US, and 100.0% for AFP + US. • In Sherman *et al.*, the NPV for AFP was 99.5% and the NPV for US was 99.6%. Noted: For Sherman *et al.* trial, sensitivity etc. unknown for both the AFP and the AFP + US trial arms; and the outcome of screening is difficult to know exactly because the data are not clearly given for the two screening groups, and the figures are not always correctly given.
General comments	

Evidence Table : Diagnostic accuracy
Question : What is the diagnostic accuracy of AFP and/or US in detecting HCC?

Bibliographic citation	2. Singal A, Volk M. L, Waljee A *et al.* Meta-analysis: surveillance with ultrasound for early-stage hepatocellular carcinoma in patients with cirrhosis. Alimentary Pharmacology & Therapeutics. 2009; 30: 37-47
Study Type / Methods	Systematic review with meta-analysis Using the MEDLINE and SCOPUS databases through 1 July 2007 yielded six studies that evaluated the accuracy (performance characteristics) of ultrasound (US) for HCC at any stage and 13 studies that were specific to early HCC. Aim: To determine the pooled sensitivity, specificity and diagnostic OR with 95% CI of US and AFP for the detection of HCC, particularly early HCC, during surveillance. Random effect model was used. Studies were selected by two of the authors based on inclusion criteria. Two authors independently assessed the selected trials.
LE	I
Number of patients & Patient characteristics	6,555 **cirrhotic patients** (from 19 studies)
Intervention	i. US ii. AFP and US
Comparison	
Length of follow up (if applicable)	6 to 12 months surveillance interval
Outcome measures/ Effect size	<u>US for detecting HCC at any stage</u> • Pooled sensitivity was 94.0% (95% CI: 83.0% to 98.0%), a pooled specificity of 94.0% (95% CI: 89.0% to 97.0%), and a pooled diagnostic OR of 232.7 (95% CI: 105.9 to 511.2). • A summary receiver operator characteristics curve (SROC) analysis demonstrated an area under the curve (AUC) of 0.98 (95% CI: 0.96 to 0.99) suggesting high diagnostic accuracy. <u>US for detecting early HCC</u> • Pooled sensitivity was 63.0% (95% CI: 49.0% to 76.0%) Studies with surveillance intervals of < 6 months had a pooled sensitivity of 70.1% (95% CI: 55.6% to 84.6%), while the studies with surveillance intervals between 6 and 12 months had a pooled sensitivity of 50.1% (95% CI: 40.0% to 59.2%) <u>AFP and US for detecting early HCC</u> • Pooled sensitivity increased to 69.0% (95% CI: 53.0% to 81.0%; P=0.65) • **A wide range of AFP cut-offs (15 to 400 ng/mL) were used to diagnose HCC in the included studies, although the cut-off level did not appear to affect the utility of AFP (P=0.95)** Authors conclusion: Surveillance with US demonstrates limited sensitivity for early HCC, although this may be improved by testing at 6-month intervals. AFP provided no additional benefit to US. Currently available evidence evaluating surveillance US has significant limitations such as verification bias and are of suboptimal quality.
General comments	

Evidence Table : Diagnostic accuracy
Question : What is the diagnostic accuracy of AFP and/or US in detecting HCC?

Bibliographic citation	3. Colli A, Fraquelli M, Casazza G *et al.* Accuracy of ultrasonography, spiral CT, magnetic resonance, and alpha-fetoprotein in diagnosing hepatocellular carcinoma: a systematic review. American Journal of Gastroenterology. 2006; 101: 513-523
Study Type / Methods	Systematic review with meta-analysis Pertinent cross-sectional studies having as **reference standard pathological examinations** of the explanted liver or resected segment(s), biopsies of focal lesion(s), and/or a period of follow-up, were identified using MEDLINE, EMBASE, Cochrane Library, and CancerLit. Pooled sensitivity, specificity, and likelihood ratios (LR) were calculated using the random effect model. Studies were selected by four reviewers based on inclusion criteria, and independently assessed the selected trials.
LE	I
Number of patients & Patient characteristics	Patients with **chronic liver disease (cirrhosis or chronic hepatitis)** assessed with the aim of detecting the possible presence of HCC
Intervention	i. US ii. Spiral CT iii. MRI iv. AFP
Comparison	Reference standard: Pathological examinations of the explanted liver or resected segment(s), biopsies of focal lesion(s), and/or a period of follow-up
Length of follow up (if applicable)	
Outcome measures/ Effect size	The pooled estimates of the **14 US studies** were 60.0% (95% CI: 44.0% to 76.0%) for sensitivity, 97.0% (95% CI: 95.0% to 98.0%) for specificity, 18.0 (95% CI: 8.0 to 37.0) for LR+, and 0.5 (95% CI: 0.4 to 0.6) for LR-. For the **10 CT studies**, pooled sensitivity was 68.0% (95% CI: 55.0% to 80.0%), specificity 93.0% (95% CI: 89.0% to 96.0%), LR+ 6.0 (95% CI: 3.0 to 12.0), and LR- 0.4 (95% CI: 0.3 to 0.6). For the **9 MRI studies**, pooled sensitivity was 81.0% (95% CI: 70.0% to 91.0%), specificity 85.0% (95% CI: 77.0% to 93.0%), LR+ 3.9 (95% CI: 2.0 to 7.0), and LR- 0.3 (95% CI: 0.2 to 0.5). From the **9 AFP studies**, the sensitivity and specificity varied widely, and this could not be entirely attributed to the threshold effect of the different cut-off levels used. Authors conclusion: • US is highly specific but insufficiently sensitive to detect HCC in many cirrhotics or to support an effective surveillance program. • The role of CT and MRI remains to be defined, particularly as additional tests aimed at confirming positive US results • High-quality prospective studies are needed to define the actual diagnostic role of AFP
General comments	

Evidence Table : Diagnostic accuracy
Question : What is the diagnostic accuracy of AFP and/or US in detecting HCC?

Bibliographic citation	4. Gebo K, Chander G, Jenckes M *et al.* Screening test for **hepatocellular carcinoma in patients with chronic hepatitis C**: a systematic review. Hepatology. 2002; 36: S84-S92
Study Type / Methods	Systematic review The search strategy involved searching MEDLINE and other electronic databases between January 1985 and March 2002 on two key questions: 1. What is the efficacy of using screening test for HCC to improve clinical outcomes? 2. What are the sensitivity and specificity of screening tests for HCC? Studies were selected by two reviewers based on inclusion criteria, and independently assessed the selected trials.
LE	I
Number of patients & Patient characteristics	Patients with **chronic hepatitis C**
Intervention	i. AFP ii. US
Comparison	No screening
Length of follow up (if applicable)	
Outcome measures/ Effect size	<u>Answer key Question 1</u> **1 prospective** cohort surveillance study was identified, suggested that HCC was detected earlier and was more often resectable in patients who had twice yearly screening with AFP and hepatic US than in patients who had usual care: <u>Answer key Question 2</u> • **24 studies**, which included patients with chronic hepatitis C or B or both, addressed the sensitivities and specificities of screening tests. • The **sensitivity of AFP** for detecting HCC was moderately high at at 60.0% to 80.0%, with a specificity of 70.0% to 90.0%, for a threshold value decreases from 400.0 ng/mL to 10.0 ng/mL • The sensitivity and specificity of US were limited in that some were designed to assess the incidence of HCC and not to assess the performance characteristics of US Authors conclusion: Screening of patients with chronic hepatitis C with AFP and US may improve detection of HCC, but studies are needed to determine whether screening improves clinical outcome.
General comments	

Evidence Table : **Diagnostic accuracy**
Question : **What is the diagnostic accuracy of AFP and/or US in detecting HCC?**

Bibliographic citation	5. Gupta S, Bent S, and Kohlwes J. Test characteristics of α-fetoprotein for detecting hepatocellular carcinoma in patients with hepatitis C. Annals of Internal Medicine. 2003; 139(1): 46-51
Study Type / Methods	Systematic review MEDLINE search was performed from 1966 through December 2002 for English and non-English articles examining the test characteristics of AFP for identifying HCC. Relevant articles were evaluated for quality of evidence by three authors.
LE	I
Number of patients & Patient characteristics	**HCV-infected patients with or without cirrhosis**
Intervention	AFP
Comparison	CT, MRI, histopathology, and disease-free time greater than 2 years were considered adequate gold standards.
Length of follow up (if applicable)	
Outcome measures/ Effect size	**5 studies** met all inclusion criteria and were analyzed included: • **2 studies were prospective cohort** design and **3 were case-control** designs that potentially overestimate diagnostic accuracy. By using the most commonly reported cut-off value of a positive test result for HCC (**AFP level > 20 µg/l**), • The sensitivity ranged from 41.0% to 65.0%, while specificity ranged from 80.0% to 94.0%. LR+ ranged from 3.1 to 6.8 and LR- ranged from 0.4 to 0.6. 4 of the 5 studies reported sensitivity and specificity for an **AFP cut-off value > 200 µg/l**, a value that is frequently reported to be specific for the diagnosis of HCC. • The range of specificities was very high at this cut-off value (99.0% to 100.0%), but the sensitivity was very low (20.0% to 45.0%) Authors conclusion: Current studies examining the test characteristics of AFP for diagnosing HCC in patients with HCV have substantial methodological limitations, making it difficult to define clear estimates of sensitivity and specificity for this test. A prospective study done with careful attention to limitation of bias is clearly needed to define whether any screening strategy can provide clinically important benefits.
General comments	

Evidence Table : Diagnostic accuracy
Question : What is the diagnostic accuracy of AFP and/or US in detecting HCC?

Bibliographic citation	6. Chen JG, Parkin DM, Chen QG *et al.* Screening for liver cancer: results of a randomised controlled trial in **Qidong, China**. Journal of Medical Screening. 2003; 10(4): 204-209
Study Type / Methods	Randomised controlled trial To investigate the effectiveness of screening for liver cancer in reducing mortality from the disease in a high-risk population in China during the period 1989 to 1995. Screening was planned to be 6 monthly AFP assays, with follow-up of subjects having an abnormal test (≥ 20 µg/l). All subjects were followed up for liver cancer and/or death until 31 December 1995. Case confirmation: US (majority), US + CT (4.3%), pathological (10.4%), autopsy, or aspiration of liver tissue post mortem (12.0%)
LE	I
Number of patients & Patient characteristics	5,581 men aged 30 to 69 years (**HBsAg positive**) were identified by population screening and randomly assigned to a screening group (group A, 3,712), and controls (group B, 1,869).
Intervention	AFP
Comparison	No screening
Length of follow up (if applicable)	6 years
Outcome measures/ Effect size	The overall (on 1 or more occasion) sensitivity and specificity of the screening programme for liver cancer using AFP was **55.3% and 86.5%**, respectively. These figures include cases of liver cancer among subjects who did not attend for screening at some or all of the scheduled examinations. In subjects who complied with all scheduled screening tests, the sensitivity was **80.0% and specificity 80.9%**.
General comments	

Evidence Table : Diagnostic accuracy
Question : What is the diagnostic accuracy of AFP and/or US in detecting HCC?

Bibliographic citation	7. Zhang B and Yang B. Combined α fetoprotein and ultrasonography as a screening test for primary liver cancer. J Med Screen. 1999; 6: 108-110
Study Type / Methods	Randomised controlled trial To assess the validity and cost of a screening test for primary liver cancer (PLC) using combined serum AFP and US. Controls received no screening whereas screening group were invited to have an AFP test and US examination every 6 months. A screening test was defined as positive when AFP > 20 µg/l, and when US demonstrated a new solid lesion in the liver. Outcome measures were detection rate, false positive rate, PPV, the cost for each PLC detected, and the average cost of detecting each additional PLC by the combined method. The number of small PLC detected was used for the economic evaluation.
LE	I
Number of patients & Patient characteristics	18,816 people, aged 35 to 59 years with **hepatitis B virus infection or a history of chronic hepatitis** in urban Shanghai, China. Subjects were randomly allocated to a screening (n=9,373) or control (n=9,443) group.
Intervention	i. AFP + US ii. AFP alone iii. US alone (Gold Standard: CT, MRI & biopsy)
Comparison	No screening
Length of follow up (if applicable)	5 years
Outcome measures/ Effect size	Primary liver cancer (PLC) was detected in 51 subjects, 36 of whom had small PLC. i. AFP + US Detection rate: 92.0% False positive rate: 7.5% PPV: 3.0% ii. AFP alone Detection rate: 69.0% False positive rate: 5.0% PPV: 3.3% iii. US alone Detection rate: 84.0% False positive rate: 2.9% PPV: 6.6% Authors conclusion: The results of this study show that the use of US is not only more effective, but cheaper than AFP testing alone. The combined test increases the efficacy of screening, but also increases the false positive rate and the cost. Therefore, combined AFP and US as a screening test is suitable for relatively developed areas of China, but otherwise, US alone is the method of choice.
General comments	

Evidence Table : Diagnostic accuracy
Question : Is HCC screening using AFP and/or US effective in term of detection rate, mortality rate, survival rate, QOL, QALY?

Bibliographic citation	1. Wun YT, and Dickinson JA. Alpha-fetoprotein and/or liver ultrasonography for liver cancer screening in patients with chronic hepatitis B. Cochrane Database of Systematic Review 2003, Issue 2. Edited (no change to conclusion), published in issue 1, 2009.
Study Type / Methods	Systematic review Included two RCTs: Yang *et al.* 1999 • Shanghai, China • Cluster randomised trial • 300 factories, enterprises, or school were listed: a random sample from these units were drawn and allocated into a screening and control group • No information on generation of the allocation sequence and allocation concealment • No blinding • Used intention to treat analysis Sherman *et al.* 1995 • Toronto, Canada • Randomised clinical feasibility trial • No information on generation of the allocation sequence and allocation concealment • No blinding • Used intention to treat analysis Both randomised trials use the threshold of 20 µg/L for AFP.
LE	I
Number of patients & Patient characteristics	Yang *et al.* 1999 18,816 Shanghai urban residents aged 35 to 55 years with **HBsAg+ve or chronic hepatitis** in either screened (n=9,373) or control group (n=9,443) Sherman *et al.* 1995 1,069 confirmed **chronic hepatitis B** aged over 18 years (mean 39 ± 12 years). They were randomised in two arms: AFP arm (n=531) and AFP + US (n=538)
Intervention	Yang *et al.* AFP + US every six months Sherman *et al.* AFP + US every six months
Comparison	Yang *et al.* No screening Sherman *et al.* AFP
Length of follow up (if applicable)	Yang *et al.* 5 years Sherman *et al.* 26 months
Outcome measures/ Effect size	Yang *et al.* i. Number of detected HCC • A total of 86 HCC were detected in the screened group and 51 HCC in the control group after five years (**OR=1.69**, 95% CI: 1.20 to 2.36). ii. Survival rate • For year-1 to year-5 were given by the authors as: (a) screened group (86 HCC) 65.0%, 60.2%, 52.7%, 52.7%, 52.7% (b) control group (51 HCC) 30.0%,6.5%, 0.0%, 0.0%, 0.0%. iii. HCC mortality • 41/9,373 (0.44%) screened patients and 51/9,443 (0.54%) control patients died from HCC. This difference was not significant (OR=0.81, 95% CI: 0.54 to 1.22). Sherman *et al.* i. Number of detected HCC • Identified six HCC in the AFP + US arm and eight HCC in the AFP arm (OR=0.74, 95% CI: 0.26 to 2.12). ii. Survival rate • 5 patients survived 6 to 38 months after surgery, without further information to compare the AFP + US group versus the AFP group iii. HCC mortality • 5/1,069 (0.46%) but no data were given about the number of deaths in the two intervention arms of the trial.
General comments	

Evidence Table : Diagnostic accuracy
Question : Is HCC screening using AFP and/or US effective in term of detection rate, mortality rate, survival rate, QOL, QALY?

Bibliographic citation	2. Chen JG, Parkin DM, Chen QG *et al.* Screening for liver cancer: results of a randomised controlled trial in Qidong, China. Journal of Medical Screening. 2003; 10(4): 204-209
Study Type / Methods	Randomised controlled trial To investigate the effectiveness of screening for liver cancer in reducing mortality from the disease in a high-risk population in China during the period 1989 to 1995. Screening was planned to be six monthly AFP assays, with follow-up of subjects having an abnormal test (≥ 20 µg/l). All subjects were followed up for liver cancer and/or death until 31 December 1995. Case confirmation: US (majority), US + CT (4.3%), pathological (10.4%), autopsy, or aspiration of liver tissue post mortem (12.0%)
LE	I
Number of patients & Patient characteristics	5,581 men aged 30 to 69 years (**HBsAg positive**) were identified by population screening and randomly assigned to a screening group (group A, 3,712), and controls (group B, 1,869).
Intervention	AFP
Comparison	No screening
Length of follow up (if applicable)	6 years
Outcome measures/ Effect size	Overall, 374/5,581 (6.7%) liver cancer were diagnosed. The incidence was higher (1,341.7 per 100,000) in the screened group A, than in the group B controls (1,195.6 per 100,000), but the difference was not significant (p=0.30). 327/374 cases with primary liver cancer died due to this cancer during the follow-up period. The mortality rate from liver cancer was 1,138.1 per 100,000 in the screened group A and 1,113.9 per 100,000 in the group B controls, but the difference was not statistically significant (p=0.86). The percentage of cases in clinical stage I was significantly higher in group A (29.6%) than in group B (6.0%). The one-, three, and five-year relative survival rates were 23.7%, 7.0%, and 4.0% in group A, and 9.7%, 4.0%, and 4.1% in group B, respectively, with no difference in five-year survival between the groups. A Poisson regression model showed that the probability of death (rate ratio) in the screening group was 0.83 (95% CI: 0.68 to 1.03) relative to the control group. Authors conclusion: Screening with AFP resulted in earlier diagnosis of liver cancer, but the gain in lead time did not result in any overall reduction in mortality, because therapy for the patients found by screening was ineffective. Further studies using improved methods of screening, diagnosis and treatment are indicated.
General comments	

Evidence Table : Diagnostic accuracy
Question : Is HCC screening using AFP and/or US effective in term of detection rate, mortality rate, survival rate, QOL, QALY?

Bibliographic citation	3. Zhang BH, Yang BH, and Tang ZY. Randomized controlled trial of screening for hepatocellular carcinoma. J Cancer Res Clin Oncol. 2004; 130: 417-422
Study Type / Methods	Randomised controlled trial Aim: To assess the effect of biannual screening with a combination of AFP and US on HCC mortality in people at increased risk. Screening group participants were invited to have an AFP test and US examination every 6 months. Controls received no screening and continued to use health-care facilities. Screening was stopped in December 1997 and all participants were followed up until December 1998. The cut-off value of AFP was 20 µg/l. An abnormality detected by US was a solid lesion in the liver. All analyses were on an intention-to-treat basis.
LE	I
Number of patients & Patient characteristics	18,816 people, aged 35 to 59 years with **hepatitis B virus infection or a history of chronic hepatitis** in urban Shanghai, China. Participants were randomly allocated to a screening (n=9,373) or control (n=9,443) group.
Intervention	Combination of AFP and US
Comparison	No screening
Length of follow up (if applicable)	5 years
Outcome measures/ Effect size	The screened group completed 58.2% of the screened offered. When the screened group was compared to the control group: • The number of HCC was 86 versus 67 • Subclinical HCC being 52/86 (60.5%) versus 0 • Small HCC 39/86 (45.3%) versus 0 • Resection achieved 40/86 (46.5%) versus 5/67 (7.5%) • 1-, 3-, and 5-year survival rate 65.9%, 52.6%, 46.4% versus 31.2%, 7.2%, 0.0%, respectively 32 people died from HCC in the screened group versus 54 in the control group. The HCC mortality rate was significantly lower in the screened group than in controls, being 83.2 per 100,000 and 131.5 per 100,000, respectively. The rate ratio for mortality from HCC was 0.63 (95% CI: 0.41 to 0.98). These results reveal a significant reduction in mortality at 5-year follow-up in the screened group compared to the control group. Authors conclusion: Biannual screening with a combination of AFP and US reduced HCC mortality after 5-year follow-up by 37.0%.
General comments	

Evidence Table : **Diagnostic accuracy**
Question : **Is HCC screening using AFP and/or US effective in term of detection rate, mortality rate, survival rate, QOL, QALY?**

Bibliographic citation	4. McMahon B, Bulkow L, Harpster A *et al.* Screening for hepatocellular carcinoma in Alaska natives infected with chronic hepatitis B: a 16-year population-based study. Hepatology. 2000; 32(4)
Study Type / Methods	Cohort study Prospective 16-year, population-based cohort study to determine the impact of screening for HCC in 1,487 HBsAg-positive Alaska native carriers with AFP determinations every 6 months. Men and non-pregnant women with an elevated AFP level were evaluated for the presence of HCC by US examination. The long-term survival rate for patients whose HCC was detected by the screening programme was compared with a historical control group of Alaska native patients with HCC from the same population who were clinically diagnosed with HCC between 1969 and October 1982, through a National Cancer Institute-sponsored Cancer Registry.
LE	II-2
Number of patients & Patient characteristics	1,487 patients chronically infected with HBV (HBsAg-B positive for 12 months or longer)
Intervention	AFP followed by US
Comparison	
Length of follow up (if applicable)	16 years
Outcome measures/ Effect size	Between October 1982 and December 1998, 26,752 AFP determinations in HbsAg carriers were performed. During the 16-year period, one or more AFP elevations were found in 61 men and 39 nonpregnant woman During follow-up, HCC was diagnosed in 32 patients (24 men and 8 women). HCC tumors less than 6 cm were found in 23 patients; 22 patients had resections, and 1 patient refused a resection. Compared with 12 patients with hepatitis B virus (HBV)-related HCC diagnosed from 1969 to October 1982, before this programme, the 5- and 10-year survival rate for the 32 patients with HCC were 42.0% (P=0.008) and 30.0% (P=0.07), respectively. 5- and 10-year tumor-free survival rates for carriers who had a normal AFP level on initial screening and subsequently developed HCC were 29.0% (P=0.004) and 24.0% (P=0.024), respectively. Authors conclusion: This population-based screening programme using AFP determination showed that in most HbsAg-positive carriers, HCC can be detected at an early, potentially resectable stage and showed significant survival benefits at 5 and 10 years in screened carriers compared with historical controls.
General comments	

Evidence Table : Diagnostic accuracy
Question : What is the cost/cost-effectiveness of using AFP and/or US for HCC screening?

Bibliographic citation	1. Zhang B and Yang B. Combined α fetoprotein and ultrasonography as a screening test for primary liver cancer. J Med Screen. 1999; 6: 108-110
Study Type / Methods	Direct cost analysis To assess the validity and cost of a screening test for primary liver cancer (PLC) using combined serum AFP and US. Controls received no screening whereas screening group were invited to have an AFP test and US examination every 6 months. A screening test was defined as positive when AFP > 20 µg/l, and when US demonstrated a new solid lesion in the liver. The cost was 5 RMB ($0.60 ($1 = 8.3 RMB)) for an AFP test, 10 RMB ($1.20) for an US examination, and about 500 RMB ($60) for investigating a screen positive patient. Detection of a small PLC (diameter < 5 cm) was used as a measure of the effectiveness in the cost-effectiveness analysis.
LE	I
Number of patients & Patient characteristics	18,816 people, aged 35 to 59 years with **hepatitis B virus infection or a history of chronic hepatitis** in urban Shanghai, China. Subjects were randomly allocated to a screening (n=9,373) or control (n=9,443) group.
Intervention	i. AFP + US ii. AFP alone iii. US alone (Gold Standard: CT, MRI & biopsy)
Comparison	No screening
Length of follow up (if applicable)	5 years
Outcome measures/ Effect size	Primary liver cancer (PLC) was detected in 51 subjects, 36 of whom had small PLC. i. <u>AFP + US</u> Small PLC (n): 36 Total cost (RMB): 1,087,410 Cost for each small PLC (RMB): 30,206 (USD $3,639) ii. <u>AFP alone</u> Small PLC (n): 25 Total cost (RMB): 628,470 Cost for each small PLC (RMB): 25,139 (USD$3,029) iii. <u>US alone</u> Small PLC (n): 32 Total cost (RMB): 526,440 Cost for each small PLC (RMB): 16,451(USD$1,982) *RMB (Chinese currency) (1 RMB = $0.1205) Authors conclusion: The results of this study show that the use of US is not only more effective, but cheaper than AFP testing alone. The combined test increases the efficacy of screening, but also increases the false positive rate and the cost. Therefore, combined AFP and US as a screening test is suitable for relatively developed areas of China, but otherwise, US alone is the method of choice.
General comments	

Evidence Table : Diagnostic accuracy
Question : What is the cost/cost-effectiveness of using AFP and/or US for HCC screening?

Bibliographic citation	2. Lin O, Keeffe E, Sanders G *et al.* Cost-effectiveness of screening for hepatocellular carcinoma in patients with cirrhosis due to chronic hepatitis C. Aliment Pharmacol Ther. 2004; 19: 1159-1172
Study Type / Methods	Cost-effectiveness analysis Using Markov decision model, several screening strategies with abdominal US or CT and AFP at 6 to 12 month intervals in 40-year-old patients with chronic hepatitis C and compensated cirrhosis were simulated from a societal perspective. Results were expressed in terms of discounted costs, LYs, QALYs, and ICERs for QALYs gained. Only direct medical costs were considered. Discounting was implemented at 3% per year for QALYs.
LE	
Number of patients & Patient characteristics	Base case subject: 40-year-old patients with **chronic hepatitis C** and compensated cirrhosis
Intervention	3 plausible strategies were modelled: i. AFP and US every 6 months (biannual AFP/US) ii. AFP and US every 12 months (annual AFP/US) iii. AFP every 6 months with US every 12 months (biannual AFP/annual US) They also modelled screening using AFP and CT, instead of US
Comparison	No screening
Length of follow up (if applicable)	
Outcome measures/ Effect size	For the least efficacious strategy, annual AFP/US, the ICER (versus no screening) was USD$23,043 per QALY For the strategy most commonly used in United States (biannual AFP/annual US), the ICER was USD$33,083 per QALY (versus annual AFP/US). The most efficacious strategy (biannual AFP/US) entailed a higher ICER of USD$73,789 per QALY (versus biannual AFP/annual US). Screening using CT and AFP resulted in better survival compared with the corresponding strategy using US and AFP, with ICER ranging from approximately USD$23,000 to USD$96,000 per QALY. Authors conclusion: The best screening protocol has yet to be formally defined, but based on the results, US at 12-month intervals and AFP at 6-month intervals was a reasonable strategy, offering the greatest gain in life-expectancy while still maintaining an ICER < USD$50,000 per QALY. More frequent screening with US provides some additional benefit, but is more expensive. Screening with CT instead of US is more efficacious and appears to be cost-effective, and deserves further study.
General comments	

Evidence Table : Diagnostic accuracy
Question : What is the cost/cost-effectiveness of using AFP and/or US for HCC screening?

Bibliographic citation	3. Thompson Coon J, Rogers G, Hewson P *et al.* Surveillance of cirrhosis for hepatocellular carcinoma: a cost-utility analysis. British Journal of Cancer. 2008; 98: 1166-1175
Study Type / Methods	Cost-utility analysis Evaluated the effectiveness and cost-effectiveness of surveillance strategies for HCC in individuals with cirrhosis using a decision-analytic model (Markov). Separate cohorts with cirrhosis due to alcoholic liver disease (ALD), HBV, and HCV were simulated. Results were also combined to approximate a mixed aetiology population. Comparisons were made between varieties of surveillance algorithms using AFP assay and/or US at 6- and 12-monthly intervals. Parameter estimates were obtained from comprehensive literature reviews. Uncertainty was explored using one-way and probabilistic sensitivity analyses. Cost and utility values are attached to each state, and the differences between the aggregated costs and health outcomes in each simulation are used to estimate the cost-effectiveness of surveillance, expressed as incremental cost per quality-adjusted life-year (QALY). Cost and QALYs were discounted at 3.5% per year.
LE	
Number of patients & Patient characteristics	Simulated populations: People with compensated cirrhosis aged ≤ 70 years with no pre-existing medical conditions that might preclude treatment with liver transplantation (OLT) or hepatic resection (including current alcohol or intravenous drug use). The model considers 3 cirrhosis aetiologies (ALD, HBV, and HCV). Results were also combined to simulate a mixed aetiology cohort.
Intervention	• Annual AFP-triage • Annual US • Annual AFP + US • 6-monthly AFP-triage • 6-monthly US • 6-monthly AFP + US
Comparison	No surveillance
Length of follow up (if applicable)	
Outcome measures/ Effect size	Effectiveness of surveillance The 6-monthly AFP + US was most effective across all outcomes when compared with no surveillance: • Operable HCC (16.9%) • Dying of HCC (10.8%) • NNS to prevent one death from HCC (11) However, the cheapest strategy, annual AFP-triage, still achieved substantial gains: more than doubling the number of operable HCC found (11.9%), and increasing the number of small tumours found more than six-fold (1.9%). <u>Cost-utility of surveillance</u> In an incremental analysis, neither of the US-only strategies would be considered (since they are both slightly less effective and more costly than surveillance at the same frequency with AFP-triage). Therefore, in the mixed aetiology cohort, the cheapest surveillance strategy is annual AFP-triage, with ICER of £20,700 per QALY. The addition of US to this strategy increased the ICER to £60,100 per QALY gained.
General comments	

Evidence Table : Diagnostic accuracy
Question : What is the cost/cost-effectiveness of using AFP and/or US for HCC screening?

Bibliographic citation	3. Thompson Coon J, Rogers G, Hewson P *et al.* Surveillance of cirrhosis for hepatocellular carcinoma: a cost-utility analysis. British Journal of Cancer. 2008; 98: 1166-1175
Study Type / Methods	Cost-utility analysis Evaluated the effectiveness and cost-effectiveness of surveillance strategies for HCC in individuals with cirrhosis using a decision-analytic model (Markov). Separate cohorts with cirrhosis due to alcoholic liver disease (ALD), HBV, and HCV were simulated. Results were also combined to approximate a mixed aetiology population. Comparisons were made between varieties of surveillance algorithms using AFP assay and/or US at 6- and 12-monthly intervals. Parameter estimates were obtained from comprehensive literature reviews. Uncertainty was explored using one-way and probabilistic sensitivity analyses. Cost and utility values are attached to each state, and the differences between the aggregated costs and health outcomes in each simulation are used to estimate the cost-effectiveness of surveillance, expressed as incremental cost per quality-adjusted life-year (QALY). Cost and QALYs were discounted at 3.5% per year.
LE	
Number of patients & Patient characteristics	Simulated populations: People with compensated cirrhosis aged ≤ 70 years with no pre-existing medical conditions that might preclude treatment with liver transplantation (OLT) or hepatic resection (including current alcohol or intravenous drug use). The model considers 3 cirrhosis aetiologies (ALD, HBV, and HCV). Results were also combined to simulate a mixed aetiology cohort.
Intervention	• Annual AFP-triage • Anual US • Annual AFP + US • 6-monthly AFP-triage • 6-monthly US • 6-monthly AFP + US
Comparison	No surveillance
Length of follow up (if applicable)	
Outcome measures/ Effect size	Probabilistic sensitivity analysis At willingness-to-pay (WTP) threshold of £30,000 per QALY: • 6-monthly AFP + US is only likely to be considered cost-effective in individuals with HBV-related cirrhosis. • In those with HCV-related cirrhosis, 6-monthly AFP-triage is more likely to be considered cost-effective. • In individuals with ALD-related cirrhosis, there is uncertainty about which strategy would be most cost-effective • In the mixed aetiology cohort, 6-monthly AFP-triage appears to be the most cost-effective surveillance protocol, and remains the foremost option until WTP reaches very high levels Authors conclusion: From the analysis, in a UK NHS context, surveillance of individuals with cirrhosis for HCC should be considered effective and cost-effective. The economic efficiency of different surveillance strategies is predicted to vary markedly according to cirrhosis aetiology.
General comments	

APPENDIX 6

LIST OF EXCLUDED STUDIES

1. Snowberger N, Chinnakotla S, Lepe R *et al.* Alpha fetoprotein, ultrasound, computerized tomography and magnetic resonance imaging for detection of hepatocellular carcinoma in patients with advance cirrhosis. Alimentary Pharmacology & Therapeutics. 2007; 26: 1187-1194

2. Kodo M, Zheng RQ, Kim SR *et al.* Diagnostic accuracy of Imaging for liver cirrhosis compared to histologically proven liver cirrhosis. Intervirology. 2008; 1: 17-26

3. Arrieta O, Cacho B, Espinosa D *et al.* The progressive elevation of alpha fetoprotein for the diagnosis of hepatocellular carcinoma in patients with liver cirrhosis. BMC Cancer. 2007; 7: 28

4. Colombo M, Franchis R, Ninno E *et al.* Hepatocellular carcinoma in Italian patients with cirrhosis. The New England Journal of Medicine. 1991; 325 (10): 675-680

5. Bolondi L, Sofia S, Siringo S *et al.* Surveillance programme of cirrhotic patients for early diagnosis and treatment of hepatocellular carcinoma: a cost effectiveness analysis. Gut. 2001; 48: 251-259

6. Yuen MF, Cheng CC, Lauder I *et al.* Early detection of hepatocellular carcinoma increases the chance of treatment: Hong Kong experience. Hepatology. 2000; 2: 330-335

7. Caturelli E, Pompili M, Bartolucci *et al.* Hemangio-like lesions in chronic liver disease: Diagnostic evaluation in patients. Radiology. 2001; 220 (2): 337-342

8. Takashima T, Matsui O, Suzuki M *et al.* Diagnosis and screening of small hepatocellular carcinomas. Radiology. 1982; 145: 635-638

9. Takayasu K, Moriyama N, Muramatsu Y *et al.* The diagnosis of small hepatocellular carcinomas: Efficacy of various imaging procedures in 100 patients. AJR. 1990; 155: 49-54

10. Mita K, Kim SR, Kudo M *et al.* Diagnostic sensitivity of imaging modalities for hepatocellular carcinoma smaller than 2 cm. World Journal of Gastroenterology. 2010; 16 (33): 4187-4192

11. Li CX, Xu GL, Jiang ZY *et al.* Analysis of clinical effect of high-intensity focused ultrasound on liver cancer. World Journal of Gastroenterology. 2004; 10 (15): 2201-2204

12. Fung KT, Li FT, Raimondo ML *et al.* Systematic review of radiological imaging for hepatocellular carcinoma in cirrhotic patients. The British Journal of Radiology. 2004; 77: 633-640

13. Tangkijvanich P, Anukulkarnkusol N, Suwangool P *et al.* Clinical characteristics and prognosis of hepatocellular carcinoma. J Clin Gastroenterol. 2000; 31 (4): 302-308

14. Ando E, Kuromatsu R, Tanaka M *et al.* Surveillance program for early detection of hepatocellular carcinoma in Japan: Results of specialized department of liver disease. J Clin Gastroenterol. 2006; 40 (10): 942-948

PERTINENT DETAILS OF DIAGNOSTIC ACCURACY/EFFECTIVENESS STUDIES

Authors	Study design	Year	Country / region	Patient characteristics	Test methods	Cut-off levels (ng/mL)	Screening intervals (month)	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV %	Detection no / per 100,000 screened (95% CI)	Mortality rate / per 100,000 screened (95% CI)	Survival rate (%) 5-years
Yang *et al.*	SR	1999	Hong Kong	HBsAg+ve	AFP	> 20	6	68.6	95.0	3.3	-	-	-
					US	-		84.3	97.1	6.6	-	-	-
					AFP+US	> 20		92.2	95.0	3.0	86 vs. 51	0.44 vs. 0.54	52.7 vs. 0.0
Singal A *et al.*	SR with MA	2009	United States	Cirrhotic	US at any HCC stage	-	-	94.0	94.0	-	-	-	-
					US at early HCC	-	-	63.0	-	-	-	-	-
							< 6	70.1	-	-	-	-	-
							6 - 12	50.1	-	-	-	-	-
					AFP+US at early HCC	15 - 400	-	69.0	-	-	-	-	-
Colli A *et al.*	SR with MA	2006	Italy	Chronic liver disease (cirrhosis/hepatitis)	AFP	10 - 200	1 - 23	-	-	-	-	-	-
					US	-		60.0	97.0	-	-	-	-
					Spiral CT	-		68.0	93.0	-	-	-	-
					MRI	-		81.0	85.0	-	-	-	-
Gebo et al.	SR	2002	United States	Chronic HCV	AFP	10 – 400	-	60.0 – 80.0	70.0 – 90.0	-	-	-	-
					US	-		-	-		-	-	-
Gupta S, Bent S, Kohlwes J	SR	2003	United States	HCV-infected with or without cirrhosis	AFP	> 20	-	41.0 – 65.0	80.0 – 94.0	-	-	-	-
						> 200		20.0 – 45.0	99.0 – 100.0	-	-	-	-
Chen *et al.*	RCT	2003	Qidong, China	HBsAg+ve	*AFP (not complied)	≥ 20	6	55.3	86.5	-	-	1,138 vs. 1,114	4.0 vs. 4.1
					*AFP (complied)			80.0	80.9	-	-	-	-
Zhang B, Yang B	RCT	1999	Shanghai, China	HBV infection or history of chronic hepatitis	AFP	> 20	6	-	-	3.3	69.0	-	-
					US	-		-	-	6.6	84.0	-	-
					AFP+US	> 20		-	-	3.0	92.0	-	-
Zhang BH, Yang BH, Tang ZY	RCT	2004	Shanghai, China	HBV infection or history of chronic hepatitis	AFP+US	> 20	6	-	-	-	86 vs. 67	83.2 vs. 131.5	46.4 vs. 0.0
McMahon *et al.*	Cohort	2000	Alaska	HBsAg+ve	AFP	15	6	-	-	-	-	-	42.0

SR: Systematic review
MA: Meta-analysis
RCT: Randomised control trial

HCC: Hepatocellular carcinoma
AFP: Alpha-fetoprotein
US: Ultrasound

CT: Computed tomography
MRI: Magnetic resonance imaging
HBsAg; Hepatitis B surface antigen

HBV: Hepatitis B virus
HCV: Hepatitis C virus
PPV: Positive predictive value

* complied: attend for screening scheduled examination

Malaysian Health Technology Assessment Section (MaHTAS)
Medical Development Division, Ministry of Health Malaysia,
Level 4, Block E1, Complex E, Precinct 1,
Federal Government Administrative Centre
62590, Putrajaya, Malaysia
Tel: 03 - 8883 1246/47
Fax: 03 - 8883 1230

ISBN 978-967-0399-36-2