FUNDED BY
NIHR | National Institute for Health and Care Research

MANCHESTER 1824
The University of Manchester

NIHR134670

The benefits, harms and costs of surveillance for hepatocellular carcinoma in people with cirrhosis: synthesis of observational and diagnostic test accuracy data and cost–utility analysis

PROJECT PROTOCOL

Version number: 1.0
Date: 27 May 2022

University of BRISTOL
UNIVERSITY OF EXETER
UNIVERSITY OF PLYMOUTH
NHS University Hospitals Plymouth NHS Trust
BRITISH LIVER TRUST
NIHR | Nottingham Biomedical Research Centre
NIHR | Applied Research Collaboration South West Peninsula
NHS Nottingham University Hospitals NHS Trust

Table of contents

Funding acknowledgement and disclaimer

This project is funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (project reference NIHR134670). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Version control

Version	Date	Description
1.0	27 May 2022	Original protocol based on Detailed Research Plan approved by the NIHR HTA board with the following changes: • Removal of application text solely focusing on the merits of the bid as a proposal • Reincorporating detail on patient and public involvement (specified separately in application process); including flowchart in main doc • Minor edits to tenses and framing

Abbreviations used in this document

Abbreviation	Definition
AASLD	American Association for the Study of Liver Diseases
AFP	Alpha-Fetoprotein
AFP-L3	Lens Culinaris Agglutinin-Reactive Fraction of AFP
ARLD	Alcohol-Related Liver Disease
BASL	British Association for the Study of the Liver
BLT	British Liver Trust
BRC	Biomedical Research Centre
CHEERS	Consolidated Health Economic Evaluation Reporting Standards
DCP	Des-γ-Carboxyprothrombin
EASL	European Association for the Study of the Liver
ELISA	Enzyme-Linked Immunosorbent Assay
GALAD	Gender, Age, AFP-L3, AFP and DCP
HBV	Hepatitis B Virus
HCC	Hepatocellular Carcinoma
HCV	Hepatitis C Virus
HSROC	Hierarchical Summary Receiver Operating Characteristic
HTA	Health Technology Assessment
IPDAS	International Patient Decision-aid Standards
LI-RADS	Liver Imaging Reporting and Data System
NAFLD	Non-Alcoholic Fatty Liver Disease
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health and Care Research
PPIE	Patient and Public Involvement and Engagement
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROBAST	Prediction model Risk Of Bias Assessment Tool
PROSPERO	International Prospective Register of Systematic Reviews
QALYs	Quality-Adjusted Life-Years
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
REML	Restricted Maximum Likelihood
ROC	Receiver Operating Characteristic
WHO	World Health Organisation
WP	Work Package

1. Summary of research

Background

Hepatocellular carcinoma (HCC) is relatively unusual among cancers in having a precursor condition that is identifiable in a substantial majority of cases – cirrhosis[1]. As a result, people with cirrhosis represent an obvious target for periodic follow-up ('surveillance') for HCC. If it is to save lives, surveillance must fulfil a three-stage hypothesis: (1) regular monitoring should enable us to detect HCCs when they are smaller and less advanced than they will be by the time they come to attention symptomatically, meaning (2) they will be more amenable to treatment, leading to (3) improved survival in the surveilled population.

An influential 2007 NIHR HTA assessed the cost effectiveness of HCC surveillance in people with cirrhosis. However, the liver disease environment has changed in many ways since then. Our project will account for all these changes in providing an up-to-date assessment of whether and how surveillance should be offered to people with cirrhosis.

Methods

There will be 4 work-packages (WPs). WP1 is a systematic review and meta-analysis of cohort studies comparing HCCs found under surveillance with those diagnosed incidentally or symptomatically. Outcomes of interest are tumour characteristics at diagnosis (size, stage, focality, invasiveness), treatment received (transplantation, treatment with curative intent) and survival (all-cause and HCC-specific). As all such evidence is subject to a high risk of bias, we will assess methods carefully and evaluate any attempts to adjust for known biases.

WP2 is a systematic review and synthesis of diagnostic accuracy data relating to tests for HCC in people with cirrhosis. We will assess the accuracy of imaging (ultrasound, CT, MRI), conventional biomarkers (AFP, DCP) and genomic analytes (microRNA, circulating tumour cells). We will estimate the accuracy of different tests alone or in combination, using Bayesian statistical techniques to provide new insight into how tests work at all possible thresholds and in comparison with each other, especially when it comes to picking up liver cancer at an early stage.

In WP3, we will develop a mathematical decision-analytic model to estimate the lifetime costs, benefits and harms of different HCC surveillance regimens – compared with each other and no surveillance – in people with cirrhosis. It will simulate the natural history, diagnosis and treatment of HCC using the outputs of WP2 and other best-available evidence. We will incorporate accuracy as a function of test threshold, to establish the thresholds that maximise the net benefit of surveillance. As well as benefits, we will simulate harms associated with surveillance (biopsy, contrast-induced kidney injury, exposure to ionising radiation, psychological and iatrogenic morbidity associated with overdiagnosis). We will perform extensive subgroup analysis, to identify whether the balance of costs, benefits and harms varies according to patient-level characteristics (for example, aetiology of cirrhosis, presence of other long-term conditions and eligibility for HCC treatment).

In WP4, we will develop a print-based patient decision-aid, using what we have learned in WPs1–3 to quantify the expected benefits and harms of surveillance. It will aim to support shared decision-making for people with cirrhosis about whether to start – or stop – surveillance. We will

summarise expected outcomes for those strategies that emerge as good value for money at a population level. Even if it is not amongst these options, we will include 6-monthly ultrasound +/- AFP, to ensure that our research is immediately useful for liver services as they are most likely to be configured at present.

Patient and public involvement

Our PPI strategy, devised with the support of the British Liver Trust (BLT), comprises 3 elements: (1) we will recruit 2 team members as co-investigators and co-authors; (2) we will hold 2 online workshops with 10–15 participants, identified from existing groups of liver patients, including BLT's patient community; at the second, we will alpha-test the decision-aid developed in WP4; (3) we have secured the participation of representatives from BLT to sit on our advisory group. We can also use BLT's online patient forum (c25,000 users) to access the views of a wider cross-section of the community.

Dissemination

BLT are going to help us share our findings with patients and professionals; they will be invited to badge and publicise the decision-aids from WP4. We will work with NICE and national liver organisations to maximise the impact of our research for population-level decision-making. We will publish at least 3 papers in scientific journals and present findings at at least 2 conferences. We will make the statistical synthesis code from WP2 and the decision-model from WP3 freely available for future researchers.

Project plan

The project will take 21 months. The project team includes clinicians with expertise in hepatology and radiology, statisticians, health economists, evidence reviewers and patient experts, including several members of the 2007 HTA team. We have recruited an advisory group of methods, clinical and patient experts who will provide oversight and advice for all aspects of the project.

2. Background and rationale

In 2007, several of the investigators were among the authors of an NIHR HTA exploring the cost-effectiveness of surveillance for hepatocellular carcinoma (HCC) in people with cirrhosis[2,3]. That research was influential in the UK and internationally: it was central to NICE's guidelines on hepatitis B (HBV; CG165) and cirrhosis (NG50), and it was also cited in guidance from the WHO[4], and European[5] and American[6] liver associations. However, the liver disease environment reflected in the evidence reviewed – and simulated in the world that was modelled – in 2007 has changed beyond recognition:

- Above all, we are amid a liver disease epidemic. Incidence of cirrhosis rose by 50% in the first decade of the 21st century[7], and age-standardised incidence of HCC approximately doubled between 2005 and 2017[8].
- The aetiology of cirrhosis is also changing. The 2007 HTA accounted for 3 causes: HBV, hepatitis C (HCV) and alcohol-related liver disease (ARLD). It noted that obesity-related cases – non-alcoholic fatty liver disease (NAFLD) – were on the rise, although there were

insufficient data to include them in our analyses. Now, NAFLD is the second-commonest cause of cirrhosis in the UK[9].

- Conversely, prevalence of HCV infection has reduced over the last decade[10]. Moreover, since 2010, NICE has approved 8 combinations of direct-acting antiviral medicines for HCV. A sustained response to these agents reduces but does not eliminate risk of HCC[11]. There are also 3 positive NICE appraisals of antivirals for HBV, although evidence that they prevent HCC is equivocal[12].
- These shifts in underlying liver disease have knock-on effects on the epidemiology of HCC. The typical person with HCC is now older[13,14] and more likely to have type 2 diabetes and other comorbidities[1].
- For people who develop HCC, the range of treatment options has grown. The 2007 HTA only modelled resection and transplantation. In current practice, locoregional therapies including radiofrequency and microwave ablation are used as curative options for small tumours. Transarterial chemoembolisation may be used as 'bridging' or 'downstaging' therapy (keeping people within transplantation criteria or reducing tumour size so they are amenable to treatment)[15,16]. These techniques may also be used for inoperable HCCs, as may systemic chemotherapy: in the past 5 years, NICE has recommended 3 options for untreated HCC and 1 for second-line use. Selective internal radiation therapy has recently been added to the list of approved options.
- Diagnostic technologies have also advanced. The 2007 HTA modelled 1 biochemical assay (alphafetoprotein; AFP) and 1 imaging test (ultrasound) as surveillance tools. There is now a greatly enhanced evidence-base for each, supplemented by diagnostic accuracy data for a range of conventional and genetic biomarkers[17] and increasing interest in cross-sectional imaging[18]. Investigators have also proposed diagnostic prediction models that could improve on these measures in isolation[19].
- The same period has seen an ever-increasing focus on empowering people to make decisions about their healthcare that are consistent with their values and goals. Up until 2007, MEDLINE and Embase indexed fewer than 1,000 publications referring to a 'shared decision'; since then, there have been over 10,000 such papers.

Clearly, these advances could affect the balance of benefits, harms and costs of surveillance for HCC, and provide opportunities to personalise care in a way that reflects the circumstances and preferences of people with cirrhosis. However, despite everything that has changed in the surrounding landscape, UK centres undertake surveillance in much the same way as 15 years ago.

National and international guidelines tend to recommend 6-monthly surveillance[20–22]. Other authorities urge caution, emphasising shortcomings in the available evidence[23,24]. This uncertainty promotes ambivalence among UK liver clinicians: a 2015 survey revealed that provision of surveillance is 'poor overall', and found that 'doubts over effectiveness' and uncertainty about cost were key barriers to implementation[25].

Even where surveillance is recommended, there is ambiguity about the optimal mode and frequency of follow-up. The status of AFP assay is especially controversial[26]; hence, NICE

recommends ultrasound 'with or without AFP' (NG50). Furthermore, despite clear heterogeneity in risk of HCC and accuracy of tests in different groups, none of the available guidance stratifies by aetiology of cirrhosis, or other patient-specific factors.

2.1. Evidence explaining why this research is needed now

The NHS Long-Term Plan contains an ambition that 75% of cancers will be diagnosed at stage 1 or 2 by 2028. But fewer than 30% of liver cancer diagnoses currently meet this target (only pancreas and lung cancers have worse rates of early detection)[27]. The Long-Term Plan suggests 'personalised and risk-stratified screening' is key to improving such statistics. This project will provide substantive new evidence to establish whether and how this challenge can be met for liver cancer. In a similar spirit, the 'Less Survivable Cancers Taskforce' – a coalition of patient charities including the British Liver Trust (BLT) – has called for the robust implementation of surveillance as one of its 7 key recommendations to 'close the deadly cancer gap'[28].

However, these challenges cannot be met by an indiscriminate increase in surveillance activity. Having experienced the disruption of routine care associated with COVID-19, patients have a right to presume that any face-to-face appointments that are reintroduced come with evidence-based expectation of benefit. There is also increasing attention on the potential harms of surveillance[29]. Positive tests always cause anxiety. They also lead to diagnostic investigations that may be harmful – biopsy is associated with pain and bleeding; cross-sectional imaging has the potential for contrast-medium toxicity and exposure to ionising radiation. Even in the case of true-positive findings, overdiagnosis of HCCs that would not have affected the person's quality or length of life becomes an importunate consideration in an ageing population facing multiple competing hazards of death[30]. Our research aims to provide an accessible way of weighing up the expected benefits and harms of surveillance for people in different circumstances, and we will summarise our findings with an explicit intention of facilitating shared decision-making. We intend these outputs to be of value not only for people facing a choice to start surveillance, but also to those for whom it might be reasonable to stop – e.g. people facing lower risks of HCC (such as those responding to antiviral therapy for HCV) or experiencing comorbidities that may limit their access to treatment should they develop HCC. This goal aligns well with recent NICE guidance, which contains a strong recommendation that patient decision-aids should be used as part of shared decision-making, where available and reliable (NG197).

Our research is also timely in view of the rapidly increasing focus on genomic biomarkers for cancer detection. We will assess current evidence on liquid biopsy techniques and, if sufficient data are available, simulate their use in surveillance programmes. Moreover, to ensure that our research continues to provide value in a field in which evidence is likely to evolve at pace, we will establish parameters that hypothetical future tests would need to meet – in terms of combination of cost and accuracy – in order to supplant today's technologies as an optimal use of NHS resources (see ¶4.3.4).

2.2. Published evidence and ongoing research

Using our draft search strategy (*Table 1*) and BMJ Knowledge Centre's systematic review filter[31], we searched MEDLINE (August 2021) for systematic reviews of surveillance for HCC in people with cirrhosis.

The most comprehensive systematic review of cohort studies comparing surveillance with none includes 47 studies published before 2014[32]. It finds that HCC surveillance is associated with significant improvements in early tumour detection, receipt of curative therapy and 3-year overall survival in patients with cirrhosis[32].

There are multiple systematic reviews of the diagnostic accuracy of tests for HCC. A recently published Cochrane review covers ultrasound and AFP alone or in combination[33], including 373 studies (although fewer than 250 provided data for quantitative synthesis). The authors find that, at a threshold of 20 ng/ml, AFP is less sensitive than ultrasound for the diagnosis of all HCC (60% [58% to 62%] -v- 72% [63% to 79%]), and also less specific. However, when it comes to a small subsample of studies reporting accuracy in detecting 'resectable' HCCs, AFP could be more sensitive than ultrasound (65% [62% to 68%] -v- 53% [38% to 67%]). Evidence from a small number of studies (n=6) suggested that combining the 2 tests increases true-positive detection-rates (relative sensitivity versus ultrasound alone 1.28 [1.03 to 1.53]) with little or no penalty for specificity. There are also systematic reviews addressing the diagnostic accuracy of cross-sectional imaging (n=33)[18], including abbreviated protocols (n=15)[34], as well as conventional (n=13)[35] and genomic biomarkers (n=67)[36]. There are additional Cochrane protocols on 2 imaging topics[37,38]; however, these envisage assessing accuracy in a confirmatory setting where other tests have already raised suspicion of HCC, whereas we are interested in these tools in the setting of first-line detection.

A recent systematic review of economic analyses[39] confirms that there have been no attempts to assess the cost-effectiveness of surveillance from a UK perspective since the 2007 HTA[2]. However, the authors identify several publications from other jurisdictions; while these are not directly helpful for NHS decision-making, they can inform the design of our analysis.

There is 1 UK-focused visualisation of benefits and harms of surveillance[29]; this will be a helpful precedent as we design our decision-aid, though it is based on calculations that our project will supersede.

2.3. How the research will add to the body of knowledge

This will be the most comprehensive health technology assessment of surveillance for HCC so far undertaken. It will build on the strengths of the 2007 analysis, while extending the assessment to include additional systematic reviews and provide outputs that can inform shared decision-making. We believe that a multicomponent evidence synthesis of the type we propose is the best approach to inform population- and individual-level decision-making. Randomised trials of surveillance versus none would undoubtedly help resolve uncertainty; however, recruitment appears unfeasible[40] and the trials would need to be impossibly large to estimate the effects of all possible programmes across different strata of the population. In contrast, our proposed research design will enable us to estimate the lifetime costs, benefits and harms of a variety of surveillance strategies, stratified according to a range of patient characteristics.

The most comprehensive systematic review of cohort studies comparing surveillance with none includes 47 studies published before 2014[32]. Our scoping suggests there are at least as many newer publications, so we may analyse up to 100 studies in WP1. As well as adding the new studies to the evidence-base, our research will add to the body of knowledge by paying more

heed to biases that may explain some or all of the differences between cohorts, where previous authors have been relatively uncritical in attributing causal effect to surveillance status.

Although, as summarised in ¶2.2, there have been many attempts to synthesise evidence on the diagnostic accuracy of tests for HCC, each individual study has nontrivial shortcomings. What is more, the piecemeal fashion in which the collected evidence-base has accumulated means that it is hard to compare between tests. Our research will substantially move the state of knowledge forward, by (a) identifying, appraising and synthesising evidence across the full range of possible surveillance tests using uniform standards, (b) stratifying results in consistent ways (differentiating between aetiology of underlying liver disease and size/stage of tumour detected), and (c) using advanced synthesis methods to account for multiple thresholds and comparative accuracy (see ¶4.2.5, below). Our synthesis methods will also provide results that are directly applicable to our decision model (WP3) – for example, enabling us to identify test thresholds that maximise the net benefit of surveillance – which would not be possible if we attempted to base our model on published reviews.

The research will continue to generate knowledge after its completion, as we will make our decision-model freely available to future researchers.

3. Overarching research plan / methods

The following section details methods that are common to all WPs; ¶4 provides specific objectives and methods for each WP in turn.

3.1. Aims and objectives

Our overarching aims are to establish the people with cirrhosis to whom surveillance for HCC should be offered, to define what that surveillance should comprise (test(s), thresholds and frequency), and to quantify the expected benefits and harms for people deciding whether to enrol.

To achieve this, we will use formal evidence reviews to inform and complement the development of an original decision-analytic model, and to distil our findings into patient-focused materials to support shared decision-making (see *Figure 1*). Our project comprises 4 linked work packages:

- **WP1** Systematic review and synthesis of cohort studies comparing HCCs found under surveillance with those diagnosed incidentally and/or symptomatically
- **WP2** Systematic review and synthesis of diagnostic accuracy of tests for HCC
- **WP3** Health economic decision model estimating benefits, harms and costs of different surveillance strategies compared with each other and no surveillance
- **WP4** Developing a patient decision-aid quantifying the expected benefits and harms of surveillance

3.2. Target population

Adults with cirrhosis of any aetiology who have never had HCC.

Wherever data allow, we will stratify our analyses to account for patient characteristics: age; sex; cause of cirrhosis; liver disease severity; comorbidity and frailty (competing hazard of death from

other causes); amenability to treatment (e.g. people who would be eligible for a transplant if HCC is detected compared with those who would not have this option).

For cause of cirrhosis, we anticipate there will be enough data to stratify by ARLD, HBV, HCV and NAFLD. It will inevitably be harder to find data for rarer causes of cirrhosis (primary biliary cholangitis, haemochromatosis, autoimmune hepatitis, Wilson's disease); however, these aetiologies will not be excluded from the reviews.

3.3. Health technologies being assessed

We define surveillance as the repeated application at specified intervals of 1 or more diagnostic tests intended to identify HCC. The tests may include

- Imaging, including
  - Ultrasound (including B-mode and contrast-enhanced techniques)
  - CT (multiphase HCC-specific protocols, where possible distinguishing between reporting standards, e.g. LI-RADS)
  - MRI (distinguishing between dynamic contrast-enhanced, abbreviated and noncontrast protocols and, where possible, further distinguishing between reporting standards, e.g. LI-RADS)
- Conventional biomarkers. As the literature includes a very wide range of markers, many of which are studied in small populations and few papers, we will restrict our attention to the 3 that, as agreed by our advisers, might feasibly be measured routinely in the NHS:
  - AFP (we will treat plasma and serum samples and those quantified by ELISA and chemiluminescence as interchangeable)
  - Lens culinaris agglutinin-reactive fraction of AFP (AFP-L3)
  - Des-gamma-carboxyprothrombin (DCP); also known as protein induced by vitamin K absence or antagonists II (PIVKA-II)
- Genomic biomarkers, which may include
  - MicroRNA
  - Circulating tumour DNA
- Validated diagnostic prediction models incorporating 1 or more of the above (NB we will only include studies reporting these in evidence reviews where investigators have used them as binary classifiers with a prespecified threshold; we will exclude derivation studies and those where threshold is manipulated post hoc to maximise some accuracy criterion). Possible examples include:
  - GALAD[41]
  - Multitarget HCC blood test (mt-HBT)[42]
  - Doylestown algorithm[43]

3.4. Common methods for systematic reviews (WPs 1 and 2)

We will follow standard guidance for systematic reviews[44–46]. We will register our protocols in PROSPERO. We will report our findings using relevant standards: MOOSE[47] for WP1 and PRISMA-DTA[48] for WP2.

3.4.1. Identifying evidence

An experienced information specialist (MR) will develop, document and deliver searches.

We will search MEDLINE and Embase in Ovid and the Cochrane Database of Systematic Reviews. *Table 1* shows our draft MEDLINE strategy to identify the population and interventions of interest. See ¶¶4.1.3 and 4.2.3, below, for our plans for searches for individual reviews. In addition, we will check reference lists and included studies in related reviews. We will also carry out forward citation-chasing of relevant past reviews using Scopus.

We will run update searches to identify relevant evidence published while we are conducting the project. This will be from no more than 6 months from our final deadline.

Two reviewers will undertake title and abstract screening and full-text selection independently, with disagreements resolved by consensus.

Table 1: Draft population and intervention search strategy (MEDLINE)

#	Term	Results
1	(hepatocellular and surveillance).ti.	511
2	(hepatocellular and screening).ti.	480
3	(hepatocellular and detection).ti.	917
4	or/1-3	1,848
5	exp Liver Cirrhosis/	93,474
6	Liver Diseases, Alcoholic/	5,685
7	exp Fatty Liver/	37,061
8	cirrhosis.ti,ab.	95,010
9	hepatitis.ti,ab.	225,266
10	liver disease*.ti,ab.	106,467
11	subclinical hepatocellular carcinoma.ti,ab.	4
12	subclinical HCC.ti,ab.	11
13	steatohepat*.ti,ab.	12,736
14	or/5-13	399,806
15	Carcinoma, Hepatocellular/di, dg	16,114
16	hepatocellular carcinoma.ti,ab.	96,029
17	hepatocarcinoma.ti,ab.	3,991
18	hepatoma.ti,ab.	27,264
19	or/15-18	126,341
20	Population Surveillance/	61,329
21	Mass Screening/	109,064
22	(screening and hepatocellular).ti,ab.	3,729
23	surveillance.ti,ab.	195,802
24	ultrasound.ti,ab.	265,290
25	AFP.ti,ab.	14,750
26	f?etoprotein*.ti,ab.	19,427
27	ultrasonography.ti,ab.	93,239
28	sonography.ti,ab.	32,786
29	MRI.ti,ab.	263,144
30	CT.ti,ab.	369,161
31	or/20-30	1249,189
32	14 and 19 and 31	9,791
33	4 or 32	10,750

3.4.2. Common exclusion criteria for reviews

We will exclude studies based on populations predominantly comprising participants with noncirrhotic liver disease; studies in children; editorials, narrative reviews, letters, case reports, and preclinical studies. We will follow the guidance of Scherer and Saldanha regarding the inclusion of conference abstracts[49]: we will not include such evidence in the first instance, but will consider broadening our criteria to include it if the evidence we assemble from full publications is sparse or conflicting – if necessary, with the aid of supplementary targeted searches. We do not have resources to translate non-English publications, but we will list all foreign-language studies with English abstracts that appear relevant so readers can judge the potential impact of their exclusion.

3.4.3. Extracting data and assessing risk of bias

One reviewer will undertake data extraction and assess risk of bias; a second will check for accuracy and completeness. We will extract descriptive data for study characteristics, including

design, population size, geographical location, year(s) of recruitment, baseline population characteristics (age, sex, aetiology of cirrhosis, Child–Pugh score), surveillance tests evaluated and definition of comparators/reference standard.

Where studies present outcome data for subgroups of interest (e.g. by aetiology of cirrhosis), we will extract these separately for use in stratified syntheses.

3.4.4. Presenting results

We will provide PRISMA flow diagrams depicting how studies were identified. We will present a descriptive summary of included studies. We will tabulate study characteristics (see ¶3.4.3) and summary measures of risk of bias. We will include commentary on the major methodological problems or biases that affect the evidence.

3.5. Identification of future research priorities

In all WPs, we will identify remaining uncertainties and make recommendations for future research. Where possible, we will express unanswered questions in a structured format (e.g. PICOS – population, intervention, comparison, outcomes and study design). We will consult with our expert advisory group to ensure that we give appropriate priority to the research questions we identify, and that the research we suggest is feasible.

4. Work packages

Our project comprises 4 linked work packages, as outlined in *Figure 1*.

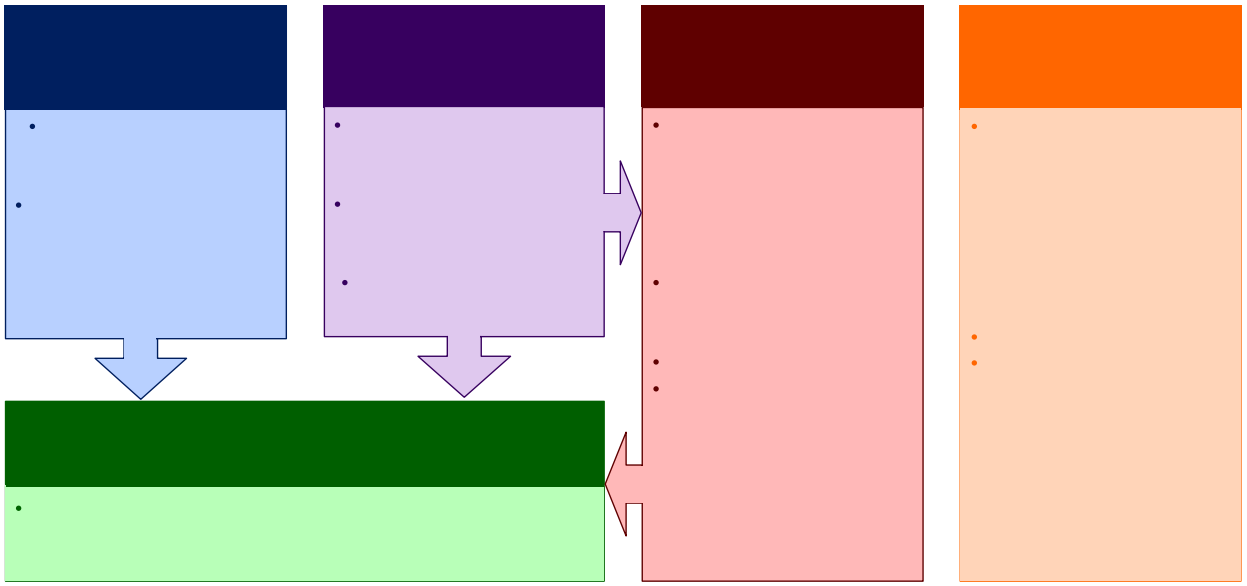

Figure 1: Design of the project, showing interrelation of work packages

4.1. WP1: systematic review of observational evidence

4.1.1. Objective

WP1 will identify, appraise and synthesise published observational evidence comparing characteristics and outcomes of HCCs found under surveillance with those diagnosed incidentally and/or symptomatically.

4.1.2. Eligibility criteria

We will include prospective and retrospective cohort studies comparing HCCs found under surveillance with at least 1 relevant control group. Ideally, these comparators will be limited to HCCs found in people with known cirrhosis who were not undergoing surveillance; however, we will not exclude studies reporting all HCCs. We will exclude case–control studies and all noncomparative research.

4.1.3. Identifying evidence

Our literature searches will combine population terms (*Table 1*) with specific terms for the technologies of interest (see ¶3.3) and a bespoke search strategy for identifying cohort studies that will be based on tested filters and tested on a body of relevant evidence before use.

4.1.4. Extracting data and assessing risk of bias

Outcomes of interest are:

- tumour characteristics at diagnosis: size (continuous and dichotomous [proportion under given thresholds]), stage (e.g. Barcelona stage; meeting Milan criteria), focality (multifocal -v- uninodular), invasiveness (vascular invasion, portal thrombosis)
- treatment received: proportion of people receiving transplantation, resection, any treatment with curative intent
- survival: all-cause and HCC-specific (difference in time-to-event, typically reported as a hazard ratio)

Where necessary for estimating time-to-event effect-measures, we will generate synthetic patient-level data from published Kaplan–Meier curves[50].

We will assess risk of bias using a bespoke instrument incorporating elements of ROBINS-I[51] and QuEENS[52]. We are mindful that the studies we will find are at high risk of multiple biases. Especially important considerations are

- Selection bias: there are certain to be systematic differences other than surveillance status between people who attend surveillance and those who do not
- Lead-time bias: in nonrandomised data, HCCs detected by surveillance will always appear to benefit from longer survival simply because they were found earlier; some studies attempt to analyse the possible impact of this

Therefore, we will assess risk of bias carefully and evaluate any methods authors have used to adjust for known biases.

4.1.5. Synthesising data

We will perform pairwise meta-analyses comparing outcomes for HCCs found under surveillance with those diagnosed incidentally and/or symptomatically. In a top-level analysis, we will consider all surveillance regimens as interchangeable – that is, we will pool all data to estimate differences associated with 'some regular surveillance' versus 'no regular surveillance', in common with previous meta-analyses[32]. Additionally, using stratified analyses and/or meta-regression (see below), we will attempt to identify differences in observed effect that may be attributable to frequency of surveillance and/or test(s).

We will use random-effects models for all meta-analyses, irrespective of statistical heterogeneity. This is because heterogeneity of effect is to be expected where surveillance regimens and patient characteristics differ, and placing proportional weight on larger studies is not a desirable property of syntheses of observational data.

If sufficient data are available, we will explore whether between-study heterogeneity can be explained by study-level characteristics, using random-effects meta-regression. Potential effect-modifiers include aetiology of cirrhosis, severity of underlying liver disease, geographical setting and characteristics of surveillance (modality and/or frequency).

4.2. WP2: systematic review of diagnostic accuracy

4.2.1. Objective

WP2 will systematically identify, appraise and synthesise evidence on the diagnostic accuracy of tests for HCC in people with cirrhosis.

4.2.2. Eligibility criteria

- **Patients:** As per the target population for the whole project (see ¶3.2)
- **Index tests:** Any of the health technologies being assessed for the whole project (see ¶3.3), either alone or in combination
- **Target condition:** HCC of any size (where possible, stratifying according to size and/or stage; see below)
- **Reference standard:** Our ideal reference standard is explant pathology where transplantation is unrelated to HCC. However, such studies are relatively unusual (Colli et al. 2021[33] only found 17), so we will also include studies that use histology of resected or biopsied lesions or radiological (CT or MRI) follow-up to define true disease status, so long as all participants have at least 6 months' follow-up following the index test.

We will only include studies comprising participants with and without HCC – that is, we will need to be able to construct at least one 2×2 table from the data (i.e. the cross-classification of cases found to be positive and negative on index and reference tests, equating to true-positive, false-negative, false-positive and true-negative results). Cross-sectional analyses of people with cirrhosis and unknown HCC status are ideal (so-called '1-gate' designs[53]), but we will also include case–control ('2-gate') studies comparing people with cirrhosis and HCC with people with cirrhosis alone, accounting for this factor when assessing risk of bias. We will exclude case–

control studies using healthy controls, as this design is likely to lead to inflated estimates of both sensitivity and specificity[54,55].

We will include studies reporting the diagnostic accuracy of trends as well as absolute values of continuous markers; in this case, the index test becomes 'change in marker *x*' (whether absolute or relative) rather than 'value of marker *x*'. Where studies report accuracy of combinations or sequences of tests, or the comparative accuracy of 2 or more index tests, we will attempt to recover the fully cross-classified (2×2×2) tables, which will allow estimation of inter-test correlations for use in WP3.

4.2.3. Identifying evidence

We do not intend to perform full database searches to identify evidence on the diagnostic accuracy of two of the index tests, AFP and/or ultrasound. This is because Cochrane has recently published a systematic review on this topic[33], which has a very broad, sensitivity-maximising search strategy (retrieving over 35,000 references) and similar eligibility criteria to ours. We cannot use this publication instead of performing our own review, as it does not explore factors that are critical to our project, including the performance of AFP at all possible thresholds (they only report and pool data at 2 thresholds), diagnostic accuracy for identifying tumours at various sizes and comparative accuracy of these 2 tests when considered in the network of all possible approaches (see ¶4.2.5). However, it is extremely useful as a catalogue of relevant evidence: we judge it unlikely that their review will have missed studies we would want to include. Therefore, we will compare the Cochrane review's list of 373 included studies against our eligibility criteria and also check the 219 references they excluded at full-text screening. We will also perform an update search to identify any studies that have been published since their search (June 2020).

For all other tests, our literature searches will combine population terms (*Table 1*) with specific terms for the technologies of interest (see ¶3.3) and a bespoke design filter for identifying relevant diagnostic accuracy data that we will test by assessing its ability to identify the ultrasound and AFP studies found by Cochrane[33].

4.2.4. Extracting data and assessing risk of bias

Where available, we will extract raw count data showing the number of cases and non-cases correctly classified (and numbers incorrectly classified) by index tests. Most characteristically, such data will appear as 2×2 tables. Where test accuracy is reported at more than one diagnostic threshold, we will extract the 2×2 table for each threshold. We will also add other dimensions as necessary (to account for, e.g., multiple index tests, different sizes of tumours detected). In any case where count data are not directly available, we will attempt to reconstruct them from reported measures of diagnostic accuracy.

We will assess risk of bias using QUADAS-2[56] (or QUADAS-C[57] for comparative accuracy studies).

4.2.5. Synthesising data

For each test or test combination, where each study reports only one 2×2 table of test results we will synthesise data using bivariate meta-analysis of sensitivity and specificity[58], which accounts

for between-study correlation between true- and false-positive results. We will fit binomial likelihoods to the test counts to avoid problems associated with normal approximations[59]. We will present paired forest plots of sensitivity and specificity, and plots of study-specific estimates and meta-analysis results in receiver operating characteristic (ROC) space. When diagnostic thresholds do not vary substantially across studies, we will present pooled summary estimates of sensitivity and specificity from these meta-analyses, with 95% credible ellipses representing joint uncertainty. Where there is variation in thresholds across studies, we will use the equivalence of the bivariate and hierarchical summary ROC (HSROC) models to draw summary ROC curves[60,61].

For continuous biomarkers (e.g. AFP), we anticipate that many studies will report accuracy at more than one diagnostic threshold. In this situation, we will instead synthesise the data using an advanced meta-analysis approach, which uses all available test accuracy data to produce pooled estimates of sensitivity and specificity across the full range of numerical thresholds[62].

Where sufficient studies are available comparing two or more index tests, we will perform separate meta-analyses including only these 'comparative' studies, to enable unbiased inference about how well tests work compared with each other[63]. Additionally, if sufficient fully cross-classified data (2×2×2 tables) are available, we will explore advanced meta-analysis models that synthesise these data to produce unbiased estimates of the accuracy of tests used in sequence and appropriately precise estimates of comparative accuracy[64]. We also aim to extend these approaches to jointly synthesise evidence on the accuracy of many or all tests using network meta-analysis (NMA)[65,66], if we identify a connected network of test comparisons reporting at similar diagnostic thresholds.

We will stratify all evidence syntheses by size of HCC detected, as the accuracy of tests for detecting smaller, earlier tumours is of primary importance for surveillance (we also need separate estimates for the decision model in WP3; see below). It may also be instructive to present sensitivity analyses stratifying by stage of HCC (e.g. Barcelona Clinic Liver Cancer categories or those meeting / not meeting Milan criteria). We will use subgroup analysis and meta-regression to explore other sources of heterogeneity.

We will also illustrate results by calculating expected results in natural frequencies for hypothetical cohorts of 1,000 people, at 1 or more indicative estimates of HCC prevalence. This will also enable us to calculate positive and negative predictive values.

We will take a Bayesian approach to statistical analysis, computed using WinBUGS[67] and/or JAGS[68] in R. We will use vague prior distributions across all analyses and will check for sensitivity to choice of vague priors. We will make the data and code underpinning our analyses freely available on an open-source platform (e.g. GitHub).

4.3. WP3: health economic decision model

4.3.1. Objective

WP3 will estimate the costs, benefits and harms of surveillance for HCC using a decision-analytic model simulating natural history, diagnosis and treatment of HCC in people with cirrhosis.

4.3.2. Structure

The model will adopt a similar approach to the 2007 HTA (*Figure 2*) – a state-transition model simulating natural history of cirrhosis and HCC (with transitions reflecting decompensation and HCC development), overlaid on which are states representing HCC detection (through surveillance or symptoms) and treatment.

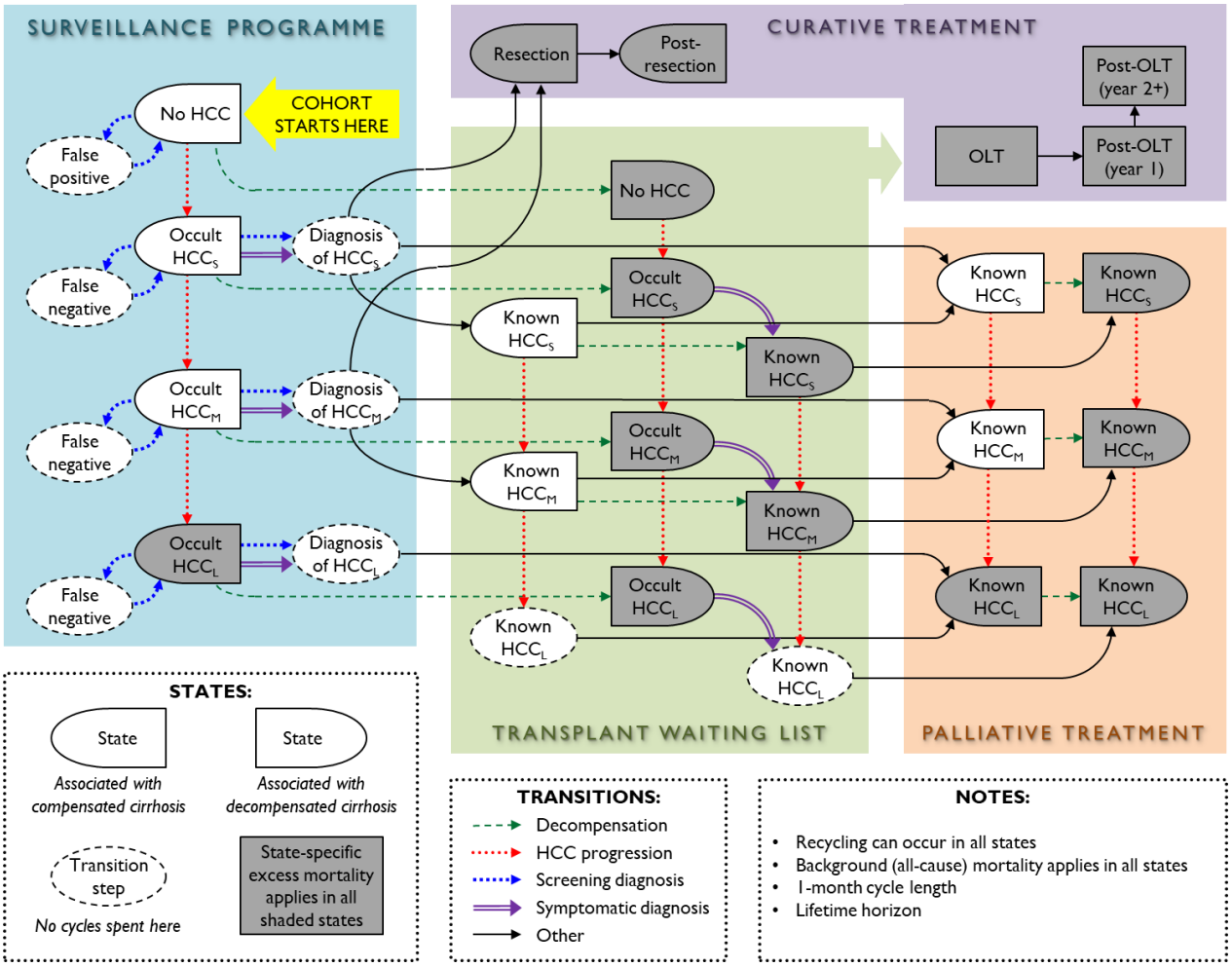

Figure 2: Structure of 2007 model

However, we will extend the structure to reflect the present-day pathway. The extent of modification will be mediated by the availability of suitable data to characterise new aspects of the pathway. As a minimum, we will explore

- Providing more detailed characterisation of underlying liver disease:
  - allowing for recompensation of cirrhosis (e.g. following antiviral therapy)
  - modelling cirrhosis progression for people who are not eligible for transplantation
- Incorporating a wider range of HCC treatment options
  - radiofrequency/microwave ablation with curative intent
  - locoregional therapies that may delay or reverse HCC growth for people awaiting transplantation
    - systemic therapies for inoperable disease that are recommended by NICE.
- Explicit consideration of potential harms of surveillance, which may include:

  - pain and less common sequelae of biopsy
  - contrast-induced kidney injury
  - exposure to ionising radiation
  - psychological and iatrogenic morbidity associated with overdiagnosis of liver pathology
  - psychological and iatrogenic morbidity associated with incidental (nonhepatic) abdominal findings
- Better handling of competing causes of mortality

These modifications will provide an updated representation of the clinical pathway that is fit for present-day analysis. Moreover, they will enable us to generate finely specified, risk-stratified outputs. For example, the 2007 HTA assumed everyone was eligible for transplantation and subject to homogeneous risks of non-HCC death, whereas we intend the revised model to explore how the benefits, harms and costs of surveillance may vary for people in a heterogeneous range of circumstances (see subgroup analyses in ¶4.3.4, below).

4.3.3. Parameters

WP2 will provide critical inputs. We will incorporate accuracy as a function of test threshold, to establish the thresholds that maximise the net benefit of surveillance.

We will identify other model inputs (e.g. natural history of cirrhosis and HCC; effects of treatment for HCC; costs and quality of life) using informal but transparent methods that aim to satisfy the principle of 'saturation' – that is, to 'identify the breadth of information needs relevant to a model and sufficient information such that further efforts to identify more information would add nothing to the analysis'[69].

The costs we take into account will include: surveillance appointments and tests, work-up of suspicious findings, treatment for underlying liver disease, curative and palliative treatment of HCC. We will estimate quality of life by attaching utilities to health-states (compensated and decompensated cirrhosis; presence of HCC); we will also seek to account for physical harms of diagnosis (biopsy, contrast-induced kidney injury, exposure to ionising radiation) as well as psychological morbidity associated with (true- and false-) positive findings.

4.3.4. Analytic approach

The analysis will conform to the NICE reference case (PMG9).

We will simulate strategies defined in 4 dimensions:

- Frequency (e.g. 6-monthly)
- Test(s) (e.g. AFP assay)
- Threshold(s) (e.g. 20 ng/ml)
- Stopping rule (e.g. discontinue surveillance if negative at age 80)

The model will estimate expected lifetime costs and QALYs for each strategy, comparing them with each other and with no surveillance in fully incremental cost–utility analysis.

The model will be fully probabilistic; base-case results will be the mean of at least 1,000 iterations and we will generate cost-effectiveness acceptability and/or expected loss curves. We will perform

thorough one-way sensitivity analysis, to understand how our analyses are sensitive to changes in individual parameter values. Multi-way deterministic analyses are also likely to be helpful, to explore, e.g., the interaction between age and competing risk of death in attenuating the potential benefits of surveillance.

We will undertake extensive subgroup analysis, to explore how the balance of benefits, harms and costs varies among people with different characteristics – e.g. aetiology of liver cirrhosis, eligibility for treatment (see ¶3.2). We will also perform scenario analyses reflecting particular clinical uncertainties. For example, to simulate the decision-problem for people whose BMI makes them ineligible for MRI, we can assess whether removing MRI from the diagnostic algorithm impacts the effectiveness of surveillance with other tools; for people whose livers cannot be adequately imaged by ultrasound, we might assess whether the benefits of CT-led surveillance outweigh its costs and harms in a world without ultrasound.

We will also meet Cancer Research UK's challenge to define characteristics and costs that hypothetical future tests would have to achieve to represent an effective use of NHS resources[70].

We will report our findings using guidance on best practice (CHEERS[71]). We will make the data and code underpinning our analyses freely available on an open-source platform (e.g. GitHub).

4.4. WP4: patient decision-aid

4.4.1. Objective

In WP4, we will develop a print-based patient decision-aid, using what we have learned in WPs1–3 to quantify the expected benefits and harms of surveillance. It will aim to support shared decision-making for people with cirrhosis about whether to start – or stop – surveillance.

Our motivation for developing this resource is threefold:

- First, and most simply, people facing a decision about starting or continuing surveillance have a right to the best possible estimate of likely outcomes, presented in a way that they are most likely to find comprehensible. NICE's recent guideline on Shared Decision-Making (NG197) contains a strong recommendation that patient decision-aids should be used, where available and reliable, based on 'strong evidence' that they increase participation in decision-making, improve knowledge and risk perception and reduce decisional conflict.
- More particularly, there is evidence that people are more likely to adhere to surveillance for HCC if they report feeling involved in the decision process[72]. Hence, informed shared decision-making not only ensures that people with cirrhosis understand the intended benefits of surveillance; it actively increases the likelihood that those benefits will be realised in people who elect to undergo it.
- Lastly, based on team-members' direct experience of NICE guideline development, we believe it is very likely that future updates of liver guidelines (CG165; NG50) would consider the decision to offer surveillance for HCC a 'preference-sensitive decision point'. Under these circumstances, NICE asks guideline developers to provide outputs that 'make it easy for professionals and practitioners to compare the options and discuss them with the person' (PMG20). It may also develop its own decision-aids. However, rather than expecting guideline

developers to assimilate every nuance of our research some time after it is complete, we think it makes sense for the people who know the evidence synthesis best to lay out what it implies for individual-level decision-making and to test that with a broad cross-section of patients. That way, policymakers like NICE will have access to ready-made resources that meet their needs, rather than having to piece them together after the event.

4.4.2. Methods

We will bring together findings from WPs1–3 to summarise expected outcomes with and without surveillance. The decision-aids will illustrate increase in early diagnosis of HCC, treatment received and impact on overall survival, set against likelihood of false-positive diagnoses, potentially harmful investigations, and overdiagnosis (which may include incidental findings of unrelated pathology).

We will develop the aid following the content and process standards set out by NICE in its recently published framework for patient decision-aids (ECD8). This, in turn, references the recommendations of the International Patient Decision-Aid Standards (IPDAS) collaboration[73,74]. As this is an evidence synthesis project, we will stop short of the kind of primary evaluative research the full IPDAS guidelines envisage, but will aim to produce outputs that are fit for that purpose. The process will comprise:

1. Understanding user needs. We will orientate the WP at PPI workshop #1 by seeking the views of people who are currently facing this decision or have faced it in the past. This step precedes the collection and synthesis of evidence, so that we can ensure that the outputs of our work address the dimensions of the decision about which patients would find it useful to receive information. We will also obtain feedback from the clinicians on our advisory group (see ¶8.2).
2. Developing a prototype. The technical team will work iteratively with patient and clinician co-investigators to produce a provisional design. We will follow guidance from IPDAS[75] and NICE's guideline on Shared Decision-Making (NG197) about presenting quantitative information; for example, we are likely to present event probabilities numerically and with pictographs ('icon arrays') and we will depict the whole population under surveillance, rather than focusing only on people who experience events.
3. Alpha-testing with patients and clinicians not involved in the development process to check acceptability, comprehensibility and usability. We will elicit patient views at PPI workshop #2 and obtain feedback from the clinicians on our advisory group (see ¶8.2).
4. Using feedback from step 3, producing a final version fit for beta-testing in 'live' settings and/or formal evaluation outside this project (e.g. in a randomised trial).

If we can, we will tailor the decision-aid to reflect the varying balance of benefits and harms in different populations. We will document the process using NICE's patient decision-aid self-assessment framework (ECD8) and guidance from IPDAS[74].

4.4.3. Surveillance strategies to include in WP4 (handling dependence on WP3)

We will summarise expected findings for those strategies that emerge as good value for money at a population level and also – in the event that it is not already included in this group – 6-monthly ultrasound +/- AFP (6moUS–AFP), the approach NICE currently recommends (CG165; NG50). Generically speaking, there are 3 broad outcomes that could possibly emerge from WP3. Our plans for each eventuality are as follows:

1. WP3 confirms that 6moUS–AFP is optimal for all people with cirrhosis.
    - In this case, we will only produce a decision-aid summarising the benefits and harms of 6moUS–AFP.
2. WP3 suggests that surveillance strategies other than 6moUS–AFP are optimal for some or all people with cirrhosis.
    - In this case, we will produce versions of our decision-aid for such strategies. If there are multiple similarly cost-effective options, we will make a judgement (with the input of our expert advisory group) about a restricted number for which it may be helpful to summarise expected outcomes, based on which are most likely to be implemented.
    - Regardless of whether it is among the cost-effective options for some or all people, we will produce a decision-aid summarising the benefits and harms of 6moUS–AFP. This will ensure that our outputs are immediately useful for liver services as they are most likely to be configured at present, even if our findings suggest that a reconfigured approach would ultimately provide greater net benefit.
3. WP3 finds that no surveillance programme is cost effective for some or all people with cirrhosis.
    - In this case, we will still produce a decision-aid summarising the benefits and harms of 6moUS–AFP. Again, our motivation is to provide something that helps people make decisions for as long as they are being offered surveillance, even if we find that the offer does not represent an effective use of NHS resources.

Once we have agreed a prototype template (see above), we plan to automate the generation of quantitative summaries from cost-effectiveness model outputs, which will make it easy to produce outputs for relevant strategies in a fast and flexible way.

5. Patient and public involvement

Patient and Public Involvement and Engagement (PPIE) has already formed – and will continue to form – a central part of this project, with a robust PPIE plan and commitment to ensure patients and the public are partners throughout the lifecycle of the research.

5.1. How were patients and the public involved in planning the project?

At the time of the stage 1 application, we gained support from the British Liver Trust (BLT), and Nottingham University Hospitals NHS Trusts (NUH) PPIE infrastructure. This consultation has led to the idea for WP4, a decision-aid summarising expected benefits and harms of surveillance with a view to supporting shared decision-making for people with cirrhosis. This received support

from carers and patients living with cirrhosis: "*We need to be proactive not reactive, assisting the individual patient in the decision-making process is crucial.*"

For the stage 2 application, we continued to build on initial engagement through strategic direction from Head of PPIE Kate Frost, NUH, who is PPI Lead for the project. We gained feedback on our plans and how we express them from members of the Nottingham NIHR Biomedical Research Centre (BRC) Liver Patient Advisory Group, which is a heterogeneous group of patients living with liver conditions. The detailed research plan and PPI section were reviewed by the British Liver Trust and members of the NIHR BRC Liver Advisory Group and Research Volunteers as part of NUH Research and Innovation, all with a diverse range of lived experience and expertise. NUH Research Volunteers advised on the PPIE plan for the study, emphasising the importance of continued representation of under-served communities in line with NUH R&I's equality, diversity and inclusivity strategy when recruiting to future involvement sessions.

We also presented the project plans to the NUH Research Volunteer expert panel, including 2 people with experience of cirrhosis and HCC (themselves or their immediate family). The proposed research was well received, and the debate that followed highlighted the substantial heterogeneity in preferences and attitudes that may exist between people. Some people said they would want to have access to surveillance if it provides even a small chance of detecting treatable disease that would otherwise be missed; others expressed the view that, as for any screening test, we should be confident that the benefits outweigh the harms before promoting any surveillance programme. However, participants agreed that providing patients with accurate and understandable information about the possible outcomes of surveillance is critical.

5.2. How will patients and the public be involved in delivering the project?

Our PPI strategy comprises 3 elements:

1. **Patient co-investigators**. We will recruit 2 team members at PPI workshop #1 (see below). Ideally, 1 will currently be eligible for surveillance and 1 will have experienced treatment for HCC. They will attend monthly meetings (¶8.1), contribute to the design and interpretation of analyses and co-author outputs (including, but not limited to, those intended for nonclinical audiences). They will co-create the decision-aid for WP4, and help to present it for alpha-testing at PPI workshop #2 (¶4.4). The PPI team at Nottingham BRC will provide support. We will encourage patient co-Is to join their research volunteer programme, gaining access to additional training and development.

2. **Direct engagement with the patient community** in 2 online workshops with 10–15 participants, identified from an NIHR-funded cohort of people with cirrhosis in Nottingham[76] and BLT's patient community, targeted to reflect diversity of liver disease community. Workshop #1, at the start of the project, will present the project plan, and seek suggestions to ensure the relevance and usefulness of outputs for patients. Workshop #2, at the end, will present findings and alpha-test the decision-aid, seeking feedback to maximise usability and effective communication.

3. **Representation on advisory group**. We have secured the participation of representatives from BLT to sit on our advisory group. They have in-depth knowledge of the broad range of

views held by the people they represent. They will also provide important policy context and advise on effective dissemination, with both patients and strategic audiences in mind.

In addition, we have access to BLT's patient forum throughout the project. We will use this to test ideas, when our patient co-Is advise that the views of a wider cross-section of the community would be helpful. We have followed NIHR standards on costing PPI involvement.

In recruiting collaborators, we will aim to reflect the diversity of people with cirrhosis. Hepatitis (especially HBV) is more common in migrant populations[77] and all forms of liver disease have higher prevalence in people facing greater socioeconomic deprivation[78]. We will ensure the inclusion of underserved communities by partnering with organisations with specific expertise in relevant areas. To target people experiencing socioeconomic deprivation, we will work with the Rebalancing the Outer Estates Foundation, a small charity working with socioeconomically deprived areas of North Nottingham. We have made early contact with multiple other groups that will help us to target immigrant communities.

6. Dissemination, outputs and anticipated impact

6.1. Intended outputs

As a minimum, we intend to produce the following outputs:

- **Publications:** In addition to the NIHR monograph, we will submit at least 3 open-access papers to internationally recognised academic journals (1 for each of WPs1–3).
- **Conferences:** We will present our findings at at least 2 scientific conferences (we have provisionally targeted the BASL annual meeting and the EASL congress).
- **Targeted outputs:** As noted below, we have made links with patient and professional groups with established channels for reaching their members. We will partner with them to prepare materials tailored for dissemination through these channels – e.g. visual summaries to share on social media, focused updates for clinicians, briefings for policymakers.
- **Decision-aids:** As a special case of the above, the materials we develop in WP4 will be made freely available, and we will work with our partners to make sure they are accessible to people who would find them useful – this implies targeting clinical as well as patient audiences, as the onus should not be on newly diagnosed people to be aware of the resources that are available to support them.
- **Open science:** We will make the statistical synthesis code from WP2 and the decision-model from WP3 freely available on an open-source platform (e.g. GitHub).

6.2. Informing and engaging patients/service users, carers, NHS, social care organisations and the wider population

We will use our partnership with BLT to target patient, professional and policy audiences. BLT has over 1.2m unique website visitors per year, an online patient forum with 24,000 members, Facebook and Twitter accounts with 16,000 followers, plus quarterly newsletters targeted at patients and healthcare professionals.

The NIHR CRN East Midlands, in which our PPI lead's Trust is a partner, run a research champions programme (based on the NIHR scheme). We will seek opportunities to engage with this resource to identify ambassadors for our work.

6.3. How our outputs will enter the health and care system or society as a whole

NICE is planning a minor update of its guideline on cirrhosis (NG50); this will not cover surveillance for HCC, but we will register as stakeholders. We will alert NICE's event-tracking system to our work, and ensure that our outputs are optimised to inform guideline development processes. As we are committed to developing a version of the decision-aid reflecting expected outcomes for the surveillance regimen currently recommended by NICE (see ¶4.4.2), we will immediately be able to submit that output for endorsement as a 'quality standard support resource' for the relevant NICE products (QS152; see PMG29 for process).

We will invite BLT to 'badge' the decision-aid we develop in WP4 (though their involvement in the project does not compel them to do so).

Our advisory group includes the current chair and other members of HCC-UK, a subgroup of BASL engaging clinicians of all disciplines with an interest in HCC.

6.4. Possible barriers for further research, development, adoption and implementation

As noted in ¶2.1, implementation of surveillance is patchy in the NHS, despite NICE guidance that it should be offered (CG165; NG50). We believe that our project can address some of the cited barriers[80]: doubts about effectiveness and costs (if surveillance is shown to be a good use of resources, these should be allayed); poor patient adherence (our decision-aid will address this both by increasing patient activation among people who want surveillance and by giving people who are unlikely to adhere an explicit opportunity to decline follow-up).

There are also logistical barriers to implementing surveillance (e.g. effective liaison between hepatology and radiology departments) that we cannot directly address. However, we anticipate that, if they emphasise the value of surveillance, our findings will provide impetus to treat establishing or maintaining effective systems as a priority. Our dissemination activities will provide materials that could be used to underpin local cases for investment. Equally, if our findings cast doubt on the net benefit of surveillance (either for everyone or for clearly identifiable groups of people), it would be appropriate to use them as a basis for disinvestment decisions. In that event, the current existence of NICE guidance recommending surveillance would be a barrier to adoption; however, we envisage that our evidence would be central to future reconsideration of such guidance.

6.5. Further funding or support that may be required beyond the project

Before the decision-aid we develop in WP4 could be formally recommended as a routine tool in the NHS, it should be quality assured (in line with NICE NG197) to establish that it is likely to improve the decision-making process and increase decision quality[79]. To be certain of these outcomes, it should be evaluated in a randomised trial.

It is also easy to envisage interactive versions of the decision-aid, which could tailor risk estimation to individual circumstances, like the Predict websites developed by the University of

Cambridge's Winton Centre. Such resources would require funding and technical expertise beyond the scope of our project to develop.

7. Project / research timetable

The project will run for 21 months from July 2022 until March 2024. *Figure 3* illustrates the month-by-month schedule. Key intermediate milestones will be: completion of systematic reviews and syntheses (WP 1 & 2) – end May 2023; completion of decision model (WP3) – end Nov 2023.

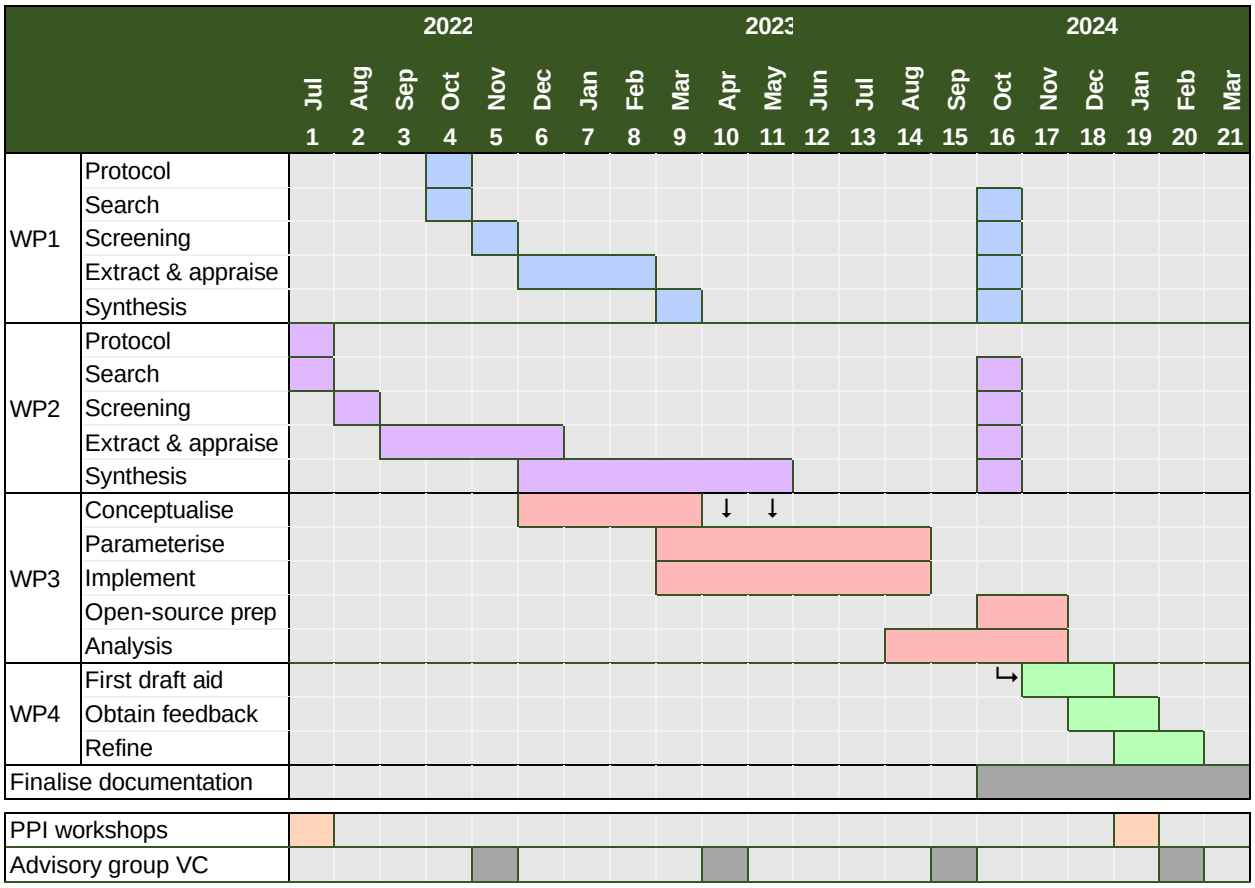

Figure 3: Project timetable

8. Project management

8.1. Day-to-day

We will hold monthly research team meetings to maintain connections between work packages. While each work package is active (see ¶6.5), the people involved will have weekly team meetings. As investigators are based in several different institutions, all meetings will be held remotely via videoconference.

GR will provide overall project management, with administrative assistance from Manchester Centre for Health Economics.

8.2. Expert advisory group

We have assembled a multidisciplinary advisory group (see *Table 2*), comprising clinical, methods and patient experts. The group will provide oversight and advice for all aspects of the project, to ensure its successful completion. A wide range of clinical specialties is involved in the pathway we will be exploring and simulating; therefore, it is vital that we have access to expertise from disciplines not represented among the investigators (interventional radiology, surgery, oncology, nursing) and it is equally important that the whole team is subject to critical challenge from peers. The group includes members of national organisations (HCC-UK, British Liver Trust)

Table 2: Expert advisory group

Area of expertise	Member
Hepatology	A Marshall (Royal Free)
	V Athwal (Manchester)
Radiology	C Hammond (Leeds)
	C Clarke (Nottingham)
Surgery	D Manas (Newcastle)
Oncology	R Hubner (Manchester)
Nursing	A Clements (Plymouth)
PPI	TBC (British Liver Trust)
Methods (health economics)	L Claxton (NICE)
Methods (statistics)	A Sutton (U of Leicester)

There will be 4 advisory group meetings over the course of the project, at which the research team will give an update on progress and seek critical feedback. We will also ask the group to comment on protocols and outputs for each WP, as they are completed. Advisory group members have agreed to act as first point of contact for any ad hoc queries beyond the research team's direct expertise.

9. Ethics

The project does not need ethical approval, as it is based on scientific literature and other data in the public domain. The planned PPI activity does not require specific ethical approval in accordance with NIHR guidance.

10. Project / research expertise

The project team will comprise clinical, methods and patient experts:

- **Gabriel Rogers** Senior research fellow in health economics. Extensive experience in both systematic review (e.g. HTAs including the 2007 one[2,81,82]) and decision modelling (e.g. as lead health economist on around 30 NICE guidelines). Experience at NICE working with professionals and patients to develop decision-aids; development team for Shared Decision Making guideline (NG197). Now a member of the NICE Technology Appraisals committee.
  - Chief investigator. Project management (coordinating across WPs, liaising with project partners and expert advisers, reporting to NIHR, delivery of final outputs); lead WP3; co-lead WP1 and WP4; supervise health economist.
- **Kris Bennett** Currently completing an NIHR Academic Clinical Fellowship with a focus on hepatology. When this project begins, he will be a full-time student – his time is provided as support-in-kind.
  - Co-lead for WP1; second reviewer for WP2. Contribute to conceptualising and parameterising model in WP3.

- **Matthew Cramp** Professor of hepatology. Senior hepatologist with experience of organising and delivering specialist liver services at a local, regional and national level. Past president of the British Association for the Study of the Liver, a member of the Hepatopancreatico-biliary Clinical Reference Group and a member of the Lancet Commission on Liver Disease.
  - Hepatology input to all WPs.
- **Kate Frost** Head of patient and public involvement for Nottingham BRC, providing strategic direction for all PPIE activity across Nottingham University Hospitals and their centres of excellence.
  - PPI lead. Facilitating recruitment and support of patient co-investigators; organisation of PPIE workshops and recruitment of attendees; contributing to patient-focused dissemination materials.
- **Hayley Jones** Senior lecturer in medical statistics. Expertise in advanced methods for meta-analysis of test accuracy, in particular having developed models for meta-analysis of continuous tests across diagnostic thresholds[62]. She also currently holds an MRC-NIHR NIRG award to develop methods for meta-analysis of comparative test accuracy (MR/T044594/1) which will feed into this project.
  - Co-lead WP2; synthesis of diagnostic data; supervise statistician
- **Morwenna Rogers** Information specialist with extensive experience in systematic reviews and expertise in search methods for evidence syntheses.
  - Design and conduct searches for WP1 and WP2; ad-hoc support for other information needs (e.g. model parameterisation)
- **Steve Ryder** Consultant physician in hepatology and gastroenterology. Director of research and innovation at Nottingham University Hospitals NHS Trust. Trustee of the British Liver Trust. Former national lead for the Hepatology Clinical Research Network
  - Hepatology input to all WPs; liaison with British Liver Trust.
- **Ken Stein** Professor of public health. Public health physician with substantial expertise in HTA and communicating risk to patients. Past vice-chair of the NICE Technology Appraisal committee. Holder of NIHR Senior Investigator award.
  - Strategic oversight; advice on shared decision-making; co-lead WP4; mentoring first-time chief investigator
- **Kelsey Watt** Consultant gastrointestinal radiologist. Particular interest in diagnostic imaging of chronic liver disease.
  - Radiology input to all WPs.
- **Nicky Welton** Professor in statistical and health economic modelling. Extensive experience delivering and supervising applied HTA projects, providing expertise in evidence synthesis and economic modelling. She is director of the NICE Guidelines Technical Support Unit, and a member of the NICE Technology Appraisals committee.
  - Advice and oversight: evidence synthesis and economic modelling
- **Penny Whiting** Professor of clinical epidemiology. Over 20 years' experience of leading systematic reviews, particularly of diagnostic accuracy studies. Led or contributed to the development of risk of bias tools including QUADAS-2, PROBAST and ROBINS-I and contributed to reporting guidelines including PRISMA 2020 and STARD 2015.

- Oversight: systematic reviews (diagnostic test accuracy); co-lead WP2; supervise systematic reviewer

We will recruit the following research staff:

- Systematic reviewer (WP1 & WP2; 100%FTE for 9 months; supervised by PW)
- Statistician (WP2; 50%FTE for 6 months; supervised by HJ)
- Health economist (WP3; 100%FTE for 12 months; supervised by GR).

We have budgeted additional time for all 3 to contribute to WP4.

We will also recruit 2 patient co-investigators, for whom we have budgeted payment and resources (see PPI plan).

11. Success criteria

Our primary success criteria are that our research should prove its worth

- … from a population-level perspective. We will know we have achieved this if decision-makers such as those that cited the 2007 HTA (NICE, WHO, EASL, AASLD) rely on our new work as they update their guidance on HCC surveillance strategies. Our plan to engage such decision-makers is central to this goal; see ¶¶6.1 and 6.2.
- … from an individual-level perspective. We will consider our work a success if people with cirrhosis are able to access and make use of our outputs (for example, we would be delighted if our partnership with BLT results in them making our decision-aid available via their patient information channels; see ¶6.2).
- … from a scientific perspective. We will know our work has been successful if it influences future research, and we will have direct evidence of this if future researchers adopt and develop our open-source resources (primarily those interested in liver disease, but potentially more broadly as well)

12. References

1. Dyson J, Jaques B, Chattopadyhay D, Lochan R, Graham J, Das D, et al. Hepatocellular cancer: The impact of obesity, type 2 diabetes and a multidisciplinary team. J Hepatol 2014;60(1):110–7.
2. Thompson Coon J, Rogers G, Hewson P, Wright D, Anderson R, Cramp M, et al. Surveillance of cirrhosis for hepatocellular carcinoma: systematic review and economic analysis. Health Technol Assess [Internet] 2007;11(34). Available from: https://www.journalslibrary.nihr.ac.uk/hta/hta11340/
3. Thompson Coon J, Rogers G, Hewson P, Wright D, Anderson R, Jackson S, et al. Surveillance of cirrhosis for hepatocellular carcinoma: a cost–utility analysis. Br J Cancer 2008;98(7):1166–75.
4. World Health Organization. Guidelines for the prevention, care and treatment of persons with chronic hepatitis B infection [Internet]. World Health Organization; 2015 [cited 2021 Apr 10]. Available from: https://apps.who.int/iris/handle/10665/154590
5. Liver EA for the S of the, Cancer EO for R and T of. EASL–EORTC Clinical Practice Guidelines: Management of hepatocellular carcinoma. J Hepatol 2012;56(4):908–43.
6. Sherman M, Bruix J, Porayko M, Tran T. Screening for hepatocellular carcinoma: The rationale for the American Association for the Study of Liver Diseases recommendations. Hepatology 2012;56(3):793–6.
7. Ratib S, West J, Crooks CJ, Fleming KM. Diagnosis of Liver Cirrhosis in England, a Cohort Study, 1998–2009: A Comparison With Cancer. Off J Am Coll Gastroenterol ACG 2014;109(2):190–8.
8. Burton A, Tataru D, Driver RJ, Bird TG, Huws D, Wallace D, et al. Primary liver cancer in the UK: Incidence, incidence-based mortality, and survival by subtype, sex, and nation. JHEP Rep [Internet] 2021 [cited 2021 Apr 9];3(2). Available from: https://www.jhep-reports.eu/article/S2589-5559(21)00008-2/abstract
9. Sepanlou SG, Safiri S, Bisignano C, Ikuta KS, Merat S, Saberifiroozi M, et al. The global, regional, and national burden of cirrhosis by cause in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Gastroenterol Hepatol 2020;5(3):245–66.
10. Public Health England. Hepatitis C in the UK 2020: Working to eliminate hepatitis C as a major public health threat [Internet]. 2020. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/943154/HCV_in_the_UK_2020.pdf
11. Colombo M, Iavarone M. Role of antiviral treatment for HCC prevention. Best Pract Res Clin Gastroenterol 2014;28(5):771–81.
12. Thiele M, Gluud LL, Dahl EK, Krag A. Antiviral therapy for prevention of hepatocellular carcinoma and mortality in chronic hepatitis B: systematic review and meta-analysis. BMJ Open 2013;3(8):e003265.
13. Office for National Statistics. Cancer registration statistics, England, 2005 [Internet]. 2008 [cited 2021 Apr 9]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancerregistrationstatisticscancerregistrationstatisticsengland
14. Office for National Statistics. Cancer registration statistics, England, 2017 [Internet]. 2019 [cited 2021 Apr 9]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancerregistrationstatisticscancerregistrationstatisticsengland
15. Kollmann D, Selzner N, Selzner M. Bridging to liver transplantation in HCC patients. Langenbecks Arch Surg 2017;402(6):863–71.
16. Parikh ND, Waljee AK, Singal AG. Downstaging hepatocellular carcinoma: A systematic review and pooled analysis. Liver Transpl 2015;21(9):1142–52.
17. Lou J, Zhang L, Lv S, Zhang C, Jiang S. Biomarkers for Hepatocellular Carcinoma. Biomark Cancer 2017;9:1179299X16684640.
18. Roberts LR, Sirlin CB, Zaiem F, Almasri J, Prokop LJ, Heimbach JK, et al. Imaging for the diagnosis of hepatocellular carcinoma: A systematic review and meta-analysis. Hepatology 2018;67(1):401–21.
19. Berhane S, Toyoda H, Tada T, Kumada T, Kagebayashi C, Satomura S, et al. Role of the GALAD and BALAD-2 Serologic Models in Diagnosis of Hepatocellular Carcinoma and Prediction of Survival in Patients. Clin Gastroenterol Hepatol 2016;14(6):875-886.e6.
20. NICE. Cirrhosis in over 16s: assessment and management [Internet]. [cited 2021 Apr 8];Available from: https://www.nice.org.uk/guidance/ng50
21. Galle PR, Forner A, Llovet JM, Mazzaferro V, Piscaglia F, Raoul J-L, et al. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. J Hepatol 2018;69(1):182–236.
22. Marrero JA, Kulik LM, Sirlin CB, Zhu AX, Finn RS, Abecassis MM, et al. Diagnosis, Staging, and Management of Hepatocellular Carcinoma: 2018 Practice Guidance by the American Association for the Study of Liver Diseases. Hepatology 2018;68(2):723–50.
23. Lederle FA, Pocha C. Screening for Liver Cancer: The Rush to Judgment. Ann Intern Med 2012;156(5):387–9.
24. Jepsen P, West J. We need stronger evidence for (or against) hepatocellular carcinoma surveillance. J Hepatol [Internet] 2021 [cited 2021 Apr 8];0(0). Available from: https://www.journal-of-hepatology.eu/article/S0168-8278(21)00020-9/abstract
25. Cross TJS, Villaneuva A, Shetty S, Wilkes E, Collins P, Adair A, et al. A national survey of the provision of ultrasound surveillance for the detection of hepatocellular carcinoma. Frontline Gastroenterol 2016;7(2):82–9.
26. Lee SS, Sherman M. Why Won't the Alpha-fetoprotein Test Go Gentle Into the Good Night? Gastroenterology 2018;154(6):1572–3.
27. Office for National Statistics. Cancer survival in England - adults diagnosed, 2013-2017 [Internet]. 2019. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancersurvivalratescancersurvivalinenglandadultsdiagnosed
28. Less Survivable Cancers Taskforce. Closing the deadly cancer gap: detect early, diagnose fast, save lives [Internet]. 2020. Available from: http://lesssurvivablecancers.org.uk/wp-content/uploads/2020/08/Early-Diagnosis-report-final.pdf

29. Taylor EJ, Jones RL, Guthrie JA, Rowe IA. Modeling the benefits and harms of surveillance for hepatocellular carcinoma: Information to support informed choices. Hepatology 2017;66(5):1546–55.
30. Rich NE, Parikh ND, Singal AG. Overdiagnosis: An Understudied Issue in Hepatocellular Carcinoma Surveillance. Semin Liver Dis 2017;37(4):296–304.
31. Study design search filters | BMJ Best Practice [Internet]. [cited 2021 Aug 16];Available from: https://bestpractice.bmj.com/info/toolkit/learn-ebm/study-design-search-filters/
32. Singal AG, Pillai A, Tiro J. Early Detection, Curative Treatment, and Survival Rates for Hepatocellular Carcinoma Surveillance in Patients with Cirrhosis: A Meta-analysis. PLOS Med 2014;11(4):e1001624.
33. Colli A, Nadarevic T, Miletic D, Giljaca V, Fraquelli M, Štimac D, et al. Abdominal ultrasound and alpha-foetoprotein for the diagnosis of hepatocellular carcinoma in adults with chronic liver disease. Cochrane Database Syst Rev [Internet] 2021 [cited 2021 Aug 16];(4). Available from: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD013346.pub2/full
34. Gupta P, Soundararajan R, Patel A, Kumar-M P, Sharma V, Kalra N. Abbreviated MRI for hepatocellular carcinoma screening: A systematic review and meta-analysis. J Hepatol 2021;75(1):108–19.
35. Wang X, Zhang Y, Yang N, He H, Tao X, Kou C, et al. Evaluation of the Combined Application of AFP, AFP-L3%, and DCP for Hepatocellular Carcinoma Diagnosis: A Meta-analysis. BioMed Res Int 2020;2020:e5087643.
36. Chen VL, Xu D, Wicha MS, Lok AS, Parikh ND. Utility of Liquid Biopsy Analysis in Detection of Hepatocellular Carcinoma, Determination of Prognosis, and Disease Monitoring: A Systematic Review. Clin Gastroenterol Hepatol 2020;18(13):2879-2902.e9.
37. Nadarevic T, Giljaca V, Colli A, Fraquelli M, Casazza G, Miletic D, et al. Computed tomography for the diagnosis of hepatocellular carcinoma in chronic advanced liver disease. Cochrane Database Syst Rev [Internet] 2019;(6). Available from: doi.org/10.1002/14651858.CD013362
38. Fraquelli M, Nadarevic T, Giljaca V, Colli A, Miletic D, Štimac D, et al. Contrast-enhanced ultrasound for the diagnosis of hepatocellular carcinoma in advanced chronic liver disease. Cochrane Database Syst Rev [Internet] 2019;(11). Available from: doi.org/10.1002/14651858.CD013483
39. Nguyen ALT, Nguyen HTT, Yee KC, Palmer AJ, Blizzard CL, de Graaff B. A Systematic Review and Narrative Synthesis of Health Economic Evaluations of Hepatocellular Carcinoma Screening Strategies. Value Health [Internet] 2021 [cited 2021 Apr 8];Available from: https://www.sciencedirect.com/science/article/pii/S1098301521000358
40. Poustchi H, Farrell GC, Strasser SI, Lee AU, McCaughan GW, George J. Feasibility of conducting a randomized control trial for liver cancer screening: Is a randomized controlled trial for liver cancer screening feasible or still needed? Hepatology 2011;54(6):1998–2004.
41. Johnson PJ, Pirrie SJ, Cox TF, Berhane S, Teng M, Palmer D, et al. The Detection of Hepatocellular Carcinoma Using a Prospectively Developed and Validated Model Based on Serological Biomarkers. Cancer Epidemiol Prev Biomark 2014;23(1):144–53.
42. Chalasani NP, Porter K, Bhattacharya A, Book AJ, Neis BM, Xiong KM, et al. Validation of a novel multi-target blood test shows high sensitivity to detect early-stage hepatocellular carcinoma. Clin Gastroenterol Hepatol [Internet] 2021 [cited 2021 Sep 1];Available from: https://www.sciencedirect.com/science/article/pii/S1542356521008661
43. Wang M, Devarajan K, Singal AG, Marrero JA, Dai J, Feng Z, et al. The Doylestown Algorithm: A Test to Improve the Performance of AFP in the Detection of Hepatocellular Carcinoma. Cancer Prev Res (Phila Pa) 2016;9(2):172–9.
44. Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care [Internet]. York: University of York; Available from: https://www.york.ac.uk/crd/guidance/
45. Cochrane Handbook for Systematic Reviews of Interventions v6.2 [Internet]. Available from: https://training.cochrane.org/handbook/current
46. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy [Internet]. Available from: methods.cochrane.org/sdt/handbook-dta-reviews
47. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000;283(15):2008–12.
48. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, and the PRISMA-DTA Group, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA 2018;319(4):388–96.
49. Scherer RW, Saldanha IJ. How should systematic reviewers handle conference abstracts? A view from the trenches. Syst Rev 2019;8(1):264.
50. Guyot P, Ades A, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol 2012;12(1):9.
51. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016;355:i4919.
52. Faria R, Alava MH, Manca A, Wailoo AJ. NICE DSU Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness in technology appraisal: methods for comparative individual patient data [Internet]. Available from: http://nicedsu.org.uk/technical-support-documents/observational-data-tsd/
53. Rutjes AW, Reitsma JB, Vandenbroucke JP, Glas AS, Bossuyt PM. Case–Control and Two-Gate Designs in Diagnostic Accuracy Studies. Clin Chem 2005;51(8):1335–41.
54. Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JHP, et al. Empirical Evidence of Design-Related Bias in Studies of Diagnostic Tests. JAMA 1999;282(11):1061–6.

55. Whiting P, Rutjes AWS, Reitsma JB, Glas AS, Bossuyt PMM, Kleijnen J. Sources of Variation and Bias in Studies of Diagnostic Accuracy. Ann Intern Med 2004;140(3):189–202.
56. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. Ann Intern Med 2011;155(8):529–36.
57. Yang B, Whiting P, Davenport C, Deeks J, Hyde C, Mallett S, et al. Development of QUADAS-C, a risk of bias tool for comparative diagnostic accuracy studies [preprint]. 2021;Available from: https://osf.io/hq8mf/
58. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005;58(10):982–90.
59. Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. J Clin Epidemiol 2006;59(12):1331–2.
60. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med 2001;20(19):2865–84.
61. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JAC. A unification of models for meta-analysis of diagnostic accuracy studies. Biostatistics 2007;8(2):239–51.
62. Jones HE, Gatsonsis CA, Trikalinos TA, Welton NJ, Ades AE. Quantifying how diagnostic test accuracy depends on threshold in a meta-analysis. Stat Med 2019;38(24):4789–803.
63. Takwoingi Y, Leeflang MMG, Deeks JJ. Empirical Evidence of the Importance of Comparative Studies of Diagnostic Test Accuracy. Ann Intern Med 2013;158(7):544–54.
64. Trikalinos TA, Hoaglin DC, Small KM, Terrin N, Schmid CH. Methods for the joint meta-analysis of multiple tests. Res Synth Methods 2014;5(4):294–312.
65. Menten J, Lesaffre E. A general framework for comparative Bayesian meta-analysis of diagnostic studies. BMC Med Res Methodol 2015;15(1):70.
66. Nyaga VN, Aerts M, Arbyn M. ANOVA model for network meta-analysis of diagnostic test accuracy data. Stat Methods Med Res 2018;27(6):1766–84.
67. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. Stat Comput 2000;10(4):325–37.
68. Plummer M. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Proc 3rd Int Workshop Distrib Stat Comput 2003;124(25.10):1–10.
69. Kaltenthaler E, Tappenden P, Paisley S, Squires H. NICE DSU Technical Support Document 13: Identifying and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models [Internet]. Available from: http://nicedsu.org.uk/technical-support-documents/reviewing-model-parameters-tsd/
70. Cancer Research UK. Early detection and diagnosis of cancer: a roadmap to the future [Internet]. 2020. Available from: https://www.cancerresearchuk.org/sites/default/files/early_detection_diagnosis_of_cancer_roadmap.pdf
71. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. BMJ 2013;346:f1049.
72. Singal AG, Volk ML, Rakoski MO, Fu S, Su GL, McCurdy H, et al. Patient Involvement in Healthcare is Associated With Higher Rates of Surveillance for Hepatocellular Carcinoma. J Clin Gastroenterol 2011;45(8):727–32.
73. Coulter A, Stilwell D, Kryworuchko J, Mullen PD, Ng CJ, van der Weijden T. A systematic development process for patient decision aids. BMC Med Inform Decis Mak 2013;13(2):S2.
74. Witteman HO, Maki KG, Vaisson G, Finderup J, Lewis KB, Dahl Steffensen K, et al. Systematic Development of Patient Decision Aids: An Update from the IPDAS Collaboration. Med Decis Making 2021;0272989X211014163.
75. Trevena LJ, Zikmund-Fisher BJ, Edwards A, Gaissmaier W, Galesic M, Han PK, et al. Presenting quantitative information about decision outcomes: a risk communication primer for patient decision aid developers. BMC Med Inform Decis Mak 2013;13(2):S7.
76. Biomarkers in Nottingham Compensated Cirrhosis Cohort (3CN) [Internet]. Available from: https://nddcbru.org.uk/study/biomarkers-in-nottingham-compensated-cirrhosis-cohort-3cn
77. National Institute for Health and Care Excellence (NICE). Hepatitis B and C testing: people at risk of infection [PH43] [Internet]. 2012 [cited 2021 Aug 19];Available from: https://www.nice.org.uk/guidance/ph43
78. Public Health England. Liver disease profiles: statistical commentary [Internet]. GOV.UK2018 [cited 2021 Aug 19];Available from: https://www.gov.uk/government/statistics/liver-disease-profiles-april-2018-update/liver-disease-profiles-statistical-commentary-april-2018
79. Sepucha KR, Borkhoff CM, Lally J, Levin CA, Matlock DD, Ng CJ, et al. Establishing the effectiveness of patient decision aids: key constructs and measurement instruments. BMC Med Inform Decis Mak 2013;13(2):S12.
80. Cross TJS, Villaneuva A, Shetty S, Wilkes E, Collins P, Adair A, et al. A national survey of the provision of ultrasound surveillance for the detection of hepatocellular carcinoma. Frontline Gastroenterol 2016;7(2):82–9.
81. Rogers G, Elston J, Garside R, Roome C, Taylor R, Younger P, et al. The harmful health effects of recreational ecstasy: a systematic review of observational evidence. Health Technol Assess 2009;13(6):1–354.
82. Rogers G, Hoyle M, Thompson Coon J, Moxham T, Liu Z, Pitt M, et al. Dasatinib and nilotinib for imatinib-resistant or -intolerant chronic myeloid leukaemia: a systematic review and economic evaluation. Health Technol Assess [Internet] 2012 [cited 2020 Oct 4];16(22). Available from: https://www.journalslibrary.nihr.ac.uk/hta/hta16220/