

Evidence-based Practice Center Systematic Review Protocol

Project Title: Imaging Techniques for the Surveillance, Diagnosis, and Staging of Hepatocellular Carcinoma

I. Background and Objectives for the Systematic Review

Hepatocellular carcinoma (HCC) is the most common primary malignant neoplasm of the liver, usually developing in the setting of chronic liver disease and cirrhosis. Worldwide, it is the fifth most common cause of cancer and the third most common cause of cancer death.¹ According to the National Cancer Institute, there were 156,940 deaths attributed to liver and intrahepatic bile duct cancer in the United States in 2011, with 221,130 new cases diagnosed.² The National Cancer Institute's Surveillance, Epidemiology, and End Results Cancer Statistics Review found that the lifetime risk of developing liver and intrahepatic bile duct cancer is 1 in 132, with the age-adjusted incidence rate being 7.3 per 100,000 people per year.³ The highest incidence rates are found in Asian/Pacific Islanders (22.1 per 100,000 men and 8.4 per 100,000 women). The age-adjusted death rate is estimated at 5.2 per 100,000 people per year in the United States, with the highest sex-specific rates among Asian/Pacific Islander men (14.7 per 100,000) and American Indian/Alaskan Native women (6.6 per 100,000). The overall 5-year relative survival rate is 14.4 percent.

The 2011 Annual Report to the Nation on the Status of Cancer reported that for the 5-year and 10-year periods analyzed, deaths from liver cancer have significantly increased from 1998–2007 and from 2003–2007 in both men and women.⁴ These new cases are mostly traceable to cirrhosis caused by either hepatitis B virus (HBV) or hepatitis C virus (HCV) infection or long-term alcohol abuse.⁵ The American Association for the Study of Liver Diseases (AASLD) has identified the following groups as being at high risk for developing HCC and recommends that these population groups undergo surveillance: Asian male HBV carriers over age 40, Asian female HBV carriers over age 50, HBV carriers with a family history of HCC, African/North American black HBV carriers, HBV or HCV carriers with cirrhosis, all individuals with other causes for cirrhosis (including alcoholic cirrhosis), and patients with stage 4 primary biliary cirrhosis.⁶

HCC is an aggressive tumor associated with poor survival but, when diagnosed early, may be amenable to potentially curative treatments. The three phases of pretherapy management of HCC include surveillance, diagnosis, and staging. Surveillance is the use of periodic testing to monitor lesions in the liver that give rise to a clinical suspicion of HCC. The diagnosis phase involves the use of additional tests (radiological and/or histopathological) to confirm that the lesion detected in the liver is HCC. Staging of HCC is based on the size and number of lesions and helps to determine appropriate treatments. Other factors that influence treatment decisions include comorbidities and general health status.

The objectives of imaging during each of these three phases are different. In surveillance, the objective is early detection, and the use of imaging techniques for surveillance has been proposed as a means of identifying HCC at earlier stages in high-risk patients, such as those with cirrhosis.⁶ In diagnosis, the objective is to confirm the diagnosis when faced with a clinically suspicious lesion. In staging, the objective is to provide information to make decisions about and to initiate early and appropriate treatment. A number of imaging techniques are available to identify the presence of lesions, diagnose HCC, and determine the stage of the disease (Table 1). Understanding how different imaging strategies affect clinical decisionmaking and ultimately patient outcomes is challenging: imaging techniques may be used alone, in various combinations or algorithms, and/or with liver-specific biomarkers; the nature of imaging techniques is evolving; trade-offs in diagnostic accuracy (e.g., sensitivity and specificity) may be unclear; potential harms may be difficult to uncover; and limitations in the evidence may result because of sparse data on patient outcomes.

Surveillance strategies for HCC use available imaging techniques alone or in a particular sequence. For example, some centers use ultrasound (US) alternatively with either computed tomography (CT) or magnetic resonance imaging (MRI) every 6 months. Some of these strategies also make use of variations that represent evolving technologies, such as the use of liver-specific MRI contrast agents such as gadolinium ethoxybenzyl diethylenetriamine pentaacetic acid (Gd-EOB-DTPA) and superparamagnetic iron oxide (SPIO), dual energy CT, and newer fluorodeoxyglucose positron emission tomography (FDG-PET) tracers such as ¹⁸F-fluorothymidine (FLT), ¹¹C-choline, and ¹¹C-methionine.

In addition to imaging tests, biomarkers for HCC have also been used in surveillance and diagnosis. Alpha-fetoprotein is the most widely used serological marker for HCC, but it is recommended only as an adjunct to imaging.⁶ Another HCC biomarker is des-gamma-carboxy prothrombin; however, this marker has been evaluated in patients with late-stage HCC only, and its role in relation to imaging for the surveillance and early diagnosis of HCC is unknown. Other biomarkers include glypican 3, heat shock protein 70, and glutamine synthetase. The use of glypican 3, heat shock protein 70, and glutamine synthase have not been validated in the clinical setting.

There is clinical uncertainty about which imaging technique to use to diagnose and stage HCC. It is possible to confirm the diagnosis with the availability of either a combination of tests or specific sequences of tests; however, the test performance of these combinations and sequences against a single test should be evaluated before employing them in regular clinical practice. In addition, the use of different reference standards—such as those for explanted liver specimens in patients undergoing transplantation, percutaneous, or surgical biopsy, for imaging, and for clinical followup—could introduce heterogeneity and result in some misclassification due to sampling error, inadequate specimens, insufficient follow-up, or other factors. Therefore, potential variation in test performance with different reference standards also should be examined.

Finally, other factors, including risk factors for HCC and disease characteristics like etiology, tumor size, and level of liver dysfunction may impact the diagnostic accuracy or clinical utility of various imaging strategies. We propose to conduct a comprehensive review of the comparative effectiveness of different imaging strategies for HCC that addresses all of these issues in order to better inform patient and provider/clinician decisions.

Table 1. Imaging techniques used in the surveillance, diagnosis, and staging of hepatocellular carcinoma

Imaging Modality	Key Characteristics	Surveillance	Diagnosis	Staging
Transabdominal Ultrasound	This modality uses ultrasound waves and their reflection from tissue interfaces to generate images of the underlying anatomy. Conventional ultrasound has limited lesion characterization. Ultrasound characterization of a liver mass can be improved by using intravenous (IV) contrast agents (microbubbles).	●	● (IV contrast only)	
Spiral Computed Tomography (CT)	This cross-sectional imaging modality is based on x-ray exposure and acquisition of data through a set of detectors arrayed in a linear fashion. Spiral CT continuously scans the anatomy, acquiring a volume of information to generate images in multiple planes. Contrast-enhanced CT images are obtained after injecting iodinated IV contrast media. Multiple passes are performed at specific timing in order to perform a multiphase contrasted study, which is the most appropriate for diagnosing and staging hepatocellular cancer.	●	●	●
Multidetector CT (MDCT)	MDCT scanners are based on the same imaging principles as spiral CT devices but acquire data very quickly by utilizing a two-dimensional array of detectors which increases the speed of image acquisition. MDCT permits faster scanning, which decreases motion artifacts and thereby improves image quality. MDCT scanners provide high-resolution anatomic information in any plane.	●	●	●
Dual Energy CT	This modality uses x-rays of varying energy (70–140 kVp) to increase tissue contrast and detect different elements (e.g., iodine, calcium) within the liver. There are single-source (conventional) CT scanners that can obtain dual energy studies and dual source-dual energy CT scanners that have two x-ray sources of different intensity and two corresponding detector sets that permit fast and efficient dual energy studies.	●	●	●
Magnetic Resonance Imaging (MRI)	This imaging technique uses a strong magnetic field and radiofrequency pulses to obtain anatomic images of the body. MRI scanning is slower than CT scanning and requires that the patient remain still during image acquisition. Contrast-enhanced multiphase MRI of the liver provides accurate anatomic information and excellent lesion characterization. In addition, MRI can assess tissues for iron load, fat content, diffusion characteristics, and edema. Different contrast media can be used, such as gadolinium ion and iron oxide.	●	●	●
FDG-Positron Emission Tomography	This functional imaging technique uses radioisotope-tagged tracers to examine the level and type of biochemical activity in lesions suspected to be cancerous throughout the body (making it useful to study metastases). The most commonly used tracer is fluorodeoxyglucose (¹⁸ F-FDG), which detects cells exhibiting increased glucose transport and metabolism (cancer cells exhibit such metabolic activity). Alternative tracers have been investigated for liver cancer.			●

II. The Key Questions

The proposed Key Questions (KQs) for this report were posted on the Agency for Healthcare Research and Quality (AHRQ) Effective Health Care Program Web site for public comments from September 21 through October 19, 2012. No public comments were received, so no changes were made to the KQs at that time. In June 2013, the final KQs and PICOTS (population, intervention, comparator, outcomes, timing, and setting) were revised before public posting of the protocol based on input from technical experts. The following changes were made: the inclusion of biomarker levels as a potential modifier of test performance; the population for surveillance was revised to include all patients with cirrhosis, including patients with alcoholic cirrhosis; conventional CT was removed as an intervention because it is an outdated modality; 1998 was set as a cutoff date for searches to exclude outdated technologies; and subquestions regarding potential modifiers of test performance were revised to be consistent for all KQs.

Key Question 1

What is the comparative effectiveness of available imaging-based surveillance strategies (listed below under interventions for KQ 1), used singly or in sequence for detecting HCC among individuals undergoing surveillance for HCC (individuals at high risk for HCC and individuals who have undergone liver transplants for HCC)?

- a. What is the comparative test performance of imaging-based surveillance strategies for detecting HCC?
 - i. How is a particular technique's test performance modified by use of various reference standards (e.g., explanted liver samples, histological diagnosis, or clinical and imaging followup)?
 - ii. How is the comparative effectiveness modified by patient-level characteristics (e.g., body mass index, number of lesions, tumor diameter, or cause of liver disease) or other factors (e.g., technical aspects of imaging techniques, biomarker levels, test operator or interpreter skill, setting)?
- b. What is the comparative effectiveness of imaging-based surveillance strategies on intermediate outcomes like diagnostic thinking?
- c. What is the comparative effectiveness of imaging-based surveillance strategies on clinical and patient-centered outcomes?
- d. What are the adverse effects or harms associated with imaging-based surveillance strategies?

Key Question 2

What is the comparative effectiveness of imaging techniques (listed under the interventions for KQ 2), used singly, in combination, or in sequence in diagnosing HCC among individuals in whom an abnormal lesion has been detected while undergoing surveillance for HCC (individuals at high risk for HCC and individuals who have undergone liver transplants for

HCC) or through the evolution of symptoms and abdominal imaging done for other indications?

- a. What is the comparative test performance of imaging techniques for diagnosing HCC?
 - i. How is a particular technique's test performance modified by use of various reference standards (e.g., explanted liver samples, histological diagnosis, or clinical and imaging followup)?
 - ii. How is the comparative effectiveness modified by patient-level characteristics (e.g., body mass index, number of lesions, tumor diameter, or cause of liver disease) or other factors (e.g., technical aspects of imaging techniques, biomarker levels, test operator or interpreter skill, setting)?
- b. What is the comparative effectiveness of the various imaging techniques on intermediate outcomes like diagnostic thinking and use of additional diagnostic procedures such as fine-needle or core biopsy?
- c. What is the comparative effectiveness of the various imaging techniques on clinical and patient-centered outcomes?
- d. What are the adverse effects or harms (related to testing or a test-associated diagnostic workup) associated with the various imaging techniques?

Key Question 3

What is the comparative effectiveness of imaging techniques (listed under the interventions for KQ 3), used singly, in combination, or in sequence in staging HCC among patients diagnosed with HCC?

- a. What is the comparative test performance of imaging techniques to predict HCC tumor stage?
 - i. How is a particular technique's test performance modified by use of various reference standards (e.g., explanted liver samples, histological diagnosis, or clinical and imaging followup)?
 - ii. How is the comparative effectiveness modified by patient-level characteristics (e.g., body mass index, number of lesions, tumor diameter, or cause of liver disease) or other factors (e.g., technical aspects of imaging techniques, biomarker levels, test operator or interpreter skill, setting)?
- b. What is the comparative test performance of imaging techniques on diagnostic thinking?
- c. What is the comparative effectiveness of imaging techniques on clinical and patient-centered outcomes?
- d. What are the adverse effects or harms associated with using imaging techniques related to testing or test-associated diagnostic workup?

PICOTS by Key Question

- **Population(s)**

- Key Question 1
 - Patients at high risk for HCC undergoing surveillance. The population of high-risk patients is defined, as per the AASLD clinical guidelines, as composed of the following: Asian male HBV carriers over age 40, Asian female HBV carriers over age 50, HBV carriers with a family history of HCC, African/North American black HBV carriers, all individuals with cirrhosis (including alcoholic cirrhosis), HBV or HCV carriers with cirrhosis, and patients with stage 4 primary biliary cirrhosis.⁶ Other definitions of high-risk patients as defined by the primary studies will be accepted.
 - Patients who have undergone liver transplants for HCC, either with or without HCC detected in the explanted liver.
 - Both population groups will be considered separately.
- Key Question 2
 - Patients at high risk for HCC in whom a suspicious lesion(s) has been detected by surveillance or by other means.
 - Patients who have undergone liver transplants for HCC, either with or without HCC detected in the explanted liver.
 - Both population groups will be considered separately.
- Key Question 3
 - Patients diagnosed with HCC who require staging before initial treatment.
- All Key Questions
 - Patients with cholangiocarcinoma will be excluded.
- **Interventions**
 - Key Question 1
 - US, spiral CT, multidetector CT (MDCT), dual energy CT, or MRI.
 - Studies that included surveillance strategies of any other imaging test with or without additional biomarkers would also be included. The strategies could include the techniques being used singly or in a specific sequence.
 - Key Question 2
 - Imaging techniques, used singly, in combination, or in a specific sequence, including US, spiral CT, MDCT, dual energy CT, MRI (including contrast agents like Gd-EOB-DTPA and SPIO), or fluorodeoxyglucose positron emission tomography (FDG-PET)

with different tracers (including ^{18}F , fluorothymidine [FLT], ^{11}C -choline, and ^{11}C -methionine, or others).

- Key Question 3

- Imaging techniques, used singly, in combination, or in a specific sequence, including US, spiral CT, MDCT, dual energy CT, MRI with contrast (including contrast agents such as Gd-EOB-DTPA and SPIO), FDG-PET with different tracers (including ^{18}F , FLT, ^{11}C -choline, and ^{11}C -methionine, or others), or contrast CT.
- Test performance of imaging techniques will be stratified by the different staging systems used.

- All Key Questions

- Outdated imaging techniques (e.g., conventional, nonspiral/nonmultidetector CT, or imaging techniques used before 1995) will be excluded.
- Imaging techniques not available or in use in the United States (e.g., hepatic portography) will be excluded.

- **Comparators**

- For studies of diagnostic accuracy (comparative test performance), the reference standard comparators will be histopathology (based on explanted liver specimens or biopsy) or clinical and imaging followup, and the imaging comparators will be alternative imaging tests or strategies.
- For studies of comparative effectiveness, the comparators will be no imaging or alternative imaging strategies.

- **Outcomes for Each Key Question**

- Key Question 1

- Diagnostic outcomes include:
 - Detection rates of HCC lesions.
 - Types of HCC lesions detected.
 - Test performance (e.g., sensitivity and specificity, predictive values, likelihood ratios, area under the receiver operating curve, or others) for diagnosing HCC, including stage-specific accuracy.
 - For all KQs, potential modifiers of measures of test performance will be evaluated, including the reference standards used (e.g., explanted liver samples, histological diagnosis, or clinical and imaging followup), patient and tumor-level characteristics (e.g., body mass index, number of lesions, tumor diameter, or cause of liver disease), or other factors (e.g., technical aspects of the imaging techniques, biomarker levels, test operator or interpreter skill, setting).



- Intermediate outcomes include:
 - Effects on diagnostic thinking.
 - Effects on clinical decisionmaking.

- Clinical and patient-centered outcomes include:
 - Overall mortality or survival.
 - Recurrence of HCC, including rates of seeding by fine-needle aspiration.
 - Quality of life as measured with scales such as the Short-Form Health Survey (SF-36®) or EuroQol 5D (EQ-5™) or as defined by the primary studies.
 - Psychosocial effects of diagnostic testing on patients, patients' caregivers, and other family members, as measured by self-reported questionnaire instruments.
 - Resource utilization and patient burden (e.g., costs associated with the imaging procedure, access to the imaging facility, the number of imaging procedures, and other procedures conducted).

- Key Question 2
 - Diagnostic outcomes include:
 - Type of HCC lesions detected.
 - Test performance (e.g., sensitivity and specificity, predictive values, likelihood ratios, area under the receiver operating curve, or others) for diagnosing HCC. As in KQ 1, potential modifiers of measures of test performance will be evaluated, including the reference standards used (e.g., explanted liver samples, histological diagnosis, or clinical and imaging followup), patient and tumor-level characteristics (e.g., body mass index, number of lesions, tumor diameter, or cause of liver disease), or other factors (e.g., technical aspects of the imaging techniques, biomarker levels, test operator or interpreter skill, setting).

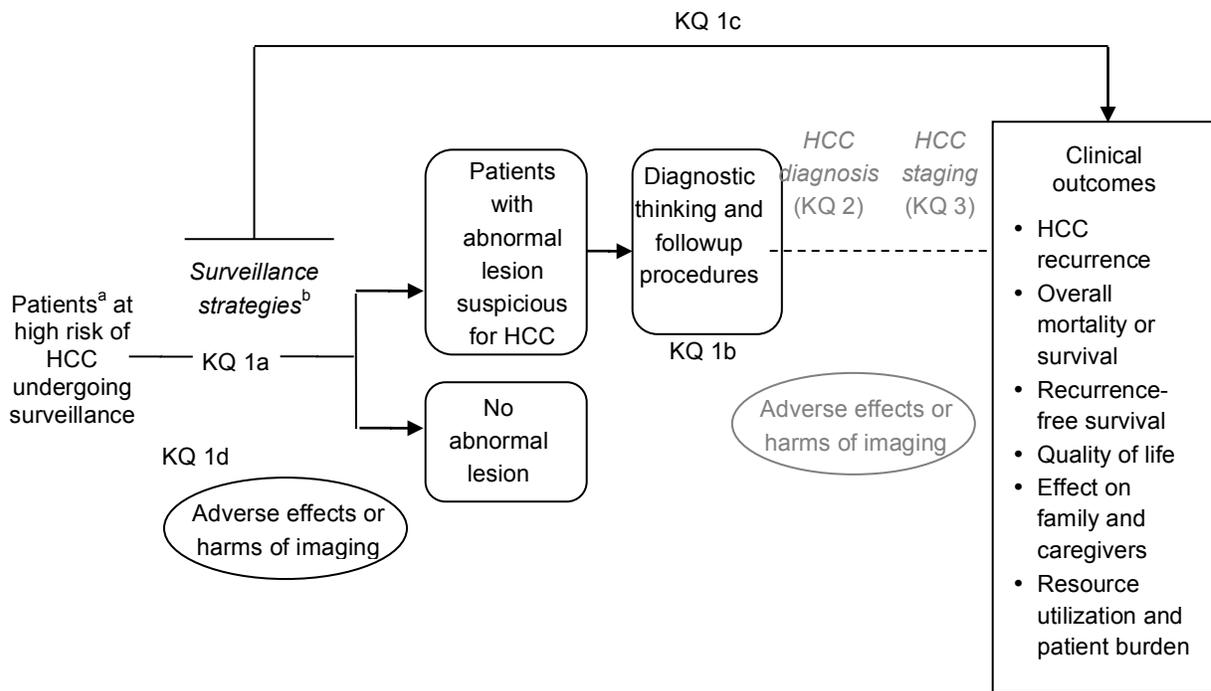
 - Intermediate outcomes include:
 - Effects on diagnostic thinking.
 - Effects on clinical decisionmaking.

 - Clinical and patient centered outcomes include:
 - Overall mortality or survival.
 - Recurrence of HCC, including rates of seeding by fine-needle aspiration
 - Quality of life as measured with scales such as the Short-Form Health Survey (SF-36®) or EuroQol 5D (EQ-5™) or as defined by the primary studies.
 - Psychosocial effects of diagnostic testing on patients, patients' caregivers, and other family members, as measured by self-reported questionnaire instruments.
 - Resource utilization and patient burden (e.g., costs associated with the imaging procedure, access to the imaging facility, the number of imaging procedures and other procedures conducted).

- Key Question 3
 - Diagnostic outcomes include:
 - Measures for stage-specific accuracy of imaging (e.g., Obuchowski method for calculating the area under the receiver operating curve, stage reclassification rates).
 - Intermediate outcomes include:
 - Effects on diagnostic thinking.
 - Effects on clinical decisionmaking.
 - Clinical and patient-centered outcomes include:
 - Overall mortality or survival.
 - Recurrence of HCC, including rates of seeding by fine-needle aspiration
 - Quality of life as measured with scales such as the Short-Form Health Survey (SF-36®) or EuroQol 5D (EQ-5™) or as defined in the primary studies.
 - Psychosocial effects of diagnostic testing on patients, patients' caregivers, and other family members as measured by self-reported questionnaire instruments.
 - Resource utilization and patient burden (e.g., costs associated with the imaging procedure, access to the imaging facility, the number of imaging procedures and additional procedures conducted).
- Key Questions 1d, 2d, and 3d (Adverse Events or Harms)
 - Adverse effects or harms associated with the imaging techniques (e.g., test-related anxiety, adverse events secondary to venipuncture, contrast allergy, exposure to radiation).
 - Adverse effects or harms associated with test-associated diagnostic workup (e.g., harms of biopsy or harms associated with workup of other incidental tumors discovered on imaging).
- **Timing**
 - No restrictions will be placed on timing.
 - For studies of comparative effectiveness, duration of followup, timing of interventions, and frequency of interventions will be recorded.
- **Settings:**
 - All relevant care settings (e.g., primary and secondary care).

III. Analytic Framework

Figure 1. Analytic framework—surveillance

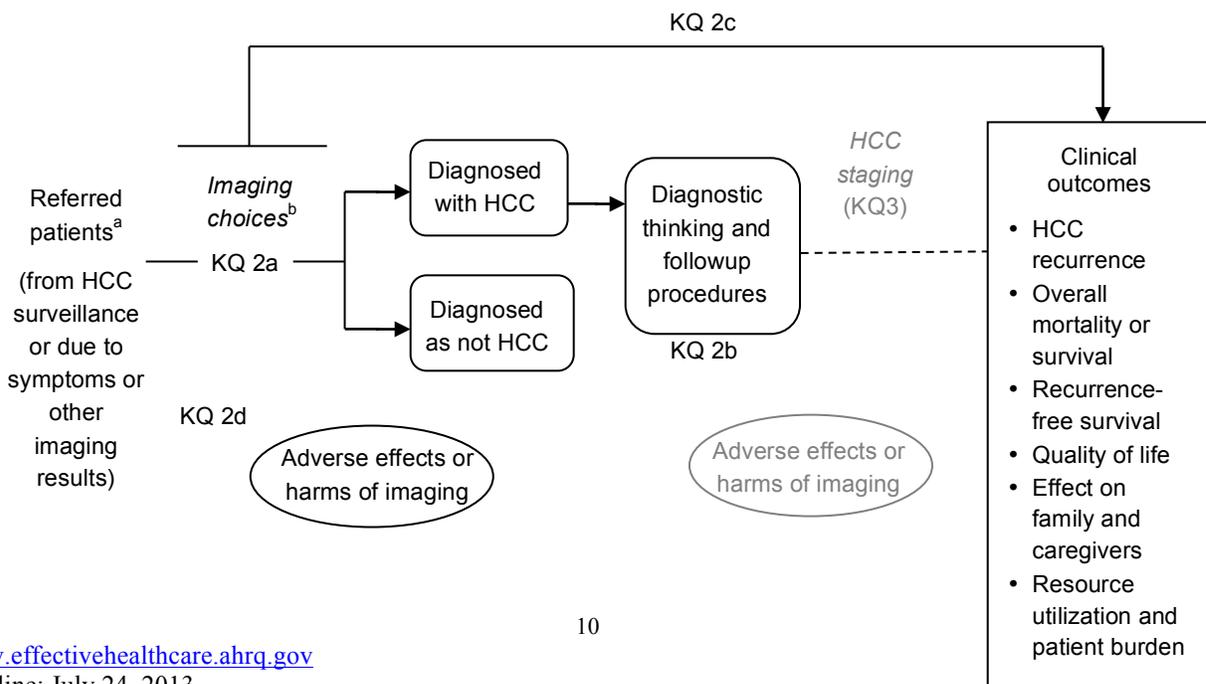


^a Patient-level characteristics (modifying factors) include body mass index, number of lesions, tumor diameter, and cause of liver disease.

^b Imaging techniques are used singly, in combination, or in sequence with or without biomarkers used as modifiers; modifying factors include the technical aspects of the imaging techniques, the skills of the test operator or interpreter, and setting.

HCC = hepatocellular carcinoma; KQ = key question

Figure 2. Analytic framework—diagnosis

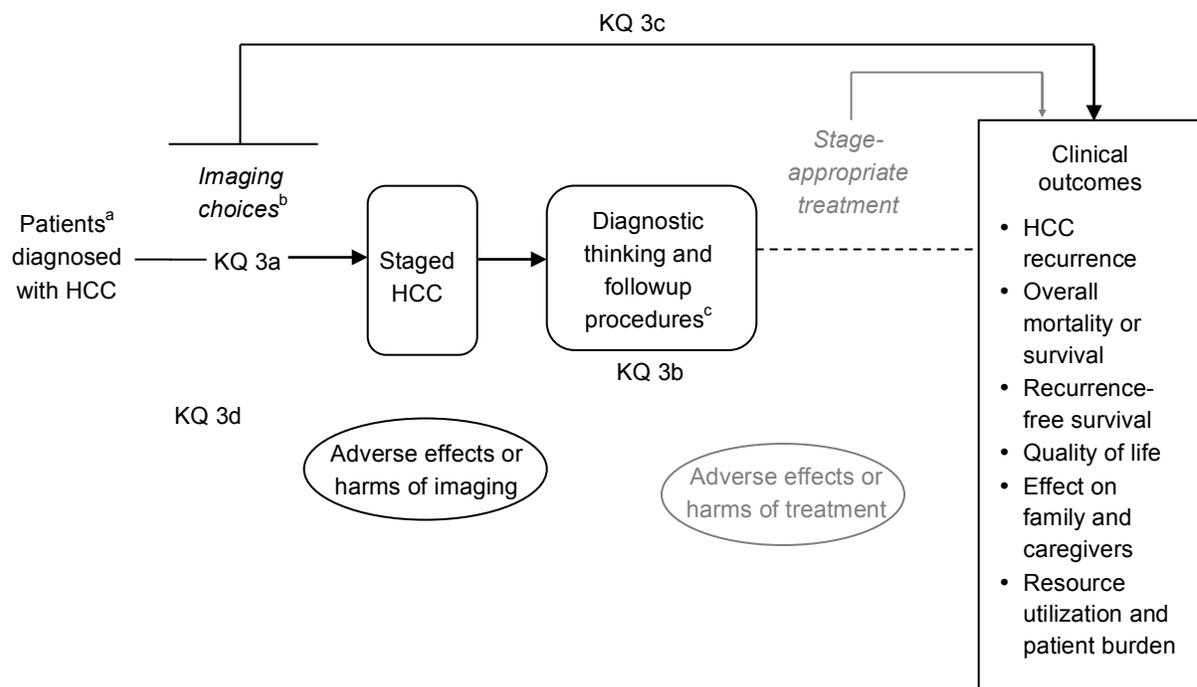


^a Patient-level characteristics (modifying factors) include body mass index, number of lesions, tumor diameter, and cause of liver disease.

^b Imaging techniques are used singly, in combination, or in sequence with or without biomarkers used as modifiers; modifying factors include the technical aspects of the imaging techniques, the skills of the test operator or interpreter, and setting.

HCC = hepatocellular carcinoma; KQ = key question

Figure 3. Analytic framework—staging



^a Patient-level characteristics (modifying factors) include body mass index, number of lesions, tumor diameter, and cause of liver disease.

^b Imaging techniques are used singly, in combination, or in sequence with or without biomarkers used as modifiers; modifying factors include the technical aspects of the imaging techniques, the skills of the test operator or interpreter, and setting.

^c Followup procedures include biopsy.

HCC = hepatocellular carcinoma; KQ = key question

IV. Methods

We will perform the systematic review in accordance with the Evidence-based Practice Center (EPC) methods guides.⁷⁻⁹

A. Criteria for Inclusion/Exclusion of Studies in the Review

The criteria for inclusion and exclusion of studies will be based on the KQs and are described in the PICOTS section above. Only English-language studies will be included, because some

more invasive and costly imaging techniques employed in other countries are not representative of current practice in the United States.

Study Designs

The following study designs will be included:

- Controlled or comparative randomized and nonrandomized trials and controlled or comparative observational studies.
- Studies of diagnostic accuracy

Systematic reviews will be used as primary sources of evidence if they address a KQ and are assessed as being at low risk of bias (as defined below in part D).

Case reports, case series, letters to the editor, and nonsystematic reviews will be excluded.

Sample Size

Studies with very small numbers of participants ($n < 20$) will not be included.

Publication Date Range

Because of changes in imaging technologies, searches will be limited by a start date of 1998.

B. Searching for the Evidence: Literature Search Strategies for Identification of Relevant Studies To Answer the Key Questions

For the primary literature, we will search Ovid MEDLINE[®], Scopus, Evidence-Based Medicine Reviews (Ovid), the Cochrane Central Register of Controlled Trials, the Cochrane Database of Systematic Reviews, the Database of Abstracts of Reviews of Effects, and the Health Technology Assessment Database from 1998 to 2013. Gray literature will be sought by soliciting input from Technical Expert Panel (TEP) members who represent various stakeholder perspectives and by searching relevant Web sites including clinical trial registries (ClinicalTrials.gov, Current Controlled Trials, ClinicalStudyResults.org, and the WHO International Clinical Trials Registry Platform), regulatory documents (Devices@FDA), and individual product Web sites. See Table 2 for a sample of the proposed search strategy. Additional studies will be identified by reviewing the reference lists of published clinical trial and review articles that our TEP suggested.

The criteria for inclusion and exclusion of studies are based on the KQs and the PICOTS approach. We will use the inclusion criteria described in Appendix A. Studies will be reviewed at both the abstract and full-text level by two reviewers to ensure accuracy of the study selection. Full-text articles will be included when consensus occurs between the reviewers. If consensus is not reached about an article by the two initial reviewers, a senior investigator will review the article and adjudicate the decision with regard to inclusion or exclusion.

Scientific information packets (SIPs) will be requested via a notice published in the *Federal Register*; SIPs will be not be requested from specific manufacturers given the widespread availability of imaging technologies from many manufacturers. Library searches will be updated while the draft report is posted for public comment and peer review to capture any new publications. Literature identified through the updated search will be assessed by using the same process of dual review as all other studies considered for inclusion in the evidence report. If any pertinent new literature is identified for inclusion in the report, it will be incorporated before the final submission of the report.

Table 2. Sample of proposed search strategy for Hepatocellular Carcinoma Imaging – Ovid MEDLINE® (1998–2013)

1. Carcinoma, Hepatocellular/
2. Liver Neoplasms/
3. ("hepatocellular cancer" or "hepatocellular carcinoma" or "HCC").ti,ab.
4. Diagnostic Imaging/
5. Ultrasonography/
6. Magnetic Resonance Imaging/
7. exp Tomography, Emission-Computed/ or exp Positron-Emission Tomography/ or exp Tomography, Spiral Computed/
8. ("CT" or "dynamic multidetector computed tomography" or "MDCT" or "spiral CT" or "dual source CT" or "contrast CT" or "MRI" or "FDG-PET").ti,ab.
9. or/1-3
10. or/4-8
11. 9 and 10
12. "Sensitivity and Specificity"/
13. "Predictive Value of Tests"/
14. ROC Curve/
15. "Reproducibility of Results"/
16. (sensitiv\$ or "predictive value" or accurac\$).ti,ab.
17. or/12-16
18. 11 and 17
19. limit 18 to yr="1998 - 2013"

C. Data Abstraction and Data Management

After studies are selected for inclusion, data will be abstracted by one researcher; second researchers will independently check the abstraction for accuracy. Data on the following categories will be abstracted, including but not limited to: study design, year, setting, country, sample size, eligibility criteria, population and clinical characteristics, intervention characteristics, and results relevant to each KQ as outlined in the PICOTS section above. Additional information on lesion size, stages, time limits, and the reference standard used will be collected when reported. Information about ablative treatments received between a diagnostic test and the reference standard will also be extracted, as this could affect measures of diagnostic

accuracy. Data extraction forms will include criteria specific to describing the imaging technology used, whether the technology is currently in use or not, and the technical aspects of each technology (such as row number, phase, contrast rate, slice, size for MDCT; scanner type, phases, contrast, section thickness, spatial resolution, and acquisition time for MRI; whether the operator was a technician, radiologist, or gastroenterologist; and if contrast was used for studies of US). If available, other information may be abstracted, such as the number of patients randomized relative to the number of patients enrolled and how similar those patients are to the target population. All study data will be verified for accuracy and completeness by a second team member.

A record of studies excluded at the full-text level, including the primary reason for exclusion, will be maintained.

D. Assessment of Methodological Risk of Bias (Quality) of Individual Studies

The quality of individual controlled trials, systematic reviews, and observational studies will be assessed with clearly defined templates and criteria, as predefined by the Effective Health Care Program.⁹ Randomized trials and cohort studies will be evaluated with criteria and methods developed by the U.S. Preventive Services Task Force.¹⁰ These criteria will be used in conjunction with the approach recommended in the chapter “Assessing the Risk of Bias of Individual Studies When Comparing Medical Interventions” in the AHRQ *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*.¹¹ Studies of diagnostic test performance will be assessed using the approach recommended in the AHRQ *Methods Guide for Medical Test Reviews*, which is based on QUADAS methods.^{7,12}

Individual studies will be rated as having “low,” “medium,” or “high” risk of bias. Studies rated “low” are considered to have the least risk of bias, and their results will be considered valid. Randomized trials and cohort studies assessed as having low risk of bias include clear descriptions of the population, setting, interventions, and comparison groups; a valid method for allocation of patients to treatment (for randomized trials); low dropout rates and clear reporting of dropouts; appropriate means for preventing bias; appropriate measurement of and analysis of confounders (for cohort studies), and appropriate measurement of outcomes. Studies of diagnostic test performance that are assessed as having low risk of bias use a reliable reference standard, apply the reference standard to all patients, use blinded interpretation of the diagnostic test and the reference standard, and use preset criteria to define a positive test.

Studies rated as having “medium” risk are susceptible to some bias, though not enough to invalidate the results. These studies may not meet all the criteria for a rating of low risk of bias but no flaw is likely to cause major bias. The study may be missing information, making it difficult to assess limitations and potential problems. The “medium” risk of bias category is broad, and studies with this rating will vary in their strengths and weaknesses; the results of some studies assessed to have medium risk of bias are likely to be valid, while others may be only possibly valid.

Studies rated as having “high” risk of bias have significant flaws that imply biases of various types that may invalidate the results. They have a serious or “fatal” flaw in design, analysis, or reporting; large amounts of missing information; discrepancies in reporting; or serious problems in the delivery of the intervention. The results of these studies are at least as likely to reflect

flaws in the study design as the true difference between the compared interventions. We will not exclude studies rated as having high risk of bias a priori, but these studies will be considered to be less reliable than studies rated as having low risk of bias when synthesizing the evidence, particularly if discrepancies between studies are present.

Each study evaluated will be dual reviewed for risk of bias by two team members. Any disagreements will be resolved by consensus.

E. Data Synthesis

We will construct evidence tables identifying the study characteristics (as described in the PICOTS), results pertinent to the KQs, and quality ratings for all included studies. We expect important clinical heterogeneity in the studies (e.g., variability in the reference standard, specific imaging techniques, geographical setting, patient characteristics, and operator experience) and will design abstraction tools to address these factors. We will review studies using a hierarchy-of-evidence approach, where the best evidence is the focus of our synthesis for each KQ. We will prioritize studies that directly compare outcomes for two or more imaging tests.

Meta-analyses of measures of diagnostic outcomes (including diagnostic accuracy) and clinical and patient-centered outcomes will be conducted when feasible to summarize data and obtain more precise estimates. Depending on the degree of clinical heterogeneity, pooling studies may be inadvisable.¹³ We will pool only studies that are clinically comparable and could provide a meaningful combined estimate and will perform sensitivity and subgroup analysis if statistical heterogeneity is present. In order to determine whether meta-analysis could be meaningfully performed, we will consider the quality of the studies; the heterogeneity among studies in design, patient population, interventions, and outcomes; and magnitude of effect size. We will conduct sensitivity analyses or meta-regression as needed. When quantitative analysis cannot be performed, the data will be summarized qualitatively in summary tables and descriptive text.

A random-effects model will be used to combine the different outcomes, except for a rare binary outcome in a comparative situation where a fixed-effects model will be used. When the between-study heterogeneity is estimated to be zero, the random-effects model produces the same results as the fixed-effects model. Measures for diagnostic accuracy and clinical and patient-centered outcomes often entail different statistical techniques in meta-analysis, and a statistician expert in quantitative synthesis will determine specific meta-analytic methods appropriate to the data characteristics. Tests of heterogeneity will be conducted using the standard χ^2 tests and I^2 statistic, when appropriate, or other measures based on the choice of specific meta-analytic methods. As outlined in the subsets of each KQ and in the PICOTS, preidentified subgroups related to setting, patient characteristics, technical aspects, and study quality will be explored to explain potential heterogeneity in effects.

F. Grading the Strength of Evidence for Individual Comparisons and Outcomes

The strength of evidence for each KQ will be assessed by one researcher for each outcome described in the PICOTS using the approach described in the AHRQ *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*.⁹ To ensure consistency and validity of the

evaluation, each grade will be reviewed by the entire team of investigators. The following categories will guide this review:

- Risk of bias (low, medium, or high)
- Consistency (consistent, inconsistent, or unknown/not applicable)
- Directness (direct or indirect)
- Precision (precise or imprecise)

We will synthesize the overall quality of each body of evidence, based on the factors above. We will also estimate publication bias by examining whether studies with smaller sample sizes tended to have positive or negative assessments outcomes.

The strength of evidence will be assigned an overall grade of high, moderate, low, or insufficient according to a four-level scale:

- **High.** High confidence that the evidence reflects the true effect. Further research is very unlikely to change our confidence in the estimate of effect.
- **Moderate.** Moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of effect and may change the estimate.
- **Low.** Low confidence that the evidence reflects the true effect. Further research is likely to change the confidence in the estimate of effect and is likely to change the estimate.
- **Insufficient.** Evidence either is unavailable or does not permit estimation of effect.

G. Assessing Applicability

Applicability will be estimated by examining:

- Characteristics of the patient populations (e.g., type and severity of the underlying disease/cause of HCC, body mass index, tumor characteristics such as size, treatment received between the test and the reference standard, age, sex, race, comorbid conditions)
- Sample size of the studies
- Settings in which the studies are performed (e.g., use of different reference standards in different settings such as academic centers and community care facilities)
- Countries (e.g., patients in developing countries)
- Characteristics of the provider (e.g., a technician, radiologist, or gastroenterologist may have conducted the test; skill level of the operator, interpreter, or pathologist)

On the advice of the TEP, to ensure the review will be applicable to current practice in the United States, the timeframe will be limited to studies with a publication year of 1998 or later and to studies that report on imaging technologies in current use in the United States. Variability in the studies may limit the ability to generalize the results to other populations and settings.

V. References

1. Parkin DM. Global cancer statistics in the year 2000. *Lancet Oncol.* 2001 Sep;2(9):533-43. PMID: 11905707.
2. National Cancer Institute. Liver Cancer. <http://www.cancer.gov/cancertopics/types/liver>. Accessed September 21, 2011.
3. Howlader N, Noone A, Neyman N, et al. SEER Cancer Statistics Review, 1975-2010. Bethesda, MD: National Cancer Institute; 2013. Available at http://seer.cancer.gov/csr/1975_2010/.
4. Kohler BA, Ward E, McCarthy BJ, et al. Annual report to the nation on the status of cancer 1975–2007, featuring tumors of the brain and other nervous system. *J Natl Cancer Inst.* 2011 May 4;103(9):1-23. PMID: 21454908.
5. McGlynn KA, London WT. Epidemiology and natural history of hepatocellular carcinoma. *Best Pract Res Clin Gastroenterol.* 2005 Feb;19(1):3-23. PMID: 15757802.
6. Bruix J, Sherman M; American Association for the Study of Liver Diseases. Management of hepatocellular carcinoma: an update. *Hepatology.* 2011 Mar;53(3):1020-2. PMID: 21374666.
7. Chang SM, Matchar DB, eds. *Methods Guide for Medical Test Reviews*. AHRQ Publication No. 12-EHC017-EF. Rockville, MD; Agency for Healthcare Research and Quality; June 2012. Available at http://effectivehealthcare.ahrq.gov/ehc/products/246/558/Methods-Guide-for-Medical-Test-Reviews_Full-Guide_20120530.pdf.
8. Matchar DB. Introduction to the *Methods Guide for Medical Test Reviews*. In: Chang SM, Matchar DB, eds. *Methods Guide for Medical Test Reviews*. AHRQ Publication No. 12-EHC017-EF. Rockville, MD; Agency for Healthcare Research and Quality; 2012: chapter 1. Available at http://effectivehealthcare.ahrq.gov/ehc/products/246/558/Methods-Guide-for-Medical-Test-Reviews_Full-Guide_20120530.pdf.9.
9. *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. AHRQ Publication No. 10(12)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; April 2012. Available at http://www.effectivehealthcare.ahrq.gov/ehc/products/60/318/MethodsGuide_Prepublication-Draft_20120523.pdf.
10. 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 Apr;20(3 Suppl):21-35. PMID: 11306229.

11. Viswanathan M, Ansari MT, Berkman ND, et al. Assessing the risk of bias of individual studies in systematic reviews of health care interventions. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 12-EHC047-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2012: chapter 5. Available at http://www.effectivehealthcare.ahrq.gov/ehc/products/60/318/MethodsGuide_Prepublication-Draft_20120523.pdf.
12. Whiting PF, Rutjes AW, Westwood ME, et al; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011 Oct 18;155(8):529-36. PMID: 22007046.
13. Fu R, Gartlehner G, Grant M, et al. Conducting quantitative synthesis when comparing medical interventions. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(12)-EHC063-EF Rockville, MD: Agency for Healthcare Research and Quality; 2012: chapter 9. Available at http://www.effectivehealthcare.ahrq.gov/ehc/products/60/318/MethodsGuide_Prepublication-Draft_20120523.pdf.

VI. Definition of Terms

Not applicable.

VII. Summary of Protocol Amendments

In the event of protocol amendments, the date of each amendment will be accompanied by a description of the change and the rationale.

VIII. Review of Key Questions

For all Evidence-based Practice Center (EPC) reviews, Key Questions were reviewed and refined as needed by the EPC with input from Key Informants and the Technical Expert Panel (TEP) to assure that the questions are specific and explicit about what information is being reviewed. In addition, the Key Questions were posted for public comment and finalized by the EPC after review of the comments.

IX. Key Informants

Key Informants are the end-users of research, including patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others with experience in making health care decisions. Within the EPC Program, the Key Informant role is to provide input into identifying the Key Questions for research that will inform health care decisions. The EPC solicits input from Key Informants when developing questions for systematic review or when identifying high-priority research gaps and needed new research. Key Informants are not involved in analyzing the evidence or writing the report and have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Key Informants must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals are invited to serve as Key Informants and those who present with potential conflicts may be retained. The Task Order Officer (TOO) and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

X. Technical Experts

Technical Experts comprise a multidisciplinary group of clinical, content, and methodological experts who provide input in defining populations, interventions, comparisons, or outcomes, as well as identifying particular studies or databases to search. They are selected to provide broad expertise and perspectives specific to the topic under development. Divergent and conflicted opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore study questions, design, and/or methodological approaches do not necessarily represent the views of individual technical and content experts. Technical Experts provide information to the EPC to identify literature search strategies and recommend approaches to specific issues as requested by the EPC. Technical Experts do not do analysis of any kind nor contribute to the writing of the report and have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Technical Experts must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Technical Experts and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

XI. Peer Reviewers

Peer Reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodological expertise. Peer review comments on the preliminary draft of the report are considered by the EPC in preparation of the final draft of the report. Peer Reviewers do not participate in writing or editing of the final report or other products. The synthesis of the scientific literature presented in the final report does not necessarily represent the views of individual reviewers. The dispositions of the peer review comments are documented

and will, for Comparative Effectiveness Reviews and Technical Briefs, be published 3 months after the publication of the Evidence Report.

Potential reviewers must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Invited Peer Reviewers may not have any financial conflict of interest greater than \$10,000. Peer Reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through the public comment mechanism.

XII. EPC Team Disclosures

EPC core team members must disclose any financial conflicts of interest greater than \$1,000 and any other relevant business or professional conflicts of interest. Related financial conflicts of interest that cumulatively total greater than \$1,000 will usually disqualify EPC core team investigators.

XIII. Role of the Funder

This project was funded under Contract No. HHSA 290-2012-00014-I from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The Task Order Officer reviewed contract deliverables for adherence to contract requirements and quality. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

Appendix A. Inclusion/Exclusion Codes and Criteria

Full-text Review Codes

Include:

1. = Include 1SR = Systematic Review, not directly used, but all studies checked for inclusion

Exclude:

2. = Wrong population
3. = Wrong intervention
4. = Wrong comparator (for studies of clinical efficacy) or wrong reference standard (for studies of diagnostic accuracy)
5. = Wrong outcome (for studies of clinical efficacy) or does not report measures of diagnostic accuracy
6. = Wrong setting
7. = Wrong study design for Key Question
8. = Wrong publication type (review article, editorial, results reported elsewhere, no original data, case report)
9. = Not English language but otherwise relevant
10. = Wrong duration of followup
11. = Sample size too small
12. = Not human population, animal study
13. = Inadequate reference standard (nonpathologically based reference standard)

Table 1. Inclusion / Exclusion Criteria for Key Question 1

Category	Inclusion Criteria	Exclusion Criteria
Population and Patient Characteristics	Patients at high risk for hepatocellular carcinoma (HCC) undergoing surveillance. The population of high risk patients is defined, as per the American Association for the Study of Liver Diseases (AASLD) clinical guidelines, as composed of the following: Asian male hepatitis B virus (HBV) carriers over age 40; Asian female HBV carriers over age 50; HBV carriers with family history of HCC; African/North American black HBV carriers; all individuals with cirrhosis, including alcoholic cirrhosis, cirrhotic HBV or HCV carriers; and patients with stage 4 primary biliary cirrhosis. Other definitions of high risk patients as defined by the primary studies will be accepted. Patients who have undergone liver transplants for HCC, either with or without HCC detected in the explant liver.	Children and adolescents, patients with cholangiocarcinoma
Interventions and Comparators	<p>Ultrasound (U/S), computed tomography (CT), including spiral CT, multidetector CT (MDCT), or dual source CT, or magnetic resonance imaging (MRI). Studies that included surveillance strategies of any other imaging test with or without additional biomarkers would also be included. The strategies could include the techniques being used singly, or in a specific sequence.</p> <p>Comparators:</p> <ul style="list-style-type: none"> • For studies of diagnostic accuracy (comparative test performance), the reference standard comparators will be histopathology (based on explanted liver specimens or biopsy) or clinical and imaging followup, and the imaging comparators will be alternative imaging tests or strategies. 	CT arteriography, CT portography and interventions used in treatment (imaging guided radiofrequency ablation), outdated imaging techniques (e.g., conventional, non-spiral/non-multidetector CT or imaging performed prior to 1995) and imaging techniques not available or in use in the U.S. (e.g., hepatic

Category	Inclusion Criteria	Exclusion Criteria
	<ul style="list-style-type: none"> For studies of comparative effectiveness, the comparators will be no imaging or alternative imaging strategies. 	portography).
Outcomes	<p>Diagnostic outcomes include:</p> <ul style="list-style-type: none"> Detection rates of tumors identified by surveillance. Type of lesions detected—the detected tumors undergo followup imaging or pathological diagnosis; the comparative frequency of various tumors that have been discovered by different imaging techniques is of interest. Test performance (e.g., sensitivity and specificity) of surveillance in identifying HCC that is eventually diagnosed by pathological diagnosis or clinical and imaging followup as the reference standard. For all key questions, potential modifiers of measures of test performance will be evaluated, including the reference standards used (e.g., explanted liver samples, histological diagnosis, or clinical and imaging follow-up), patient and tumor-level characteristics (e.g. body mass index, number of lesions, tumor diameter, or cause of liver disease) or other factors (e.g., technical aspect of imaging techniques, biomarker levels, test operator or interpreter skill, setting). Frequency of various tumor stages and stage-appropriate treatment provided, conditional on diagnosis of HCC. <p>Intermediate outcomes include:</p> <ul style="list-style-type: none"> Effects on diagnostic thinking. Effects on clinical decisionmaking. <p>Clinical and patient-centered outcomes include:</p> <ul style="list-style-type: none"> Overall mortality or survival. Recurrence of HCC, including rates of seeding by fine needle aspiration. Quality of life, measured using scales like Short-Form Health Survey (SF-36[®]) or EuroQol 5D (EQ-5[™]), or as defined by the primary studies. Psychosocial effects of diagnostic testing on patients, patients' caregivers, and other family members, as measured by self-reported questionnaire instruments. Resource utilization and patient burden (for example, increased cost associated with access to imaging facility, the number of imaging procedures conducted). 	Treatment response
Settings	All relevant care settings (e.g., primary and secondary care)	NA
Study Designs	Randomized controlled trials, cohort studies, and other observational studies. Pull systematic reviews to check for included studies.	Case-control studies, case studies, literature reviews Studies that do not report sensitivity, cost effectiveness modeling studies

Table 2. Inclusion / Exclusion Criteria for Key Question 2

Category	Inclusion Criteria	Exclusion Criteria
Population and Patient Characteristics	<p>Patients at high risk for HCC in whom a suspicious lesion(s) has been detected at surveillance or by other means</p> <p>Patients who have undergone liver transplants for HCC, either with or without HCC detected in the explant liver</p>	Children and adolescents, patients with cholangiocarcinoma

Category	Inclusion Criteria	Exclusion Criteria
Interventions and Comparators	<p>Imaging techniques, used singly, in combination or in a specific sequence, including MDCT, spiral CT, dual source CT, U/S, and MRI (including contrast agents like Gd-EOB-DTPA and SPIO), and FDG-PET with different tracers (including 18F, fluorothymidine (FLT), 11C choline, and 11C methionine, or others). The imaging techniques that are used in combination with biomarkers (such as alpha-fetoprotein [AFP], des-gamma-carboxy prothrombin) will be considered in the comparisons. Studies where biomarkers such as glypican 3, heat shock protein 70, and glutamine synthetase are combined with imaging techniques to define the reference standard will be considered in the context of the reference standard used.</p> <p>Comparators:</p> <ul style="list-style-type: none"> For studies of diagnostic accuracy (comparative test performance), the reference standard comparators will be histopathology (based on explanted liver specimens or biopsy) or clinical and imaging followup, and the imaging comparators will be alternative imaging tests or strategies. For studies of comparative effectiveness, the comparators will be no imaging or alternative imaging strategies. 	CT arteriography, CT portography and interventions used in treatment (imaging guided RFA), outdated imaging techniques (e.g., conventional, non-spiral/non-multidetector CT or imaging performed prior to 1995) and imaging techniques not available or in use in the U.S. (e.g., hepatic portography).
Outcomes	<p>Diagnostic outcomes include:</p> <ul style="list-style-type: none"> Test performance (e.g., sensitivity and specificity) for diagnosis of HCC using pathological diagnosis or clinical and imaging follow-up as the reference standard. <p>Intermediate outcomes include:</p> <ul style="list-style-type: none"> Rates of followup tests conducted, including further imaging and invasive procedures like core biopsy. <p>Clinical and patient-centered outcomes include:</p> <ul style="list-style-type: none"> Overall mortality or survival. Recurrence of HCC, including rates of seeding by fine needle aspiration Quality of life, measured using scales like Short-Form Health Survey (SF-36[®]) or EuroQol 5D (EQ-5[™]), or as defined by the primary studies. Psychosocial effects of diagnostic testing on patients, patients' caregivers, and other family members, as measured by self-reported questionnaire instruments. Resource utilization and patient burden (e.g., costs associated with travel to imaging facility, the number of imaging procedures conducted) 	Treatment response
Settings	All relevant care settings (e.g., primary and secondary care)	NA
Study Designs	Randomized controlled trials, cohort studies, and other observational studies. Pull systematic reviews to check for included studies.	Case-control studies, case studies, literature reviews Studies that do not report sensitivity, cost effectiveness modeling studies

Table 3. Inclusion / Exclusion Criteria for Key Question 3

Category	Inclusion Criteria	Exclusion Criteria
Population and Patient Characteristics	Patients diagnosed with HCC who require staging prior to initial treatment	Children and adolescents, patients with cholangiocarcinoma
Interventions and Comparators	Imaging techniques, used singly, in combination, or in a specific sequence, including MDCT, spiral CT, dual source CT, U/S, contrast CT, and MRI with contrast (including	CT arteriography, CT portography and interventions used in

Category	Inclusion Criteria	Exclusion Criteria
	<p>contrast agents like Gd-EOB-DTPA and SPIO), and FDG-PET with different tracers (including 18F, fluorothymidine (FLT), 11C choline, and 11C methionine, or others). Test performance of imaging techniques will be stratified by the different staging systems used.</p> <p>Comparators:</p> <ul style="list-style-type: none"> For studies of diagnostic accuracy (comparative test performance), the reference standard comparators will be histopathology (based on explanted liver specimens or biopsy) or clinical and imaging followup, and the imaging comparators will be alternative imaging tests or strategies. For studies of comparative effectiveness, the comparators will be no imaging or alternative imaging strategies. 	<p>treatment (imaging guided radiofrequency ablation), outdated imaging techniques (e.g., conventional, non-spiral/non-multidetector CT or imaging performed prior to 1995) and imaging techniques not available or in use in the U.S. (e.g., hepatic portography).</p>
Outcomes	<p>Diagnostic outcomes include:</p> <ul style="list-style-type: none"> Measures for accuracy of staging (stage reclassification). <p>Intermediate outcomes include:</p> <ul style="list-style-type: none"> Effects on diagnostic thinking. Effects on clinical decisionmaking. <p>Clinical and patient-centered outcomes include:</p> <ul style="list-style-type: none"> Overall mortality or survival. Recurrence of HCC, including rates of seeding by fine needle aspiration Quality of life, measured using scales like Short-Form Health Survey (SF-36) or EuroQol 5D, or as defined in the primary studies. Psychosocial effects of diagnostic testing on patients, patients' caregivers, and other family members, as measured by self-reported questionnaire instruments. Resource utilization and patient burden (e.g., increased cost associated with access to imaging facility, the number of imaging procedures done). 	Treatment response
Settings	All relevant care settings (e.g., primary and secondary care)	NA
Study Designs	Randomized controlled trials, cohort studies, and other observational studies. Pull systematic reviews to check for included studies.	<p>Case-control studies, case studies, literature reviews</p> <p>Studies that do not report sensitivity , cost effectiveness modeling studies</p>

Table 4. Inclusion / Exclusion Criteria for Key Questions 1d, 2d, 3d

Category	Inclusion Criteria	Exclusion Criteria
Adverse Events or Harms	<p>Adverse effects or harms associated with the imaging techniques (e.g., test-related anxiety, adverse events secondary to venipuncture, contrast allergy, exposure to radiation)</p> <p>Adverse effects or harms associated with test-associated diagnostic workup (e.g., followup with other incidental tumors discovered on imaging)</p>	

In general, exclude

- Studies that do not use a pathologic reference standard (code as 13 but set aside)
- Non-English-language publications at the full text level (include relevant citations at the abstract level).
- Studies that don't report measures of diagnostic accuracy (sensitivity/specificity)
- CT arteriography
- CT portography



- Conventional CT
- Interventions used in treatment (imaging guided radiofrequency ablation, etc.)
- Studies of treatment response
- Studies looking at liver metastases
- Studies with imaging conducted prior to 1995
- Children and adolescents
- Patients with cholangiocarcinoma
- Studies of nonhuman populations

Note: Minimum sample size to be determined after full text review.