Surveillance Imaging and Alpha Fetoprotein for Early Detection of Hepatocellular Carcinoma in Patients With Cirrhosis: A Meta-analysis

Kristina Tzartzeva[1],[*], **Joseph Obi**[1],[*], **Nicole E. Rich**[1], **Neehar D. Parikh**[2], **Jorge A. Marrero**[1], **Adam Yopp**[3], **Akbar Waljee**[2], and **Amit G. Singal**[1],[4]

[1]Department of Internal Medicine, UT Southwestern Medical Center, Dallas, TX

[2]Department of Internal Medicine, University of Michigan, Ann Arbor, MI

[3]Department of Surgery, UT Southwestern Medical Center, Dallas, TX

[4]Department of Clinical Sciences, University of Texas Southwestern, Dallas, TX

Abstract

Background & Aims—Society guidelines differ in their recommendations for surveillance to detect early-stage hepatocellular carcinoma (HCC) in patients with cirrhosis. We compared the performance of surveillance imaging, with or without alpha fetoprotein (AFP), for early detection of HCC in patients with cirrhosis

Methods—Two reviewers searched MEDLINE and SCOPUS from January 1990 through August 2016 to identify published sensitivity and specificity of surveillance strategies for overall and early detection of HCC. Pooled estimates were calculated and compared using the DerSimonian and Laird method for a random effects model. The study was conducted in accordance with Preferred Reporting Items for Systematic Review and Meta-analysis guidelines.

Correspondence: Amit G. Singal, M.D., M.S., Dedman Scholar of Clinical Care, Associate Professor of Medicine, Division of Digestive and Liver Diseases, University of Texas Southwestern, 5959 Harry Hines Blvd, POB 1, Suite 420, Dallas TX 75390-8887, Tel: 214-645-6111, Fax: 214-645-6114.
[*]Drs Tzartzeva and Obi contributed equally and are co-first authors on this manuscript.

Conflicts of Interests: None of authors have relevant conflicts of interest

Author Contributions
<u>Kristina Tzartzeva</u> was involved in acquisition of data, interpretation of data, and critical revision of manuscript for important intellectual content.
<u>Joseph Obi</u> was involved in acquisition of data, interpretation of data, and critical revision of manuscript for important intellectual content.
<u>Nicole Rich</u> was involved in interpretation of data and critical revision of manuscript for important intellectual content.
<u>Neehar Parikh</u> was involved in interpretation of data and critical revision of manuscript for important intellectual content.
<u>Jorge Marrero</u> was involved in interpretation of data and critical revision of manuscript for important intellectual content.
<u>Adam Yopp</u> was involved in interpretation of data and critical revision of manuscript for important intellectual content.
<u>Akbar Waljee</u> was involved in interpretation of data and critical revision of manuscript for important intellectual content.
<u>Amit G. Singal</u> was involved in study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript, critical revision of manuscript for important intellectual content, and study supervision. He is the guarantor of the article.

Results—Thirty-two studies (comprising 13367 patients) characterized sensitivity of imaging with or without AFP measurement for detection of HCC in patients with cirrhosis. Ultrasound detected any stage HCC with 84% sensitivity (95% CI, 76%–92%), but early-stage HCC with only 47% sensitivity (95% CI, 33%–61%). In studies comparing ultrasound with vs without AFP measurement, ultrasound detected any stage HCC with a lower level of sensitivity than ultrasound plus AFP measurement (relative risk [RR], 0.88; 95% CI, 0.83–0.93) and early-stage HCC with a lower level of sensitivity than ultrasound plus AFP measurement (RR, 0.81; 95% CI, 0.71–0.93). However, ultrasound alone detected HCC with a higher level of specificity than ultrasound plus AFP measurement (RR, 1.08; 95% CI, 1.05–1.09). Ultrasound with vs without AFP detected early-stage HCC with 63% sensitivity (95% CI, 48%–75%) and 45% sensitivity (95% CI, 30%–62%), respectively (*P*=.002). Only 4 studies evaluated computed tomography or magnetic resonance image-based surveillance, which detected HCC with 84% sensitivity (95% CI, 70%–92%).

Conclusions—In a meta-analysis of publications, we found ultrasound alone to detect early-stage HCC with a low level of sensitivity in patients with cirrhosis. Addition of AFP to ultrasound analysis significantly increases the sensitivity of HCC detection in clinical practice.

INTRODUCTION

Hepatocellular carcinoma (HCC) is one of the leading causes of death in patients with cirrhosis, with prognosis largely driven by tumor stage.[1] Patients detected at an early stage are eligible for curative therapies and can achieve 5-year survival rates approaching 70% with liver transplantation or surgical resection.[2] Conversely, those with more advanced tumors are only eligible for palliative treatments and have a poor prognosis, with median survival of 1–2 years.[3]

Several cohort studies have demonstrated an association between HCC surveillance and early tumor detection and improved survival in patients with cirrhosis.[4, 5] However, there is uncertainty regarding what surveillance strategy is most effective for early tumor detection in clinical practice.[6–8] Choice of surveillance modality must balance sensitivity to optimize early HCC detection, specificity to minimize surveillance-related harms, and costs to remain cost-effective. Professional society guidelines from the American Association for the Study of Liver Diseases (AASLD) and European Association for the Study of the Liver (EASL) both recommend surveillance using ultrasound every 6 months in patients with cirrhosis; however, they disagree about the value of adding the serum biomarker alpha fetoprotein (AFP) as an adjunct surveillance test.[9, 10] Additionally, there is increasing use of computed tomography (CT) and magnetic resonance imaging (MRI) for surveillance purposes in clinical practice given concerns about ultrasound's accuracy.[7, 11]

Given few direct comparative studies, we are forced to primarily rely on indirect comparisons across studies. The aim of this systematic review is to characterize and

compare the performance of surveillance modalities including ultrasound with or without AFP, CT, and MRI for the detection of HCC.

METHODS

Literature search and study selection

We searched the MEDLINE and SCOPUS databases from January 1990 through August 2016 using search terms described in Supplemental Methods. Additional studies that may have been missed by the electronic search were identified through manual searching of reference lists from applicable studies and consultation with experts in the field. Two investigators (K.T. and J.O.) independently reviewed publication titles identified by the search strategy. If the applicability of an article could not be determined by title or abstract alone, the full text was reviewed. Articles were independently evaluated for possible inclusion and any disagreements were resolved through consensus with a third reviewer (A.S.). The study was conducted in accordance with Preferred Reporting Items for Systematic Review and Meta-analysis (PRISMA) guidelines.

Study inclusion and exclusion criteria

Studies were included for analysis if they evaluated abdominal imaging (ultrasound, CT, or MRI) with or without AFP for HCC surveillance in a cohort of patients with cirrhosis. Surveillance was defined as the repeated use of the test at a regular interval over time to detect a previously undiagnosed lesion. Studies evaluating imaging for screening or diagnostic purposes instead of surveillance were not included in the analysis. Studies performed exclusively in a non-cirrhotic cohort, such as patients with chronic hepatitis, were excluded. If the study cohort included both patients with cirrhosis and chronic hepatitis, only data regarding cirrhosis patients were included when possible. If data could not be extracted for the subset of patients with cirrhosis, we only included those studies in which a majority of patients had cirrhosis. Studies in which <50% of patients had cirrhosis, or those in which the proportion of patients with cirrhosis was not detailed, were excluded. Studies using sequential test combinations, such as ultrasound testing in patients based on AFP levels, were excluded because information bias from the initial study could have unpredictable effects on the ultrasound operating characteristics. Studies were required to report the number of discovered HCC and number of missed HCC for each surveillance test, as lack of data for false negative results (i.e. patients with missed lesions) precluded sensitivity calculations. Studies that reported the proportion of HCC discovered by surveillance, but not stratified by test, were excluded. Additional exclusion criteria included non-English language, non-human data, and lack of original data. If duplicate publications used the same cohort of patients, the data from the most recent manuscript were included.

Data extraction

Three authors (K.T., J.O., A.S.) independently reviewed and extracted required information from eligible studies using standardized forms. A fourth investigator (N.R.) was available to resolve any discrepancies between the sets of extracted data. The data extraction form included the following study design items: geographical location and date of study, characteristics and size of study cohort, inclusion and exclusion criteria, surveillance

methods, surveillance interval, duration of follow-up and 'gold-standard' methods for diagnostic confirmation of HCC. In addition, the extraction form recorded the following primary data: number of HCC discovered during surveillance (true positives), number of false positives, number of missed lesions (false negatives) and number of true negatives for each surveillance test. We also recorded the proportion of HCC discovered at an early stage, as defined by Milan criteria: one nodule <5 cm or 2–3 nodules, each <3 cm in diameter, without gross vascular invasion or extra-hepatic metastases.[12] We defined early stage HCC by Milan Criteria instead of the Barcelona Clinic Liver Cancer (BCLC) staging system because many studies did not report necessary data for liver dysfunction or patient performance status. Two authors (K.T. and A.S.) independently assessed study quality by a modified checklist based upon the Quality Assessment Tool for Diagnostic Accuracy (QUADAS2) guidelines with discrepancies resolved by consensus.[13]

Statistical analysis

The aim of this study was to compare the performance of surveillance strategies (ultrasound, CT, and MRI with or without AFP) to detect HCC, particularly at an early stage. For each individual study, per-patient sensitivity, per-patient specificity and diagnostic ORs with 95% confidence intervals were calculated as possible. Pooled estimates of each calculation were computed using STATA 14 (StataCorp, College Station, TX, USA). Estimates of effect were pooled using the DerSimonian and Laird method for a random effects model. The heterogeneity of diagnostic test parameters was initially evaluated graphically by examination of forest plots and statistically by the inconsistency index, with values >50% consistent with the possibility of substantial heterogeneity.[14] Sensitivity analysis, in which one study is removed at a time from the model, was performed to determine if there was possible undue influence of a single study. Publication bias was initially evaluated graphically by funnel plot analysis and then statistically using Begg's test. A summary receiver operator characteristics (SROC) curve was constructed to illustrate the distribution of sensitivities and specificities. The area under the curve (AUC) was computed, with perfect tests having an AUC of 1 and poor tests having an AUC close to 0.5. Subgroup analyses were planned for HCC detection for predefined subsets of studies based on (i) prospective vs. retrospective study design, (ii) location of study; (iii) study period; and (iv) inclusion of patients without cirrhosis.

RESULTS

Literature Search

Upon review of the 10260 titles identified by the search strategies, 615 abstracts were further examined. 155 publications underwent full-text review to determine their eligibility for the meta-analysis and 126 were excluded. One study was excluded because they did not use imaging, 30 articles evaluated imaging but not as a surveillance tool, 11 studies were not conducted among patients with cirrhosis, 16 studies were excluded for lack of original data and 68 studies had insufficient data for extraction. The remaining 29 studies were selected after meeting all applicable inclusion criteria (Supplemental Figure). Finally, recursive literature searches identified 3 additional articles that met inclusion criteria, producing a total of 32 studies.[15–46]

Study Characteristics

Characteristics of the 32 studies evaluating the performance of imaging with or without AFP to detect HCC are described in Table 1.[15–46] Among a total of 13367 patients,1877 developed HCC. Fifteen studies (n=4480 patients) reported data on early HCC; of 516 patients who developed HCC, 319 (61.8%) were detected at an early stage. [15,20,22,23,27–32,36,39,40,43,46] Twenty-eight studies (n=10743 patients) exclusively included patients with cirrhosis, with four studies including some patients with significant fibrosis. Most studies (n=23) were prospective in design, although 9 collected data on surveillance test performance retrospectively. Seven studies were conducted in the United States, 14 in Europe, 7 in Asia, and 4 in other countries. Most studies evaluated ultrasound as the surveillance imaging modality; however, two evaluated CT-based surveillance and two evaluated MRI-based surveillance. There was no evidence of publication bias by Begg's test (p=0.85).

Quality Assessment

Quality assessment of included studies is provided in Table 2. Although most studies had appropriate patient selection, 4 studies only enrolled patients listed for liver transplantation potentially introducing selection bias and overestimating ultrasound performance. Several studies excluded patients with limited life expectancy due to comorbidities or Child C cirrhosis, but this was not felt to introduce selection bias given surveillance is not recommended in Child C cirrhosis patients outside of candidates for liver transplantation. Nine studies had applicability concerns about patient selection given inclusion of patients with advanced fibrosis but not cirrhosis.[16, 21–24, 30, 33, 37, 44,] Most studies used standard reference tests to confirm HCC diagnosis, including CT, MRI, and biopsy; however, 10 studies used AFP > 400 ng/mL, so reference tests were not independent of the surveillance tests. Further, elevated AFP is no longer recommended for diagnosis by AASLD guidelines for HCC diagnosis given imperfect specificity. Lack of independence between surveillance and reference tests was also a concern, with 2 studies evaluating CT and MRI-based surveillance, potentially overestimating surveillance test performance.[29, 46] Finally, most (n=26) studies failed to perform reference tests in those with negative surveillance tests, introducing verification bias and potentially overestimating surveillance test performance.

Ultrasound alone for HCC Detection

Thirty-one studies, with 12977 patients, reported sensitivity of ultrasound for detection of HCC at any stage[15–33, 35–46], of which 15 studies (n=4400) evaluated detection of HCC at an early stage. [15,20,22,23,27–32,36,39,40,43,46] There was a wide range in sensitivities for any stage detection (28% to 100%), as well as early HCC detection, which ranged from 21% to 89%. The pooled sensitivity of ultrasound was 84% (95%CI 76% – 92%) for HCC detection at any stage but significantly lower at 47% (95%CI 33% – 61%) for early HCC detection (Figure 1A and 1B). In subgroup analyses, there was no significant difference in ultrasound sensitivity for early HCC detection by prospective vs. retrospective study design (p=0.12), or inclusion of patients without cirrhosis (p=0.89). We found no difference in ultrasound sensitivity for early detection by study location (p=0.12), although sensitivity was numerically lower in U.S. studies[23, 29, 36] (36%, 95%CI 27% – 47%) compared to studies

conducted in Europe[15, 20, 27, 31, 32, 39, 40, 43] (47%, 95%CI 28% – 67%). Similarly, there was some improvement in ultrasound sensitivity for early HCC detection over time from 21% (95%CI 5%-51%) in studies conducted prior to 1990[27] to 45% (95%CI 15% – 77%) among studies conducted primarily in 1990s[15, 20, 22, 31, 32, 39] to 50% (35% – 66%) for studies conducted primarily after 2000[23, 28–3036, 40, 43, 46]; however, this difference did not reach statistical significance (p=0.17).

Fifteen studies (n=8519) characterized overall accuracy (i.e. both sensitivity and specificity) of ultrasound for detection of HCC at any stage [15–17, 22, 25, 27, 29, 30, 36, 40, 41, 43–46], of which 9 studies (n=3426) evaluated ultrasound performance for early HCC detection[15, 22, 27, 29, 30, 36, 40, 43, 46]. Summary ROC curves for any stage HCC and early HCC detection are shown in Figures 2A and 2B, respectively. Although the area under the curve for any stage HCC was 0.96 (95%CI 0.94 – 0.98), it was significantly lower at 0.88 (95%CI 0.85 – 0.90) for early stage HCC. In this subset, pooled sensitivity and specificity of ultrasound for early stage detection were 53% (95%CI 35% – 70%) and 91% (95%CI 86 – 94%), respectively. The positive and negative likelihood ratios of ultrasound for early stage HCC were 5.8 (95%CI 3.7 – 9.2) and 0.51 (95%CI 0.35 – 0.75), respectively, with a diagnostic odds ratio of 11 (95%CI 5 – 24).

Ultrasound with AFP for HCC Detection

In the 18 studies (n=8526) that compared sensitivity of ultrasound with or without AFP for detection of HCC at any stage, sensitivity of ultrasound alone was 78% (95%CI 67% – 86%) compared to 97% (95%CI 91% – 99%) for ultrasound plus AFP. [20–27, 30, 33, 35–37, 40–42, 45, 46] The sensitivity of ultrasound alone was significantly lower compared to ultrasound plus AFP (RR 0.88, 95%CI 0.83 – 0.93).

In the 7 studies (n=2770) that compared sensitivity of ultrasound with or without AFP for early HCC detection, ultrasound alone had significantly lower sensitivity than ultrasound plus AFP (RR 0.81, 95%CI 0.71 – 0.93) (Table 3, Figure 3).[20, 23, 27, 30, 36, 40, 46] Pooled sensitivities of ultrasound with and without AFP for early HCC detection were 63% (95%CI 48% – 75%) and 45% (95%CI 30% – 62%), respectively. The benefit of AFP was consistent across subgroups including prospective studies (RR 0.78, 95%CI 0.66 – 0.92), [20, 23, 27, 36, 40, 46] studies conducted in the United States (RR 0.59, 95%CI 0.41 – 0.85),[23, 36] studies exclusively enrolling patients with cirrhosis (RR 0.76, 95%CI 0.60 – 0.95), [20, 27, 30, 36, 40, 46] and studies conducted after year 2000 (RR 0.79, 95%CI 0.66 – 0.95). [23, 30, 36, 40, 46] Although there was not statistical heterogeneity (I^2=0%), we performed a sensitivity analysis in which the study by Singal and colleagues[36] was removed given it appeared to be a potential outlier on visual inspection of forest plots. Ultrasound alone continued to have significantly lower sensitivity for early HCC detection compared to ultrasound plus AFP (RR 0.84, 95%CI 0.73 – 0.96). We also performed a sensitivity analysis removing studies by Qian and Lok[23, 30] given inclusion of patients without cirrhosis, and ultrasound alone continued to have lower sensitivity for early HCC detection than ultrasound plus AFP (RR 0.76, 95%CI 0.60 – 0.95).

Six studies (n=4782) compared accuracy of ultrasound (i.e. both sensitivity and specificity) with and without AFP for HCC detection,[25, 27, 36, 40, 45, 46] of which 4 studies (n=2245)

reported accuracy for early HCC detection.[27, 36, 40, 46] As above, sensitivity of ultrasound alone was lower than ultrasound plus AFP for any stage and early stage HCC detection. However, ultrasound alone had higher specificity than ultrasound plus AFP (RR 1.08, 95%CI 1.05 – 1.09). The specificity for ultrasound alone was 92% (95%CI 85% – 96%) compared to 84% (95%CI 77% – 89%) for ultrasound plus AFP. Overall, the diagnostic OR of ultrasound for early HCC detection (7, 95%CI 3 – 15) was similar to the diagnostic odds ratio for ultrasound plus AFP (8, 95%CI 3 – 23).

Subgroup Analysis of Prospective Studies

In prospective cohort studies, the pooled sensitivity of ultrasound was 86% (95%CI 75% – 95%) for detection of any stage HCC but only 42% (95%CI 27% – 58%) for early HCC detection. [15–16, 18–21, 23–29, 31–32, 36–40, 46] In the 6 studies reporting both sensitivity and specificity of ultrasound for early HCC detection, pooled sensitivity, specificity, and AUROC were 48% (28% – 68%), 92% (90% – 94%), and 0.91 (0.89 – 0.94), respectively. [15, 27, 29, 36, 40, 46] Compared to ultrasound and AFP, ultrasound had significantly lower sensitivity for detection of any stage HCC (RR 0.80, 95%CI 0.72 – 0.88) and early stage HCC (0.78, 95%CI 0.66 – 0.92). [20, 23, 27, 36, 40, 46] In this subset of studies, the sensitivity of ultrasound and AFP for any stage and early stage HCC were 95% (83% – 100%) and 60% (95%CI 45% – 74%), respectively, compared to only 72% (56% – 86%) and 40% (22% – 58%), respectively, for ultrasound only.

CT and MRI for HCC Detection

Four studies (n=897) characterized performance of cross-sectional surveillance imaging for detection of HCC - two evaluated CT-based surveillance[29, 41] and two evaluated MRI-based surveillance.[34, 46] Pocha and colleagues[29] performed a single-center randomized trial comparing CT and ultrasound-based surveillance in patients with cirrhosis. The sensitivity and specificity of CT for any stage detection were 87.5% (95%CI 50.8% – 99.9%) and 87.5% (95%CI 77.7% – 93.5%), respectively; however the sensitivity of CT for early HCC detection was lower at 62.5% (95%CI 30.4% – 86.5%) and did not significantly differ from ultrasound. Van Thiel and colleagues[41] reported sensitivity and specificity for any stage detection of 70.0% (95%CI 45.7% – 88.1%) and 100% (95%CI 96.3% – 100%), respectively, but did not report performance characteristics for early HCC detection.

The two studies evaluating MRI had a pooled sensitivity and specificity for any HCC detection of 83.1% (95%CI 72.0% – 90.5%) and 89.1% (95%CI 86.5% – 91.3%), respectively.[34, 46] Kim and colleagues[46] compared MRI and ultrasound in a cohort of 407 patients with cirrhosis who underwent 1100 surveillance exams and found MRI had a sensitivity of 83.7% (95%CI 69.7% – 92.2%) for early HCC detection, which was significantly higher than ultrasound (25.6%, 95%CI 14.8% – 40.4%).

DISCUSSION

Ultrasound currently forms the backbone of professional society recommendations for HCC surveillance; however, our meta-analysis highlights its suboptimal sensitivity for detection of HCC at an early stage. Ultrasound's sensitivity and specificity for any-stage HCC detection

exceed 90%, but it detects less than half of HCC patients at an early stage. Using ultrasound in combination with AFP appears to significantly improve sensitivity for detecting early HCC with a small, albeit statistically significant, trade-off in specificity. There are few studies evaluating other alternative surveillance strategies including CT or MRI-based surveillance, particularly at an early stage; however, available data suggest MRI retains high sensitivity.

There has been considerable debate regarding the potential benefit of adding AFP to ultrasound-based HCC surveillance programs.[8, 47, 48] It is clear the best way to address this debate would be a randomized clinical trial given inherent limitations of observation studies with indirect comparisons including biases with patient selection, flow and timing of tests, and potential confounders including the proportion of patients with obesity or advanced liver dysfunction, which are known to impact ultrasound sensitivity. However, a previously attempted randomized controlled trial comparing ultrasound alone to ultrasound with AFP was not possible given high rates of AFP contamination in the ultrasound alone study arm, highlighting providers' reluctance to not perform AFP and rely on ultrasound alone.[40] This also underlines the importance of our study's findings, as this meta-analysis of cohort studies may represent the highest level of evidence possible comparing the two surveillance tests.

A prior meta-analysis of 13 prospective cohort studies concluded AFP was not of significant additional value compared to ultrasound alone[48]; however, only prospective studies conducted prior to 2007 were included. Four subsequent prospective studies[23, 36, 40, 46] and one retrospective study[30] directly compared the sensitivity of ultrasound with or without AFP for early HCC detection, with a consistently observed benefit of adding AFP across these studies. The most recent AASLD guidelines state surveillance should be performed using ultrasound with or without AFP, as it is "not possible to determine which type of surveillance test, US alone or the combination of US plus AFP, leads to greater improvement in survival".[10] This recommendation represents a departure from prior AASLD guidelines and current EASL guidelines, which recommend against use of AFP during surveillance given insufficient sensitivity and specificity.[9, 49] Our findings support this change and, in fact, suggest surveillance should routinely be done using ultrasound with AFP. Although studies did not evaluate the potential impact of surveillance strategies on overall survival, using ultrasound in combination with AFP significant improves early HCC detection – the mediating pathway for surveillance improving survival – compared to ultrasound alone.

Although the pooled sensitivity of ultrasound for early HCC detection was low, there was wide variation in performance between studies. We explored this heterogeneity in subgroup analyses including study design, study location, study period, and inclusion of patients with cirrhosis; however, no single factor was able to fully explain differences in ultrasound performance. Although we noted numerical improvement in ultrasound sensitivity over time from 21% in studies conducted prior to 1990 to 50% among those conducted after 2000, this difference did not reach statistical significance. Further, the benefit of using AFP in combination with ultrasound was still observed in the subset of studies conducted after 2000. Several other factors that potentially affect ultrasound performance could not be evaluated given insufficient reporting of these details in included studies. Although we could not

directly evaluate the effect of ultrasound expertise, we found no difference in results by study location – a potential surrogate for ultrasound protocol. Whereas physicians often perform ultrasounds in Asia and parts of Europe, they are typically conducted by ultrasound technicians in the United States. In addition to operator dependency, ultrasound is reported to have lower quality imaging and sensitivity in patients with obesity, non-alcoholic steatohepatitis, or increased liver nodularity.[50, 51] These associations suggest the efficacy of ultrasound alone for HCC surveillance may become more limited in the future given increasing incidences of obesity and NAFLD-related cirrhosis. Centralization of HCC surveillance to high-volume/tertiary centers, in which expert operators are more likely available to perform ultrasounds, is likely not a feasible solution given the large number of patients with cirrhosis and the frequent nature of requiring surveillance imaging every 6 months.[52, 53] Although it may still be possible to define study settings (e.g. high volume centers with expert operators) and patient populations (e.g. non-obese patients with viral-related compensated cirrhosis) in whom ultrasound alone can be effective for early HCC detection, our findings suggest using AFP in combination with ultrasound is likely beneficial in most settings and patient populations.

Criticism of AFP has largely centered on its insufficient sensitivity and specificity for early HCC detection when used alone; however, our findings demonstrate adding AFP to ultrasound surveillance is associated with significantly improved sensitivity. It should be noted the increased sensitivity for early HCC by using AFP with ultrasound was associated with a trade-off in decreased specificity. However, most studies evaluated AFP at a single-threshold value of 20 ng/mL, with any elevated AFP level counted as a false positive and thereby decreasing specificity. Studies suggest AFP false positives are more likely in the setting of hepatitis C and/or elevated ALT, so this may be less problematic as HCC epidemiology shifts to NASH-related.[54, 55] Further, increasing data suggest monitoring longitudinal patterns of AFP over time can increase biomarker accuracy and may better reflect how AFP is interpreted in clinical practice.[56, 57] Increasing AFP levels, even if below a cut-off of 20 ng/mL, can be a sign of HCC and stable or decreasing AFP levels, even if above 20 ng/mL, are reassuring. A recent study evaluating HCC surveillance harms demonstrated physical harms related to false positive AFP results are mitigated by clinical interpretation.[58] In contrast, ultrasound-related physical harms may be underestimated if solely relying on false positive results given high rates of diagnostic testing for indeterminate ultrasound results.[58, 59] Future studies comparing surveillance strategies should incorporate evaluation of physical, psychological, and financial harms instead of simply reporting specificity.[60]

Our meta-analysis is the first to include comparative studies including CT and MRI versus ultrasound. We found only four studies have evaluated CT- and MRI-based surveillance strategies despite increasing use of these modalities in clinical practice. A trial comparing CT and ultrasound found no significant improvement in early HCC detection between the two tests, although it only included only 163 patients.[29] In addition to lack of demonstrated benefits, CT-based surveillance is limited by potential physical harms including radiation exposure and potential contrast-induced nephrotoxicity.[61, 62] The trials evaluating MRI-based surveillance report high specificity and sensitivity, including for early HCC detection; however, this strategy likely cannot be expanded to all cirrhosis patients given high costs

(including co-payments), contraindications to MRI (implanted hardware, claustrophobia), long scan times compared to ultrasound, and limited MRI availability (particularly in more remote areas).[63] Early data suggest an abbreviated MRI protocol may be able to address some of these issues, while retaining high accuracy, but evaluation in cohort studies is still needed.[64, 65]

Results from our study must be interpreted within the limitations of the included studies. First, most studies only reported detection of HCC at any stage, instead of early HCC detection, which over-estimates surveillance test performance. Further, no studies compared differences in downstream outcomes including curative treatment receipt or overall survival. Second, most studies did not include additional follow-up or apply a "gold standard' reference test in patients with normal surveillance tests to confirm absence of HCC so were limited by verification bias, which may also result in over-estimation of surveillance test performance. Third, many studies did not report factors that affect ultrasound quality (e.g. operator experience or proportion of obese patients) so we could not identify subgroups in whom ultrasound alone may be sufficient. Although these limitations may affect point estimates, they should affect ultrasound and AFP equally so our analysis comparing ultrasound with and without AFP should be unaffected.

In summary, we demonstrated ultrasound has suboptimal sensitivity for early HCC detection, highlighting the need for alternative surveillance strategies. There are currently insufficient data to support routine use of CT- or MRI-based surveillance in all patients with cirrhosis. Using AFP in combination with ultrasound significantly increases early HCC detection, suggesting this may be the preferred surveillance strategy for patients with cirrhosis until superior surveillance strategies are available.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Abbreviations

AASLD	American Association for the Study of Liver Diseases
AFP	alpha fetoprotein
AUC	area under the curve
BCLC	Barcelona Clinic Liver Cancer
CT	computed tomography
EASL	European Association for the Study of the Liver

HCC	hepatocellular carcinoma
MRI	magnetic resonance imaging
PRISMA	Preferred Reporting Items for Systematic Review and Meta-analysis
QUADAS2	Quality Assessment Tool for Diagnostic Accuracy
SROC	summary receiver operator characteristics

References

1. El-Serag HB. Epidemiology of viral hepatitis and hepatocellular carcinoma. Gastroenterology. 2012; 142:1264–1273 e1. [PubMed: 22537432]

2. Llovet JM, Fuster J, Bruix J. Intention-to-treat analysis of surgical treatment for early hepatocellular carcinoma: resection versus transplantation. Hepatology. 1999; 30:1434–40. [PubMed: 10573522]

3. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, de Oliveira AC, Santoro A, Raoul JL, Forner A, Schwartz M, Porta C, Zeuzem S, Bolondi L, Greten TF, Galle PR, Seitz JF, Borbath I, Haussinger D, Giannaris T, Shan M, Moscovici M, Voliotis D, Bruix J. Sorafenib in advanced hepatocellular carcinoma. N Engl J Med. 2008; 359:378–90. [PubMed: 18650514]

4. 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:e1001624. [PubMed: 24691105]

5. Kansagara D, Papak J, Pasha AS, O'Neil M, Freeman M, Relevo R, Quinones A, Motu'apuaka M, Jou JH. Screening for hepatocellular carcinoma in chronic liver disease: a systematic review. Ann Intern Med. 2014; 161:261–9. [PubMed: 24934699]

6. Dalton-Fitzgerald E, Tiro J, Kandunoori P, Halm EA, Yopp A, Singal AG. Practice Patterns and Attitudes of Primary Care Providers and Barriers to Surveillance of Hepatocellular Carcinoma in Patients with Cirrhosis. Clin Gastroenterol Hepatol. 2014

7. Joshi K, Mendler M, Gish R, Loomba R, Kuo A, Patton H, Kono Y. Hepatocellular carcinoma surveillance: a national survey of current practices in the USA. Dig Dis Sci. 2014; 59:3073–7. [PubMed: 25027206]

8. Marrero JA, El-Serag HB. Alpha-fetoprotein should be included in the hepatocellular carcinoma surveillance guidelines of the American Association for the Study of Liver Diseases. Hepatology. 2011; 53:1060–1. author reply 1061–2. [PubMed: 21374678]

9. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. Eur J Cancer. 2012; 48:599–641. [PubMed: 22424278]

10. Heimbach J, Kulik LM, Finn R, Sirlin CB, Abecassis M, Roberts LR, Zhu A, Murad MH, Marrero J. Aasld guidelines for the treatment of hepatocellular carcinoma. Hepatology. 2017

11. Singal AG, Nehra M, Adams-Huet B, Yopp AC, Tiro JA, Marrero JA, Lok AS, Lee WM. Detection of hepatocellular carcinoma at advanced stages among patients in the HALT-C trial: where did surveillance fail? Am J Gastroenterol. 2013; 108:425–32. [PubMed: 23337478]

12. Mazzaferro V, Regalia E, Doci R, Andreola S, Pulvirenti A, Bozzetti F, Montalto F, Ammatuna M, Morabito A, Gennari L. Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. N Engl J Med. 1996; 334:693–9. [PubMed: 8594428]

13. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011; 155:529–36. [PubMed: 22007046]

14. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003; 327:557–60. [PubMed: 12958120]

15. Bolondi L, Sofia S, Siringo S, Gaiani S, Casali A, Zironi G, Piscaglia F, Gramantieri L, Zanetti M, Sherman M. Surveillance programme of cirrhotic patients for early diagnosis and treatment of hepatocellular carcinoma: a cost effectiveness analysis. Gut. 2001; 48:251–9. [PubMed: 11156649]

16. Caturelli E, Bartolucci F, Biasini E, Vigliotti ML, Andriulli A, Siena DA, Attino V, Bisceglia M. Diagnosis of liver nodules observed in chronic liver disease patients during ultrasound screening for early detection of hepatocellular carcinoma. Am J Gastroenterol. 2002; 97:397–405. [PubMed: 11866279]

17. Chalasani N, Horlander JC Sr, Said A, Hoen H, Kopecky KK, Stockberger SM Jr, Manam R, Kwo PY, Lumeng L. Screening for hepatocellular carcinoma in patients with advanced cirrhosis. Am J Gastroenterol. 1999; 94:2988–93. [PubMed: 10520857]

18. Chen TH, Chen CJ, Yen MF, Lu SN, Sun CA, Huang GT, Yang PM, Lee HS, Duffy SW. Ultrasound screening and risk factors for death from hepatocellular carcinoma in a high risk group in Taiwan. Int J Cancer. 2002; 98:257–61. [PubMed: 11857416]

19. Giardina MG, Matarazzo M, Morante R, Lucariello A, Varriale A, Guardasole V, De Marco G. Serum alpha-L-fucosidase activity and early detection of hepatocellular carcinoma: a prospective study of patients with cirrhosis. Cancer. 1998; 83:2468–74. [PubMed: 9874450]

20. Henrion J, Libon E, De Maeght S, Schapira M, Ghilain JM, Maisin JM, Heller FR. Surveillance for hepatocellular carcinoma: compliance and results according to the aetiology of cirrhosis in a cohort of 141 patients. Acta Gastroenterol Belg. 2000; 63:5–9. [PubMed: 10907311]

21. Izzo F, Cremona F, Ruffolo F, Palaia R, Parisi V, Curley SA. Outcome of 67 patients with hepatocellular cancer detected during screening of 1125 patients with chronic hepatitis. Ann Surg. 1998; 227:513–8. [PubMed: 9563539]

22. Larcos G, Sorokopud H, Berry G, Farrell GC. Sonographic screening for hepatocellular carcinoma in patients with chronic hepatitis or cirrhosis: an evaluation. AJR Am J Roentgenol. 1998; 171:433–5. [PubMed: 9694470]

23. Lok AS, Sterling RK, Everhart JE, Wright EC, Hoefs JC, Di Bisceglie AM, Morgan TR, Kim HY, Lee WM, Bonkovsky HL, Dienstag JL. Des-gamma-carboxy prothrombin and alpha-fetoprotein as biomarkers for the early detection of hepatocellular carcinoma. Gastroenterology. 2010; 138:493–502. [PubMed: 19852963]

24. Luo K, Liu Z, Karayiannis P. Effect of antiviral treatment on alfa-fetoprotein levels in HBV-related cirrhotic patients: early detection of hepatocellular carcinoma. J Viral Hepat. 2010; 17:511–7. [PubMed: 19818003]

25. Mok TS, Yeo W, Yu S, Lai P, Chan HL, Chan AT, Lau JW, Wong H, Leung N, Hui EP, Sung J, Koh J, Mo F, Zee B, Johnson PJ. An intensive surveillance program detected a high incidence of hepatocellular carcinoma among hepatitis B virus carriers with abnormal alpha-fetoprotein levels or abdominal ultrasonography results. J Clin Oncol. 2005; 23:8041–7. [PubMed: 16258102]

26. Oka H, Kurioka N, Kim K, Kanno T, Kuroki T, Mizoguchi Y, Kobayashi K. Prospective study of early detection of hepatocellular carcinoma in patients with cirrhosis. Hepatology. 1990; 12:680–7. [PubMed: 1698703]

27. Pateron D, Ganne N, Trinchet JC, Aurousseau MH, Mal F, Meicler C, Coderc E, Reboullet P, Beaugrand M. Prospective study of screening for hepatocellular carcinoma in Caucasian patients with cirrhosis. J Hepatol. 1994; 20:65–71. [PubMed: 7515408]

28. Paul SB, Sreenivas V, Gulati MS, Madan K, Gupta AK, Mukhopadhyay S, Panda SK, Acharya SK. Incidence of hepatocellular carcinoma among Indian patients with cirrhosis of liver: an experience from a tertiary care center in northern India. Indian J Gastroenterol. 2007; 26:274–8. [PubMed: 18431010]

29. Pocha C, Dieperink E, McMaken KA, Knott A, Thuras P, Ho SB. Surveillance for hepatocellular cancer with ultrasonography vs. computed tomography -- a randomised study. Aliment Pharmacol Ther. 2013; 38:303–12. [PubMed: 23750991]

30. Qian MY, Yuwei JR, Angus P, Schelleman T, Johnson L, Gow P. Efficacy and cost of a hepatocellular carcinoma screening program at an Australian teaching hospital. J Gastroenterol Hepatol. 2010; 25:951–6. [PubMed: 20546449]

31. Sangiovanni A, Del Ninno E, Fasani P, De Fazio C, Ronchi G, Romeo R, Morabito A, De Franchis R, Colombo M. Increased survival of cirrhotic patients with a hepatocellular carcinoma detected during surveillance. Gastroenterology. 2004; 126:1005–14. [PubMed: 15057740]

32. Santagostino E, Colombo M, Rivi M, Rumi MG, Rocino A, Linari S, Mannucci PM. A 6-month versus a 12-month surveillance for hepatocellular carcinoma in 559 hemophiliacs infected with the hepatitis C virus. Blood. 2003; 102:78–82. [PubMed: 12649165]

33. Sato T, Tateishi R, Yoshida H, Ohki T, Masuzaki R, Imamura J, Goto T, Kanai F, Obi S, Kato N, Shiina S, Kawabe T, Omata M. Ultrasound surveillance for early detection of hepatocellular carcinoma among patients with chronic hepatitis C. Hepatol Int. 2009; 3:544–50. [PubMed: 19669240]

34. Shah TU, Semelka RC, Pamuklar E, Firat Z, Gerber RD, Shrestha R, Russo MW. The risk of hepatocellular carcinoma in cirrhotic patients with small liver nodules on MRI. Am J Gastroenterol. 2006; 101:533–40. [PubMed: 16542290]

35. Shimauchi Y, Tanaka M, Kuromatsu R, Ogata R, Tateishi Y, Itano S, Ono N, Yutani S, Nagamatsu H, Matsugaki S, Yamasaki S, Tanikawa K, Sata M. A simultaneous monitoring of Lens culinaris agglutinin A-reactive alpha-fetoprotein and des-gamma-carboxy prothrombin as an early diagnosis of hepatocellular carcinoma in the follow-up of cirrhotic patients. Oncol Rep. 2000; 7:249–56. [PubMed: 10671666]

36. Singal AG, Conjeevaram HS, Volk ML, Fu S, Fontana RJ, Askari F, Su GL, Lok AS, Marrero JA. Effectiveness of hepatocellular carcinoma surveillance in patients with cirrhosis. Cancer Epidemiol Biomarkers Prev. 2012; 21:793–9. [PubMed: 22374994]

37. Solmi L, Primerano AM, Gandolfi L. Ultrasound follow-up of patients at risk for hepatocellular carcinoma: results of a prospective study on 360 cases. Am J Gastroenterol. 1996; 91:1189–94. [PubMed: 8651169]

38. Tong MJ, Blatt LM, Kao VW. Surveillance for hepatocellular carcinoma in patients with chronic viral hepatitis in the United States of America. J Gastroenterol Hepatol. 2001; 16:553–9. [PubMed: 11350553]

39. Tradati F, Colombo M, Mannucci PM, Rumi MG, De Fazio C, Gamba G, Ciavarella N, Rocino A, Morfini M, Scaraggi A, Taioli E. A prospective multicenter study of hepatocellular carcinoma in italian hemophiliacs with chronic hepatitis C. The Study Group of the Association of Italian Hemophilia Centers. Blood. 1998; 91:1173–7. [PubMed: 9454746]

40. Trinchet JC, Chaffaut C, Bourcier V, Degos F, Henrion J, Fontaine H, Roulot D, Mallat A, Hillaire S, Cales P, Ollivier I, Vinel JP, Mathurin P, Bronowicki JP, Vilgrain V, N'Kontchou G, Beaugrand M, Chevret S. Ultrasonographic surveillance of hepatocellular carcinoma in cirrhosis: a randomized trial comparing 3- and 6-month periodicities. Hepatology. 2011; 54:1987–97. [PubMed: 22144108]

41. Van Thiel DH, Yong S, Li SD, Kennedy M, Brems J. The development of de novo hepatocellular carcinoma in patients on a liver transplant list: frequency, size, and assessment of current screening methods. Liver Transpl. 2004; 10:631–7. [PubMed: 15108254]

42. Mancebo A, Gonzalez-Dieguez ML, Cadahia V, Varela M, Perez R, Navascues CA, Sotorrios NG, Martinez M, Rodrigo L, Rodriguez M. Annual incidence of hepatocellular carcinoma among patients with alcoholic cirrhosis and identification of risk groups. Clin Gastroenterol Hepatol. 2013; 11:95–101. [PubMed: 22982095]

43. Frey RS, Boldanova T, Heim M. Ultrasound surveillance for hepatocellular carcinoma: real-life performance in a hepatology outpatient clinic. Swiss Med Wkly. 2015; 145:w14200. [PubMed: 26588016]

44. Pinero F, Marciano S, Anders M, Orozco F, Zerega A, Cabrera CR, Bana MT, Gil O, Andriani O, de Santibanes E, McCormack L, Gadano A, Silva M. Screening for liver cancer during transplant waiting list: a multicenter study from South America. Eur J Gastroenterol Hepatol. 2015; 27:355–60. [PubMed: 25563142]

45. Chang TS, Wu YC, Tung SY, Wei KL, Hsieh YY, Huang HC, Chen WM, Shen CH, Lu CH, Wu CS, Tsai YH, Huang YH. Alpha-Fetoprotein Measurement Benefits Hepatocellular Carcinoma Surveillance in Patients with Cirrhosis. Am J Gastroenterol. 2015; 110:836–44. quiz 845. [PubMed: 25869392]

46. Kim SY, An J, Lim YS, Han S, Lee JY, Byun JH, Won HJ, Lee SJ, Lee HC, Lee YS. MRI With Liver-Specific Contrast for Surveillance of Patients With Cirrhosis at High Risk of Hepatocellular Carcinoma. JAMA Oncol. 2017; 3:456–463. [PubMed: 27657493]

47. Mehta A, Singal AG. Hepatocellular Carcinoma Surveillance: Does Alpha-Fetoprotein Have a Role? Gastroenterology. 2015; 149:816–7. [PubMed: 26231605]

48. Singal A, Volk ML, Waljee A, Salgia R, Higgins P, Rogers MA, Marrero JA. Meta-analysis: surveillance with ultrasound for early-stage hepatocellular carcinoma in patients with cirrhosis. Aliment Pharmacol Ther. 2009; 30:37–47. [PubMed: 19392863]

49. Bruix J, Sherman M. Management of Hepatocellular Carcinoma: An Update. Hepatology. 2010; 53:1–35.

50. Simmons O, Fetzer DT, Yokoo T, Marrero JA, Yopp A, Kono Y, Parikh ND, Browning T, Singal AG. Predictors of adequate ultrasound quality for hepatocellular carcinoma surveillance in patients with cirrhosis. Aliment Pharmacol Ther. 2017; 45:169–177. [PubMed: 27862091]

51. Del Poggio P, Olmi S, Ciccarese F, Di Marco M, Rapaccini GL, Benvegnu L, Borzio F, Farinati F, Zoli M, Giannini EG, Caturelli E, Chiaramonte M, Trevisani F. Factors that affect efficacy of ultrasound surveillance for early stage hepatocellular carcinoma in patients with cirrhosis. Clin Gastroenterol Hepatol. 2014; 12:1927–33 e2. [PubMed: 24582947]

52. Mokdad AA, Lopez AD, Shahraz S, Lozano R, Mokdad AH, Stanaway J, Murray CJ, Naghavi M. Liver cirrhosis mortality in 187 countries between 1980 and 2010: a systematic analysis. BMC Med. 2014; 12:145. [PubMed: 25242656]

53. Singal AG, Yopp A, CSS, Packer M, Lee WM, Tiro JA. Utilization of hepatocellular carcinoma surveillance among American patients: a systematic review. J Gen Intern Med. 2012; 27:861–7. [PubMed: 22215266]

54. Gopal P, Yopp AC, Waljee AK, Chiang J, Nehra M, Kandunoori P, Singal AG. Factors that affect accuracy of alpha-fetoprotein test in detection of hepatocellular carcinoma in patients with cirrhosis. Clin Gastroenterol Hepatol. 2014; 12:870–7. [PubMed: 24095974]

55. Yang JD, Dai J, Singal AG, Gopal P, Addissie BD, Nguyen MH, Befeler AS, Reddy KR, Schwartz M, Harnois DM, Yamada H, Gores GJ, Feng Z, Marrero JA, Roberts LR. Improved Performance of Serum Alpha-Fetoprotein for Hepatocellular Carcinoma Diagnosis in HCV Cirrhosis with Normal Alanine Transaminase. Cancer Epidemiol Biomarkers Prev. 2017; 26:1085–1092. [PubMed: 28258053]

56. Lee E, Edward S, Singal AG, Lavieri MS, Volk M. Improving screening for hepatocellular carcinoma by incorporating data on levels of alpha-fetoprotein, over time. Clin Gastroenterol Hepatol. 2013; 11:437–40. [PubMed: 23247324]

57. Tayob N, Lok AS, Do KA, Feng Z. Improved Detection of Hepatocellular Carcinoma by Using a Longitudinal Alpha-Fetoprotein Screening Algorithm. Clin Gastroenterol Hepatol. 2016; 14:469–475 e2. [PubMed: 26260109]

58. Atiq O, Tiro J, Yopp AC, Muffler A, Marrero JA, Parikh ND, Murphy C, McCallister K, Singal AG. An assessment of benefits and harms of hepatocellular carcinoma surveillance in patients with cirrhosis. Hepatology. 2017; 65:1196–1205. [PubMed: 27775821]

59. Verma, A., Konerman, MA., Zhao, B., Lok, AS., Parikh, ND. Frequency and work-up of indeterminate nodules during surveillance of hepatocellular carcinoma in patients with cirrhosis. The Liver Meeting; 2016 Nov 11–15; Boston MA. AASLD; 2016. Abstract 434

60. Harris RP, Sheridan SL, Lewis CL, Barclay C, Vu MB, Kistler CE, Golin CE, DeFrank JT, Brewer NT. The harms of screening: a proposed taxonomy and application to lung cancer screening. JAMA Intern Med. 2014; 174:281–5. [PubMed: 24322781]

61. Lee CI, Haims AH, Monico EP, Brink JA, Forman HP. Diagnostic CT scans: assessment of patient, physician, and radiologist awareness of radiation dose and possible risks. Radiology. 2004; 231:393–8. [PubMed: 15031431]

62. Barrett BJ, Carlisle EJ. Metaanalysis of the relative nephrotoxicity of high- and low-osmolality iodinated contrast media. Radiology. 1993; 188:171–8. [PubMed: 8511292]

63. Dill T. Contraindications to magnetic resonance imaging: non-invasive imaging. Heart. 2008; 94:943–8. [PubMed: 18552230]

64. Besa C, Lewis S, Pandharipande PV, Chhatwal J, Kamath A, Cooper N, Knight-Greenfield A, Babb JS, Boffetta P, Padron N, Sirlin CB, Taouli B. Hepatocellular carcinoma detection: diagnostic performance of a simulated abbreviated MRI protocol combining diffusion-weighted and T1-

weighted imaging at the delayed phase post gadoxetic acid. Abdom Radiol (NY). 2017; 42:179–190. [PubMed: 27448609]

65. Goossens N, Singal AG, King LY, Andersson KL, Fuchs BC, Besa C, Taouli B, Chung RT, Hoshida Y. Cost-Effectiveness of Risk Score-Stratified Hepatocellular Carcinoma Screening in Patients with Cirrhosis. Clin Transl Gastroenterol. 2017; 8:e101. [PubMed: 28640287]

Figure 1.
Sensitivity of ultrasound alone to detect hepatocellular carcinoma at a) any stage and b) an early stage (within Milan Criteria)

Author, Year		Risk Ratio (95% CI)
Pateron 1994		0.60 (0.18 - 2.04)
Henrion 2000		0.69 (0.38 - 1.25)
Lok 2010		0.70 (0.42 - 1.18)
Qian 2010		0.88 (0.61 - 1.27)
Trinchet 2011		0.87 (0.74 - 1.03)
Singal 2012		0.50 (0.30 - 0.83)
Kim 2016		0.79 (0.40 - 1.53)
Pooled Risk Ratio I-squared 0%		0.81 (0.71 - 0.93)

0.17   1   5.67

Figure 2.
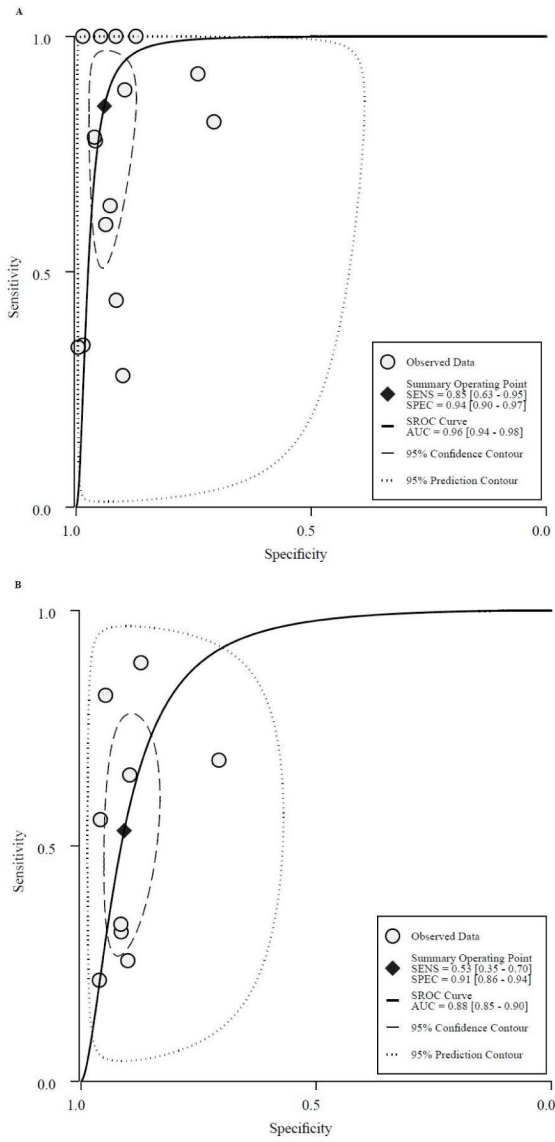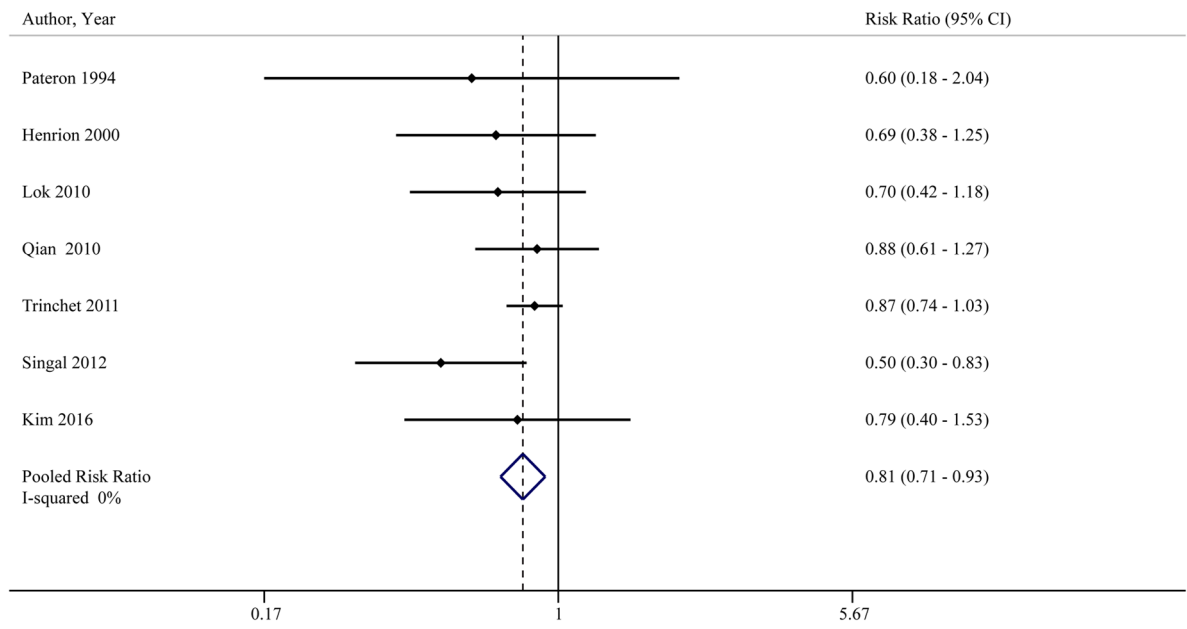Overall accuracy of ultrasound alone to detect hepatocellular carcinoma at a) any stage and b) an early stage (within Milan Criteria)
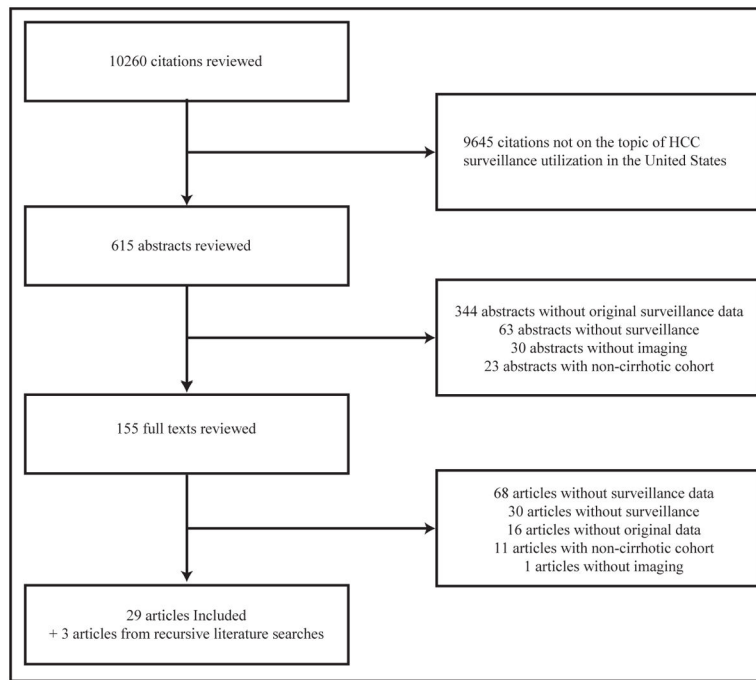
Figure 3.
Sensitivity of ultrasound with and without AFP to detect HCC at an early stage

Table 1

Characteristics of Studies

Author Year	Study Period	Study location	Study Design	Number patients (% cirrhosis)	Number HCC (% early HCC)	Imaging modality	Sensitivity any stage HCC (sensitivity early HCC)	Specificity HCC	Follow-up (months)
Oka 1990 *	1983–1988	Japan	Prospective	140 (100%)	40 (NR)	Ultrasound	82.5% (NR)	NR	Median 36 (2–72)
Pateron 1994 **	1986–1990	France	Prospective	118 (100%)	14 (35.7%)	Ultrasound	78.6% (21.4%)	96.2%	Median 36 (4–48)
Solmi 1996 *	1988–1993	Italy	Prospective	360 (70.6%)	24 (NR)	Ultrasound	100% (NR)	NR	Mean 56 (18–72)
Tradati 1998	1992–1996	Italy	Prospective	40 (100%)	6 (33.3%)	Ultrasound	100% (33.3%)	NR	48
Giardina 1998	1988–1995	Italy	Prospective	132 (100%)	19 (NR)	Ultrasound	100% (NR)	NR	Mean 64 (6–86)
Larcos 1998 *	Not reported	Australia	Retrospective	232 (>50%)	6 (33.3%)	Ultrasound	100% (33.3%)	91.6%	96
Izzo 1998 *	1993–1996	Italy	Prospective	325 (>50%)	67 (NR)	Ultrasound	86.6% (NR)	NR	NR
Chalassani 1999	1994–1997	United States	Retrospective	285 (100%)	27 (NR)	Ultrasound	59.3% (NR)	92.8%	Median 15 (6–42)
Shimauchi 2000 *	1990–1995	Japan	Retrospective	78 (100%)	21 (NR)	Ultrasound	76.2% (NR)	NR	Mean 42 ± 1.6
Henrion 2000 **	1995–1998	France	Prospective	141 (100%)	6 (100%)	Ultrasound	66.7% (66.7%)	NR	Median 34 (3–42)
Tong 2001	1991–1999	United States	Prospective	173 (100%)	31 (NR)	Ultrasound	100% (NR)	NR	Mean 35 (12–103)
Bolondi 2001	1989–1998	Italy	Prospective	313 (100%)	61 (82.0%)	Ultrasound	100% (82.0%)	94.8%	Mean 56 (6–100)
Caturelli 2002	1992–1997	Italy	Prospective	1599 (>50%)	269 (NR)	Ultrasound	100% (NR)	98.6%	84
Chen 2002	1991–1998	Taiwan	Prospective	292 (100%)	49 (NR)	Ultrasound	61.2% (NR)	NR	Mean 84
Santagostino 2003	1996–2001	Italy	Prospective	66 (100%)	8 (25.0%)	Ultrasound	100% (25%)	NR	72
Sangiovanni 2004	1987–2001	Italy	Prospective	417 (100%)	112 (24.1%)	Ultrasound	100% (24.1%)	NR	Mean 148 (1–213)

Author Year	Study Period	Study location	Study Design	Number patients (% cirrhosis)	Number HCC (% early HCC)	Imaging modality	Sensitivity any stage HCC (sensitivity early HCC)	Specificity HCC	Follow-up (months)
Van Thiel 2004 *	1998–2003	United States	Retrospective	100 (100%)	20 (NR)	CT	60% (NR)	100%	Mean 96
						Ultrasound	70% (NR)	93.8%	
Mok 2005 *	1997–2004	Hong Kong	Prospective	940 (100%)	32 (NR)	Ultrasound	34.4% (NR)	98.6%	Median 49
Shah 2006	2001–2004	United States	Prospective	310 (100%)	22 (NR)	MRI	77.3% (NR)	82.6%	Mean 22
Paul 2007	2001–2004	India	Prospective	194 (100%)	9 (44.4%)	Ultrasound	100% (44.4%)	NR	Median 26 (0 – 181)
Sato 2009 *	1994–2004	Japan	Retrospective	1431 (>50%)	243 (NR)	Ultrasound	90.9% (NR)	NR	Mean 73
Luo 2010 *	Not reported	China	Prospective	93 (100%)	16 (NR)	Ultrasound	75% (NR)	NR	NR
Qian 2010 **	1998–2004	Australia	Retrospective	268 (91.4%)	22 (81.8%)	Ultrasound	81.8% (68.2%)	70.7%	Mean 36
Lok 2010 **		United States	Prospective	116 (56.9%)	39 (61.5%)	Ultrasound	56.4% (35.9%)	NR	46
Trinchet 2011 **	2000–2009	France	Prospective	1278 (100%)	123 (74.8%)	Ultrasound	88.6% (65.0%)	89.7%	Mean 47
Singal 2012 **	2004–2006	United States	Prospective	442 (100%)	41 (73.2%)	Ultrasound	43.9% (31.7%)	91.5%	Median 42 (7 – 79)
Pocha 2013	2002–2011	United States	Prospective	163 (100%)	17 (58.8%)	CT	77.8% (62.5%)	95.9%	Mean 31
						Ultrasound	88.9% (55.5%)	87.5%	
Mancebo 2013 *	1992–2010	Spain	Prospective	450 (100%)	62 (66.1%)	Ultrasound	88.7% (NR)	NR	Median 42
Frey 2015	2011–2012	Switzerland	Retrospective	285 (87.0%)	9 (88.9%)	Ultrasound	100% (88.9%)	87.3%	24
Pinero 2015	2005–2011	Argentina	Retrospective	572 (100%)	56 (NR)	Ultrasound	33.9% (NR)	99.6%	Median 3 (1 – 7)
Chang 2015 *	2002–2010	Taiwan	Retrospective	1597 (100%)	363 (58.7%)	Ultrasound	92.0% (NR)	74.1%	Median 57 (17 – 144)
Kim 2016 **	2011–2014	Korea	Prospective	407 (100%)	43 (97.7%)	MRI	86.0% (83.7%)	94.2%	18
						Ultrasound	27.9% (25.6%)	90.1%	

HCC – hepatocellular carcinoma

NR – not reported

*
Comparative studies of ultrasound *vs.* ultrasound + AFP for any stage HCC detection

**
Comparative studies of ultrasound *vs.* ultrasound + AFP for early stage HCC detection

Table 2

Quality Assessment of Studies by QUADAS-2 Checklist

Author Year	Patient Selection		Index Test		Reference Standard		Flow and Timing
	Risk of Bias	Applicability Concerns	Risk of Bias	Applicability Concerns	Risk of Bias	Applicability Concerns	Risk of Bias
Oka 1990	*	*	*	*	High	*	High
Pateron 1994	*	*	*	*	*	*	High
Solmi 1996	*	High	*	*	High	*	High
Tradati 1998	*	*	*	*	High	*	High
Giardina 1998	*	*	*	*	*	*	High
Larcos 1998	*	High	*	*	*	*	High
Izzo 1998	*	High	*	*	*	*	High
Chalassani 1999	High	*	*	*	*	*	*
Shimauchi 2000	*	*	*	*	*	*	*
Henrion 2000	*	*	*	*	*	*	High
Tong 2001	*	*	*	*	*	*	High
Bolondi 2001	*	*	*	*	*	*	High
Caturelli 2002	*	High	*	*	High	*	High
Chen 2002	*	*	*	*	High	*	High
Santagostino 2003	*	*	*	*	*	*	High
Sangiovami 2004	*	*	*	*	High	*	High
Van Thiel 2004	High	*	*	*	*	*	*
Mok 2005	*	*	*	*	*	*	High
Shah 2006	High	*	*	*	*	*	*
Paul 2007	*	*	*	*	High	*	High
Sato 2009	*	High	*	*	*	*	High
Luo 2010	High	High	High	*	*	*	High
Qian 2010	*	High	*	*	High	*	High
Lok 2010	High	High	*	*	High	*	High
Trinchet 2011	*	*	*	*	High	*	High
Singal 2012	*	*	*	*	High	*	*

Author Year	Patient Selection		Index Test		Reference Standard		Flow and Timing
	Risk of Bias	Applicability Concerns	Risk of Bias	Applicability Concerns	Risk of Bias	Applicability Concerns	Risk of Bias
Pocha 2013	*	*	*	*	High	*	High
Mancebo 2013	*	*	*	*	High	*	High
Frey 2015	*	High	*	*	*	*	High
Pinero 2015	High	*	*	*	*	*	*
Chang 2015	*	*	*	*		*	High
Kim 2016	*	*	*	*	High	*	*

High = high risk of bias;

*
= low risk of bias

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Performance of ultrasound with and without AFP for early HCC detection

Author Year	AFP Cut-off	Sensitivity early HCC		Specificity	
		Ultrasound alone	Ultrasound plus AFP	Ultrasound alone	Ultrasound plus AFP
Pateron 1994	15 ng/mL	21.4% (4.7% – 50.8%)	35.7% (12.8% – 64.9%)	96.2% (90.4% – 98.9%)	82.7% (74.0% – 89.4%)
Henrion 2000	200 ng/mL	66.7% (22.3% – 95.7%)	100% (60.7% – 100%)	Not reported	Not reported
Qian 2010	Not reported	68.2% (45.1% – 86.1%)	77.3% (54.6% – 92.2%)	Not reported	Not reported
Lok 2010	Not reported	35.9% (21.2% – 52.8%)	51.3% (34.8% – 67.6%)	Not reported	Not reported
Trinchet 2011	20 ng/mL	65.0% (55.9% – 73.4%)	74.8% (66.2% – 82.2%)	89.7% (87.8% – 91.4%)	82.9% (80.6% – 85.0%)
Singal 2012	20 ng/mL	31.7% (18.1% – 48.1%)	63.4% (46.9% – 77.9%)	91.5% (88.4% – 94.1%)	83.3% (79.3% – 86.8%)
Kim 2016	20 ng/mL	25.6% (13.5% – 41.2%)	32.6% (19.1% – 48.5%)	90.1% (86.6% – 93.0%)	88.2% (84.4% – 91.3%)

AFP – alpha fetoprotein

HCC – hepatocellular carcinoma