DR. NEEHAR DILIP PARIKH (Orcid ID : 0000-0002-5874-9933)

Article type   : Original Article

GALAD Demonstrates High Sensitivity for HCC Surveillance in a Cohort of Patients with Cirrhosis

Amit G. Singal MD MS[1,2], Nabihah Tayob PhD[3], Anand Mehta PhD[4], Jorge A. Marrero MD MS[1,2], Hashem El-Serag MD MPH[5], Qingchun Jin MPH[3], Cristian Saenz de Viteri BS[6], Austin Fobar MS[6], Neehar D. Parikh MD MS[6]

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

[2]Harold C. Simmons Cancer Center, UT Southwestern Medical Center, Dallas, TX

[3]Department of Biostatistics, Dana Farber Cancer Center, Boston MA

[4]Department of Internal Medicine, Medical University of South Carolina, Charleston, SC

[5]Department of Internal Medicine, Baylor College of Medicine, Houston, TX

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

Correspondence: Amit G. Singal, M.D., M.S.

Division of Digestive and Liver Diseases

UT Southwestern Medical Center

5959 Harry Hines Blvd, POB 1, Suite 420

Dallas TX 75390-8887

Tel: 214-645-6111

Fax: 214-645-6114

Email: amit.singal@utsouthwestern.edu

ABSTRACT

Background: Most patients with hepatocellular carcinoma (HCC) are diagnosed at a late stage, highlighting the need for more accurate surveillance tests. Although biomarkers for HCC early detection have promising data in phase II case-control studies, evaluation in cohort studies is critical prior to adoption in practice.

Methods: We leveraged a prospective cohort of patients with Child Pugh A or B cirrhosis who were followed until incident HCC, liver transplantation, death, or lost to follow-up. We used a prospective specimen-collection, retrospective-blinded-evaluation (PRoBE) design for biomarker evaluation of GALAD, longitudinal GALAD and the HES algorithm –compared to alpha fetoprotein (AFP) – using patient-level sensitivity and screening-level specificity.

Results: Of 397 patients with cirrhosis, 42 patients developed HCC (57.1% early-stage) over a median of 2.0 years. Longitudinal GALAD had the highest c-statistic for HCC detection (0.85, 95%CI 0.77 – 0.92), compared to single-timepoint GALAD (0.79, 95%CI 0.71 – 0.87), AFP (0.77, 95%CI 0.69 – 0.85), and HES (0.76, 95%CI 0.67 – 0.83). When specificity was fixed at 90%, the sensitivity for HCC of single-timepoint and longitudinal GALAD was 54.8% and 66.7%, respectively, compared to 40.5% for AFP. Sensitivity for HCC detection was higher when restricted to patients with biomarker assessment within 6 months prior to HCC diagnosis, with the highest sensitivities observed for single-timepoint (72.0%) and longitudinal GALAD (64.0%), respectively. Sensitivity of single-timepoint and longitudinal GALAD for early-stage HCC was 53.8% and 69.2%, respectively.

Conclusion: GALAD demonstrated high sensitivity for HCC detection in a cohort of patients with cirrhosis. Validation of these results are warranted in large phase III datasets.

Lay Summary: A blood-based panel including age, sex, and three biomarkers was able to accurately detect liver cancer in at-risk patients with cirrhosis. These data highlight the potential value of blood-based screening tests to improve early detection of liver cancer.

Summary Box

What is already known about this subject?
1. Ultrasound and AFP are currently recommended for HCC surveillance but miss over one-third of HCC at an early stage
2. Several biomarkers have demonstrated promising early data in case-control studies but require validation in phase III cohort studies

What are the new findings?
1. Single-biomarker strategies fail to achieve sufficient sensitivity for early-stage HCC detection
2. In phase III biomarker evaluation leveraging a cohort of patients with cirrhosis, GALAD demonstrated high sensitivity for early-stage HCC detection.

3. Longitudinal GALAD values may further increase sensitivity for early-stage HCC detection

How might it impact on clinical practice in the forseeable future?

1. GALAD demonstrates high sensitivity for early-stage HCC and is promising as an alternative surveillance strategy for HCC in patients with cirrhosis

INTRODUCTION

Hepatocellular carcinoma (HCC) is the fastest increasing cause of cancer-related death in the U.S. and one of the leading causes of death in patients with compensated cirrhosis.[1] However, prognosis is driven by tumor stage, with curative options available if patients are detected at an early stage, affording long-term survival.[2-3] In contrast, patients detected with more advanced tumor burden are only eligible for palliative therapies and have a median survival of 1-2 years. The close association between early detection and improved survival underlie the recommendation of HCC surveillance in patients with cirrhosis by multiple professional societies.[4-5]

HCC surveillance is performed using semi-annual abdominal ultrasound, with or without alpha fetoprotein (AFP); however, data have shown that these tests typically have a sensitivity of only 60% for early stage HCC detection.[6] Further, poor ultrasound performance may be increasingly problematic as the epidemiology of cirrhosis changes given inadequate ultrasound visualization in patients with obesity and nonalcoholic fatty liver disease (NAFLD).[7-8] This poor performance can also lead to increased risk of screening-related harms due to false positive or indeterminate screening results.[9-10] Finally, the dependence on ultrasound-based surveillance often requires patients to come in for separate radiology appointments, which can drive underuse of HCC surveillance in clinical practice.[11-12] Overall, these limitations highlight the need for more effective HCC surveillance tests, which can increase early HCC detection.

Several serum-based biomarkers and biomarker panels have promising data suggesting high sensitivity and specificity for HCC in patients with cirrhosis.[13] Alpha fetoprotein (AFP) remains the only serum biomarker to undergo all five phases of biomarker validation for HCC surveillance[14]; however single biomarkers have suboptimal performance for early HCC detection, likely related to tumor heterogeneity. More recent algorithms, such as GALAD and HES, combine multiple biomarkers, with or without demographic and clinical features, and have achieved higher sensitivity for early HCC detection, exceeding 70%; however, prior data are largely limited to phase II case-control studies, which can overestimate biomarker performance.[13] An expert panel guidance document from the International Liver Cancer Association recently stressed the critical importance of longitudinal cohort studies (phase III validation) to determine if biomarkers can detect cancer early before it becomes clinically evident; however, this has been limited by a dearth of cohorts with available clinical samples.[14] Further, most prior phase II studies have evaluated a single biomarker or biomarker panel, and there are few data comparing several HCC early detection biomarkers in a single cohort. Finally, recent data suggest incorporation of longitudinal data may improve biomarker performance by identifying earlier increases in biomarker values suggesting early stage HCC as well as reducing the risk of false positive results; however, most data evaluate biomarkers at a single time point (i.e., single threshold).[15] Therefore, the aim of our study was to compare multiple biomarkers – including single threshold AFP, AFP-L3%, des-gamma-carboxy prothrombin (DCP), the HCC Early Detection Screening (HES) algorithm, single threshold GALAD, and longitudinal GALAD– in a longitudinal cohort of patients with cirrhosis.

METHODS

Study Population

We leveraged a previously described cohort of patients with cirrhosis from the University of Michigan who were enrolled into a surveillance program between January 2004 and September 2006.[16] In brief, all patients had Child Pugh A or B cirrhosis, without known HCC or suspicious liver lesions, at enrollment. Cirrhosis was defined based on compatible histology or imaging

showing a cirrhotic appearing liver with signs of portal hypertension. Other exclusion criteria included significant hepatic decompensation (refractory ascites, grade 3-4 encephalopathy, or hepatorenal syndrome), co-morbid medical conditions with a life expectancy of less than one year, prior solid organ transplant, and known extrahepatic primary tumor.

Patients were prospectively followed with semi-annual ultrasound-based HCC surveillance until incident HCC, liver transplantation, death, or lost to follow-up. HCC was defined using AASLD criteria, i.e., histology or characteristic imaging in lesions ≥1 cm, and early-stage HCC was defined as Barcelona Clinic Liver Cancer (BCLC) stage 0 or A. The Social Security Death File and the State of Michigan Death Records were used to ascertain date of death for any patients lost to follow-up. Patients were not directly involved in the design, conduct, or reporting of the research. The study was approved by IRBs at the University of Michigan (HUM00046376) and UT Southwestern Medical Center (STU 082017-013).

Data and Blood Collection

Demographic and clinical data were collected at enrollment, including age, gender, race/ethnicity liver disease etiology, and Child Pugh score. Liver disease etiology was classified as hepatitis C virus (HCV)-related (presence of HCV antibody or RNA), hepatitis B virus (HBV)-related (presence of HBV surface antigen), alcohol-related liver disease (alcohol intake >40 gm/day for >10 years), non-alcoholic steatohepatitis (NASH) (absence of other etiologies with metabolic syndrome), other (e.g., hereditary hemochromatosis, primary sclerosing cholangitis, primary biliary cirrhosis), or cryptogenic. Serum and plasma were collected from all patients at each visit and stored at -80C, without interval thawing.

Biomarker Evaluation

Biomarker evaluation was performed using a prospective-specimen-collection, retrospective-blinded-evaluation (PRoBE) design.[17] For this study, biomarkers were evaluated at multiple time points during follow-up, so we are able to evaluate algorithms that incorporate serial biomarker measurements into the screening decision. We compared these algorithms to approaches that only consider biomarker levels measured at a single visit in screening decisions.

Serum from each visit for cases and controls was transferred to Wako Diagnostics lab for AFP,

AFP-L3 and DCP measurements.[18] AFP, AFP-L3%, and DCP were performed using a microchip capillary electrophoresis and liquid-phase binding assay on a μTASWako i30 auto analyzer (Wako Pure Chemical Industries, Ltd. Osaka, Japan). For this study, we evaluated highly sensitive AFP-L3 (hs-AFP-L3), which can be measured at lower AFP levels and lower AFP-L3 percentages than the older AFP-L3 assay.[19] All assays were performed blinded to HCC vs. non-HCC status.

AFP, AFP-L3, and DCP were assessed based on the biomarker levels at each single visit. HES is an AFP-adjusted algorithm that combines current AFP and change in AFP over the last year with age, platelets, ALT, and interaction terms.[22-23] The algorithm was recently updated to include etiology of cirrhosis and this version was used for this analysis.[23]

GALAD was calculated at each time point using the equation: $Z = -10.08 + (0.09 \times age) + (1.67 \times male\ sex) + (2.34 * \log AFP) + (0.04 \times AFP\text{-}L3) + (1.33 \times \log DCP)$.[20] In addition, we utilized the parametric empirical Bayes algorithm to evaluate longitudinal GALAD where a personalized threshold that incorporates GALAD screening history was used for every patient at each screening occasion.[15,21] In this analysis, a weighted average of mean GALAD in the non-HCC population and average of prior GALAD values for each patient defines a patient-specific threshold for defining an abnormal screening result. If a patient has no screening history, the longitudinal and single threshold rules are equivalent; however, longitudinal GALAD depends more on an individual's screening history as he/she accumulates more screening tests.

Statistical Analysis

We evaluated single-timepoint biomarkers and longitudinal biomarkers using the entire cohort (**Supplemental Figure 1**). To incorporate longitudinal screening history for GALAD, we used the PEB estimate, a weighted average of mean biomarker values in the non-HCC population and average of prior biomarker values for each patient to obtain a subject-specific threshold.[21] Parameters of the PEB model were estimated using only data from non-HCC patients to define the control population mean. In those without prior results, the PEB algorithm reduces to a standard fixed threshold approach; while for those with prior results, the PEB algorithm depends more on an individual's history as he/she accumulates more results.

Biomarker performance was evaluated using three complementary measures: 1) area under the

receiver operating characteristic (AUROC) curve, 2) sensitivity, i.e., patient-level true-positive rate (TPR), and screening-level specificity (1-false positive rate (FPR)) at established cut-offs from prior literature, and 3) sensitivity at a cut-off fixing screening-level specificity at 90%. Specificity was assessed at the screening level because each false positive can lead to diagnostic evaluation resulting in physical, financial and psychological harms. We estimated sensitivity at the patient level, so this was defined as the proportion of HCC cases with at least 1 positive screening during the pre-diagnostic period. The pre-diagnostic period was also separated into two windows (0-6 months and 7-12 months) prior to HCC diagnosis. Our primary outcome was any-stage HCC detection, and a secondary outcome was early-stage HCC detection. To calculate 95% confidence intervals for biomarker performance, we used a bootstrap procedure among 2000 datasets, each of which was constructed by randomly sampling patients with replacement. All analyses were conducted using R v4.0.3.

RESULTS

Patient Characteristics

Of 397 eligible patients with cirrhosis, 42 developed HCC over a median follow-up of 2.0 years (IQR 0.5 – 4.25 years). Patient characteristics are detailed in **Table 1**. Median age of the cohort was 52.0 years, 59.9% were male, and majority (>90%) were non-Hispanic White. The most common etiologies of liver were hepatitis C infection (47.4%), alcohol-related liver disease (15.4%), and NAFLD (19.6%). Median Child Pugh score was 7, with 40.1% having Child Pugh A cirrhosis and 55.2% Child Pugh B cirrhosis. Of the 42 patients who developed HCC, 57.1% had BCLC stage 0/A HCC, with the majority (59.5%) having unifocal HCC and median tumor size of 2.4 cm. Patients had a median of 3 (range 1-9) longitudinal visits in the analysis dataset for biomarker evaluation.

Biomarker Performance

Overall Discrimination

The overall accuracy for each biomarker at any time prior to diagnosis is illustrated in **Figure 1A**. DCP and HES had the lowest AUROCs, 0.71 and 0.76 respectively, single-timepoint GALAD had an AUROC of 0.79 and longitudinal GALAD had the highest AUC of 0.85, although these differences were not statistically significant. When considering performance for early-stage HCC, similar results were seen, with DCP and HES continuing to have the lowest AUROCs (0.70 – 0.71), single-timepoint GALAD having an intermediate AUROC (0.78), and longitudinal GALAD having the highest AUROC (0.83) (**Figure 1B**).

Accuracy using Established Cut-offs

The performance for each biomarker using previously reported cut-offs is shown in **Table 2**. AFP, DCP and HES each had sensitivity below 50% for any-stage and early-stage HCC but maintained specificity around 90%. Higher sensitivity was observed with AFP-L3% (66.7%; 95%CI 52.0-81.3%) and single-timepoint GALAD (57.1%; 95%CI 41.9-72.4%), although AFP-L3% had lower specificity than GALAD (82.7% vs. 86.5%). Both AFP-L3% and GALAD had lower specificity than AFP, DCP, and HES. Subgroup analyses, stratified by sex, liver disease etiology, and Child Pugh score are described in **Supplemental Table 1**. Sensitivity of both AFP-L3% and GALAD notably improved when restricted to results within six months of HCC diagnosis, with both demonstrating a sensitivity of 73.7% (95%CI 52.6-93.3%) for early-stage HCC.

Accuracy with Specificity at 90%

We next explored biomarker thresholds that correspond to an acceptable threshold of 90% for screening-level specificity in our cohort. A higher threshold of -0.33 (compared to the previously reported threshold of -0.63) was identified for single-timepoint GALAD**,** whereas estimated thresholds for other biomarkers were comparable to previously published thresholds (AFP: 17.4ng/ml vs 20ng/ml; AFP-L3: 11.9% vs 10%; DCP: 5.0ng/ml vs 7.5ng/ml, respectively). At these cut-offs AFP, DCP, AFP-L3%, and HES each had sensitivity below 50% for any-stage HCC, whereas the highest sensitivity was observed with single-timepoint GALAD (54.8%; 95%CI 39.5 - 70.2%) and longitudinal GALAD (66.7%; 95%CI 51.3 – 80.8%) (**Table 3**). DCP, AFP-L3%, and HES also had the lowest sensitivity for early-stage HCC detection, with each demonstrating lower sensitivity than that of AFP (50.0%; 95%CI 28.0 - 69.0%). Single-timepoint GALAD had a

sensitivity at 53.8% (95%CI 33.3% - 73.3%) for early HCC detection, although this was higher at 73.7% (95%CI 52.0 – 93.3%) when restricted to results within 6 months of HCC diagnosis. Longitudinal GALAD appeared to have preserved high sensitivity for early HCC detection, exceeding 65%, independent of time frame.

DISCUSSION

Abdominal ultrasound with or without AFP have served as the backbone of HCC surveillance testing for over two decades. Increasing data demonstrating suboptimal sensitivity for early HCC detection highlight the importance of novel surveillance strategies. Our study extends prior literature by evaluating several biomarkers in a cohort of patients with cirrhosis, serving as a transition from phase II to pilot phase III biomarker evaluation. Our results first highlight the need to improve upon AFP's performance given sensitivity of only 50% for early-stage HCC detection. We found other single-biomarker strategies, such as AFP-L3% and DCP, also fail to achieve sufficient sensitivity, and the highest sensitivity was observed with biomarker panels. Our results demonstrate that GALAD is a promising biomarker panel, with high sensitivity for early HCC detection – whether used in a single-timepoint or longitudinal manner. However, we show that GALAD thresholds may require further adjustments when validated in larger longitudinal cohort to optimize performance.

GALAD incorporates three biomarkers (AFP, AFP-L3%, and DCP) and has demonstrated promising accuracy in several case-control studies – higher than the each of the biomarkers alone.[20,24-25] GALAD also incorporates two demographic risk factors for HCC – gender and age – which are readily available and increase performance compared to biomarkers alone. The increased accuracy of a panel including several biomarkers, compared to a single biomarker, is not surprising given the observed heterogeneity of HCC.[26-27] For example, GALAD, at a cut-off of -0.63, had a sensitivity and specificity of 79% each for early HCC detection in the multi-center EDRN case-control dataset from the US.[26] Similarly, GALAD demonstrated sensitivity and specificity exceeding 80% for early HCC detection in a multi-national study including over 6500 patients from the UK, Germany, Japan, and Hong Kong.[27] Model performance of GALAD in this

study did not appear to significantly differ between patients with viral and non-viral etiologies of liver disease. GALAD was also shown to have high test performance in a multi-site case-control study among patients with NASH-related HCC.[28] However, case-control studies can over-estimate biomarker performance, highlighting the importance of cohort studies to evaluate the performance of a biomarker to detect preclinical disease. In this cohort study, GALAD achieved a sensitivity of 70% for HCC detection when assessed within 6 months of HCC diagnosis. This performance compares favorably to the performance of ultrasound, which has a sensitivity below 50% for early HCC detection – both as assessed in a systematic review of the literature as well as reported in the original description of this cohort.[6,16] A single-center study from Mayo Clinic suggested GALAD may be complementary to ultrasound, with an AUC of 0.97 compared to 0.92 and 0.82 for GALAD and ultrasound alone, respectively[26]; however, this strategy would still require patients to attend both ultrasound and phlebotomy visits. Our results suggest that biomarkers with sufficiently high sensitivity in larger cohorts may instead supplant imaging-based surveillance.

Our work extends this prior literature by demonstrating that single-timepoint GALAD, as evaluated in prior studies, has high sensitivity when conducted within six months of HCC diagnosis, although its sensitivity is lower at earlier time points. This limitation may be partly addressed by incorporating longitudinal changes in GALAD measurements over time, which demonstrated more consistent sensitivity to detect preclinical disease over longer periods of time prior to HCC diagnosis. Incorporation of PEB longitudinal analysis was previously shown to significantly increase sensitivity of AFP for HCC detection in a secondary analysis of HALT-C.[29] Similarly, a pilot study of a small Japanese cohort suggested that GALAD scores may increase approximately 1.5 years before HCC diagnosis.[28] Notably, longitudinal measures of a single biomarker, such as evaluated by the HES algorithm, did not achieve similar early HCC detection in our cohort as longitudinal changes in multiple biomarkers, as evaluated by longitudinal GALAD.

HCC surveillance effectiveness is driven by both test accuracy as well as utilization.[30] Several studies have shown that ultrasound is operator dependent, with large site-to-site variation in quality and test performance.[6,31] Biomarker-based surveillance may offer a path to standardize

test performance across sites. Further, ultrasound-based surveillance is underused in clinical practice due to both patient- and provider-level barriers, with only one-fourth of cirrhosis patients undergoing surveillance.[12, 32-33] For example, patients report transportation, financial, and logistical barriers to surveillance, which translate into lower adherence with surveillance recommendations.[12] These barriers appear to be particularly problematic among racial/ethnic minority and socioeconomically disadvantaged patients, which are also the populations disproportionately impacted by HCC. Blood-based biomarkers have the advantage of being easy to implement in practice, across all types of clinical settings, as they can be checked with routine labs at the time of a clinic visit. Patients also appear to prefer biomarker-based surveillance to ultrasound, if it can achieve adequate sensitivity for early-stage HCC detection.[34] Therefore, a blood-based biomarker could improve surveillance effectiveness even if it has similar sensitivity for early HCC detection as ultrasound-based surveillance.

While our results are encouraging for biomarker-based surveillance, the limited number of incident HCCs resulted in wide confidence intervals and hence preclude us from making statements about statistically significant improvements. Similarly, the potential performance of longitudinal GALAD may have been underestimated in this study given the relatively short duration of follow-up compared to larger phase III studies such as EDRN HEDS and Texas HCC Consortium. Our study's sample size also limited our ability to conduct meaningful subgroup analyses to see if biomarker performance differed by important factors such as sex and liver disease etiology. There are ongoing large phase III HCC biomarker efforts including the EDRN HEDS and Texas HCC Consortium cohorts, which should allow further evaluation of early detection biomarkers in the near future.[35-36] We also acknowledge other limitations of our study, including the older nature of our cohort with a higher proportion of active hepatitis C infection than observed in contemporary cohorts. While all patients with HBV infection were on antiviral treatment, all but two non-HCC patients with HCV infection had active viremia. Notably, prior studies have not suggested any difference in performance of GALAD by liver disease etiology[27] and we did not find any significant difference in performance by viral versus non-viral liver disease etiology. Finally, we leveraged a prospective cohort study including a standardized blood collection protocol; however, some patients did not have available samples

at each time point. We feel these limitations are outweighed by the study's strengths including its prospective nature, comparison of several biomarkers in a single cohort, and incorporation of longitudinal biomarker assessments.

In summary, we found the GALAD has high sensitivity for early HCC detection and performs favorably compared to other surveillance biomarkers, particularly when used in a longitudinal manner. While further validation in larger Phase III biomarker cohorts and Phase IV studies assessing the benefit-to-harm ratio for biomarker-based surveillance are necessary, these results show the promise of blood-based biomarker panels for early detection of HCC, addressing a significant unmet need in HCC surveillance.

REFERENCES

1. Moon AM, Singal AG, Tapper EB. Contemporary Epidemiology of Chronic Liver Disease and Cirrhosis. *Clinical Gastro and Hepatology* 2020; 18(12): 2650-66

2. Llovet J, Fuster J, Bruix J. Intention-to-treat analysis of surgical treatment for early hepatocellular carcinoma: resection versus transplantation. Hepatology 1999; 30(6): 1434-40.

3. Bangaru S, Marrero JA, Singal AG. Review article: New Therapeutic Interventions for Advanced Hepatocellular Carcinoma. *Alimentary Pharmacology and Therapeutics* 2020; 51(1): 78-89.

4. Marrero JA, Kulik LM, Sirlin CB, et al. Diagnosis, Staging, and Management of Hepatocellular Carcinoma: 2018 Practice Guidance by the American Association for the Study of Liver Diseases. *Hepatology*. 2018;68(2):723-750.

5. Galle PR, Forner A, Llovet JM, et al. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. *J Hepatol*. 2018;69(1):182-236

6. **Tzartzeva K, Obi J,** Rich NE, Parikh ND, et al. Surveillance Imaging and Alpha Fetoprotein for Early Detection of Hepatocellular Carcinoma in Patients With Cirrhosis: A Meta-analysis. *Gastroenterology*. 2018;154(6):1706-1718.

7. Simmons O, Fetzer DT, Yokoo T, Marrero J, et al. Predictors of adequate ultrasound quality for hepatocellular carcinoma surveillance in patients with cirrhosis. *Aliment Pharmacol Ther*. 2017;45(1):169-177.

8. Del Poggio P, Olmi S, Ciccarese F, et al. Factors That Affect Efficacy of Ultrasound Surveillance for Early Stage Hepatocellular Carcinoma in Patients With Cirrhosis. *Clin Gastroenterol Hepatol*. 2014;12(11):1927-1933.

9. Atiq O, Tiro J, Yopp AC, Muffler A, et al. An assessment of benefits and harms of hepatocellular

carcinoma surveillance in patients with cirrhosis. *Hepatology*. 2017;65(4):1196-1205.

10.   Singal AG, Patihandia S, Obi J, Fullington H, et al. Benefits and harms of hepatocellular carcinoma surveillance in a prospective cohort of patients with cirrhosis. *Clinical Gastroenterology Hepatology* 2021

11.   **Farvardin S, Patel J**, Khambaty M, Yerokun O, et al. Patient-Reported Barriers are Associated with Lower HCC Surveillance Rates in Patients with Cirrhosis. *Hepatology* 2017; 65(3): 875-84

12.   Singal AG, Tiro JA, Murphy CC, Blackwell J, et al. Patient-reported barriers are associated with receipt of hepatocellular carcinoma surveillance in a multi-center cohort of patients with cirrhosis. *Clinical Gastroenterology Hepatology* 2021

13.   Parikh ND, Mehta A, Singal AG, Block T, Marrero JA, Lok AS. Biomarkers for the Early Detection of Hepatocellular Carcinoma. *Cancer Epi Biomarkers Prevention* 2020; 29(12): 2495 – 2503.

14.   Singal AG, Hoshida Y, Pinato DJ, Marrero JA, et al. International Liver Cancer Association (ILCA) White Paper on Biomarker Development for Hepatocellular Carcinoma. *Gastroenterology* 2021

15.   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 Mar;14(3):469-475.e2.

16.   Singal AG**,** Conjeevaram H, Fu S, Volk ML, et al. Effectiveness of Hepatocellular Carcinoma Surveillance in Patients with Cirrhosis. *Cancer Epidemiology Biomarkers & Prevention* 2012; 21(5): 793-9.

17.   Pepe MS, Etzioni R, Feng Z, et al. Phases of biomarker development for early detection of cancer. J. Natl. Cancer Inst. 2001;93:1054-1061.

18.   Marrero JA, Feng Z, Wang Y, Nguyen M, et al. Alpha-fetoprotein, des-gamma carboxprothrombin, and lectin-bound alpha-fetoprotein in early hepatocellular carcinoma. Gastroenterology 2009; 137(1): 110-8.

19.   Toyoda H, Kumada T, Tada T, Kaneoka Y, et al. Clinical utility of highly sensitive Lens culinaris agglutin-reactive alpha-fetoprotein in hepatocellular carcinoma patients with alpha fetoprotein <20 ng/mL. Cancer Sci 2011; 102(5): 1025-31.

20.   Johnson PJ, Pirrie SJ, Cox TF, Berhane S, *et al.* The detection of hepatocellular carcinoma using a prospectively developed and validated model based on serological biomarkers. Cancer Epidemiol Biomarkers Prev 2014;23(1):144-53

21.   McIntosh, Martin W. and Urban, Nicole. (2003). A parametric empirical bayes method for cancer screening using longitudinal observations of a biomarker. Biostatistics 4(1), 27-40.

22. El-Serag HB, Kanwal F, Davila JA, Kramer J, Richardson P. A new laboratory-based algorithm to predict development of hepatocellular carcinoma in patients with hepatitis C and cirrhosis. Gastroenterology. 2014 May;146(5):1249-55.e1.

23. Tayob N, Corley DA, Christie I, Almers L, et al. Validation of the Updated Hepatocellular Carcinoma Early Detection Screening Algorithm in a Community-Based Cohort of Patients With Cirrhosis of Multiple Etiologies. Clin Gastroenterol Hepatol. 2021

24. Rich NE, John BV, Parikh, ND, Rowe I, et al. Hepatocellular carcinoma demonstrates heterogeneous growth patterns in a multicenter cohort of patients with cirrhosis. *Hepatology* 2020; 72(5): 1654-65.

25. Losic B, Craig A, Villacorta-Martin C, Martins-Filho S, et al. Intratumoral heterogeneity and clonal evolution in liver cancer. Nature Communications 2020; 11: 291

26. Yang JD, Addissie B, Mara K, Harmsen et al. GALAD Score for Hepatocellular Carcinoma Detection in Comparison with Liver Ultrasound and Proposal of GALADUS Score. Cancer Epi Biomarkers and Prevention 2019; 28(3): 531-38.

27. Berhane S, Toyoda H, Tada T, Kumada T, Kagebayashi C, Satomura S, *et al.* Role of the GALAD and BALAD-2 serologic models in diagnosis of hepatocellular carcinoma and prediction of survival in patients. Clinical Gastroenterology and Hepatology 2016;14(6):875-86. e6.

28. Best J, Bechmann LP, Sowa JP, Sydor S, Dechene A, Pflanz K, *et al.* GALAD Score Detects Early Hepatocellular Carcinoma in an International Cohort of Patients with Nonalcoholic Steatohepatitis. Clin Gastroenterol Hepatol 2020; 18(3): 728-35.

29. Tayob N, Lok AS, Do K, Feng Z. Improved detection of hepatocellular carcinoma using a longitudinal alpha-fetoprotein screening algorithm. Clin Gastro Hep 2016; 14(3): 469-75.

30. Singal AG, Lampertico P, Nahon P. Epidemiology and Surveillance for Hepatocellular Carcinoma: New Trends. *J Hepatology* 2020; 72(2): 250-61.

31. Singal AG, Nehra M, Adams-Huet B, Yopp AC, Tiro JA, Marrero JA, et al. Detection of Hepatocellular Carcinoma at Advanced Stages among Patients in the HALT-C Trial: Where Did Surveillance Fail? *American J. Gastroenterology* 2013; 108(3): 425-32.

32. Wolf E, Rich NE, Marrero JA, Parikh ND, Singal AG. Use of hepatocellular carcinoma surveillance in patients with cirrhosis: A systematic review and meta-analysis. *Hepatology* 2021; 73(2): 713-25.

33. Simmons OL, Feng Y, **Parikh ND, Singal AG**. Primary Care Provider Practice Patterns and Barriers to Hepatocellular Carcinoma Surveillance. *Clinical Gastroenterology Hepatology* 2019; 17(4): 766-73.

34. **Woolen SA, Singal AG,** Davenport MS, Troost JP, et al. Patient Preferences for Hepatocellular Carcinoma Surveillance Parameters: A MultiCenter Conjoint Study. *Clinical Gastroenterology Hepatology* 2021

35. Feng Z, Marrero JA, Khaderi S, Singal AG, et al. Design of the Texas Hepatocellular Carcinoma Consortium Cohort Study. *Am J Gastro* 2019; 114(3): 530-2.

36. Borges K, Dai J, Parikh ND, Schwartz M, et al. Rationale and Design of the Hepatocellular carcinoma Early Detection Strategy study: A multi-center longitudinal initiative of the National Cancer Institute's Early Detection Research Network. Contemp Clin Trials 2019; 76: 49-54.

Table 1. Patient Characteristics

Characteristic	Patients without HCC (n=355)	Patients who developed HCC (n=42)	p-value*
Age	52.0 (23.0 – 82.0)	53.5 (42.0 – 67.0)	0.32
Sex (% male)	208 (58.6%)	30 (71.4%)	0.13
Race/Ethnicity (%)			0.27
Non-Hispanic White	324 (91.3)	36 (85.7)	
Non-Hispanic Black	8 (2.3)	2 (4.8)	
Hispanic White	7 (2.0)	2 (4.8)	
Asian	4(1.1)	1 (2.4)	
Other/unknown	12 (3.4)	1 (2.4)	
Etiology of Liver Disease (%)			0.83
Hepatitis C	164 (46.2%)	24 (57.1%)	

Alcohol-related	55 (15.5%)	6 (14.3%)	
NASH/cryptogenic	71 (20.0%)	7 (16.7%)	
Hepatitis B	16 (4.5%)	1 (2.4%)	
Other	49 (13.8%)	4 (9.5%)	
Child Pugh Class (% Child A)	145 (40.8%)	14 (33.3%)	0.41
MELD	9 (6 – 17)	10 (6- 17)	0.32
Number of HCC lesions			N/A
1	N/A	25 (59.5%)	
2		11 (26.2%)	
3		2 (4.8%)	
>3		4 (9.5%)	
Maximum HCC diameter	N/A	2.4 (0.5 – 6.0)	N/A
Vascular invasion	N/A	9 (21.4%)	N/A
Extra-hepatic metastases	N/A	0 (0%)	N/A
BCLC Stage			N/A
Stage 0/A	N/A	24 (57.1%)	
Stage B		8 (19.0%)	
Stage C		2 (4.8%)	
Stage D		8 (19.0%)	
BMI	28.9 [17.4, 68.6]	28.6 (20.4 – 50.5)	0.32
Diabetes	82 (23.1%)	7 (16.7%)	0.44
Presence of Ascites	211 (59.4%)	28 (66.7%)	0.41
Presence of hepatic encephalopathy	114 (32.1%)	17 (40.5%)	0.30
Presence of esophageal varices	208 (58.6%)	28 (66.7%)	0.41

Table 2: Biomarker performance (and 95% bootstrap confidence intervals) using established thresholds

		Any Stage HCC		Early-Stage HCC	
		Any time prior to HCC	0-6 months prior to HCC	Any time prior to HCC	0-6 months prior to HCC
GALAD (-0.63)	Patient-level Sensitivity	57.1 (41.9-72.4)	72.0 (53.8-89.3)	53.8 (33.3-73.3)	73.7 (52.6-93.3)
	Screening-level specificity	86.5 (83.0 – 89.9)			
HES Algorithm (10.17)	Patient-level Sensitivity	45.2 (30.4-60.0)	44.0 (23.8-62.5)	34.6 (15.4-54.2)	42.1 (19.0-66.7)
	Screening-level specificity	90.5(87.7 – 93.1)			
AFP 20 ng/mL	Patient-level Sensitivity	35.7 (21.7-51.4)	48.0 (28.0-68.8)	46.2 (26.1-65.4)	57.9 (33.3-80.0)
	Screening-level specificity	91.7 (88.9 – 94.3)			
AFP-L3% 10%	Patient-level Sensitivity	66.7 (52.0-81.3)	72.0 (52.4-88.9)	73.1 (54.2-88.9)	73.7 (52.4-93.3)
	Screening-level specificity	82.7 (78.5 – 86.5)			
DCP 7.5 ng/mL	Patient-level Sensitivity	23.8 (11.6-37.5)	20.0 (5.0-37.5)	30.8 (13.6-50.0)	26.3 (6.7-48.5)
	Screening-level specificity	92.3 (89.8 – 94.6)			

Table 3: Biomarker true positive rate (and 95% confidence interval) with screening-level false positive rate fixed at 10%
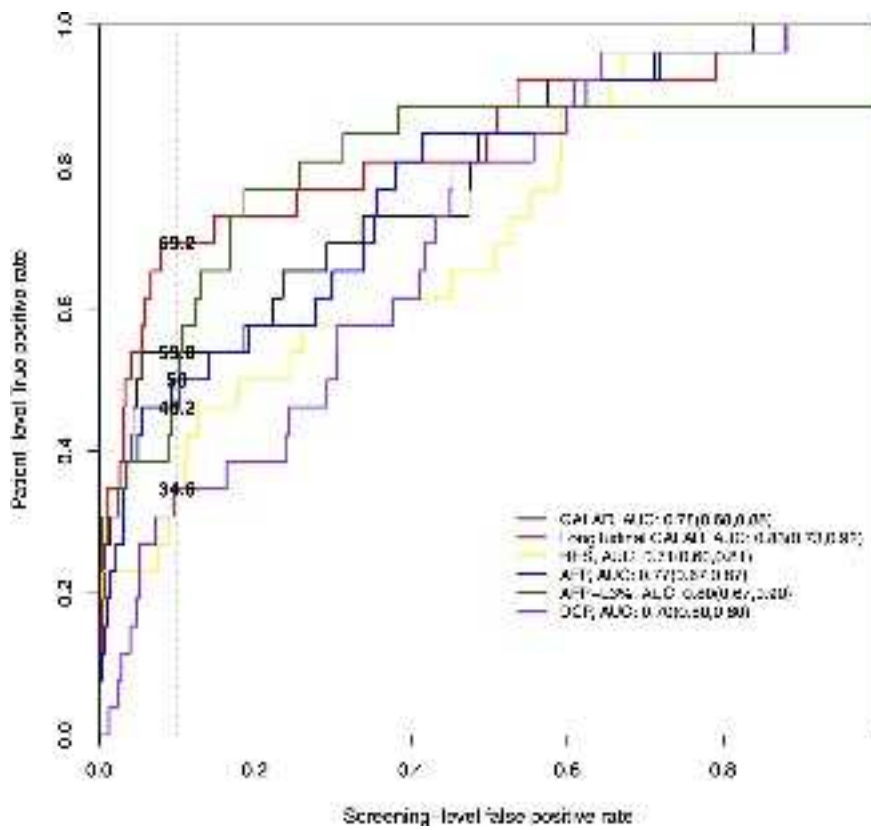
	Any Stage HCC			Early-Stage HCC		
	Any time prior to HCC	0-6 months prior to HCC	7-12 months prior to HCC diagnosis	Any time prior to HCC	0-6 months prior to HCC	7-12 months prior to HCC diagnosis
GALAD -0.33	54.8 (39.5-70.2)	72.0 (52.6-88.9)	50.0 (23.1-73.7)	53.8 (33.3-73.3)	73.7 (52.0-93.3)	44.4 (10.0-80.0)
Longitudinal GALAD	66.7 (51.3-80.8)	64.0 (42.9-83.3)	62.5 (35.3-85.7)	69.2 (50.0-86.4)	68.4 (44.4-88.9)	66.7 (28.6-100.0)
HES Algorithm 10.05	45.2 (26.1-62.1)	44.0 (21.1-62.5)	43.8 (20.0-72.2)	34.6 (14.3-56.5)	42.1 (18.2-66.7)	22.2 (0-60.0)
AFP 17.4 ng/mL	40.5 (24.2-57.1)	52.0 (30.4-75.0)	18.8 (0.0-42.9)	50.0 (28.0-69.0)	63.2 (35.3-85.7)	22.2 (0-57.1)
AFP-L3% 11.9%	45.2 (30.2-68.8)	43.8 (33.3-80.8)	43.8 (21.0-72.7)	46.2 (25.0-73.5)	33.3 (25.0-82.4)	33.3 (0-75.0)
DCP 5.9 ng/mL	26.2 (13.0-42.9)	20.0 (5.0-38.1)	37.5 (13.3-64.3)	34.6 (15.4-52.6)	26.3 (6.7-48.5)	55.6 (14.3-87.5)

Figure 1A: Receiver operating characteristic curves where patient-level true positive rate is estimated based on positive screens any time prior to HCC diagnosis in the overall cohort.

Figure 1B: Receiver operating characteristic curves where patient-level true positive rate is estimated based on positive screens any time prior to early-stage HCC diagnosis.
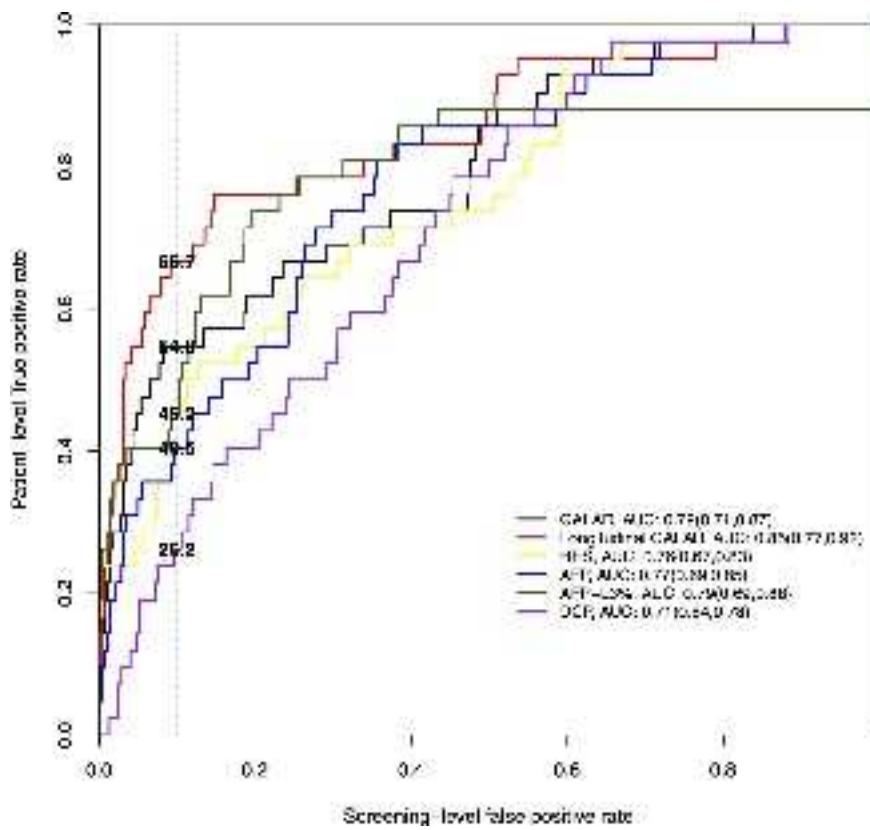
hep_32185_f1a.tiff

hep_32185_f1b.tiff