



HHS Public Access

Author manuscript

Clin Gastroenterol Hepatol. Author manuscript; available in PMC 2024 February 01.

Published in final edited form as:

Clin Gastroenterol Hepatol. 2023 February ; 21(2): 415–423.e4. doi:10.1016/j.cgh.2022.01.047.

The Performance of AFP, AFP-3, DCP as Biomarkers for Detection of Hepatocellular Carcinoma (HCC). A Phase 3 Biomarker Study in the United States

Nabihah Tayob², Fasiha Kanwal^{1,3}, Abeer Alsarraj^{1,3}, Ruben Hernaez^{1,3}, Hashem B El-Serag^{1,3}

¹Section of Gastroenterology and Hepatology, Department of Medicine, Michael E. DeBakey VA Medical Center and Baylor College of Medicine, Houston, TX

²Department of Data Science, Dana-Farber Cancer Institute (Boston, MA)

³Houston VA Health Services Research and Development Service Center of Excellence (Houston, TX), Section of Health Services Research, Michael E. DeBakey VA Medical Center (Houston, TX) and Baylor College of Medicine, (Houston, TX)

Abstract

Background: AFP, AFP L-3 and DCP in combination or in GALAD were tested for HCC surveillance in retrospective cohort and case-control studies. However, there is a paucity of prospective data and no phase 3 biomarker studies from North American populations.

Methods: We conducted a prospective-specimen collection, retrospective-blinded-evaluation (PRoBE) cohort study in patients with cirrhosis enrolled in a 6-monthly surveillance with liver imaging and AFP. Blood samples were prospectively collected every 6 months and analyzed in a retrospective blinded fashion. True positive rate (TPR) and false positive rate (FPR) for any or early HCC were calculated within 6, 12 and 24 months of HCC diagnosis based on published thresholds for biomarkers individually, in combination and in GALAD and HES scores. We calculated the area under the receiver operating curve (AUROC) and estimated TPR based on an optimal threshold at a fixed FPR of 10%.

Results: The analysis was conducted in a cohort of 534 patients; 50 developed HCC (68% early) and 484 controls with negative imaging. GALAD had the highest TPR (63.6, 73.8, 71.4% for all HCC, and 53.8, 63.3, 61.8 % for early HCC within 6, 12 and 24 months, respectively) but FPR of 21.5–22.9%. However, there were no differences in AUROC among GALAD, HES, AFP-3 or DCP. At a fixed 10% FPR, TPR for GALAD dropped (42.4, 45.2, 46.9%) and was not different from HES (36.4, 40.5, 40.8%) or AFP-3 alone (39.4, 45.2, 44.9%).

Correspondence: Hashem El-Serag, 7200 Cambridge St., Houston TX 77030; hasheme@bcm.edu; Fax: 713 798 2451.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Disclosures: All authors have no conflicts to disclose

Conclusions: In a prospective cohort phase 3 biomarker study, GALAD was associated with a considerable improvement in sensitivity for HCC detection but an increase in false positive results. GALAD performance was modest and not different from AFP-3 alone or HES.

Keywords

hepatocellular carcinoma; hepatitis C; cirrhosis

Background

The incidence rates for HCC have tripled in the past three decades, with most of the increase attributed to hepatitis C related HCC. Other major etiological risk factors for HCC include alcoholic liver disease, non-alcoholic fatty liver disease (NAFLD) and hepatitis B virus infection. Most HCC risk factors lead to the formation and progression of cirrhosis, which is present in 80–90% of HCC patients.¹

Survival of patients with HCC remains dismal, with an overall 5-year survival of less than 15%, except in patients who receive potentially curative therapy (i.e., liver transplant, surgical resection, or ablation), in whom a considerable improvement in survival has been consistently observed (5-year survival between 40% and 70%). Given this, several practice guidelines recommend HCC surveillance in high-risk patients to detect HCC at an early stage, when curative treatment can be applied.^{2, 3} Patients with cirrhosis constitute the main target for HCC surveillance.

Liver ultrasound in combination with alpha-fetoprotein (AFP) has been the cornerstone of recommended HCC surveillance. A meta-analysis of the cohort studies evaluating liver ultrasound for the detection of early-stage HCC, reported a pooled sensitivity of 45%, which increased to 63% with adding AFP.⁴ AFP when used alone has unacceptably low sensitivity and is generally not recommended for HCC surveillance. We developed and validated in several retrospective cohort studies the Hepatocellular Carcinoma Early Detection Screening (HES) algorithm that combines AFP with age and few clinical variables (ALT, platelets) with the option of including AFP change within the previous one year and/or the underlying cirrhosis etiology.^{5–9} HES was associated with 5.2–8.6% increased sensitivity over AFP alone. However, HES has not been validated in a prospective cohort study. Another emerging strategy for HCC surveillance is a combination of liver ultrasound and serum AFP, AFP *Lens culinaris* agglutinin-reactive fraction of AFP (AFP-L3), and des-gamma-carboxy prothrombin (DCP). Retrospective cohort and case-control studies suggest that the triple biomarker combination (AFP, AFP-L3, DCP) slightly improve the accuracy of early detection of HCC over AFP alone, and the use of these biomarkers for HCC surveillance is recommended and widely practiced in Japan. GALAD score derived from **G**ender, **A**ge, **A**FP-L3, **A**FP, and **D**CP was validated in several retrospective case-control studies^{10–13} but not in prospective cohort studies within the United States. A key step in biomarker validation for HCC early detection is to examine the performance of the triple biomarker in preclinical samples from a phase 3 biomarker study in a well-defined cirrhosis population that is eligible for surveillance.^{14, 15} We followed the PRoBE guidance (Prospective-specimen collection, retrospective-blinded-evaluation) for conducting a phase 3 biomarker study.¹⁶

We conducted a prospective cohort study to validate the performance of the triple biomarker panel (AFP, AFP-L3, and DCP) individually, in combination and in calculators (e.g., GALAD, HES) for identification of HCC using conventional cutoffs as well as comparing sensitivities at a fixed 90% specificity.

Methods

We carried out a prospective cohort study among patients with cirrhosis undergoing biannual HCC surveillance using liver imaging and AFP at the Michael E. DeBakey VA Medical Center in Houston, Texas from 07/23/2014 to 07/31/2017 with follow up through 7/31/2018. We closely followed the PRoBE guidance for conducting a phase 3 biomarker study.¹⁶ A written informed consent was obtained from each patient. The study was approved by the Institutional Review Board at the Michael E. DeBakey Veterans Affairs Medical Center in Houston, Texas. We followed all patients until the development of HCC diagnosis, death or end of the study. HCC diagnosis was made according to the Association for the Study of Liver Diseases (AASLD) criteria including histological or radiological diagnosis using characteristic appearance (arterial enhancement and delayed washout) on triple phase contrast CT or MRI.² Early stage HCC was defined as a single nodule no larger than 5 cm. This was a simple and straightforward determination, and one that was close to the definition to the Milan criteria in this study where most HCC (70%) was a single mass.

Cirrhosis was defined by histology (liver biopsy showing F4), elastography (FibroScan liver stiffness median >12kPa; magnetic resonance >4 kPa), endoscopic signs of portal hypertension (esophageal or gastric varices, portal hypertensive gastropathy) or radiological signs of cirrhosis or portal hypertension (splenomegaly, intraabdominal varices, collaterals, recanalized umbilical vein or ascites). HCC surveillance was defined as a combination of liver imaging (ultrasound, CT, MRI), the type of which determined by the treating clinician, and serum AFP. We excluded patients with conditions that are associated with a significant reduction in survival (e.g., invasive cancer, congestive heart failure, severe COPD, advanced neurodegenerative disease or dementia) or conditions associated with impaired ability to consent or adhere to the study procedures (e.g., uncontrolled psychosis or depression, or homelessness). We also excluded patients with known present or past HCC, those actively listed for liver transplantation because they undergo intense HCC surveillance, and patients with non-hemangioma non-cystic liver masses because they are likely to receive repeated liver imaging. We defined active HCV by positive HCV RNA at every AFP test, on treatment by receipt of HCV treatment at any 1 AFP test; and treated with SVR as those with observed SVR at 1 AFP test during study period. We defined chronic hepatitis B (HBV) by the presence of positive hepatitis B surface antigen test, alcoholic liver disease by self-reported heavy current or past alcohol use, and non-alcoholic fatty liver disease by the presence of one or more major metabolic dysfunction trait in the absence of active HCV, HBV or alcoholic liver disease. We defined the findings of abnormal nodule on surveillance imaging as definitive for HCC if a diagnosis of HCC is made during the study period; negative for HCC if a diagnosis of “not HCC” is made during the study period or at least one year of follow up (through 7/31/2018), and indeterminate or non-negative if HCC diagnosis cannot be ruled out or in during the study period. All HCC diagnoses were also adjudicated post hoc by one experienced hepatologist (RH). at the end of the study using

printed deidentified reports (not the entire EMR) while blinded to the clinical diagnosis and the biomarker testing results.

We prospectively collected blood and isolated stored serum samples within 2 months of each surveillance liver imaging. We performed retrospective assays for AFP L-3, AFP and DCP in serum samples using Microchip capillary electrophoresis and liquid-phase binding assay on an uTASWako i30 automated analyzer (Wako Life Sciences, Inc. in Mountain View, CA USA). Testing was done on batches of stored serum blinded to case and control status, and the levels of biomarkers were not made available or used for clinical decision making. None of the samples were thawed or used for any purpose before the study specific assays.

Statistical Analyses

We examined the performance of the biomarkers individually, in combination (AFP, AFP-3, DCP) and in scores (GALAD, HES) in distinguishing cirrhosis cases with HCC from cirrhosis controls without HCC. The primary analysis was conducted among patients without coumadin use and without inconclusive liver nodule imaging. A secondary analysis reflecting the overall clinical effectiveness of biomarkers was conducted among all patients in the cohort including those with coumadin use or inconclusive liver nodule imaging.

We estimated patient-level sensitivity or true positive rate (TPR) as the probability of 1 positive screen test within 6, 12 or 24 months prior to HCC diagnosis. Given that the exact time of HCC onset cannot be determined in this study, to fully understand the performance of the algorithms, we have examined multiple time intervals. We estimated 1- specificity or false positive rate (FPR) as the cumulative probability of a positive screen at any time during the study in cirrhosis controls without HCC or a positive screen that occurred earlier than 6, 12 or 24 months before HCC diagnosis. TPR is estimated at a patient-level since any positive screening result during the preclinical period could lead to an earlier diagnosis of HCC, and FPR is estimated at the screening-test level since each positive screening result leads to further cost and possible complications. The 95% bias-corrected bootstrap percentile confidence intervals were estimated for all measures.¹⁷

We calculated TPR and FPR for detecting HCC based on the following published thresholds: AFP>20 ng/mL, AFP-L3>10% and DCP>7.5 ng/mL. We calculated GALAD score as $-10.08 + 0.09 * \text{Current age} + 1.67 * \text{Male} + 2.34 * \log_{10}(\text{AFP}) + 0.04 * \text{AFP-L3} + 1.33 * \log_{10}(\text{DCP})$ and used GALAD score > -0.63 as a threshold for positive score.¹³ HES is an AFP-adjusted algorithm that includes age, platelets, ALT, change in AFP over the last year, underlying cirrhosis etiology and interaction terms. We used 10.17 cutoff for HES based on a community-based cohort validation study.⁵

In addition to using these thresholds, and to facilitate comparisons among the different biomarkers and scores, we also calculated and compared the area under the ROC curve (AUROC), and estimated thresholds and TPR corresponding to fixed 10% FPR (reported specificity of AFP alone at 20ng/ml) on the receiver operating characteristic curve (ROC).¹⁸ Lastly, we calculated the concordance between the results of the GALAD and HES scores among patients that developed HCC.

All analyses were conducted using R version 3.6.1.

Results

Study Cohort

During the study period, we enrolled 741 subjects with cirrhosis in whom 59 HCC cases were diagnosed by July 31st 2017. Figure S1 shows the study flow and the exclusion of patients who withdrew or had no research blood sample or incomplete data to calculate GALAD or HES. We limited the primary analysis to patients not taking coumadin and only controls with clearly negative imaging results. This resulted in 50 incident HCC cases and 484 patients with cirrhosis and no HCC who contributed a total of 1,362 HCC surveillance episodes. The mean size of HCC at the time of diagnosis was 2.09 cm (standard deviation: 1.28). Most patients with HCC had one nodule (70%), 8 (16%) had 2 nodules, 3 (6%) had 3 nodules whereas the rest had more than three nodules. 34/50 (68%) had early HCC. The demographic and clinical characteristics are listed in Table 1.

Comparison of Biomarker Performance Using Fixed Threshold Based on Conventional Cutoffs for HCC Detection

The sensitivity for the biomarkers and scores were generally slightly higher at 12–24 months than within 6 months. Of the individual biomarkers, AFP-3 had the highest sensitivity/TPR, while AFP and DCP had the lowest. Conversely, DCP had the highest specificity (i.e., lowest FPR). The sensitivity/TPR of the three biomarkers combined was higher but the FPR was also higher than any of the individual markers. The FPR estimates were cumulative for the entire time the patients were under surveillance (Table 2). For example, within 6 months prior to HCC, combined triple biomarker sensitivity was 39.4% as compared to 18.2%–33.3% for the individual biomarkers; however, the specificity also declined as FPR was 9.7% for the combination compared with 1.4–5.5% for the individual markers. GALAD had the highest sensitivity/TPR (63.6%–73.8%) of any of the biomarkers, but also significantly higher FPR (FPR 21.5%–22.9%) than any of the individual biomarkers, combined biomarkers or HES (Table 2). Compared to AFP alone, HES had higher sensitivity (increased 2%–6%) but also slightly higher FPR (5.3%–5.9% compared with 3.7%–4.1%) but confidence intervals for all estimates overlapped.

For early stage HCC, the sensitivity of all biomarkers was lower than that observed for any HCC, but we observed similar trends (Table 3); similar to Table 2, the FPR estimates were cumulative for the entire time the patients were under surveillance. GALAD had a significantly higher sensitivity/TPR (53.8%–63.3%) than any of the individual biomarkers, the biomarker combination or HES, but GALAD was also associated with the highest FPR (21.5%–22.1%).

Comparison of Biomarker Performance Using AUROC and Threshold Based on Cutoffs Corresponding to 10% FPR for HCC Detection.

The AUROC for 6, 12 and 24 months were similar for GALAD and HES (for example, GALAD AUROC was 0.79 and 0.75 for any HCC and early HCC within 6 months and HES had 0.78 and 0.76 for the respective measurements) but higher than the individual

biomarkers, although none of the differences were significant given that confidence intervals for estimates overlapped (Figures 1, 2 and 3). We established cutoffs defining positivity of biomarkers by fixing screening test level specificity at 90% in the ROC curves of biomarker within 6, 12 or 24 months prior to HCC diagnosis. For example, within 12 months prior to diagnosis, the cutoff for AFP, AFP-L3%, and DCP were 11.2 ng/mL, 8.4%, and 1.52 ng/mL, respectively (Figure 2), compared to conventional cutoffs (20 ng/mL, 10%, and 7.5 ng/mL, for AFP, AFP-L3%, and DCP, respectively).

Using these new cutoffs and unlike the fixed threshold cutoffs, the patient-level sensitivity/TPR of GALAD (45.2%, 95% CI: 26.2–59.0) dropped for all HCC to a comparable to that of AFP-L3 (45.2%, 95% CI: 29.3–59.0) or HES algorithm (40.5%, 23.5–53.8) but higher than either AFP or DCP alone (33.3%, 95% CI: 17.1–46.8 and 31.0%, 95% CI: 16.3–45.5, respectively).

For early stage HCC patients (Figure 1–3), GALAD sensitivity/TPR at 10% screening-level FPR dropped even further (compared with Table 3 using conventional cutoffs). AFP-L3% alone had the highest patient-level sensitivity/TPR compared to the other biomarkers, GALAD or HES at 12 months prior to diagnosis and comparable to HES at 24 months prior to diagnosis. At 6 months prior to HCC diagnosis, HES has slightly higher patient-level sensitivity/TPR than either AFP-L3% or AFP in early stage HCC patients.

Concordance.—The agreement between GALAD and HES was 75.8%, 80.9% and 73.1% for measurements within 6, 12 and 24 months prior to HCC diagnosis. (Table S1). For all HCC, the incremental positive yield of GALAD over HES was 15.2%, 11.9%, 16.3%, respectively, whereas the incremental positive yield of HES over GALAD was 9.1%, 7.1%, 10.2%, respectively within 6, 12 and 24 months prior to HCC diagnosis. In patients with early stage HCC, the concordance between GALAD and HES was 84.6%, 86.7% and 79.4% for measurements within 6, 12 and 24 months prior to HCC diagnosis and the incremental yield of HES over GALAD was 11.5%, 10.0% and 14.7%, respectively.

Secondary Analysis

We analyzed a larger group of 53 HCC cases and 564 controls (Table S2) that includes coumadin users and controls with indeterminate imaging. The findings were generally similar to the primary analysis in direction and magnitude (Table S2 and Figure S2–S4). Specifically, GALAD had highest sensitivity of 62.9–73.3% for all HCC and 51.9–62.5% for early HCC, and highest FPR 24.7–25.9% for all HCC and 24.7–25.2% for early HCC using conventional cutoffs. GALAD sensitivity markedly dropped (37.8–40.0%) and became similar to AFP-3 (37.1–42.3%) and HES (34.3–38.5%) when using ROC derived cutoffs corresponding to 10% FPR.

Discussion

We carried out this prospective multiethnic multi etiology cirrhosis cohort study according to PRoBE guidance,¹⁶ and evaluated the performance of triple biomarkers (AFP, AFP-3, DCP) individually, in combination, and in scores using both conventional cutoffs and ROC derived cutoffs at a fixed 10% specificity. None of these biomarkers or scores had acceptable

overall performance. Scores that include clinical and demographic factors (GALAD, HES) offer higher performance than their respective biomarker components. The triple biomarkers especially in GALAD score had a significantly higher sensitivity of HCC detection but this came with a significant increase in false-positive cases. The cost-effectiveness of this added GALAD benefit versus the added harm and cost is unclear and needs to be examined before wider implementation can be recommended.

We compared the sensitivity of biomarkers using a cutoff derived from the ROC curves while fixing the specificity at 90% (i.e., 10% FPR); the latter is the estimate of AFP specificity in previous studies and a generally acceptable benchmark. The large drop in GALAD sensitivity in this analysis compared to those obtained in the analyses performed using conventional cutoffs (from 53.8%–63.3 to 42.4–46.9%) is explained and to be expected given the high FPRs (21.5%–22.9%) using conventional cutoffs and highlight the overall modest test performance. The incremental positive yield in HCC cases related to higher sensitivity is potentially clinically important because it could lead to increase use of curative therapy and reduced mortality. However, the increase in FPR is also clinically significant and is associated with increased testing, anxiety and cost and diverting valuable resources from the surveillance program.^{19, 20} The relatively low GALAD overall performance while quite different from results of several previous phase 2 biomarker studies was not dissimilar from a recent smaller and more heterogeneous multi center phase 3 biomarker study that reported higher sensitivity of GALAD but also higher false positive results than individual biomarkers.²¹

The increased sensitivity conveyed by GALAD and to a lower extent HES, over their respective component biomarkers alone shows the importance of demographic and clinical factors in predicting HCC risk independent of the current biomarkers. Such framework is likely to be useful for current, as well future, blood-based biomarkers. Because GALAD is a predefined score, where the relative weight of each component is fixed, it was not possible to determine which components were the biggest driver of the high FPR for the overall score. The low FPRs for each of or any of the three biomarkers at the known thresholds (Table 2) suggest that gender and/or age are driving the high FPR of GALAD. HES incorporates only AFP but not AFP3 or DCP combined with age, disease etiology if available, ALT and platelets, and one previous AFP within one year if available was developed and validated in retrospective cohorts at national VA and Kaiser Permanente Northern California.¹ This is the first external validation in a prospective cohort. The results confirm a modest gain in sensitivity compared with AFP alone at no added cost or a large increase in FPR; therefore, overall performance as reflected by AUROC and sensitivity at fixed 10% were similar to GALAD. The concordance between GALAD and HES was modest; this indicates further potential for improved sensitivity if elements of these models are combined. HES incorporates several variables that affect HCC risk such ALT, platelets indicative of underlying hepatic function and disease etiology (HCV, HBV, alcohol, NAFLD) and therefore has the potential of producing high performance if AFP3 and DCP are added. The dataset was too small for a robust fitting of the HES with triple biomarkers. Additional large studies are needed to produce a well-fitted model of HES with triple biomarkers.

The study had few limitations. Despite the large cohort and 648 person-years of follow up in the analysis cohort, only 50 patients developed HCC cases that can be fully evaluated for biomarker analysis. While this was our target sample size based on power calculation, the analytic power was reduced due to an unanticipated issue, namely, the time of blood collection for the biomarker assays varied before HCC diagnosis between <6 months to 24 months; therefore the 6- and 12-month analyses did not incorporate all 50 patients. Therefore, many of the estimates have overlapping 95% confidence intervals with the exception of the significantly higher sensitivity for any stage HCC within 12 and 24 months prior to diagnosis but also significantly higher FPR estimates for GALAD compared to each individual biomarker. Another limitation is that the results of this study are not generalizable to a non-veteran's population. The study cohort is mostly male and VA populations in general have higher rates of comorbidities. Since AFP was part of the clinical surveillance strategy, it could have triggered evaluation for HCC diagnosis. However, for the study analyses, we did not use AFP values obtained in the clinical practice lab but rather relied on AFP analyzed in the research blood sample that was analyzed retrospectively.

Our study design is a major strength. The prospective uniform nature of specimen collection for all subjects from a single-center cohort in the PRoBE design mitigates systematic biases by ensuring that specimens for case patients and control subjects are collected in the same way. A common problem with retrospective studies is that knowledge of the subject's outcome status may affect the interpretation of an assay result or the care with which the specimen is handled. This bias is avoided in the PRoBE design by storing specimens before outcome ascertainment and by blinding specimens for retrieval and assay procedures. The biomarker values were ascertained for all subjects in a cohort, but the biomarker measurement occurred after outcome assessment. The PRoBE design avoids spectrum bias by identifying including only cirrhosis patients at risk of HCC in the cohort and subsequently identifying them as either case patients that have developed HCC or control subjects that have cirrhosis but no incidence of HCC during the study.

In summary, in a prospective cohort phase 3 biomarker study, we found an equivalent overall performance of AFP, AFP-3, and DCP individually, in combination or in GALAD or HES scores. Using the triple biomarkers combined especially in GALAD score is associated with a large increase in sensitivity of HCC detection; this advantage is offset by a large increase in false positive results. Better biomarker surveillance strategies are required

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Grant support:

This research was supported by an NIH NCI Grants (R01CA190776) and NIDDK (P30DK056338) and Wako Inc. to Dr El-Serag.

List of Abbreviations:

AFP	α -Fetoprotein
-----	-----------------------

ALT	alanine aminotransferase
CT	computed tomography
FPR	false positive rate
HCC	hepatocellular carcinoma
HCV	hepatitis C virus
HES	hepatocellular carcinoma early detection screening
ICD-9	international classification of diseases, 9th revision
MRI	magnetic resonance imaging
ROC curve	receiver operating characteristic curve
VA	U.S. Department of Veterans Affairs Healthcare System

References

1. McGlynn KA, Petrick JL, El-Serag HB. Epidemiology of Hepatocellular Carcinoma. *Hepatology*. Apr 22 2020;doi:10.1002/hep.31288
2. 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*. Aug 2018;68(2):723–750. doi:10.1002/hep.29913 [PubMed: 29624699]
3. European Association for the Study of the Liver. Electronic address eee, European Association for the Study of the L. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. *J Hepatol*. Jul 2018;69(1):182–236. doi:10.1016/j.jhep.2018.03.019 [PubMed: 29628281]
4. Tzartzeva K, Obi J, Rich NE, et al. Surveillance Imaging and Alpha Fetoprotein for Early Detection of Hepatocellular Carcinoma in Patients With Cirrhosis: A Meta-analysis. *Gastroenterology*. May 2018;154(6):1706–1718 e1. doi:10.1053/j.gastro.2018.01.064 [PubMed: 29425931]
5. Tayob N, Corley DA, Christie I, 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*. Aug 5 2020;doi:10.1016/j.cgh.2020.07.065
6. Tayob N, Christie I, Richardson P, et al. Validation of the Hepatocellular Carcinoma Early Detection Screening (HES) Algorithm in a Cohort of Veterans With Cirrhosis. *Clin Gastroenterol Hepatol*. Aug 2019;17(9):1886–1893 e5. doi:10.1016/j.cgh.2018.12.005 [PubMed: 30557738]
7. Tayob N, Richardson P, White DL, et al. Evaluating screening approaches for hepatocellular carcinoma in a cohort of HCV related cirrhosis patients from the Veteran’s Affairs Health Care System. *BMC Med Res Methodol*. Jan 4 2018;18(1):1. doi:10.1186/s12874-017-0458-6 [PubMed: 29301497]
8. White DL, Richardson P, Tayoub N, Davila JA, Kanwal F, El-Serag HB. The Updated Model: An Adjusted Serum Alpha-Fetoprotein-Based Algorithm for Hepatocellular Carcinoma Detection With Hepatitis C Virus-Related Cirrhosis. *Gastroenterology*. Dec 2015;149(7):1986–7. doi:10.1053/j.gastro.2015.10.004 [PubMed: 26519622]
9. 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*. May 2014;146(5):1249–55 e1. doi:10.1053/j.gastro.2014.01.045 [PubMed: 24462733]
10. Singal AG, Tayob N, Mehta A, et al. Doylestown Plus and GALAD Demonstrate High Sensitivity for HCC Detection in Patients With Cirrhosis. *Clin Gastroenterol Hepatol*. Apr 20 2021;doi:10.1016/j.cgh.2021.04.018

11. Best J, Bechmann LP, Sowa JP, et al. GALAD Score Detects Early Hepatocellular Carcinoma in an International Cohort of Patients With Nonalcoholic Steatohepatitis. *Clin Gastroenterol Hepatol*. Mar 2020;18(3):728–735 e4. doi:10.1016/j.cgh.2019.11.012 [PubMed: 31712073]
12. Best JBH, Heider D, Schotten C, Manka P, Bedreli S, Gorray M, Ertle J, van Grunsven LA, Dechêne A. The GALAD scoring algorithm based on AFP, AFP-L3, and DCP significantly improves detection of BCLC early stage hepatocellular carcinoma. *Gastroenterology*. 2016;54(12):1296–1305. doi:10.1055/s-0042-119529.
13. Berhane STH, Tada T, Kumada T, Kagebayashi C, Satomura S, Schweitzer N, Vogel A, Manns MP, Benckert J, Berg T, Ebker M, Best J, Dechêne A, Gerken G, Schlaak JF, Weinmann A, Wörns MA, Galle P, Yeo W, Mo F, Chan SL, Reeves H, Cox T, Johnson P. Role of the GALAD and BALAD-2 Serologic Models in Diagnosis of Hepatocellular Carcinoma and Prediction of Survival in Patients. *Clin Gastroenterol Hepatol*. 2016;14(6):875–886. [PubMed: 26775025]
14. Pepe MS, Etzioni R, Feng Z, et al. Phases of biomarker development for early detection of cancer. *J Natl Cancer Inst*. Jul 18 2001;93(14):1054–61. doi:10.1093/jnci/93.14.1054 [PubMed: 11459866]
15. Pepe MS, Feng Z. Improving biomarker identification with better designs and reporting. *Clin Chem*. Aug 2011;57(8):1093–5. doi:10.1373/clinchem.2011.164657 [PubMed: 21666069]
16. Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. *J Natl Cancer Inst*. Oct 15 2008;100(20):1432–8. doi:10.1093/jnci/djn326 [PubMed: 18840817]
17. Bradley Efron RJT. *An Introduction to the Bootstrap*. 1st Edition ed. Chapman and Hall/CRC; 1994:456 Pages.
18. Marrero JA, Feng Z, Wang Y, et al. Alpha-fetoprotein, des-gamma carboxyprothrombin, and lectin-bound alpha-fetoprotein in early hepatocellular carcinoma. *Gastroenterology*. Jul 2009;137(1):110–8. doi:10.1053/j.gastro.2009.04.005 [PubMed: 19362088]
19. Singal AG, Patibandla S, Obi J, et al. Benefits and Harms of Hepatocellular Carcinoma Surveillance in a Prospective Cohort of Patients With Cirrhosis. *Clin Gastroenterol Hepatol*. Sep 10 2020;doi:10.1016/j.cgh.2020.09.014
20. Petrasek J, Singal AG, Rich NE. Harms of hepatocellular carcinoma surveillance. *Curr Hepatol Rep*. Dec 2019;18(4):383–389. doi:10.1007/s11901-019-00488-8 [PubMed: 33214987]
21. Singal AG, Tayob N, Mehta A, Marrero JA, El-Serag HE, Jin O, de Viteri CZ, Fobar A, Parikh ND. GALAD demonstrates high sensitivity for HCC surveillance in a cohort of patients with cirrhosis. *Hepatology* Oct 2021 doi: 10.1002/hep.32185. Online ahead of print.

What you need to know

BACKGROUND

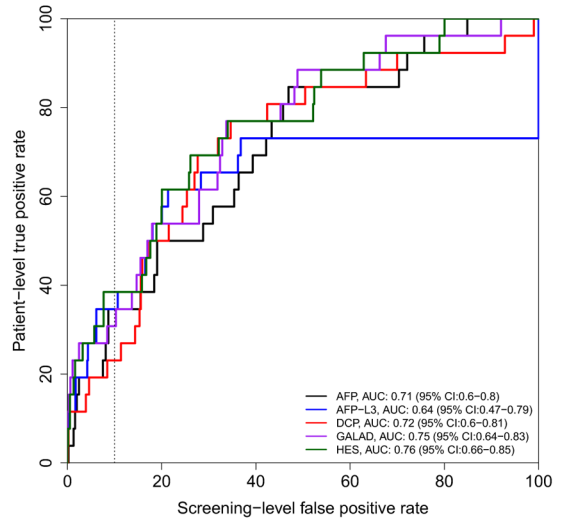
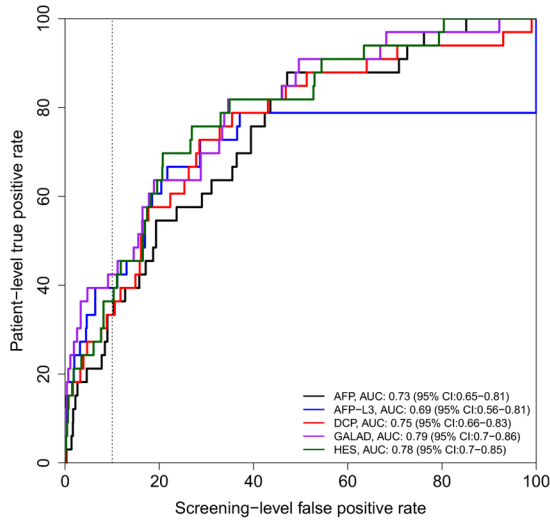
AFP, AFP L-3 and DCP in combination or in GALAD were evaluated for HCC surveillance in in retrospective cohort and case-control data. We conducted a prospective cohort study and followed the PRoBE guidance (Prospective-specimen collection, retrospective-blinded-evaluation) to validate the performance of the triple biomarker panel individually, in combination and in calculators (GALAD, HES) for HCC identification.

FINDINGS

We found a low but equivalent overall performance of AFP, AFP-3, and DCP individually, in combination or in GALAD or HES scores. Using the triple biomarkers in GALAD score is associated with a large increase in sensitivity of HCC detection; this advantage is offset by a large increase in false positive results.

IMPLICATIONS FOR PATIENT CARE

The cost-effectiveness of the GALAD benefit versus the added harm and cost is unclear and needs to be examined before wider implementation can be recommended.



Any stage HCC

Early-stage HCC

AFP>11.3ng/ml: 33.3 (95% CI:16.7-50.0)

AFP>11.2ng/ml: 34.6 (95% CI:18.2-55.0)

AFP-L3>8.4%: 39.4 (95% CI:22.9-55.0)

AFP-L3>8.4%: 34.6 (95% CI:16.0-52.6)

DCP>1.5ng/ml: 33.3 (95% CI:16.1-48.6)

DCP>1.4ng/ml: 23.1 (95% CI:7.4-40.9)

GALAD>0.05: 42.4 (95% CI:26.7-59.4)

GALAD>0: 30.8 (95% CI:12.9-47.8)

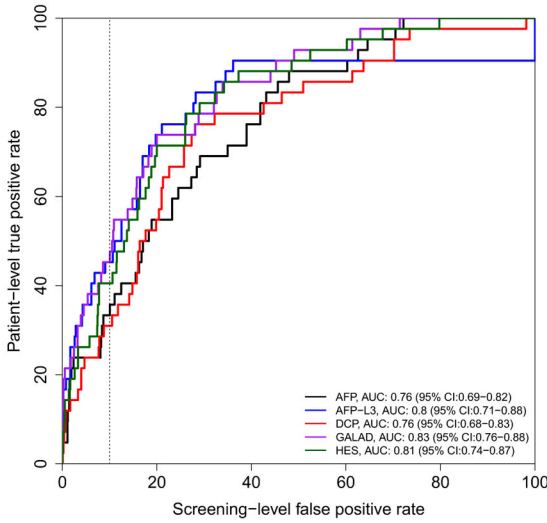
HES>8.13: 36.4 (95% CI:18.5-51.6)

HES>8.03: 38.5 (95% CI:20.7-56.7)

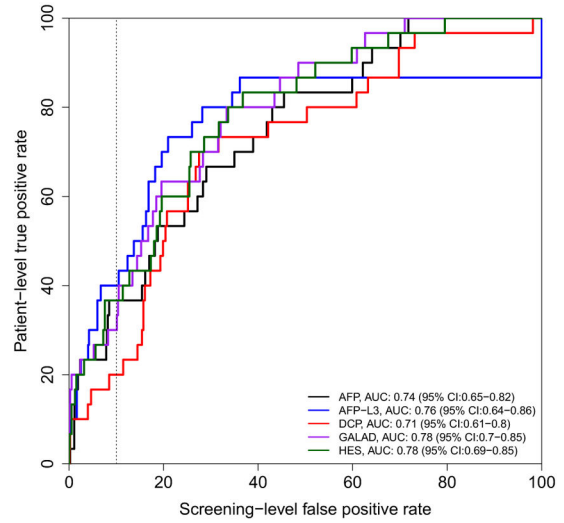
Figure 1:

Receiver operating characteristics (ROC) curve within 6 months prior to HCC diagnosis.

Below the figure, we describe patient-level true positive rate at optimal cutoff corresponding to 10% screening-level false positive rate on the ROC curves.



Any stage HCC



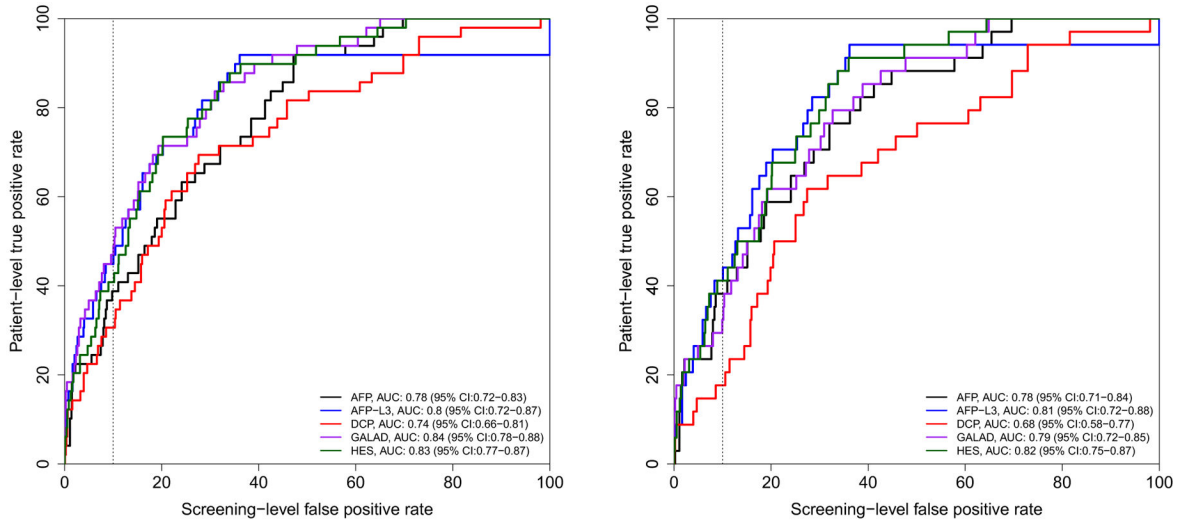
Early-stage HCC

AFP>11.2ng/ml: 33.3 (95% CI:17.1-46.8)
 AFP-L3>8.4%: 45.2 (95% CI:29.3-59.0)
 DCP>1.52ng/ml: 31.0 (95% CI:16.3-45.5)
 GALAD>0.02: 45.2 (95% CI:26.2-59.0)
 HES>8.03: 40.5 (95% CI:23.5-53.8)

AFP>10.9ng/ml: 36.7 (95% CI:20.0-55.6)
 AFP-L3>8.3%: 40.0 (95% CI:20.6-55.2)
 DCP>1.45ng/ml: 20.0 (95% CI:6.1-36.0)
 GALAD>0: 30.0 (95% CI:11.4-45.5)
 HES>8.00: 36.7 (95% CI:19.4-53.3)

Figure 2:

Receiver operating characteristics (ROC) curve within 12 months prior to HCC diagnosis. Below the figure, we describe patient-level true positive rate at optimal cutoff corresponding to 10% screening-level false positive rate on the ROC curves.



Any stage HCC

Early-stage HCC

AFP>10.8ng/ml: 38.8 (95% CI:25.0-54.7)

AFP>10.8ng/ml: 38.2 (95% CI:21.7-55.6)

AFP-L3>8.3%: 44.9 (95% CI:28.9-57.7)

AFP-L3>8.3%: 41.2 (95% CI:23.3-56.2)

DCP>1.5ng/ml: 30.6 (95% CI:15.4-43.6)

DCP>1.5ng/ml: 17.6 (95% CI:3.8-32.1)

GALAD>0: 46.9 (95% CI:28.6-60.7)

GALAD>-0.03: 32.4 (95% CI:14.3-50.0)

HES>7.96: 40.8 (95% CI:24.3-53.8)

HES>7.94: 41.2 (95% CI:23.5-57.1)

Figure 3:

Receiver operating characteristics (ROC) curve within 24 months prior to HCC diagnosis.

Below the figure, we describe patient-level true positive rate at optimal cutoff corresponding to 10% screening-level false positive rate on the ROC curves.

Table 1:

Demographic characteristics of the analysis cohort of patients with cirrhosis.

	Full Analysis Cohort	Incident HCC cases during follow-up	Cirrhosis patients with no HCC during follow-up	P-value (HCC vs Control)*
Number of subjects	534	50	484	
Age at study consent (years)	63.2 (6.6)	64.7 (5.7)	63.0 (6.7)	0.13
Race				0.13
White	263 (49%)	21 (42%)	242 (50%)	
Black	206 (39%)	18 (36%)	188 (39%)	
Latino	52 (10%)	9 (18%)	43 (9%)	
Other	13 (2%)	2 (4%)	11 (2%)	
Gender				0.62
Male	522 (98%)	50 (100%)	472 (98%)	
Female	12 (2%)	0	12 (2%)	
Number of study visits				
1	161 (30%)	17 (34%)	144 (30%)	
2	132 (25%)	12 (24%)	120 (25%)	
3	108 (20%)	12 (24%)	96 (20%)	
4	70 (13%)	3 (6%)	67 (14%)	
5 or more	63 (13%)	6 (12%)	57 (12%)	
Baseline Lab Tests, mean (SD)				
ALT (U/L)	49.2 (48.8)	46.1 (32.1)	49.5 (50.3)	0.70
AST (U/L)	56.4 (48.4)	60.6 (43.7)	56.0 (48.9)	0.09
ALP (U/L)	95.4 (62.1)	110.6 (48.6)	93.8 (63.2)	<0.001
GGT (U/L)	118.0 (152.6)	129.0 (130.3)	116.8 (154.9)	0.11
Platelets (K/cmm)	149.0 (67.4)	115.5 (45.2)	152.5 (68.4)	<0.001
Albumin (g/L)	3.7 (0.6)	3.5 (0.9)	3.7 (0.6)	<0.001
T-Bilirubin (umol/L)	1.4 (5.4)	1.3 (0.8)	1.4 (5.7)	0.04
Creatinine (umol/L)	1.0 (0.7)	0.9 (0.2)	1.0 (0.7)	0.07
Etiology				N/A (non-mutually exclusive etiology categories)
HCV	381 (71%)	42 (84%)	339 (70%)	
Active HCV	92 (24%)	19 (45%)	73 (22%)	
Receiving treatment	147 (39%)	18 (43%)	129 (38%)	
Treated with SVR	239 (63%)	13 (31%)	226 (67%)	
HBV	12 (2%)	0	12 (2%)	
Alcoholic liver disease	326 (61%)	36 (72%)	290 (60%)	
NAFLD/NASH	114 (21%)	10 (20%)	104 (21%)	
Others	24 (4%)	4 (8%)	20 (4%)	

*Data reported are means and standard deviations for continuous variables and counts and percentages for categorical variables.

* P-values from Mann-Whitney U-test for continuous variables and Fisher's exact test for categorical variables

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2:

Patient-level sensitivity or true positive rate (TPR) and screening test-level false positive rate (FPR) for detecting any stage HCC at published thresholds for 6, 12, and 24 months prior to HCC diagnosis and the associated 95% bootstrap percentile intervals.

	HCC within 6 months (HCC cases: 33, Non-HCC: 526)		HCC within 12 months (HCC cases: 42, Non-HCC: 517)		HCC within 24 months (HCC cases: 49 Non-HCC: 497)	
	TPR (%)	FPR (%)	TPR (%)	FPR (%)	TPR (%)	FPR (%)
AFP>20 ng/mL	18.2 (6.2–32.1)	4.1 (2.8–5.5)	23.8 (11.4–37.5)	3.8 (2.6–5.2)	22.4 (11.4–34.1)	3.7 (2.5–5.1)
AFP-L3%>10%	33.3 (16.7–48.5)	5.5 (3.9–7.3)	35.7 (21.1–50.0)	5.2 (3.6–7.0)	32.7 (19.6–45.6)	5.0 (3.4–6.9)
DCP>7.5 ng/mL	15.2 (4.2–29.4)	1.4 (0.7–2.1)	11.9 (3.0–22.7)	1.4 (0.8–2.2)	12.2 (4.3–22.2)	1.3 (0.7–2.0)
AFP>20 ng/mL, AFP-L3%>10% or DCP>7.5 ng/mL	39.4 (23.3–56.2)	9.7 (7.8–11.9)	47.6 (31.9–62.8)	9.3 (7.4–11.3)	46.9 (32.6–60.3)	8.8 (6.8–10.9)
GALAD> -0.63	63.6 (46.9–78.8)	22.9 (19.8–26.5)	73.8 (60.0–86.4)	22.2 (19.1–25.8)	71.4 (58.8–83.3)	21.5 (18.5–25.1)
HES>10.17%	24.2 (10.0–40.0)	5.9 (4.3–7.8)	26.2 (12.8–39.6)	5.6 (4.0–7.5)	26.5 (14.0–39.1)	5.3 (3.7–7.1)

Table 3:

Patient-level sensitivity or true positive rate (TPR) and screening test-level false positive rate (FPR) at published thresholds for detecting early-stage HCC within 6, 12, and 24 months prior to HCC diagnosis and the associated 95% bootstrap percentile intervals.

	HCC within 6 months (HCC cases: 26, Non-HCC: 513)		HCC within 12 months (HCC cases: 30, Non-HCC: 507)		HCC within 24 months (HCC cases: 34, Non-HCC: 494)	
	TPR (%)	FPR (%)	TPR (%)	FPR (%)	TPR (%)	FPR (%)
AFP>20 ng/mL	19.2 (5.0–36.4)	3.9 (2.6–5.3)	23.2 (9.1–39.3)	3.8 (2.6–5.2)	23.5 (10.0–38.2)	3.7 (2.4–5.0)
AFP-L3%>10%	26.9 (10.7–44.4)	5.3 (3.7–7.1)	30.0 (14.3–46.4)	5.1 (3.6–7.0)	36.5 (12.1–41.7)	5.0 (3.5–6.9)
DCP>7.5 ng/mL	11.5 (0.0–26.7)	1.3 (0.7–2.0)	10.0 (0.0–22.6)	1.3 (0.7–2.1)	8.8 (0.0–20.0)	1.3 (0.7–2.0)
AFP>20 ng/mL, AFP-L3%>10% or DCP>7.5 ng/mL	30.8 (13.3–48.1)	9.4 (7.4–11.4)	40.0 (22.2–57.1)	9.1 (7.2–11.2)	38.2 (21.9–54.8)	8.7 (6.8–10.9)
GALAD> -0.63	53.8 (34.4–71.9)	22.1 (18.9–25.7)	63.3 (45.8–80.0)	21.8 (18.7–25.4)	61.8 (45.5–77.8)	21.5 (18.4–25.1)
HES>10.17%	26.9 (10.5–45.5)	5.5 (3.9–7.4)	23.3 (9.1–39.3)	5.5 (3.8–7.3)	23.5 (10.0–38.2)	5.2 (3.6–7.1)

Table 4:

Concordance between GALAD and HES algorithms results (based on thresholds corresponding to 10% screening-level false positive rate) within 6, 12 and 24 months prior to diagnosis in patients with any stage HCC.

Any Stage HCC						
	6 months		12 months		24 months	
	HES positive	HES negative	HES positive	HES negative	HES positive	HES negative
GALAD positive	16 (48.5%)	3 (9.1%)	20 (47.6%)	3 (7.1%)	21 (42.5%)	5 (10.2%)
GALAD positive	5 (15.2%)	9 (27.3%)	5 (11.9%)	14 (33.3%)	8 (16.3%)	15 (30.6%)
Early Stage HCC						
	HES positive	HES negative	HES positive	HES negative	HES positive	HES negative
GALAD positive	15 (57.7%)	3 (11.5%)	18 (60.0%)	3 (10.0%)	18 (52.9%)	5 (14.7%)
GALAD positive	1 (3.8%)	7 (26.9%)	1 (3.3%)	8 (26.7%)	2 (5.9%)	9 (26.5%)