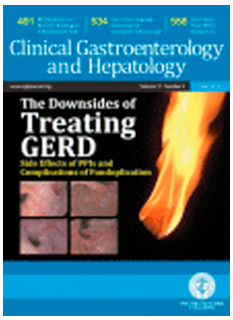
Journal Pre-proof



GALAD Score Detects Early Hepatocellular Carcinoma in an International Cohort of Patients With Nonalcoholic Steatohepatitis

Jan Best, Lars P. Bechmann, Jan-Peter Sowa, Svenja Sydor, Alexander Dechêne, Kristina Pflanz, Sotiria Bedreli, Clemens Schotten, Andreas Geier, Thomas Berg, Janett Fischer, Arndt Vogel, Heike Bantel, Arndt Weinmann, Jörn M. Schattenberg, Yvonne Huber, Henning Wege, Johann von Felden, Kornelius Schulze, Dominik Bettinger, Robert Thimme, Friedrich Sinner, Kerstin Schütte, Karl Heinz Weiss, Hidenori Toyoda, Satoshi Yasuda, Takashi Kumada, Sarah Berhane, Marc Wichert, Dominik Heider, Guido Gerken, Philip Johnson, Ali Canbay

PII: S1542-3565(19)31267-4
DOI: https://doi.org/10.1016/j.cgh.2019.11.012
Reference: YJCGH 56852

To appear in: *Clinical Gastroenterology and Hepatology*
Accepted Date: 3 November 2019

Please cite this article as: Best J, Bechmann LP, Sowa J-P, Sydor S, Dechêne A, Pflanz K, Bedreli S, Schotten C, Geier A, Berg T, Fischer J, Vogel A, Bantel H, Weinmann A, Schattenberg JM, Huber Y, Wege H, von Felden J, Schulze K, Bettinger D, Thimme R, Sinner F, Schütte K, Weiss KH, Toyoda H, Yasuda S, Kumada T, Berhane S, Wichert M, Heider D, Gerken G, Johnson P, Canbay A, GALAD Score Detects Early Hepatocellular Carcinoma in an International Cohort of Patients With Nonalcoholic Steatohepatitis, *Clinical Gastroenterology and Hepatology* (2019), doi: https://doi.org/10.1016/j.cgh.2019.11.012.

GALAD Score Detects Early Hepatocellular Carcinoma in an International Cohort of Patients With Nonalcoholic Steatohepatitis

Short title: Detection of early HCC in NASH

Jan Best[1], Lars P. Bechmann[1], Jan-Peter Sowa[1], Svenja Sydor[1], Alexander Dechêne[2,3], Kristina Pflanz[2], Sotiria Bedreli[2], Clemens Schotten[2], Andreas Geier[4], Thomas Berg[5], Janett Fischer[5], Arndt Vogel[6], Heike Bantel[6], Arndt Weinmann[7], Jörn M. Schattenberg[7], Yvonne Huber[7], Henning Wege[8], Johann von Felden[8], Kornelius Schulze[8], Dominik Bettinger[9,10], Robert Thimme[9], Friedrich Sinner[1], Kerstin Schütte[1], Karl Heinz Weiss[11], Hidenori Toyoda[12], Satoshi Yasuda[12], Takashi Kumada[13], Sarah Berhane[14], Marc Wichert[15], Dominik Heider[16], Guido Gerken[2], Philip Johnson[14], Ali Canbay[1]

[1]*University Hospital Magdeburg, Department of Gastroenterology, Hepatology and Infectious Diseases, 39120 Magdeburg, Germany;* [2]*University Hospital Essen, University Duisburg-Essen, Department of Gastroenterology and Hepatology, Hufelandstr. 55, 45122 Essen, Germany;* [3]*Paracelsus Medical University, Dept. of Gastroenterology, Hepatology and Endocrinology, Klinikum Nürnberg, 90419 Nürnberg;* [4] *Division of Hepatology, Department of Internal Medicine II, University Hospital Würzburg, Würzburg, Germany;* [5]*University Hospital Leipzig, Section of Hepatology, Division of Gastroenterology and Rheumatology, Liebigstraße 20, 04103 Leipzig, Germany;* [6]*Clinic of Gastroenterology, Hepatology and Endocrinology, Hannover Medical High School, Carl Neuberg Str. 1, 30625 Hannover, Germany;* [7]*Department of Medicine I, University Medical Center Johannes Gutenberg University, Mainz, Germany;* [8]*Department of Medicine I, University Medical Center Hamburg-Eppendorf, Germany;* [9]*Department of Medicine II, Medical Center University of Freiburg, Faculty of Medicine, University of Freiburg, Hugstetter Str. 55, D-79106, Freiburg, Germany;* [10]*Berta-Ottenstein-Programme, Faculty of Medicine, University of Freiburg, Freiburg, Germany;* [11]*Department of Internal Medicine IV, University Hospital Heidelberg, Heidelberg, Germany;* [12] *Department of Gastroenterology, Ogaki Municipal Hospital, Ogaki, Japan;* [13]*Department of Nursing, Gifu Kyoritsu University, Ogaki, Japan;* [14]*Department of Molecular and Clinical Cancer Medicine, University of Liverpool, Liverpool, UK,* [15]*University Hospital Essen, University Duisburg-Essen, Central Laboratory, Hufelandstr. 55, 45122 Essen, Germany,* [16]*Philipps-University Marburg, Department of Mathematics and Computer Science, Hans-Meerwein-Str. 6, 35037 Marburg, Germany.*

Grant support: This work was supported by the Deutsche Forschungsgemeinschaft (CA267/14-1) and by the Wilhelm-Laupitz Foundation to AC as well as the Werner Jackstaedt Foundation to LPB.

Abbreviations: AFP: alpha-fetoprotein; AFP-L3: isoform of AFP with an α 1-6 fucose residue attached at the N-acetylglucosamine terminus; DCP: des-gamma-carboxy prothrombin; GALAD: gender, age, AFP-L3, AFP and DCP; HCC: Hepatocellular carcinoma; US: ultrasound; USS: ultrasound scanning.

Correspondence:
Ali Canbay, MD
Otto-von-Guericke University
University Hospital Magdeburg,
Department of Gastroenterology, Hepatology, Infectiology
Leipziger Str. 44,
39120 Magdeburg

Germany
Phone: +49 391 67 13100
Fax: +49 391 67 13105
E-mail: ali.canbay@med.ovgu.de

Conflict of interest statement: The authors declare they have no financial or personal conflicts to declare.

Author Contributions
Study concept and design: JB, LPB, AD, PJ, GG, AC
Acquisition of data: JB, LPB, KP, SB, CS, TB, JF, AV, HB, AW, JMS, YH, HW, JvF, KS, DB, RT, FS, KS, KHW, MW, HT, SY,TK
Analysis and interpretation of data: JB, LPB, JPS, CP, SB, DH, PJ, AC, SS, AG, HT, SY,TK
Drafting of the manuscript: JB, PJ, AG, SS
Statistical analysis: JPS, SB, PJ, DH, HT, SY,TK
Obtained funding: AC, LPB
Technical, or material support: JB, LPB, KP, SB, CS, TB, JF, AV, HB, AW, JMS, YH, HW, JvF, KS, DB, RT, FS, KS, KHW, SB, MW, HT, SY,TK
Study supervision: AC
Critical revision of the manuscript for important intellectual content: all authors
Approval of the final version of the manuscript: all authors

Electronic Word count: 3995

Number of Figures and Tables: 3 Figures + 2 Tables + 2 Supplementary Figures + 2 Supplementary Tables

ABSTRACT

Background & Aims: The prevalence of nonalcoholic steatohepatitis (NASH) associated hepatocellular carcinoma (HCC) is increasing. However, strategies for detection of early-stage HCC in patients with NASH have limitations. We assessed the ability of the GALAD score, which determines risk of HCC based on patient sex; age; and serum levels of alpha-fetoprotein (AFP), AFP isoform L3 (AFP-L3), and des-gamma-carboxy prothrombin (DCP), to detect HCC in patients with NASH.

Methods: We performed a case-control study of 126 patients with HCC (20% within Milan Criteria) and 231 patients without HCC (NASH controls) from 8 centers in Germany. We compared the performance of serum AFP, AFP-L3, or DCP vs GALAD score to identify patients with HCC using receiver operating characteristic curves and corresponding area under the curve (AUC) analyses. We also analyzed data from 389 patients with NASH under surveillance for HCC in Japan, followed for a median 167 months. During the 5-year screening period, 26 patients developed HCC. To compensate for irregular intervals of data points, we performed locally weighted scatterplot smoothing, linear regression, and a non-linear curve fit to assess development of GALAD before HCC development.

Results: The GALAD score identified patients with any stage HCC with an AUC of 0.96 — significantly greater than values for serum levels of AFP (AUC, 0.88), AFP-L3 (AUC, 0.86) or DCP (AUC, 0.87). AUC values for the GALAD score were consistent in patients with cirrhosis (AUC, 0.93) and without cirrhosis (AUC, 0.98). For detection of HCC within Milan Criteria, the GALAD score achieved an AUC of 0.91, with a sensitivity of 68% and specificity of 95% at a cut-off of -0.63. In a pilot Japanese cohort study, the mean GALAD score was higher in patients with NASH who developed HCC than in those who did not develop HCC as early as 1.5 years before HCC diagnosis. GALAD scores were above -0.63 approximately 200 days before the diagnosis of HCC.

Conclusions: In a case–control study performed in Germany and a pilot cohort study in Japan, we found the GALAD score may detect HCC with high levels of accuracy in patients with NASH, with and without cirrhosis. The GALAD score can detect patients with early-stage HCC, and might facilitate surveillance of patients with NASH, who are often obese, which limits the sensitivity of detection of liver cancer by ultrasound.

KEY WORDS: NAFLD; tumor; diagnostic; serological assay; test

Need to Know:

BACKGROUND AND CONTEXT: The performance of ultrasound screening for hepatocellular carcinoma (HCC) is limited, especially in patients with nonalcoholic steatohepatitis (NASH).

NEW FINDINGS: In a retrospective analysis of patients with NASH in Germany, the GALAD score (based on patient features and laboratory test results) detected HCC (even early-stage tumors) with greater than 90% sensitivity and specificity—higher values than for any biomarker alone. In a prospective study, the GALAD score identified patients who developed HCC as early as 1.5 y before their diagnosis.

LIMITATIONS: The number of patients with early-stage HCCs was small and the frequency of liver cirrhosis in our HCC population might not be typical.

IMPACT: The GALAD score might be used in surveillance for HCC in patients with NASH.

Lay Summary: Analysis of several factors, including patient sex, age, and blood levels of several proteins, can identify patients with NASH who are at high risk for HCC.

INTRODUCTION

In industrialized countries, non-alcoholic steatohepatitis (NASH) is seen as a major risk factor for increasing incidence of hepatocellular carcinoma (HCC)[1, 2]. The current American Association for the Study of Liver Diseases (AASLD) HCC guidelines recommend surveillance by ultrasound (USS) with or without alpha-fetoprotein (AFP)[3], given the inherent limitations of USS-based surveillance[4-8]. A recent meta-analysis reported that using alpha fetoprotein (AFP) in combination significantly improves USS sensitivity to detect HCC at an early stage in patients with cirrhosis[9].

To address the insufficient performance of USS in HCC detection the GALAD score (gender, age, AFP-L3, AFP, DCP) was developed[10]. GALAD exhibited promising results for detection of early HCC in an initial prospective study[10]. A subsequent multicentre case-control study demonstrated that the GALAD model was superior to USS in early HCC detection, independent of etiology[11]. Although the latest AASLD guidelines note the potential of the GALAD score for HCC surveillance, it remarks on the need for phase III validation prior to adoption in routine clinical practice[3].

The limitations of USS alone for early detection of HCC are particularly evident in patients with NASH[6]. Serum–based biomarkers might be more effective, with or without USS, for HCC surveillance in NASH patients, although data in this patient population are currently lacking. The current study assessed the performance of the GALAD score for early HCC detection in patients with NASH-related liver disease.

PATIENTS AND METHODS

Ethics

The study was approved by the institutional review board / ethics committee of each German center and the Japanese center and carried out in accordance with the 1964 Helsinki Declaration.

Retrospective case-control population

The performance of the GALAD model was assessed in a case-control dataset from 8 German centers including 357 patients with NASH. NASH was defined according to histological features, when available, or by presence of metabolic syndrome and absence of any history of significant alcohol intake, viral hepatitis, or other possible causes of liver disease[12]. Significant alcohol intake was defined as consumption of more than two drinks daily or more than six drinks daily on weekends for at least 5 years[13]. The presence of HBV was excluded by qualitative sero-negativity for HbsAg and HCV exclusion was made by negative anti-HCV-IgG or HCV RNA. Cirrhosis was diagnosed either by histology or by overt clinical findings as portal hypertension in known chronic liver diseases. Details on this cohort are given in **Supplementary Material**.

Prospective cohort population

We conducted a pilot phase III evalution of GALAD using data from a prospective cohort recruited at a single Japanese centre. NAFLD was diagnosed based on the latest AASLD guidelines[14]: 1) fatty change of the liver observable by imaging; 2) no marked alcohol drinking habit present (ethanol intake of <210g per week for men and <140g per week for women); 3) no other causes for fatty change of the liver present; and 4) no other chronic liver disease (HCV, HBV, primary biliary cholangitis, or autoimmune hepatitis). Details on this cohort are given in **Supplementary Material.**

Clinical assessment

In both datasets, cirrhosis was diagnosed by histology or overt clinical findings including portal hypertension in the setting of known chronic liver disease. HCC was diagnosed according to the European Association for the Study of the Liver (EASL) guidelines via histology or by two different imaging modalities (dynamic contrast CT or MRI of the liver)[15, 16]. The Barcelona Clinic Liver Cancer (BCLC) staging system was used for HCC stage[17]. The GALAD score was calculated as described previously[18].

Statistical methods

Statistical analyses were performed with R (https://www.r-project.org) or Prism (GraphPad, La Jolla, CA, USA). Patient characteristics were compared between patients with and without HCC using Wilcoxon signed rank tests (continuous variables) or χ^2-tests (dichotomous variables). The GALAD model and single biomarker models were compared using receiver operating characteristic (ROC) curves and corresponding area under the curves (AUC) calculated using the R package ROCR[19]. AUC were compared according to DeLong et al.[20] with pROC[21]. The Youden's Index was calculated to identify the ideal cut-off for specific comparisons[22].

Longitudinal changes of GALAD prior HCC development were visualized via LOWESS (Locally Weighted Scatterplot Smoothing) plot. Since biomarker levels were assessed at irregular intervals before HCC diagnosis for each patient, linear regression models for each patient (square and cubic terms) were fitted. GALAD was used as a dependent variable and time to diagnosis as the explanatory variable. For each patient, model-predicted values at various time points were generated. Median

GALAD value of all patients at each time point was then calculated and 95% confidence intervals were produced by the bootstrap method. A non-linear fit model (with three parameters) was calculated over all available GALAD values prior HCC development with Prism.

RESULTS

German multicenter case-control study

Patient Characteristics

Three hundred fifty-seven patients with NASH were enrolled in the German multicenter case-control study, including 126 with and 231 without HCC **(Supplementary Figure 1A)**. All demographic and clinical details on this cohort are given in **Table 1.** The patients without HCC had a median age of 52 years and a mean BMI of 29kg/m². NASH patients without HCC were male in 51.9% of cases; 38.5 % presented with diabetes and 20.9 % with cirrhosis. NASH-HCC patients were significantly older with median age of 70.5 but had similar BMI of 29kg/m². The proportions of male sex (67.2%), diabetes (71.4%), and cirrhosis (76.2%) were significantly higher in NASH-HCC patients.

GALAD performance for HCC detection at any stage

Median levels of AFP, AFP-L3, DCP and the GALAD score were significantly higher in the NASH-HCC patients compared to NASH controls **(Table 1)**. Common cut-offs of individual markers reached only moderate sensitivity for detection of HCC at any stage (AFP 56.8% cut-off: 10ng/ml, AFP-L3 56.8% cut-off: 10%, DCP 81.6% cut-off:

0.76ng/ml) at high respective specificities (98.7%; 96.5%; 87.88%). Overall diagnostic performance, as assessed by AUC, was significantly higher (p<0.0005 for all) for GALAD than each individual biomarker **(Figure 1A)**. At the established cut-off of -0.63, he GALAD score achieved superior sensitivity of 84.8% at a specificity of 95.2%[23,24]. For the current study a optimal cut-off of -1.334 (sensitivity: 91.2%, specificity: 90.9%) was identified by the highest Youden's Index **(Supplementary Table 1)**.

GALAD performance for HCC detection at an early stage

Performance of GALAD and each biomarker in isolation were analyzed for early HCC detection, defined as BCLC stage A HCC (n=28). GALAD reached an AUC of 0.92 for early HCC detection, which was superior to AFP (*p*=0.0021), AFP-L3 (*p*<0.0001), and DCP (*p*>0.5) **(Figure 1B).** The cut-off -0.63 achieved 72.4% sensitivity and 95.2% specifiticity, whereas the newly calculated cut-off -1.334 achieved 86.2% sensitivity and 90.9% specificity.

Results were similar when early HCC was defined using Milan criteria (n=25; 20% of all NASH-HCC), with GALAD achieving an AUC of 0.90 (**Figure 1C**). Sensitivity of 68.0% and specificity of 95.2% were achieved at the common cut-off -0.63. The cut-off -1.334 reached 84.0% sensitivity and 90.9% specificity.

Performance of GALAD by cirrhosis status

In the subgroup analysis of patients with cirrhosis (n=95 HCC and 49 controls), GALAD achieved an AUC of 0.93 for HCC detection at any stage, which was significantly higher than AFP (AUC 0.79; p=0.0003) , AFP-L3 (AUC 0.75; p<0.0001), and DCP (0.83; p=0.0033) (**Figure 2A**). For early, cirrhotic HCC detection (n=22

HCC), GALAD achieved an AUC of 0.85, with 68.2% sensitivity and 91.8% specificity at -0.63 cut-off. Among those with non-cirrhotic NASH (n = 30 HCC and 182 controls), GALAD reached an AUC of 0.98 for HCC detection at any stage (**Figure 2B**), with 93.3% sensitivity and 96.1% specificity at the cut-off -0.63. For early, non-cirrhotic HCC detection (n=7 HCC), GALAD achieved an AUC of 0.94, with 85.7% sensitivity and 96.2% specificity at -0.63 cut-off.

Japanese prospective cohort study

Patient Characteristics

392 patients with NAFLD were prospectively followed under surveillance for HCC. Among these 28 patients developed HCC after a median of 10.3 [range: 3.1-21.3] years from start of surveillance. GALAD values prior to HCC diagnosis were available for 26 patients, with median time to HCC from the earliest available GALAD of 31 [range: 0-172] months. The 363 patients who did not develop HCC during surveillance had a median age of 68 and mean BMI of 25kg/m². Of these patients 49.6% were male, 38.6% had diabetes, and 19.6% had cirrhosis. The 26 patients developing HCC under surveillance were 61.5 y old (median) and had a mean BMI of 26.95kg/m². Patients with HCC development had a proportion of 80.8% males, 80.8% diabetic patients, and 34.6% with cirrhosis. For 17 patients more than one GALAD value was available prior HCC diagnosis with a median follow up of 9.8 [range: 3.2-19.2] years until HCC developed.

Pilot Phase III Evaluation of GALAD

To estimate the time frame before HCC diagnosis, in which the GALAD could indicate HCC, available GALAD values of 28 patients who developed HCC were

plotted by date prior HCC diagnosis (**Figure 3C**). A non-linear fit through these data points suggested a cut-off -0.63 would have been exceeded 200 days prior HCC diagnosis (**Figure 3D**).

Among 17 patients who had more than one set of biomarkers prior to HCC diagnosis **(Supplementary Table 2; Supplementary Figure 1B)** and 363 NASH controls (without HCC development during surveillance), changes in the GALAD score before HCC development are illustrated by LOWESS plot **(Figure 3A)** and aggregate curves generated by linear regression **(Figure 3B)**. While incremental increase of GALAD over time also occurred in NAFLD controls, patients who developed HCC had significantly higher GALAD scores several years prior to HCC diagnosis and a strong increase within a shorter time frame prior to diagnosis.

DISCUSSION

At present, NAFLD affects 25% of the world population[25]. NAFLD patients exhibit a high rate of disease progression to NASH-cirrhosis and 5-10% of HCC cases are attributed to NASH[26-28]. No effective screening strategy for NASH-derived HCC exists[29] with insufficient sensitivity of current USS and AFP based approaches to detect early HCC[15, 17, 5]. Surveillance of NASH patients for HCC is inadequate[30, 31] in particular as hepatocarcinogenesis may occur even in the absence of cirrhosis[2]. In our international multicenter study the GALAD score was successfully tested for detection of early and non-cirrhotic HCC to address current limitations in early HCC detection in NASH.

The GALAD model includes AFP, AFP-L3 and DCP, which combined are superior in HCC detection in Asian patients[32, 33], and incorporates gender and age, since older age and male sex represent independent HCC risk factors[34]. The GALAD model

exhibited excellent performance in HCC detection in multiple validation studies, including very large cohorts and early stage HCC (BCLC 0/A) of various etiologies[24, 35]. In addition the GALAD model separated HCC from pancreatic adenocarcinoma/cholangiocarcinoma (AUROC: 0.95)[24], suggesting that GALAD is specific for HCC. In a recent, large American cohort study by Yang et al., including patients from the Early Detection Research Network Phase 2 HCC Study, the performance of GALAD over all etiologies was significantly superior to USS for HCC detection[11]. However, Yang et al. found only an AUC of 0.89 for NASH-specific HCC detection by GALAD, which was not significantly better than USS[11]. In the present multicenter study of NASH-HCC patients we demonstrate an excellent performance of the GALAD for detection of NASH HCC (AUC 0.96). The lower performance in the US-American cohort might be due to smaller group sizes (30 NASH-HCC, 49 non-HCC NASH) in this etiology. Interestingly Yang et al.[11] identified the ideal GALAD cut-off to detect NASH-HCC at -0.86, which is between the established cut-off and the one identified by us.

GALAD can detect HCC independent of etiology, though insufficient data is available on a possible application for screening or detection of early stage HCC, in particular in NASH. Only 22.3% of the presented cohort comprised early stage HCC, similar to a previous study[36], supporting the demand for an effective screening approach, to detect HCC at earlier stages. Despite the relatively low number of BCLC 0/A patients the GALAD score achieved a significantly better AUC (0.92) than individual biomarkers. We could also demonstrate that the performance of GALAD was independent of cirrhosis, as similar AUCs were achieved for patients without cirrhosis (AUC 0.98) or with cirrhosis only (AUC 0.93). When limiting the analysis further to HCC patients without cirrhosis and at BCLC stage 0/A (n=7) the GALAD score could

still separate NASH patients without cirrhosis with high sensitivity and specificity. Furthermore, GALAD could separate NASH HCC within Milan criteria from controls with good sensitivity and specificity. These data on NASH-HCC patients demonstrate that GALAD can detect HCC independent of cirrhosis or stage of HCC. Indeed, even early non-cirrhotic NASH-HCC seems clearly separable from NASH controls, as even small groups resulted in robust performance. The previously established, etiology spanning cut-off -0.63 reached good sensitivities and specifities in all analyses, while in our cohort a cut-off of -1.334 reached very high sensitivities and the highest sum of sensitivity and specificity, especially for early stage NASH HCC (**Supplementary Table 1**). Thus, we propose to apply the -1.334 cut-off in prospective multicentre studies for further validation.

In addition to the retrospective data, GALAD was also tested in a prospectively recruited cohort of NAFLD patients under HCC surveillance. The GALAD was significantly higher in patients developing HCC during the observation period than in those without HCC formation. Different models were calculated for the surveillance data, demonstrating a sharp rise of the GALAD within few months prior HCC diagnosis. Depending on chosen cut-off detection of HCC could have been possible 200 or even 560 days prior HCC diagnosis. While this specific result has to be confirmed in further prospective studies, it is a promising observation for potential use of GALAD as screening tool in NASH patients.

Limitations of this study are the relatively low number of early stage BCLC 0/A cases, which would be the screening target group. However, analysis in this subgroup still yielded robust results and it should be noted that the base population is currently not screened in most centres. USS results were not available from participating German centres on the majority of patients. Frequency of liver cirrhosis in this HCC population

might not be typical, though subgroup analysis demonstrated that performance of the GALAD was independent of cirrhosis. The lack of BCLC classification and absence of histological NASH confirmation in the NAFLD surveillance cohort are another limitation.

In conclusion, our data confirm that the GALAD score is superior to individual serum markers for detection of HCC in NASH, independent of tumor stage or cirrhosis. The findings suggest, that GALAD should be investigated as potential tool for screening of NASH individuals to detect HCC at a resectable stage in a sufficiently large prospective study to identify a cut-off.

Acknowledgement: We thank Dr. Robert Küper for the technical support.

References

1. Younossi Z, Stepanova M, Ong JP, et al. Nonalcoholic Steatohepatitis Is the Fastest Growing Cause of Hepatocellular Carcinoma in Liver Transplant Candidates. Clin Gastroenterol Hepatol. 2019;17(4):748-55 e3.
2. Ertle J, Dechene A, Sowa JP, et al. Non-alcoholic fatty liver disease progresses to hepatocellular carcinoma in the absence of apparent cirrhosis. Int J Cancer. 2011;128(10):2436-43.
3. Heimbach JK, Kulik LM, Finn RS, et al. AASLD guidelines for the treatment of hepatocellular carcinoma. Hepatology. 2018;67(1):358-80.
4. Bolondi L, Sofia S, Siringo S, et al. Surveillance programme of cirrhotic patients for early diagnosis and treatment of hepatocellular carcinoma: a cost effectiveness analysis. Gut. 2001;48(2):251-9.
5. 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):1927-33 e2.
6. Simmons O, Fetzer DT, Yokoo T, et al. Predictors of adequate ultrasound quality for hepatocellular carcinoma surveillance in patients with cirrhosis. Aliment Pharmacol Ther. 2017;45(1):169-77.
7. Singal AG, Nehra M, Adams-Huet B, et al. Detection of hepatocellular carcinoma at advanced stages among patients in the HALT-C trial: where did surveillance fail? Am J Gastroenterol. 2013;108(3):425-32.
8. Wong LL, Reyes RJ, Kwee SA, et al. Pitfalls in surveillance for hepatocellular carcinoma: How successful is it in the real world? Clin Mol Hepatol. 2017;23(3):239-48.
9. 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. 2018;154(6):1706-18 e1.
10. Johnson PJ, Pirrie SJ, Cox TF, 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.
11. Yang JD, Addissie BD, Mara KC, et al. GALAD Score for Hepatocellular Carcinoma Detection in Comparison to Liver Ultrasound and Proposal of GALADUS Score. Cancer Epidemiol Biomarkers Prev. 2018.
12. Ascha MS, Hanouneh IA, Lopez R, et al. The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. Hepatology. 2010;51(6):1972-8.
13. Dawson DA, Grant BF, Harford TC. Variation in the association of alcohol consumption with five DSM-IV alcohol problem domains. Alcohol Clin Exp Res. 1995;19(1):66-74.
14. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. Hepatology. 2012;55(6):2005-23.
15. Bruix J, Sherman M, Practice Guidelines Committee AAftSoLD. Management of hepatocellular carcinoma. Hepatology. 2005;42(5):1208-36.
16. European Association For The Study Of The L, European Organisation For R, Treatment Of C. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol. 2012;56(4):908-43.
17. Bruix J, Sherman M, American Association for the Study of Liver D. Management of hepatocellular carcinoma: an update. Hepatology. 2011;53(3):1020-2.
18. Johnson PJ, Pirrie SJ, Cox TF, et al. The detection of hepatocellular carcinoma using a prospectively developed and validated model based on serological biomarkers. Cancer Epidemiology Biomarkers & Prevention. 2014;23(1):144-53.

19. Sing T, Sander O, Beerenwinkel N, et al. ROCR: visualizing classifier performance in R. Bioinformatics. 2005;21(20):3940-1.
20. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837-45.
21. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011;12:77.
22. Youden WJ. Index for rating diagnostic tests. Cancer. 1950;3(1):32-5.
23. Johnson PJ. The BALAD-2 and GALAD Biomarker Models for Hepatocellular Carcinoma. Gastroenterol Hepatol (N Y). 2017;13(4):231-3.
24. Berhane S, Toyoda H, Tada T, et al. Role of the GALAD and BALAD-2 Serologic Models in Diagnosis of Hepatocellular Carcinoma and Prediction of Survival in Patients. Clin Gastroenterol Hepatol. 2016;14(6):875-86 e6.
25. Younossi ZM, Blissett D, Blissett R, et al. The economic and clinical burden of nonalcoholic fatty liver disease in the United States and Europe. Hepatology. 2016;64(5):1577-86.
26. Younossi ZM, Otgonsuren M, Henry L, et al. Association of nonalcoholic fatty liver disease (NAFLD) with hepatocellular carcinoma (HCC) in the United States from 2004 to 2009. Hepatology. 2015;62(6):1723-30.
27. Dyson J, Jaques B, Chattopadyhay D, et al. Hepatocellular cancer: the impact of obesity, type 2 diabetes and a multidisciplinary team. J Hepatol. 2014;60(1):110-7.
28. Wong RJ, Cheung R, Ahmed A. Nonalcoholic steatohepatitis is the most rapidly growing indication for liver transplantation in patients with hepatocellular carcinoma in the U.S. Hepatology. 2014;59(6):2188-95.
29. Trevisani F, D'Intino PE, Morselli-Labate AM, et al. Serum alpha-fetoprotein for diagnosis of hepatocellular carcinoma in patients with chronic liver disease: influence of HBsAg and anti-HCV status. J Hepatol. 2001;34(4):570-5. Epub 2001/06/08.
30. Aby E, Phan J, Truong E, et al. Inadequate Hepatocellular Carcinoma Screening in Patients With Nonalcoholic Steatohepatitis Cirrhosis. J Clin Gastroenterol. 2019;53(2):142-6.
31. Oda K, Uto H, Mawatari S, et al. Clinical features of hepatocellular carcinoma associated with nonalcoholic fatty liver disease: a review of human studies. Clin J Gastroenterol. 2015;8(1):1-9.
32. Choi JY, Jung SW, Kim HY, et al. Diagnostic value of AFP-L3 and PIVKA-II in hepatocellular carcinoma according to total-AFP. World J Gastroenterol. 2013;19(3):339-46. Epub 2013/02/02.
33. Koike Y, Shiratori Y, Sato S, et al. Des-gamma-carboxy prothrombin as a useful predisposing factor for the development of portal venous invasion in patients with hepatocellular carcinoma: a prospective analysis of 227 patients. Cancer. 2001;91(3):561-9. Epub 2001/02/15.
34. Lok AS, Seeff LB, Morgan TR, et al. Incidence of hepatocellular carcinoma and associated risk factors in hepatitis C-related advanced liver disease. Gastroenterology. 2009;136(1):138-48. Epub 2008/10/14.
35. Chan SL, Mo F, Johnson PJ, et al. Performance of serum alpha-fetoprotein levels in the diagnosis of hepatocellular carcinoma in patients with a hepatic mass. HPB (Oxford). 2014;16(4):366-72.
36. Weinmann A, Alt Y, Koch S, et al. Treatment and survival of non-alcoholic steatohepatitis associated hepatocellular carcinoma. BMC Cancer. 2015;15:210.

FIGURE LEGENDS:

Figure 1: Performance of the GALAD model for detection of HCC in NASH. (**A**) GALAD achieved significantly higher area under the curve (AUC) than the individual biomarkers AFP, AFP-L3, and DCP in the overall cohort (n=; *P*<0.0005) and in the subgroup of early stage HCC (BCLC A; n=28) (**B**), except DCP. (**C**) GALAD could also separate NASH-HCC within Milan criteria (n=25) from NASH controls with good performance.

Figure 2: Cirrhosis does not influence performance of GALAD for detection of HCC or early HCC. To assess if cirrhosis or absence thereof has an impact on GALAD, analysis were divided into NASH-HCC with cirrhosis (**A**; n=93) vs. NASH cirrhosis (n=47) and NASH-HCC without cirrhosis (**B**; n=30) vs. NASH without cirrhosis (n=182). GALAD achieved good performance in the cirrhotic subpopulation and excellent performance in non-cirrhotic patients. Within the subgroup of early HCC (BCLC A) with cirrhosis (**C**; n=22) acceptable performance was reached. Separation of early HCC without cirrhosis (**D**; n=7) from non-cirrhotic controls was possible with good performance.

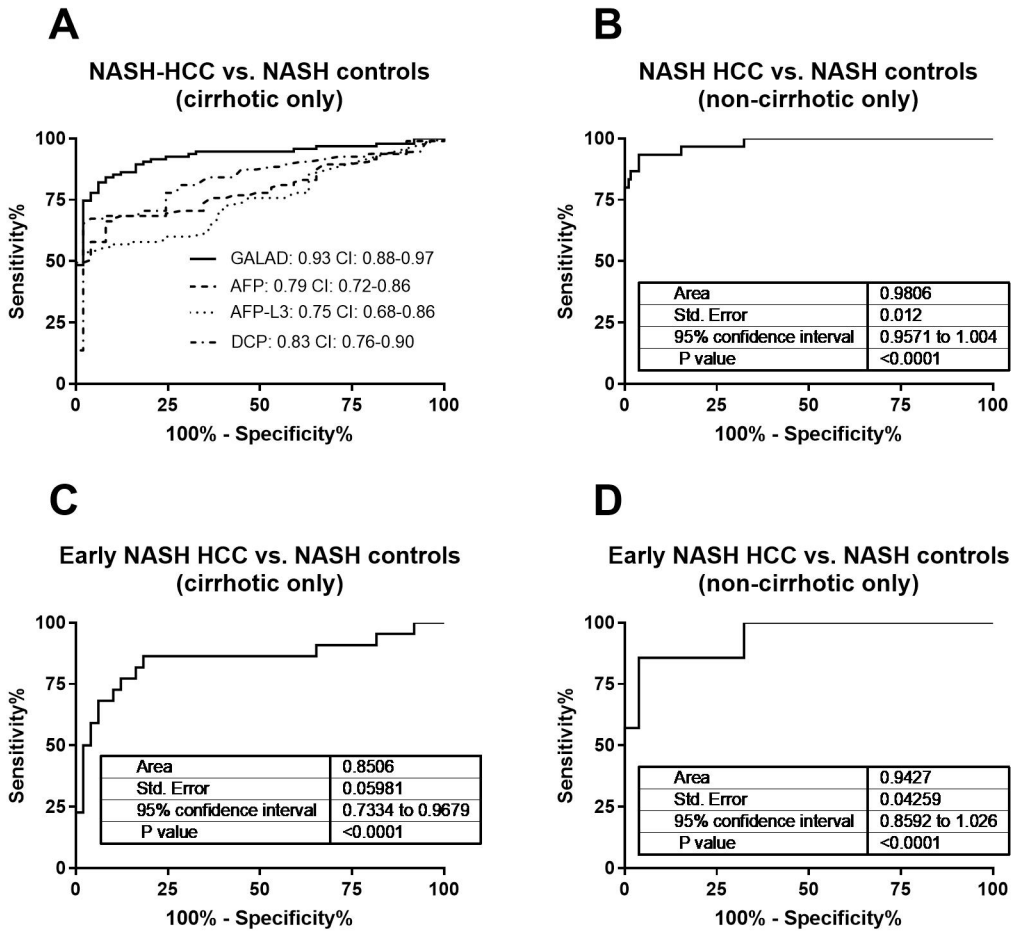
Figure 3: Rise of GALAD in NAFLD patients under surveillance occurs prior HCC diagnosis. In a surveillance cohort of NAFLD patients (n=389) GALAD was significantly higher in patients developing HCC. Rise of the GALAD score represented by (**A**) LOWESS plot and (**B**) linear regression based aggregate curves occurs months or even years before HCC development (n=17). (**C**) All available GALAD values of NAFLD patients developing HCC under surveillance (n=26) plotted against time prior HCC diagnosis resulted in a significant (p<0.0001) correlation. A non-linear (three parameter) fitting model through all datapoints resulted in an estimate curve (**D**), suggesting that the cut-offs -0.63 and -1.27 would have been surpassed 200 or even 500 days, respectively, prior HCC diagnosis. Timepoint "0" in all graphs marks date of HCC diagnosis.

Table 1: Patient characteristics					
		All German centers		Ogaki, Japan	
Variable		NASH-Non HCC (n=231)	NASH-HCC (n=126)	NASH-Non HCC (n=363)	NASH-HCC (n=26)
Age (years)		52 (44, 58.5)	70.5 (64, 75) ***	68 (59, 74)	62 (45, 75)
Sex male, n(%)		120 (51.9)	84 (67.2), n=125 ***	180 (49.6)	21 (80.8)
GALAD		-3.96 (-5.22, -2.6)	2.93 (0.77, 7.62), n=125 ***	-3.24 (-4.21, -2.20)	-0.30 (-3.29, 13.32)
AFP (ng/ml)		2.5 (1.8, 4.3)	16.1 (5.7, 497.7) ***	2.4 (1.7, 3.5)	5.9 (2.2, 35968)
AFP-L3 (%)		0.1 (0.1, 4.2)	13.2 (6.65, 46.6) ***	0.5 (0.5, 0.5)	0.85 (0, 33.5)
DCP (ng/ml)		0.37 (0.28, 0.5)	14.73 (1.14, 149.56) ***	0.20 (0.17, 0.25)	0.28 (0.13, 602.4)
Cirrhosis, n(%)		48 (20.9), n=230	96 (76.2) ***	71 (19.6), n=362 (FIB4≥3.25)	13 (81.3), n=16 (FIB4≥3.25)
Child-Pugh grade, n(%)	No cirrhosis	146 (76.4), n=191	28 (23.1), n=121 ***	Not available	Not available
	A	34 (18.3), n=191	69 (57.0), n=121 ***	Not available	Not available
	B	10 (5.2), n=191	20 (16.5), n=121 ***	Not available	Not available
	C	1 (0.5), n=191	4 (3.3), n=121 ***	Not available	Not available
BCLC stage, n(%)	A	Not applicable	29 (23.2), n=125	Not applicable	Not available
	B	Not applicable	61 (48.8), n=125	Not applicable	Not available
	C	Not applicable	30 (24.0), n=125	Not applicable	Not available
	D	Not applicable	5 (4.0), n=125	Not applicable	Not available
Albumin (g/l)		44 (40, 46.45), n=223	39 (33, 43), n=119 ***	44 (42, 46)	40.5 (29, 48)
Bilirubin (µmol/l)		8.6 (5.1, 13.2), n=225	12.1 (8.6, 20.5), n=125 ***	11.97 (8.55, 17.1), n=361	11.12 (5.13, 25.65)
BMI (kg/m^2)		29 (26.5, 33), n=219	29 (26, 33), n=119	25 (23.1, 27.1), n=237	27 (18.7, 34.7)
Diabetes, n(%)		89 (38.5)	90 (71.4) ***	140 (38.6)	21 (80.8)

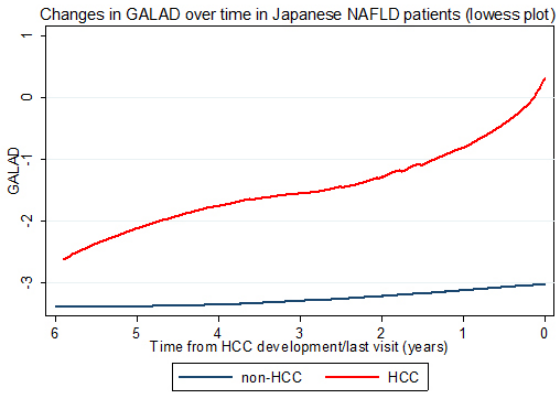
Abbreviations: AFP, alpha-fetoprotein; BCLC, Barcelona Clinic Liver Cancer; BMI body mass index; DCP, Des-gamma carboxyprothrombin; HCC, hepatocellular carcinoma; NASH, nonalcoholic steatohepatitis; *P <= .05; **P <= .01; ***P <= .001.

All continuous variables are presented as median (with interquartile range). Values are expressed as medians and range, i.e., first and third quantiles.

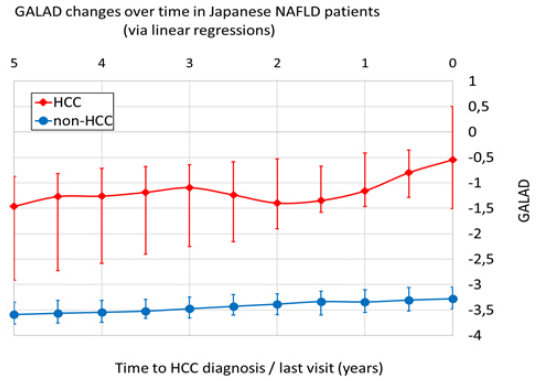
Significance is based on either Wilcoxon signed rank tests or χ^2-tests, depending on variable type.


A
NASH-HCC vs. NASH controls
(cirrhotic only)
Sensitivity%
100% - Specificity%
GALAD: 0.93 CI: 0.88-0.97
AFP: 0.79 CI: 0.72-0.86
AFP-L3: 0.75 CI: 0.68-0.86
DCP: 0.83 CI: 0.76-0.90
B
NASH HCC vs. NASH controls
(non-cirrhotic only)
Area 0.9806
Std. Error 0.012
95% confidence interval 0.9571 to 1.004
P value <0.0001
C
Early NASH HCC vs. NASH controls
(cirrhotic only)
Area 0.8506
Std. Error 0.05981
95% confidence interval 0.7334 to 0.9679
P value <0.0001
D
Early NASH HCC vs. NASH controls
(non-cirrhotic only)
Area 0.9427
Std. Error 0.04259
95% confidence interval 0.8592 to 1.026
P value <0.0001

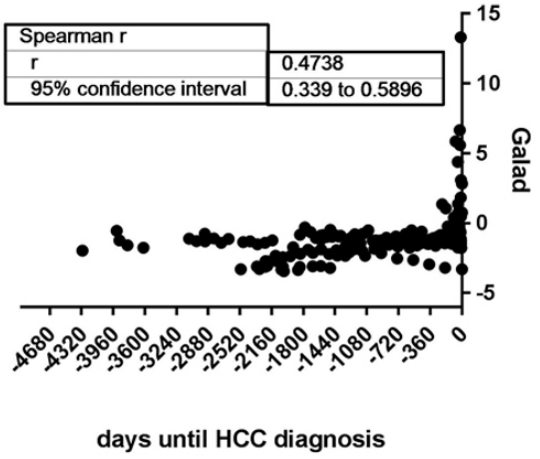
A


Changes in GALAD over time in Japanese NAFLD patients (lowess plot)
GALAD
Time from HCC development/last visit (years)
non-HCC
HCC

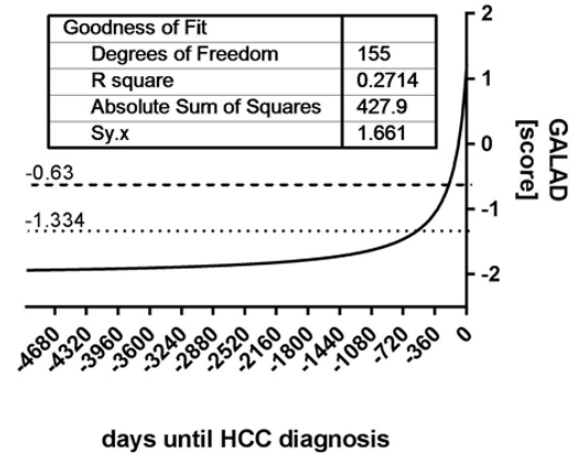
B

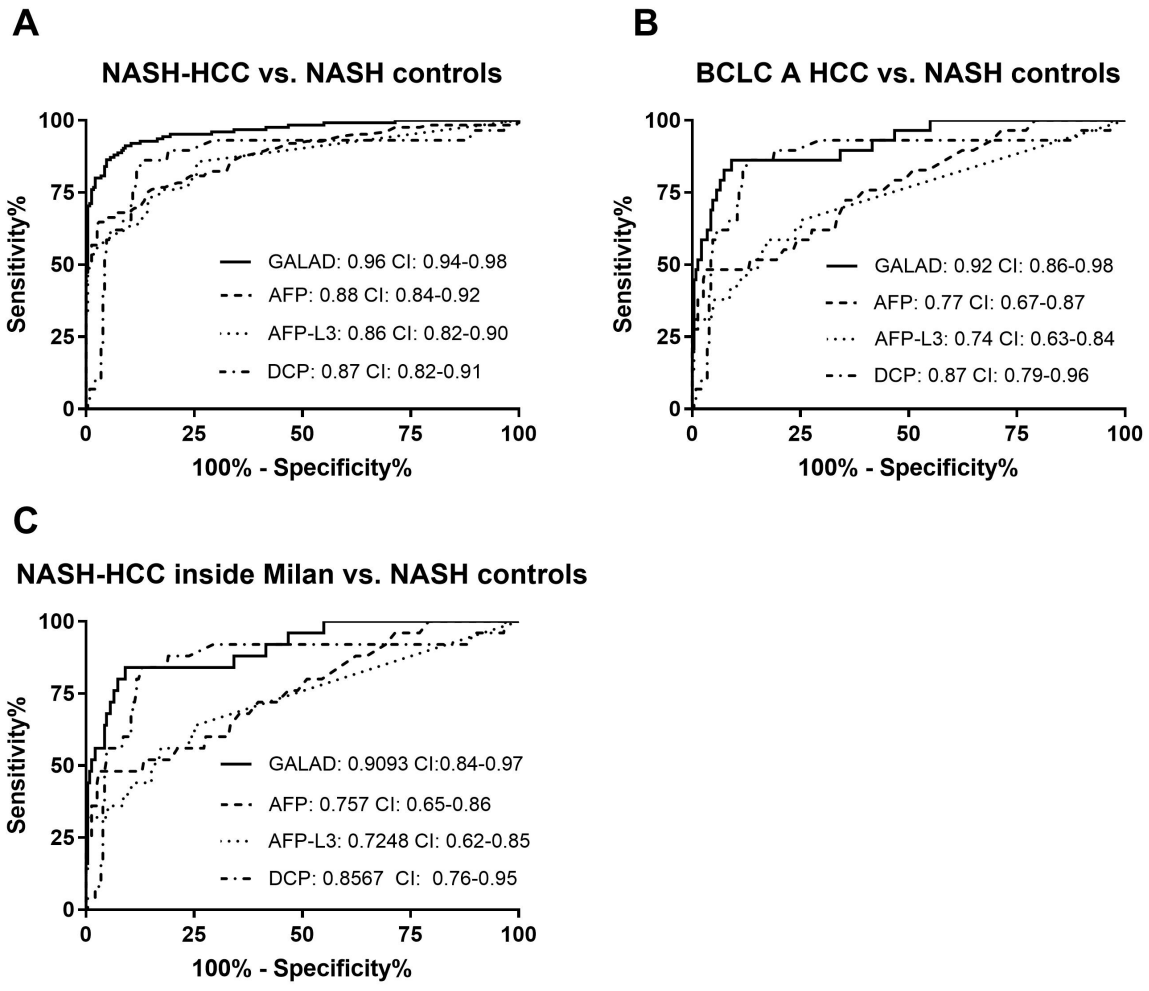

GALAD changes over time in Japanese NAFLD patients
(via linear regressions)
HCC
non-HCC
GALAD
Time to HCC diagnosis / last visit (years)

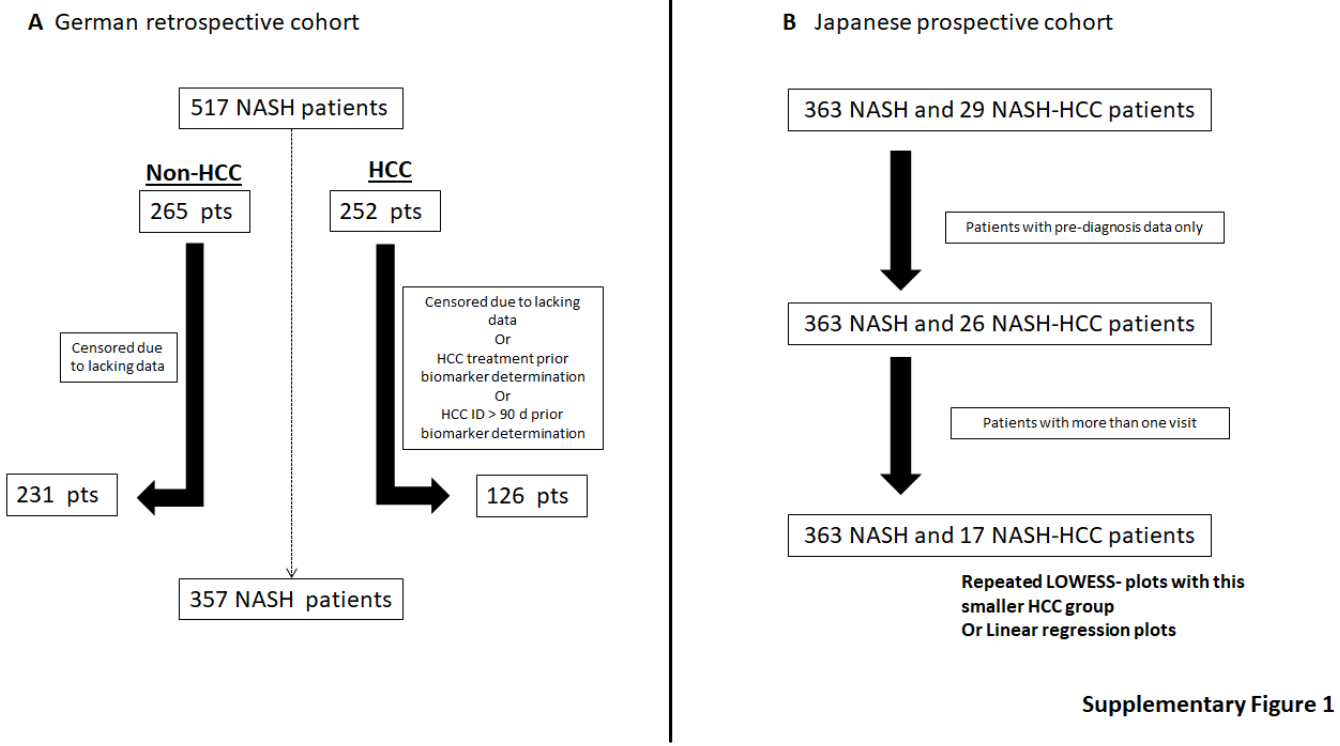
C


Spearman r
r
0.4738
95% confidence interval
0.339 to 0.5896
Galad
days until HCC diagnosis

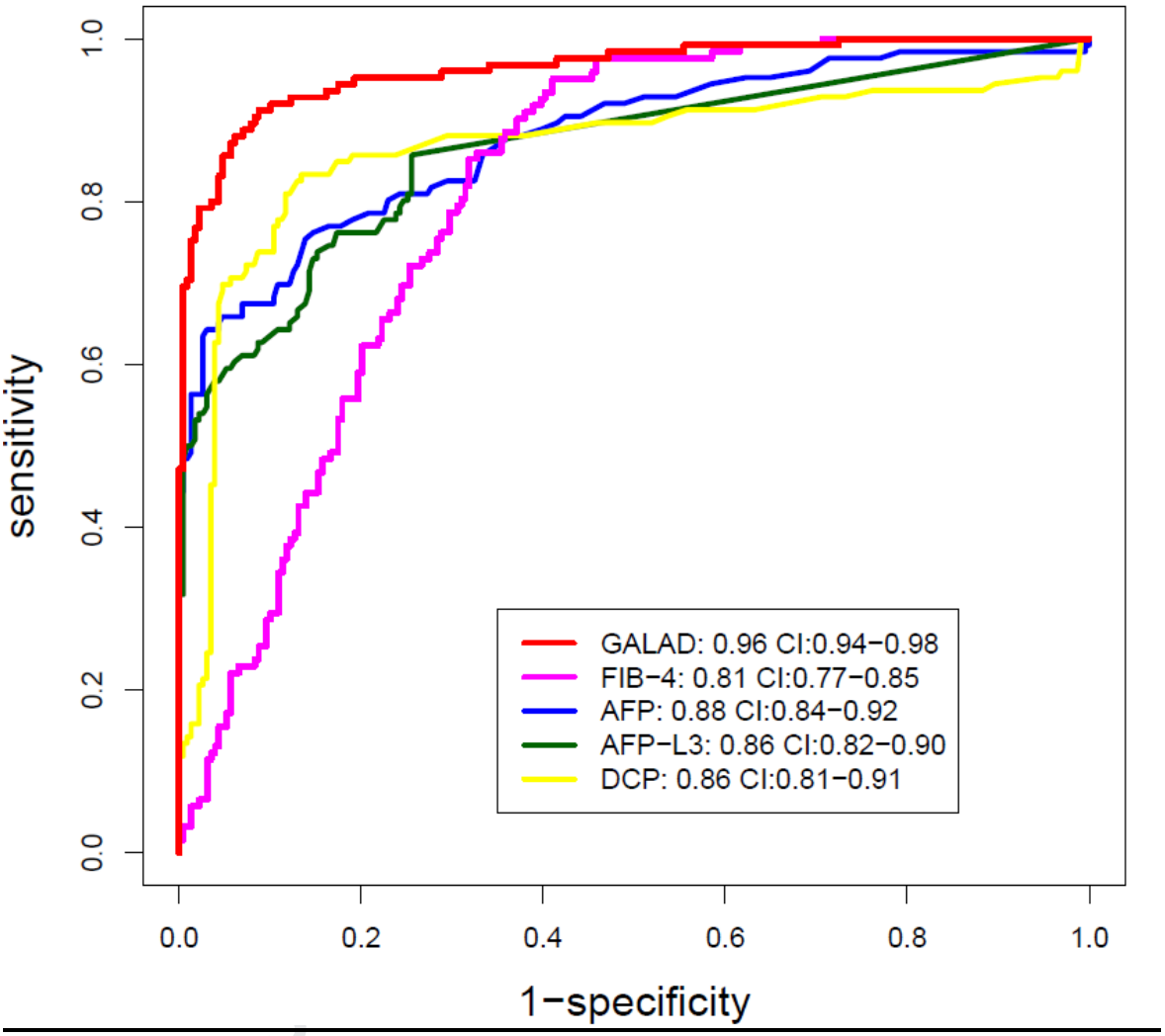
D


Goodness of Fit
Degrees of Freedom
155
R square
0.2714
Absolute Sum of Squares
427.9
Sy.x
1.661
-0.63
-1.334
GALAD [score]
days until HCC diagnosis


A
NASH-HCC vs. NASH controls
Sensitivity%
100% - Specificity%
GALAD: 0.96 CI: 0.94-0.98
AFP: 0.88 CI: 0.84-0.92
AFP-L3: 0.86 CI: 0.82-0.90
DCP: 0.87 CI: 0.82-0.91
B
BCLC A HCC vs. NASH controls
Sensitivity%
100% - Specificity%
GALAD: 0.92 CI: 0.86-0.98
AFP: 0.77 CI: 0.67-0.87
AFP-L3: 0.74 CI: 0.63-0.84
DCP: 0.87 CI: 0.79-0.96
C
NASH-HCC inside Milan vs. NASH controls
Sensitivity%
100% - Specificity%
GALAD: 0.9093 CI:0.84-0.97
AFP: 0.757 CI: 0.65-0.86
AFP-L3: 0.7248 CI: 0.62-0.85
DCP: 0.8567 CI: 0.76-0.95



Supplementary Figure 1: Patient selection and structure of presented cohorts. A Algorithm of patient selection for enrollment in the German retrospective cohort. **B** Japanese prospective surveillance program: Among 389 patients with NAFLD in a rigorous, prospective, surveillance program at the Ogaki Municipal Hospital, Japan, 29 developed HCC. Of these patients with HCC development during surveillance, 26 had GALAD values available prior HCC diagnosis, including 17 patients, who had more than one set of biomarkers recorded before HCC was diagnosed.



Supplementary Figure 3: Performance of GALAD and FIB-4 index for separation of NASH HCC from NASH controls. As cirrhosis could be a confounding factor for the predictive value of GALAD, performance of FIB-4 index was compared to GALAD (AUC: 0.96), reaching significantly lower AUC (0.81). Overall performance measured by AUC was lower for FIB-4 than for individual serum markers of the GALAD score.

Supplementary Table 1: Performance of GALAD at the cut-offs -0.63 and -1.334 for separation of NASH-HCC patients and NASH controls among various subgroups.

Subgroup	Sensitivity [%]	Specificity [%]	PPV[1]	NPV[2]	Correctly identified [%]
Cut-off -0.63 (previously established)					
All HCC (n = 125)	86.4	95.24	0.91	0.92	91.9
Within Milan criteria (n =25)	68	95.24	0.61	0.96	92.6
BCLC A (n = 29)	72.4	95.24	0.66	0.96	92.7
Non-cirrhotic BCLC A* (n = 7)	85.71	96.15	0.46	0.99	95.8
Cut-off -1.334 (identified in this study)					
All HCC (n = 125)	91.2	90.91	0.84	0.95	91.0
Within Milan criteria (n =25)	84	90.91	0.50	0.98	90.2
BCLC A (n = 29)	86.21	90.91	0.54	0.98	90.4
Non-cirrhotic BCLC A* (n = 7)	85.71	93.41	0.33	0.99	93.1

[1]: positive predictive value; [2]: negative predictive value; *: vs. non-cirrhotic NASH controls n = 182.

Supplementary Table 2: Patient data of Japanese Sub-group with more than 1 GALAD value available prior HCC diagnosis (n =17)

		Ogaki, Japan	
Variable		**NASH-Non HCC (n=363)**	**NASH-HCC (n=17)**
Age (years)		68 (59, 74)	69 (66, 77)
Sex male, n(%)		180 (49.6)	12 (70.6)
GALAD		-3.24 (-4.21, -2.20)	-0.60 (-1.49, 0.72)
AFP (ng/ml)		2.4 (1.7, 3.5)	8.2 (4.4, 13.6)
AFP-L3 (%)		0.5 (0.5, 0.5)	5.2 (0.5, 8.7)
DCP (ng/ml)		0.20 (0.17, 0.25)	0.40 (0.22, 1.03)
Cirrhosis, n(%)		71 (19.6), n=362 (FIB4≥3.25)	13 (81.3), n=16 (FIB4≥3.25)
Child-Pugh grade, n(%)	No cirrhosis	Not available	Not available
	A	Not available	Not available
	B	Not available	Not available
	C	Not available	Not available
BCLC stage, n(%)	A	Not applicable	Not available
	B	Not applicable	Not available
	C	Not applicable	Not available
	D	Not applicable	Not available
Albumin (g/l)		44 (42, 46)	38.5 (34.5, 43.5), n=16
Bilirubin (µmol/l)		11.97 (8.55, 17.1), n=361	23.09 (11.97, 25.65), n=16
BMI (kg/m^2)		25 (23.1, 27.1), n=237	24.1 (21.6, 27.3), n=7
Diabetes, n(%)		140 (38.6)	15 (88.2)

Abbreviations: AFP, alpha-fetoprotein; BCLC, Barcelona Clinic Liver Cancer; BMI body mass index; DCP, Des-gamma carboxyprothrombin; HCC, hepatocellular carcinoma; NASH, nonalcoholic steatohepatitis;

*P <= .05; **P <= .01; ***P <= .001.

All continuous variables are presented as median (with interquartile range). Values are expressed as medians and range, i.e., first and third quantiles. Significance is based on either Wilcoxon signed rank tests or χ^2-tests, depending on variable type.

The GALAD score as potential screening test for hepatocellular carcinoma in nonalcoholic steatohepatitis: An international multicenter study

Supplementary Information

Patients in the German Multicenter cohort

Patients of the retrospective German multicenter cohort were recruited from University Hospital (UH) Essen (collected between 2005 and 2016), UH Hannover (collected between 2008 and 2013), UH Leipzig (collected between 2010 and 2015), UH Mainz (collected between 2001 and 2013), UH Hamburg (collected between 2011 and 2015), UH Heidelberg (collected between 2012 and 2014), UH Magdeburg (collected between 2011 and 2013) and UH Freiburg (collected between 2010 and 2011). Cohort stratification is described in detail in **Supplementary Figure 1**. For HCC patients the biomarker measurement (AFP, AFP-L3, DCP) had to be performed within 3 months of HCC diagnosis and prior any treatment administration to be eligible for analysis.

Patients in the NASH-control group were recruited retrospectively, when all of the following three criteria were fulfilled: 1. Absence of any clinical evidence of HCC at the time the serum samples were taken. 2. Clinical presentation typical for those individuals that would be included in a screening program. 3. At least 6 months of follow up after GALAD score assessment to confirm the absence of HCC.

Patients with two or more diseases, which would predispose for cirrhosis and/or HCC, were excluded from this study.

Measurements of serological biomarkers

AFP, AFP-L3 and DCP were measured in the same serum sample (stored at -20℃) using the μTASWakoTM i30 fully automated immunoanalyzer (FUJIFILM Wako Chemicals Europe GmbH, Neuss, Germany). Liquid-phase binding assays followed by capillary electrophoresis and fluorescence detection in microchips were used for analysis[18]. Assay sensitivities were 0.3 ng/mL for AFP and 0.1 ng/mL for DCP. The percentage of AFP-L3 was determined in samples where both subfractions (AFP-L1 and AFP-L3) were >0.3 ng/mL.

GALAD

The GALAD model[25] was calculated according to the equation:

$$Z = -10.08 + 0.09 * age + 1.67 * gender + 2.34 * log_{10}(AFP) + 0.04 * AFP\_L3 + 1.33 * log_{10}(DCP)$$

with gender set as 1 for males and 0 for females.

The linear predictor (Z) is used to estimate the probability of HCC in an individual patient (ranging from 0 to 1) using the following equation: $Pr(HCC) = exp(Z)/(1 + exp(Z))$.

GALAD Score Detects Early Hepatocellular Carcinoma in an International Cohort of Patients With Nonalcoholic Steatohepatitis

Need to Know:

BACKGROUND AND CONTEXT: The performance of ultrasound screening for hepatocellular carcinoma (HCC) is limited, especially in patients with nonalcoholic steatohepatitis (NASH).

NEW FINDINGS: In a retrospective analysis of patients with NASH in Germany, the GALAD score (based on patient features and laboratory test results) detected HCC (even early-stage tumors) with greater than 90% sensitivity and specificity—higher values than for any biomarker alone. In a prospective study, the GALAD score identified patients who developed HCC as early as 1.5 y before their diagnosis.

LIMITATIONS: The number of patients with early-stage HCCs was small and the frequency of liver cirrhosis in our HCC population might not be typical.

IMPACT: The GALAD score might be used in surveillance for HCC in patients with NASH.

Lay Summary: Analysis of several factors, including patient sex, age, and blood levels of several proteins, can identify patients with NASH who are at high risk for HCC.