Published in final edited form as:

Gastroenterology. 2014 May ; 146(5): 1249–1255.e1. doi:10.1053/j.gastro.2014.01.045.

A New Laboratory Based Algorithm to Predict Development of Hepatocellular Carcinoma in Patients with Hepatitis C and Cirrhosis

Hashem B. El-Serag, Fasiha Kanwal, Jessica A. Davila, Jennifer Kramer, and Peter Richardson

¹Houston VA HSR&D Center of Excellence, Houston, Texas.

²Section of Health Services Research, Houston, Texas.

³Gastroenterology and Hepatology, Michael E. DeBakey VA Medical Center, Houston, Texas.

⁴Baylor College of Medicine, Houston, Texas.

Abstract

Background & Aims—Serum levels of α -fetoprotein (AFP) are influenced not only by the presence of hepatocellular carcinoma (HCC) but also by underlying severity and activity of liver disease, which is reflected by liver function tests. We constructed an AFP-based algorithm that included these factors to identify patients at risk for HCC, and tested its predictive ability in a large set of patients with cirrhosis.

Methods—We used the national Veterans Administration hepatitis C virus (HCV) clinical case registry to identify patients with cirrhosis, results from at least 1 AFP test, and 6 months of follow up. Our algorithm included data on age; levels of aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase, total bilirubin, albumin, creatinine, and hemoglobin; prothrombin time; and numbers of platelets and white cells. We examined the operating characteristics (calibration, discrimination, predictive values) of several different algorithms for identification of patients who would develop HCC within 6 months of the AFP test. We assessed our final model in the development and validation subsets.

© 2014 The American Gastroenterological Association. Published by Elsevier Inc. All rights reserved.

Address: Hashem B El-Serag, MD, MPH, Michael E. DeBakey VA Medical Center, 2002 Holcombe Blvd. (152), Houston, TX 77030, Phone (713) 794-8640, Fax (713) 748-7359 (hasheme@bcm.edu)..

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 citable 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.

Conflict of Interest: No conflicts of interest exist.

Author Contributions:

Hashem B. El-Serag: Funding, conception, design, interpreting results, manuscript writing, editing, decision to publish

Fasiha Kanwal Design, interpreting results, writing and editing manuscript, decision to publish

Jessica A. Davila Data collection, editing manuscript, decision to publish

Jennifer Kramer: Editing manuscript, decision to publish

Peter Richardson: Design, analysis, interpreting results, manuscript writing, editing, decision to publish

Results—We identified 11,721 patients with HCV-related cirrhosis in whom 35,494 AFP tests were performed, and 987 patients developed HCC. A predictive model that included data on levels of AFP, ALT, and platelets, along with age at time of AFP test (and interaction terms between AFP and ALT, and AFP and platelets), best discriminated between patients who did and did not develop HCC. Using this AFP-adjusted model, the predictive accuracy increased at different AFP cutoffs, compared with AFP alone. At any given AFP value, low numbers of platelets and ALT and older age were associated with increased risk of HCC, whereas high levels of ALT and normal/high numbers of platelets were associated with low risk for HCC. For example, the probabilities of HCC, based only on 20 ng/ml and 120 ng/ml AFP, were 3.5% and 11.4%, respectively. However patients with the same AFP values (20 ng/ml and 120 ng/ml) who were 70 y old, with ALT levels of 40 IU/ml and platelet counts of 100,000, had probabilities of developing HCC of 8.1% and 29.0%, respectively.

Conclusions—We developed and validated an algorithm based on levels of AFP, platelets, and ALT, along with age, which increased the predictive value for identifying patients with HCV-associated cirrhosis likely to develop HCC within 6 months. If validated in other patient groups, this model would have immediate clinical applicability.

Keywords

prognosis; liver cancer; prediction; risk factors

Introduction

Hepatocellular carcinoma (HCC) is the fastest rising cause of cancer related deaths in the United States, where hepatitis C virus (HCV) is the major underlying etiology for HCC. Survival with HCC is dismal (5 year≈10%) except in a relatively small number of patients whose cancers are detected early and receive potentially curative treatment. However, studies indicate that only 10% of patients with HCC receive these treatments.¹ Practice guidelines^{2;34} have recommended HCC surveillance for patients at high risk, such as those with HCV-related cirrhosis to increase the detection of HCC at an early stage and subsequent receipt of treatment.

Serum alpha-fetoprotein (AFP) has been extensively used as a biomarker for HCC, but its performance in HCC surveillance has been generally low. A systematic review of 5 high-quality studies that used a >20 ng/mL cutoff value for a positive AFP test reported that the sensitivity for detecting HCC was 41%–65% and specificity was 80%–94%⁵. Given these data, there have been recent calls including those of practice guidelines for completely abandoning AFP for HCC surveillance in lieu of liver ultrasound alone².

However, AFP levels and its test characteristics are likely influenced not only by the presence of HCC but also by the severity and activity of the underlying liver disease. Several U.S. based studies found an association between an elevated serum AFP level and decreased platelet count and increased AST/ALT ratio in HCV-infected patients with advanced hepatic fibrosis^{6;7}. In our previous study using a dataset of >250,000 AFP tests performed in 76,357 HCV-infected persons, we found a strong correlation (as well as an interaction) between AFP and ALT levels with the development of HCC⁸. We also found

that blood levels of platelets, AST, bilirubin, and albumin were significantly correlated with AFP levels. Incorporating these additional factors that influence AFP levels and that may also serve as independent predictors for HCC into an “adjusted” AFP-based algorithm may improve AFP's predictive ability in detecting HCC ⁹.

An adjusted AFP based algorithm may have an immediate utility and direct impact on clinical practice given the wide availability of AFP tests, high level of lab standardization, low cost, and above all, the absence of promising and readily available new biomarkers for HCC surveillance. We therefore examined the operating characteristics (such as calibration, discrimination and predictive values) of several adjusted AFP algorithms in detecting HCC within 6 months following the date of the AFP test in 11,721 patients with HCV-related cirrhosis.

Methods

Data Source and Study Population

We used data from the VA HCV Clinical Case Registry including patient demographic characteristics, laboratory test results, inpatient and outpatient visits and diagnosis, and date of death, on all HCV infected patients at 128 VA facilities. Patients were eligible for inclusion in the study if they had a positive test for HCV antibody and HCV RNA, and had cirrhosis (defined in the section below) between 1998 and 2005. Patients were followed until death, development of HCC, liver transplant or end of the study (12/31/2006). We included patients who had at least one valid serum AFP following the HCV index date. Those who developed HCC or died within 6 months of the HCV index date or had total follow up of less than 6 months were excluded from the study cohort.

Study Variables

Cirrhosis was defined by the presence of any of the three diagnostic codes 571.2, 571.5, or 571.6; we validated this definition against clinical, radiological, histological and biochemical criteria contained electronic medical records (EMR) and found it to have an 88% positive predictive value ¹⁰. The date of the first appearance of any of these codes was considered as the index date for cirrhosis. HCC was defined using two sequential steps. First, we screened the automated data for the presence of ICD-9 code 155.0 in the absence of 155.1 and identified unique patients who fulfilled these criteria. Second, among this group of likely HCC, we performed a manual structured review of EMR to confirm HCC diagnosis. The date of HCC diagnosis was defined as the date of the earliest appearance of a liver mass on ultrasound that was subsequently confirmed by CT, MRI and/or biopsy, or in the absence of a mass on ultrasound, by the first evidence on CT, MRI, or biopsy.

We collected all valid AFP tests (>0 ng/mL) for each study subject. We identified the following laboratory tests as candidates for inclusion in the adjusted AFP algorithm: AST, ALT, alkaline phosphatase, total bilirubin, albumin, prothrombin time, and platelets. We selected these laboratory tests as possible predictors because they are commonly used lab tests, can be clearly defined in EMR, and are measured using standardized and reproducible tests. We considered but excluded few additional candidate labs (e.g., ferritin and gamma

glutamyl transferase) because data were missing for >10% in our study patients suggesting that these tests are not frequently performed in clinical practice.

Analysis

To determine the predictive ability of AFP-based model (algorithm), we constructed multivariable logistic models. For each patient, the follow-up period from the HCV index date to the last follow up date was divided into 6-month intervals. Laboratory values (independent variables) from the preceding ($i-1^{\text{th}}$) 6-month period were used to predict the occurrence of HCC within the given (i^{th}) 6 month period (independent variable).

We modeled laboratory values as continuous variables to avoid loss of information, and ascertained whether each lab test could be treated as linear continuous or non-linear predictor of HCC. A combination of logarithmic transformations and spline functions were used to transform the candidate lab test value into a linear variable as appropriate³¹¹.

We started our model development with a bivariate analysis of HCC risk where AFP values were the sole predictor of HCC and unique patients were the unit of analysis. We subsequently added covariates one at a time and tested interactions between AFP and these covariates. Model-predicted HCC probabilities were generated and averaged over the set of 100 runs to obtain average probability estimates. Variables with p value <0.1 were considered for inclusion in the multivariable model.

At each step, models were assessed for discrimination (ability to separate HCC negative from HCC positive status), calibration (agreement of model-derived HCC probabilities with raw HCC probabilities) and predictive ability.

Discrimination was assessed by receiver operating curve analysis and calculation of the c-index and the Pencina-d'Agostino integrated discrimination improvement (IDI)¹². However, these conventional measures of discrimination may be insensitive to detecting small improvements in model performance when a new marker is added to a model that already includes important predictors. We therefore also measured changes in conditional variance of predicted HCC probabilities within AFP strata¹³; this is a component of the total variance (in addition the conditional means variance which is captured by c index); and the information-theoretic information gain (Kullback-Leibler divergence)¹⁴, which quantifies the increase in heterogeneity of the distribution of HCC cases within AFP-strata¹⁵. Calibration and model fit were assessed graphically by plotting model-derived probabilities against raw probabilities, and analytically by the Hosmer-Lemeshow chi-square statistic. The predictive ability was evaluated by examination of sensitivity, specificity, and positive and negative predictive values, using different thresholds of predicted HCC risk.

To internally validate the final adjusted HCC algorithm we used a split-sample approach to assess discrimination, calibration and predictive ability of the model. We randomly divided the full sample into a development subset (on which model parameter estimates and the corresponding model-based HCC probabilities were derived) and a validation subset (on which the model results were applied) and applied our final model in these two samples.

Results

We identified 11,721 patients with HCV-related cirrhosis in whom 35,494 AFP tests were performed, and 987 developed HCC by 9/30/ 2005. The median age of the study population was 52.0 years (interquartile range [IQR] 48-56), most patients (98.1%) were men, and most were non-Hispanic white (39.2%) or black (13.5%).

On average, 66.2% of the 6 month periods following HCV index date contained an AFP test; in those periods there was an average of 1.2 AFP tests per period. The median AFP value was 7.1 ng/mL (IQR: 3.9-15.3). The skewed distribution of AFP was reduced by logarithmic transformation. The association between log AFP levels and risk of HCC was expectedly strong but non-linear, and therefore log-AFP was entered as a piecewise linear spline variable in the logistic regression model. The calibration of the AFP-HCC bivariate model was high (**Table 1**), with a modified Hosmer-Lemeshow χ^2 statistic = 4.12 ($p = 0.95$) that was indicative of a good fit.

The results of the bivariate analyses for other lab values as individual predictors of HCC are shown in **Appendix 1**. 96.9% have ALT tests and 93.5% have platelet tests. 91.6% have all three tests. Guided by these findings, several lab values were examined as covariates one at a time in the model with AFP (**Table 2**). Although ALT was not a significant predictor of HCC (**Table 2**), it was included in further model building due to our previous finding of a strong 3-way association between AFP, ALT and HCC ⁸.

The increase in discrimination measures are shown in **Table 2**. In general there were small increases in the c- index and IDI measures, however larger increases were seen for additional variance and information gain for alkaline phosphatase, ALT, AST/ALT ratio, hemoglobin and age. ALT had the highest values of additional variance and information gain. The next highest covariates in additional variance were AST/ALT ratio and age; while platelet count was the next highest in information gain.

Some covariates showed varying amounts of modification of the AFP effect in predicting HCC. The largest modifying effect was seen for ALT and platelets, followed by AST/ALT ratio. The ALT modification of AFP predictive value is shown in **Figure 1** at two AFP levels (20 and 120 ng/mL). At both AFP levels, HCC probability was influenced by ALT levels with the probability for HCC being higher at lower ALT values (**Figure 1**). For example, at AFP level of 120 ng/mL (red line in Figure 1), the probability of HCC was 15% in the presence of on ALT=100 IU/ml but this increased to 30% in patients with an ALT<30 IU/ml. In contrast, bilirubin and alkaline phosphatase showed almost no effect on HCC probability at these AFP levels (**Appendix 2a and 2b**).

Based on these data, we constructed a multivariable model containing combinations of AFP, ALT, and platelets; these variables had shown independent HCC predictive ability and/or effect modification of AFP on HCC probability. The additional variance and information gain measures for these models compared with the single AFP-covariate base model are displayed in **Table 2**. We found that a model containing AFP, ALT, and platelets (and interaction terms between AFP and ALT, and AFP and platelets) had the highest additional

discrimination (on both measures) compared to all these models as well as the base AFP alone model (**Table 3**).

We selected the model containing AFP, ALT, platelets, interaction terms and age as our final model. Calibration of this model is displayed in **Appendix 3** that plots model-predicted probability levels against raw HCC frequencies (i.e. model fit). The model results only deviate from the raw frequencies at the very high risk range (>90%), where the disagreement is of no practical clinical significance.

The continuous predicted probabilities of HCC were dichotomized at threshold values and assessed by positive predictive values and sensitivities (**Figure 2**). Using model probability cut-off values of 4%, 10%, 20%, and 30%, the PPV for HCC was 14%, 34%, 52%, and 62% respectively. Higher thresholds result in fewer false negatives and more false positives which in turn results in decreasing sensitivity and increasing PPV (**Figure 2**).

Using this model, we constructed contour plots of predicted HCC probabilities at AFP = 20 ng/mL at age 50 (**Figure 3**) and age 70 (**Figure 4**), which show the joint effects of ALT and platelets on HCC probability. In general, at any given AFP value, low platelets and low ALT were predictive of increased HCC risk, whereas high ALT and normal/high platelets levels were predictive of low HCC risk, and at all lab levels the risk increases with advancing age (**Table 4**). For example, the probabilities of HCC considering only AFP values of 20 ng/ml and 120 ng/ml were 3.5% and 11.4%, respectively, but considering the same AFP values in a patient age 70, ALT 40 IU/ml, and platelets count 100,000, the actual HCC probabilities were 8.1% and 29.0%, respectively. On the other hand, HCC probabilities for AFP 20 ng/ml and 120 ng/ml in a patient age 50, ALT 80, platelets 150,000 were only 3.4% and 9.1%, respectively.

In a split-sample validation, the c-index values were 83.7% and 81.5% for the development (n=5586 subjects with 424 HCC positive) and validation (n=5760 subjects and 374 HCC positive) subsets, respectively. The calibration as indicated by χ^2 Hosmer-Lemeshow was 3.01 (p = 0.98) and 7.85 (p = 0.64)* for the two models, respectively.

In a sensitivity analysis, we excluded 67 patients with HCC in whom AFP was measured after HCC diagnosis date which was defined as the earliest date of a detected hepatic mass that was subsequently verified to be HCC. The model performed quite well in the remaining patients; fit was good with a Hosmer-Lemeshow χ^2 test value of 9.94 (p = 0.45) and discrimination was also high with a c-statistic = 82.7%.

We applied the cirrhotic-based model to 52,135 non-cirrhotic HCV+ patients with the necessary labs for model-based probability calculations; these had a mean follow-up of 237 HCC cases during 1001 days. The overall predictive ability of the model dropped considerably because the model systematically overestimated the risk of HCC (**Appendix 4**); this was expected given that the risk of HCC is low in non cirrhotics. Hosmer-Lemeshow χ^2 test value of 636.9 (p <0.001) indicating poor model fit.

Discussion

We hypothesized that common laboratory tests that are reflective of hepatic inflammation and function are associated with AFP levels and HCC risk in HCV infected patients, and that identifying these variables can improve AFP-based algorithms for HCC detection. Our final adjusted AFP algorithm included AFP in addition to age, platelet, ALT values, and their interaction terms. We found a sizeable improvement to the AFP-based risk assessment by adding age, platelets and ALT values to the predictive model. In our study, HCC risk assessment based on AFP alone is associated with low sensitivity and positive predictive value leading to potentially missed cases. The adjusted AFP model performed well in terms of calibration, discrimination as well as predictive ability in a split derivation validation sample of patients with HCV-related cirrhosis. This new algorithm could be easily applied in clinical practice using readily available information from the EMR and a web-based application. The enhanced HCC risk assessment in our model would allow the personalization of screening practices. Similar to the findings of previous studies, the NPV of a mild or no AFP elevation was high; however, the main limitation was the low PPV of AFP alone, which potentially results in “over doing” follow-up testing (e.g., CT, MRI) in the frequently encountered patients with mildly elevated AFP.

We believe that our adjusted AFP model can personalize and improve the risk estimates beyond the information provided by AFP alone. The algorithm optimizes AFP-based HCC surveillance by two ways: first, reducing false positives triggers would have considerable impact on focusing the efforts related to subsequent testing, reducing surveillance costs, and avoiding unnecessary patient/provider anxiety while optimizing approximation of individualized patient risk. Furthermore, our model incorporates the association between high HCC and normal ALT and low platelets thus reduces false negatives. The HCC risk assessments yielded by our model-based instrument can be used to create a binary decision rule for clinical action; for example, if HCC probability in the next 6 months exceeds 10% then further testing (e.g., cross-sectional imaging) can be performed. However, it is premature to adopt one fixed cutoff at this time.

There are several advantages to our modeling strategy. The average model-based probabilities represented the expected value of the labs (AFP, platelets and ALT)-HCC risk relationship at any given encounter. This modeling strategy simulated the assessment of risk of a randomly chosen patient at a randomly chosen clinical encounter. Our use of the median Wald p-values is a very conservative estimate compared with binomial calculations especially with the large number of iterations. The bivariate model based on AFP alone showed good fit with raw data (i.e., calibration). Therefore, in addition to the mean-based measures (i.e., c-index or IDI) that would be subject to a ceiling effect, we measured the enhancement to the HCC predictive model by the increase in conditional variance of predicted HCC probabilities within AFP strata¹³ and the information gain (Kullback-Leibler divergence)¹⁴, which quantifies the increase in heterogeneity of the distribution of HCC cases within AFP-strata. This method has been recently applied in biomedical sciences¹⁵⁻¹⁷. These two concepts are used to define a preferred sequence of attributes; an attribute with high mutual information should be preferred to other attributes.

Our study overcame a major difficulty in studying HCC biomarkers in the U.S., which is the dearth of suitable data sources that allow capture of a large number of patients at risk of HCC (i.e., with cirrhosis), systematic recording of surveillance tests (i.e., AFP) and a sufficient number of outcome events (i.e., HCC cases). We also made use of all the data available in the dataset; this in contrast to approaches based on matching HCC cases at the time of diagnosis with HCC negative at some “equivalent” time, which results in discarding a large amount of data among patients who never develop HCC. Furthermore, the accuracy of the main variables was very high: the labs were measured directly from the lab package of the automated database and all HCC cases were verified by structured EMR reviews.

There are limitations to our study. While we performed a split sample derivation and validation, it is crucial that future studies test the accuracy of the HCC prediction rule in patients that were not included in the development study. Having an adjusted AFP algorithm with good calibration and discrimination in the patients who were used to develop the predictive rule is not a guarantee that the model will accurately predict HCC in new patients. Indeed, most prediction rules show a reduced accuracy when applied to new patients¹⁸. The generalizability of our model to other patient groups with cirrhosis may be limited. Therefore, in addition to validation for internal validity, the model needs to be evaluated for external validity in non-veteran population, patients with liver disease other than HCV, and especially in women. If our prediction model performs inadequately in external populations, then the existing prediction model can be ‘updated’ and refined or recalibrated to improve its performance in the validation sample. This may involve adding new variables or incorporating serial changes of AFP over time¹⁹. In this iterative process, the final adjusted, or updated, model combines the information captured in the original sample with information from the new individuals from the validation cohorts²⁰. It is possible that some HCC cases were not recorded in the VA database in patients with other insurance carriers; however we previously demonstrated in the same study cohort a minimum loss of HCC diagnoses in patients with dual VA-Medicare use. We virtually eliminated an inclusive type of misclassification by manually verifying all HCC cases analyzed in this dataset, however some HCC cases may not have been captured although our previous work demonstrated very high negative predictive value for the HCC identification algorithm that we used in this study. Lastly, the study did not evaluate the value of AFP on detecting early HCC (vs. all HCC) because of lacking information on HCC stage, or the additional value of AFP over ultrasound. Our main goal was to improve the predictive value of AFP in detecting any HCC

In summary, we have constructed a predictive model of AFP, platelets, and ALT that enhanced the predictive value for detecting HCC. AFP is the only HCC biomarker that has been tested in studies that encompass all 5 phases of biomarker development²¹. Serum AFP measurement is inexpensive, simple to perform, well standardized, and is widely available. While there is active work on discovering and testing new biomarkers, none is ready for clinical use. It would be highly advantageous to have an AFP based algorithm such as the one developed in this study that can be readily calculated in the clinic in a manner similar to the current MELD calculator.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

FUNDING: This work is funded in part by National Institutes of Health (NIH) grant from the National Cancer Institute R01 116845, the Houston VA HSR&D Center of Excellence (HFP90-020), and the Texas Digestive Disease Center NIH DK58338. Dr. El-Serag is also supported by National Institute of Diabetes and Digestive and Kidney Diseases K24-04-107. The opinions expressed reflect those of the authors and not necessarily those of the Department of Veterans Affairs, the US government, the NIH or Baylor College of Medicine. The NIH had no role in the design and conduct of the study; the collection, management, analysis and interpretation of the data; or the preparation, review or approval of the manuscript.

Reference List

1. El-Serag HB, Siegel AB, Davila JA, et al. Treatment and outcomes of treating of hepatocellular carcinoma among Medicare recipients in the United States: a population-based study. *J Hepatol.* 2006; 44:158–166. [PubMed: 16290309]
2. Sherman M, Bruix J, Porayko M, et al. Screening for hepatocellular carcinoma: the rationale for the American Association for the Study of Liver Diseases recommendations. *Hepatology.* 2012; 56:793–796. [PubMed: 22689409]
3. Bruix J, Sherman M, Llovet JM, et al. Clinical management of hepatocellular carcinoma. Conclusions of the Barcelona-2000 EASL conference. European Association for the Study of the Liver. *J Hepatol.* 2001; 35:421–430. [PubMed: 11592607]
4. Poon D, Anderson BO, Chen LT, et al. Management of hepatocellular carcinoma in Asia: consensus statement from the Asian Oncology Summit 2009. *Lancet Oncol.* 2009; 10:1111–1118. [PubMed: 19880065]
5. Gupta S, Bent S, Kohlwes J. Test characteristics of alpha-fetoprotein for detecting hepatocellular carcinoma in patients with hepatitis C. A systematic review and critical analysis. *Ann Intern Med.* 2003; 139:46–50. [PubMed: 12834318]
6. Chen TM, Huang PT, Tsai MH, et al. Predictors of alpha-fetoprotein elevation in patients with chronic hepatitis C, but not hepatocellular carcinoma, and its normalization after pegylated interferon alfa 2a-ribavirin combination therapy. *J Gastroenterol Hepatol.* 2007; 22:669–675. [PubMed: 17444854]
7. Di Bisceglie AM, Sterling RK, Chung RT, et al. Serum alpha-fetoprotein levels in patients with advanced hepatitis C: results from the HALT-C Trial. *J Hepatol.* 2005; 43:434–441. [PubMed: 16136646]
8. Richardson P, Duan Z, Kramer J, et al. Determinants of serum alpha-fetoprotein levels in hepatitis C-infected patients. *Clin Gastroenterol Hepatol.* 2012; 10:428–433. [PubMed: 22155556]
9. El-Serag HB, Kanwal F. alpha-Fetoprotein in hepatocellular carcinoma surveillance: mend it but do not end it. *Clin Gastroenterol Hepatol.* 2013; 11:441–443. [PubMed: 23376321]
10. Davila JA, Weston A, Smalley W, et al. Utilization of screening for hepatocellular carcinoma in the United States. *J Clin Gastroenterol.* 2007; 41:777–782. [PubMed: 17700427]
11. Harrell, FE. Regression Modeling Strategies. Springer; New York: 2001.
12. Pencina MJ, D'Agostino RB Sr. D'Agostino RB Jr. et al. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med.* 2008; 27:157–172. [PubMed: 17569110]
13. McCullagh, P.; Nelder, J. Monographs on Statistics & Applied Probability. Second Edition.. Chapman & Hall/CRC; Generalized Linear Models. 8-1-1989
14. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. 2009
15. Fan R, Zhong M, Wang S, et al. Entropy-based information gain approaches to detect and to characterize gene-gene and gene-environment interactions/correlations of complex diseases. *Genet Epidemiol.* 2011; 35:706–721. [PubMed: 22009792]

16. Nelson JD. Finding useful questions: on Bayesian diagnosticity, probability, impact, and information gain. *Psychol Rev.* 2005; 112:979–999. [PubMed: 16262476]
17. Zhao Y, Chen MH, Pei B, et al. A Bayesian Approach to Pathway Analysis by Integrating Gene-Gene Functional Directions and Microarray Data. *Stat Biosci.* 2012; 4:105–131. [PubMed: 23482678]
18. Bleeker SE, Moll HA, Steyerberg EW, et al. External validation is necessary in prediction research: a clinical example. *J Clin Epidemiol.* 2003; 56:826–832. [PubMed: 14505766]
19. Lee E, Edward S, Singal AG, et al. Improving screening for hepatocellular carcinoma by incorporating data on levels of alpha-fetoprotein, over time. *Clin Gastroenterol Hepatol.* 2013; 11:437–440. [PubMed: 23247324]
20. Moons KG, Kengne AP, Grobbee DE, et al. Risk prediction models: II. External validation, model updating, and impact assessment. *Heart.* 2012; 98:691–698. [PubMed: 22397946]
21. 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. [PubMed: 11459866]

Figure 1. The predicted HCC probability at different AFP levels is modified by ALT levels. Results of bivariate analysis.

Figure 2.

The positive and negative predictive values at different probability cutoffs resulting from a model containing AFP, age, ALT, platelets and their interaction terms.

Figure 3.

Contour plot showing HCC probability at AFP level of 20 ng/mL as calculated from a predictive model containing AFP, age, ALT, and platelets. HCC probability at age 50 years (fixed) varies considerably over a range of ALT and platelets values. HCC probabilities are depicted in different colors.

Figure 4.

Contour plot HCC probability at AFP level of 20 ng/mL as calculated from a predictive model containing AFP, age, ALT, and platelets. HCC probability at age 70 varies considerably over a range of ALT and platelets values. HCC probabilities are depicted in different colors.

Table 1

The AFP only model based HCC probability compared with the raw HCC frequency within 6-month follow-up intervals.

AFP levels (ng/mL)	predicted HCC probability range (%)	raw HCC frequency (%)
10-20	2.27-3.78	3.18
20-30	3.78-5.07	4.58
30-50	5.07-7.28	5.61
50-70	7.28-9.20	6.49
70-90	9.20-10.92	8.49
90-120	10.92-13.23	11.83
120-150	13.23-16.93	15.77
150-200	16.93-24.01	19.20
200-250	24.01-30.76	20.20
250-300	30.76-36.98	35.01
300-400	36.98-47.65	42.64
400-500	47.65-56.11	52.06
500-750	56.11-60.46	51.56
750-1000	60.46-63.00	66.17

Table 2
Measures of predictive enhancement and change in discrimination due to addition of
covariates to the AFP-based model predicting HCC

C is the index that incorporates tied values of predictive pairs. IDI is the Integrated Discrimination Improvement

covariates (in addition to AFP)	change in C	IDI	information gain	additional variance
albumin	0.042	0.001	1.859	1.95%
alkaline phosphatase	0.036	0.092	1.608	5.53%
ALT	0.028	0.024	2.321	9.37%
AST-ALT ratio	0.036	0.090	1.794	5.71%
bilirubin	0.038	0.005	1.787	1.25%
hemoglobin	0.042	0.006	2.107	4.01%
platelets	0.033	0.002	2.267	2.17%
white blood cells	0.007	-0.002	2.148	0.97%
age	0.055	0.010	1.828	5.19%
ALT + platelets	0.045	0.040	2.355	12.92%
ALT + AST-ALT ratio	0.042	0.084	2.019	10.86%
ALT+age	0.066	0.026	2.555	11.81%
ALT+platelets+AST-ALT ratio	0.050	0.091	1.941	11.49%
ALT+platelets+age	0.069	0.049	2.590	15.55%

Table 3

Several numerical examples of changes to HCC probability that result from using the AFP-based model to predict HCC compared with AFP levels alone.

AFP	HCC probability (AFP only)	ALT	platelet	age	HCC probability (full model)
20	3.48%	40	100	50	4.52%
20	3.48%	40	100	70	8.13%
20	3.48%	40	150	50	4.22%
20	3.48%	40	150	70	7.62%
20	3.48%	80	100	50	3.65%
20	3.48%	80	100	70	6.62%
20	3.48%	80	150	50	3.41%
20	3.48%	80	150	70	6.20%
120	11.44%	40	100	50	17.93%
120	11.44%	40	100	70	29.01%
120	11.44%	40	150	50	18.93%
120	11.44%	40	150	70	30.41%
120	11.44%	80	100	50	8.55%
120	11.44%	80	100	70	14.89%
120	11.44%	80	150	50	9.09%
120	11.44%	80	150	70	15.75%