



# HHS Public Access

Author manuscript

*Gastroenterology*. Author manuscript; available in PMC 2016 October 27.

Published in final edited form as:

*Gastroenterology*. 2015 December ; 149(7): 1986–1987. doi:10.1053/j.gastro.2015.10.004.

## The Updated Model: An Adjusted Serum Alpha-Fetoprotein–Based Algorithm for Hepatocellular Carcinoma Detection With Hepatitis C Virus-Related Cirrhosis

**DONNA L. WHITE,**

Section of Gastroenterology and Hepatology, Section of Health Services Research

**PETER RICHARDSON,**

Section of Health Services Research

**NABIHA TAYOUB,**

Section of Gastroenterology and Hepatology

**JESSICA A. DAVILA,**

Section of Health Services Research

**FASIHA KANWAL,** and

Section of Gastroenterology and Hepatology

**HASHEM B. EL-SERAG**

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

---

### Dear Editor

Serum alpha-fetoprotein (AFP) has been extensively used as a biomarker for hepatocellular carcinoma (HCC), the fastest rising cause of cancer related deaths in the United States. However, its performance in HCC surveillance has been generally low; a systematic review of 5 high-quality studies using an AFP cutoff value of >20 ng/mL to indicate a positive test reported its sensitivity was 41%–65%, whereas specificity was 80%–94%.<sup>1</sup> One reason for the modest performance of AFP as a biomarker for HCC detection is that AFP levels are likely not only influenced by the presence of HCC, but also by the severity and activity of underlying liver disease. Earlier research has shown an association between AFP levels and other measures impacted by liver disease severity including levels of alanine aminotransferase (ALT) and platelets, which may themselves be associated with HCC or with false positive AFP.<sup>2,3</sup>

We previously reported on the development and validation of an adjusted AFP-based algorithm that included age, platelets, ALT values, and interaction terms (AFP and ALT, and AFP and platelets).<sup>4</sup> This study was performed in a split sample analysis of a retrospective cohort of 11,721 patients with HCV-related cirrhosis in the VA Healthcare System, and in

---

#### Conflicts of interest

The authors disclose no conflicts.

whom 642 HCCs arose. Our algorithm demonstrated sizeable improvement compared with using AFP levels alone in predicting the 6-month HCC risk in our derivation sample, and also performed well in terms of calibration (agreement of model-derived HCC probabilities with raw HCC probabilities), discrimination (ability to separate HCC negative from HCC positive status), and predictive ability in our split validation sample.

Notably, our adjusted AFP algorithm required only 1 measure of all variables, including AFP levels within a 6-month time window.<sup>4</sup> However, it is possible that serial changes in AFP levels may be more predictive of HCC risk than a single value. We, therefore, report the results of our evaluation of how an adjusted AFP algorithm that also incorporated serial changes in AFP levels would perform in HCC risk prediction.

This added analysis was performed using 6879 patients from our original cohort of patients with HCV-related cirrhosis (59% of total) who had >1 AFP test within a 365-day period before HCC. Median age of this cohort was 52 years, most (98.0%) were men, and 50.3% were non-Hispanic white or black. The median AFP value was 7.1 ng/mL (interquartile range, 3.9–15.3), with an average of 3 serial AFP tests per patient during the 365-day period. The mean duration  $\pm$  standard deviation between AFP tests was 385  $\pm$  361 days, and the median and interquartile range were 228 days (141–410).

The overall serial adjusted AFP model discrimination as assessed by the c-index (82.1%) was high with good model fit or calibration (Hosmer-Lemshow;  $\chi^2 = 1.77$ ). The c-index is a measure of how concordant higher vs lower scale values (ie, model output) are with yes versus no status of the binomial outcome (ie, HCC). A c-value of 82.1% indicates that in 82.1% of pairs of yes versus no outcomes, the yes outcome is assigned a higher scale value than the no outcome. In our split sample validation, the c-index value was 84.9% in the derivation cohort (n = 3367 patients with cirrhosis with no HCC, and 320 who developed HCC) and 75.5% validation cohort (n = 3512 patients with cirrhosis with no HCC and 322 HCC who developed HCC), respectively. Predicted probabilities of HCC deviated from raw frequencies by >10% only at the very high HCC risk range (>90%), where disagreement is of no clinical significance. Our serial AFP algorithm also performed better than our original adjusted AFP algorithm on multiple metrics, including increase in conditional variance of predicted HCC probabilities within AFP strata and the Kullback-Leibler divergence assessed information gain (additional variance, 0.194 vs. 0.174; Kullback-Leibler divergence, 2.25 vs 2.18 for the serial vs original adjusted AFP algorithms, respectively).

The performance characteristics of our serial AFP algorithm expressed as HCC probability risk thresholds are provided in Table 1. For example, using model probability output cutoff values of 4%, 10%, 20%, and 30%, the positive (and negative) predictive values for HCC were 14.5% (98.4%), 34.1% (97.9%), 51.1% (97.5%), and 61.1% (97.4%), respectively. In the hypothetical case of a 49-year-old patient with a current AFP level of 40 ng/mL, the probability of HCC in the next 6 months is 5.5%; however, if information about age, current low ALT of 20 IU/mL, and current platelets count of 100,000/mL are incorporated, the HCC probability increases to 9.0%, and if the information about a prior lower AFP level of 20 ng/mL measured 3 months ago is also incorporated, as in our new algorithm incorporating serial AFP measurements, the predicted HCC probability in the next 6 months is 11.1%.

These HCC predicted probabilities are likely to trigger different actions (eg, observation in the first scenario to possibly cross-sectional imaging in the third scenario). Conversely, our models show that with AFP level of 40 ng/mL, the presence of elevated ALT, normal platelets, and a previous AFP value of 40 are associated with considerably lower HCC probability than the ones shown in the example.

Our study is limited by the lack of an external or non-VA-based validation cohort and limited generalizability. Our cohort does not include women, Latino and Asian patients with non-hepatitis C virus-related cirrhosis, or those cured of hepatitis C virus. However, the study has several strengths including a rigorous modeling strategy that simulated HCC risk assessment in a randomly chosen patient at a randomly chosen clinical encounter, and calculation of multiple measures to assess calibration.

We have demonstrated that inclusion of serial AFP measures further improves the performance characteristics of our previously validated adjusted AFP algorithm for predicting the 6-month HCC risk among individuals with HCV-related cirrhosis, with both algorithms performing substantially better than the traditionally used AFP testing. AFP is the only HCC biomarker that has been tested in studies that encompass all 5 phases of biomarker development.<sup>5</sup> AFP is inexpensive, simple to perform, well-standardized, and widely available. It would be highly advantageous to have an AFP based algorithm that can be readily calculated in the clinic in a manner similar to the current Model for End-stage Liver Disease calculator. These calculators allow us to construct risk prediction scores that move beyond straightforward formulae, which often require simplifying assumptions, and instead allow more complicated models such as ours (which includes 4700 parameters) with piecewise linear expressions, multiple interaction terms, and averaging >100 runs to incorporate a broader range of covariates. If our models are validated in other independent cohorts, they could offer an easy-to-use clinical tool that physicians can use in real time to calculate risk scores that can both improve and personalize patient care while also minimizing unnecessary testing.

## Acknowledgments

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 funders 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.

### 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 Innovations (CIN13-413), the Texas Digestive Disease Center NIH DK58338. Drs. El-Serag and White effort is supported in part by National Institute of Diabetes and Digestive and Kidney Diseases (K24-04-107 and K01 DK081736, respectively).

## References

1. Gupta S, et al. *Ann Intern Med.* 2003; 139:46–50. [PubMed: 12834318]
2. Richardson P, et al. *Clin Gastroenterol Hepatol.* 2012; 10:428–433. [PubMed: 22155556]
3. Richardson P, et al. *Dig Dis Sci.* 2010; 55:3241–3251. [PubMed: 20844957]
4. El-Serag HB, et al. *Gastroenterology.* 2014; 146:1249–1255. [PubMed: 24462733]

5. Pepe MS, et al. J Natl Cancer Inst. 2001; 93:1054–1061. [PubMed: 11459866]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 1**

Test Performance Characteristics (Serial Alpha-fetoprotein, Age, Alanine Aminotransferase, and Platelets)  
Model Estimated 6-month HCC Probabilities at Various Risk Thresholds Generated by the Model

Predicted HCC Probability Threshold (%)	Sensitivity (%)	Specificity (%)	NPV (%)	PPV (%)
4	68.2	83.3	98.4	14.5
10	49.5	96.0	97.9	34.1
20	40.0	98.4	97.5	51.1
30	36.0	99.1	97.4	61.1
40	31.0	99.4	97.2	66.7
50	26.4	99.5	97.0	70.6
60	21.5	99.68	96.8	73.5
70	16.3	99.8	96.6	77.6
80	10.2	99.9	96.4	82.8
90	4.9	99.9	96.2	87.1

HCC, hepatocellular carcinoma; NPV, negative predictive value; PPV, positive predictive value.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript