# Predictive Algorithms: Uses and Limitations

**Janice H. Jou · Deborah A. Fisher**

Hepatocellular carcinoma (HCC) has emerged as an area in urgent need of epidemiologic and outcomes research due to the continuing rise in incidence and mortality over time [1]. The overall 5-year survival from 2003 to 2007 was 14% in the Surveillance, Epidemiology, and End Results (SEER) Program database of the National Cancer Institute [2]. Although this survival rate may be disheartening, subsets of patients with limited tumor burden at diagnosis are candidates for curative therapies including liver transplantation, resection, or tumor ablation. These patients have significantly improved survival; specifically, patients receiving liver transplantation for early-stage disease have 5-year survival of over 70% [3].

Surveillance for HCC has been recommended by multiple professional societies with the goal of identifying patients with HCC in early stages of disease who would be candidates for curative therapies [4, 5]. The recommendation from the American Association for the Study of Liver Diseases (AASLD) is for ultrasound (US) to be performed every 6 months in patients with chronic hepatitis B or cirrhosis [4]. Use of alpha fetoprotein (AFP) for surveillance or diagnostic purposes is not endorsed in these guidelines. Though clinical guidelines support HCC surveillance, there are few randomized controlled trials showing mortality benefit of surveillance for HCC [6].

Observational studies of surveillance are important because they may provide additional data to support or refute the value of surveillance in a setting where randomizing patients to a no-surveillance group would be ethically and logistically difficult. Single-center studies of HCC outcomes are underpowered to detect significant differences in outcomes for HCC, and thus large administrative databases are often used to increase power to detect potential differences.

The use of large administrative databases to assess HCC surveillance effectiveness poses several challenges. First, the data were collected for purposes other than assessment of HCC surveillance utilization and outcomes. Second, the specific test recommended for surveillance, US every 6 months, may be used for a variety of clinical indications, none of which are included in the administrative data variable for the US. The same is true for AFP, which, while not recommended in the AASLD guideline, is commonly included with US for HCC surveillance. Patients who undergo these tests for diagnostic purposes are very likely different from patients undergoing these tests for surveillance, and therefore accurate distinction is important for interpreting the association of these tests with patient outcomes. Without a specific variable to identify tests performed for surveillance versus other indications, Richardson et al. [7] developed and tested an algorithm to predict whether or not an US or AFP was for HCC surveillance.

In the classification algorithm, negative predictors of surveillance intent for AFP, meaning the intent of AFP testing was not for surveillance purposes, were alcoholism, abdominal pain, ascites, diabetes, and elevated aspartate aminotransferase (AST) levels. For ultrasound, negative predictors of surveillance intent were abdominal pain, ascites, drug dependence, and human immunodeficiency status. A predictor that an US was performed for surveillance was an AFP test within 90 days prior to US.

D. A. Fisher
Center for Health Services Research, Durham Veterans Affairs
Medical Center, Durham, NC, USA

J. H. Jou (✉) · D. A. Fisher
Duke University Medical Center, Durham, NC, USA
e-mail: janice.jou@duke.edu

Validation of the algorithm revealed that surveillance intent of AFP could be predicted reliably, whereas surveillance intent of US was more difficult to estimate. The sensitivities in the model for surveillance intent of AFP and US, respectively, were approximately 80% and 30%. US is the recommended surveillance modality for HCC, limiting the applicability of the algorithm for assessment of HCC surveillance, utilization, and outcomes. Nonetheless, strengths of the study by Richardson et al. were the methodological rigor and that their techniques may be useful to address other similar validation studies. Their findings illustrate the complexity of attempting to predict the reasons a practitioner orders a certain test from information available within administrative data.

The difficulties in using such algorithms for databases in clinical research should be a call to facilitate collaborations between clinicians, investigators, and information technology support staff to prospectively collect important variables as coded data as part of the usual clinical and administration documentation. Although this seems to be an ambitious proposition, it would potentially obviate the need for models to be utilized to determine outcomes of interest in an analysis. Additionally, natural-language programming may provide clinical details often missing in data collected for billing or productivity assessments. Templates generated by clinician at the time of an encounter can also be used to aggregate data. While an upfront investment of time and resources would be required with these methods, the advantages of these approaches would be more efficient analysis and potentially increased confidence in the data, results, and conclusions.

Undoubtedly, the impact of research using large administrative databases is immense. With a fatal and increasingly prevalent disease such as HCC, use of administrative databases has the potential to enhance our understanding of current management and ultimately to improve delivery of care. The development of predictive and classification algorithms may enhance the level of granularity of the clinical data provided. Future work should focus on developing more accurate methods of capturing clinical information required to conduct observational studies.

## References

1. Edwards BK, Ward E, Kohler BA, et al. Annual report to the nation on the status of cancer, 1975–2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. *Cancer*. 2010;116:544–573.
2. SEER stat fact sheets: liver and intrahepatic bile duct. 2010. Available from http://seer.cancer.gov/statfacts/html/livibd.html. Cited August 20, 2010.
3. Mazzaferro V, Regalia E, Doci R, et al. Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. *N Engl J Med*. 1996;334:693–699.
4. Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. *Hepatology*, July 2010.
5. 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.
6. Zhang BH, Yang BH, Tang ZY. Randomized controlled trial of screening for hepatocellular carcinoma. *J Cancer Res Clin Oncol*. 2004;130:417–422.
7. Richardson P, Henderson L, Davila JA, Kramer JR, Fitton CP, Chen GJ, El-Serag HB. Surveillance for hepatocellular carcinoma: development and validation of an algorithm to classify tests in administrative and laboratory data. *Dig Dis Sci*. 2010. doi: 10.1007/s10620-010-1387-y.