Update on Biomarkers of Hepatocellular Carcinoma

Roongruedee Chaiterakij,† Benyam D. Addissie,* and Lewis R. Roberts*

*Division of Gastroenterology and Hepatology, Mayo Clinic and Mayo Clinic Cancer Center, Rochester, Minnesota; and †Department of Medicine, Faculty of Medicine, Chulalongkorn University and King Chulalongkorn Memorial Hospital, Thai Red Cross Society, Bangkok, Thailand

New biomarkers of hepatocellular carcinoma (HCC) have been identified using advanced genomic, proteomic, and metabolomics technologies. These are being developed not only for use in diagnosis of HCC, but also in prediction of patient and treatment outcomes and individualization of therapy. Some HCC biomarkers are currently used in surveillance to detect early stage HCCs and reduce mortality. Further studies are needed to determine whether the recently identified HCC biomarkers can be used in clinical practice; most are only in phase 1 or 2 studies. The diagnostic and predictive abilities of biomarkers are limited by the heterogeneous nature of HCCs; there is no perfect single biomarker of this tumor. To improve performance, combinations of biomarkers (panels), or combinations of biomarkers and clinical parameters or laboratory test results, might be required. We describe recently discovered biomarkers of HCC and discuss challenges to their development and application.

**Keywords:** α-Fetoprotein; Liver Cancer; Prognosis; Prognostic Factor; Tumor; Early Detection.

Cancer biomarkers are molecules or substances objectively measurable in cells, tissues, or body fluids that indicate the presence of cancer or predict the risk of cancer development. Traditionally, cancer biomarkers are proteins detected in serum or plasma. Recent technologic advances in genomics, proteomics, and metabolomics have enabled better elucidation of cancer biology and characterization of key molecular events during carcinogenesis. This allows the discovery of novel hepatocellular carcinoma (HCC) biomarkers, including DNA-fusion genes, genetic mutations and epigenetic changes, messenger RNA, noncoding RNA including micro-RNAs (miRNAs), long noncoding RNAs and other species, proteins and posttranslational protein modifications (eg, phosphorylation), metabolites, and antibodies. These different analytes may be detectable in whole blood, white blood cells, serum, plasma, urine, normal tissue (eg, buccal mucosa), tumor tissue, or adjacent benign liver tissue.

**Application of HCC Biomarkers for Clinical Management**

The serum α-fetoprotein (AFP) has been the most commonly used HCC biomarker, functioning as a risk assessment tool in patients with cirrhosis, as a screening tool for early HCC detection, as a diagnostic tool, and as a prognostic tool for tumor recurrence or patient survival. Screening is the performance of a diagnostic test in individuals at risk of HCC who have no symptoms or other a priori reason to suspect the presence of HCC. Surveillance is the repeated application of a screening test. More recently, the scope of applications for HCC biomarkers has expanded beyond diagnostic and surveillance/screening purposes. HCC biomarkers can be used to identify at-risk populations, stratify patients for clinical trials, tailor therapy, and predict treatment response (Figure 1).

**Challenges to the Use of Biomarkers in Clinical Practice**

The difficulties with developing highly sensitive and specific diagnostic, predictive, and prognostic cancer biomarkers stem from two fundamental issues: the molecular heterogeneity of individual persons, and the molecular heterogeneity of cancers. There is therefore first a difficulty with establishing a baseline “normal” value of any biomarker, and second an appreciation that no unique marker is present in all cancers of a particular tissue type. Thus, from a philosophical perspective, two things are necessary to develop the perfect biomarker for any disease. First, each person has to serve as their own control subject. Ideally, a blood, urine, stool, tissue, expired air, or other sample would be collected from each person multiple times during their lifetime and these would be used to ascertain the changes in individual biomarkers over time. Second, there is a need to develop highly sensitive and specific assays for a large selection of disease-related biomarkers, including genes, messenger RNAs,

**Abbreviations used in this paper:** AFP, α-fetoprotein; AFP-L3, *Lens culinaris* agglutinin-reactive glycoform of AFP; AUROC, area under the ROC curve; CI, confidence interval; DCP, des-γ-carboxyprothrombin; EpCAM+, CTC, epithelial adhesion molecule-positive circulating tumor cells; FGF, fibroblast growth factor; GP73, golgi protein 73; GPC3, glypican-3; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; miRNA, microRNA; ROC, receiver operating characteristic; SNP, single nucleotide polymorphism.
noncoding RNAs, proteins, posttranslational protein modifications, and biochemical metabolites. This would allow clinicians to prospectively acquire multiple molecular and physiologic data points for each individual. With the anticipated advances in computing capacity it should be feasible to analyze the large amounts of data generated in a timely fashion and use it to optimize health and minimize illness for each individual.1 Currently, given the absence of these two requirements, a key strategy to optimize the information acquired from currently available biomarkers is to develop methods for using combinations of biomarkers to achieve acceptable test performance. One typical example is the fluorescence in situ hybridization test for the diagnosis of cancer in suspicious biliary strictures; no one marker provides acceptable sensitivity and specificity, but the assessment of polysomy using a combination of four markers has markedly improved sensitivity and specificity for the diagnosis of cholangiocarcinoma.2

**Phases of Biomarker Development for Early HCC Detection**

Although the scope of uses of HCC biomarkers has been broadened, the major purpose of HCC biomarkers is early HCC detection within a surveillance program, with the goal of reducing mortality from HCC. To achieve this goal, biomarkers need to be established through the following phases.3

**Phase 1: Preclinical Exploratory Studies**

The aim is to identify potential markers by comparing the differences in expression of genes, proteins, or other analytes between cancer and normal tissue, or to detect differences in the spectrum of circulating antibodies in patients with cancer compared with control individuals.

**Phase 2: Clinical Assay Development and Validation, Case-Control Studies**

A clinical assay is developed to measure the biomarkers in biospecimens that can be obtained by less invasive methods (eg, blood, urine, stool, or exhaled air). Biospecimens are obtained from established HCC cases and non-HCC control subjects representative of the target screening population. A receiver operating characteristic (ROC) curve is generated to assess the diagnostic performance of the assay. The reproducibility of the assay is also evaluated within and between laboratories.

**Phase 3: Retrospective Longitudinal Repositories Studies**

The ability of an assay to detect preclinical HCC is assessed by obtaining biospecimens at regular intervals from cohorts of individuals at risk for cancer (eg, those with cirrhosis), and following the cohort for development of cancer over time. New biomarkers can then be assessed for their ability to predict the subsequent development of cancer. If the assay can distinguish those who will subsequently develop cancer from control subjects who do not develop cancer months or years before clinical presentation, the criteria for a positive screening test are then defined for phase 4 studies.

**Phase 4: Prospective Screening Studies**

The aim is to determine the detection rate and false-referral rate (or false-positive rate) of a biomarker assay. The assay is applied for screening of a large cohort of the target population for HCC screening. Subjects testing positive are referred for further investigation to establish the diagnosis of HCC. The detection rate is the proportion of screened subjects who test positive and have HCC, whereas the false-referral rate is the proportion of screened subjects who test positive but do not have HCC. Because patients in the false-referral group may be referred for downstream investigations that may expose them to physical or psychological harm, this provides a measure of the negative consequences of the assay.

**Phase 5: Cancer Control Randomized Studies**

Finally, in the most rigorous proof of the clinical utility of a biomarker, a prospective randomized study is performed to determine whether use of the screening...
test can reduce HCC mortality in the target population. At-risk HCC subjects are randomly assigned into 2 groups, those who undergo and those who do not undergo the screening test, and are followed up to compare their survival outcomes.

**Considerations for Applying HCC Biomarkers in Clinical Practice**

New biomarkers should ideally be evaluated through all 5 phases of biomarker development before being used as a screening tool in practice. However, it may be impractical to conduct a phase 5 randomized study to prove assay efficacy.

An assay should be validated before being routinely used in a particular population because the positive and negative predictive value of the assay varies depending on the disease prevalence in each population. The key measure of the overall assay performance is the area under the ROC curve (AUROC), which is a plot of the sensitivity versus specificity. The quoted sensitivity and specificity of an assay generally vary depending on the designated cutoff. An optimal cutoff may be estimated that maximizes the assay sensitivity and specificity; however, ultimately, the selection of the best cutoff depends on the purpose for which the biomarker will be used (ie, a cutoff that provides a high sensitivity may be most appropriate for HCC screening, whereas a cutoff that provides high specificity is most appropriate for confirming the diagnosis of HCC).

Dynamic changes in biomarker levels over time may be more useful than a single measurement in isolation. However, most published studies report biomarker performance calculated from a single measurement, thus there is limited scientific proof of the utility of changes in biomarker levels over time. A recent UK study reported that elevation of AFP over time was the trigger for further investigation to confirm the diagnosis of HCC in 10% of patients with HCC. In real-life practice, experts frequently use the dynamic changes of biomarker values rather than a single measurement. The assay performance in clinical practice may therefore be better than that reported in the literature.

Biomarker performance may vary depending on the etiology of liver disease. For example, the AUROC (95% confidence interval [CI]) of AFP in differentiating cirrhosis versus HCC in patients infected with hepatitis C virus was 0.64 (0.49–0.80), whereas the AUROC (95% CI) in differentiating hepatitis B virus (HBV)–related cirrhosis or chronic HBV infection versus HBV-induced HCC was 0.90 (0.84–0.97).

**Known HCC Biomarkers**

**Biomarkers for Prediction of HCC Risk**

The HCC-4 risk score for predicting HCC risk in patients infected with hepatitis C virus with all stages of liver fibrosis was developed by combining AFP with other patient and laboratory factors, including age, gamma globulin, and platelet count. The score predicted the annual risk for HCC development with an AUROC of 0.80. For patients infected with HBV, an HCC risk predictive model was developed using viral factors, including HBV DNA, hepatitis B surface antigen level, hepatitis B e antigen status, and HBV genotype, and host factors, including age, gender, family history of HCC, and alanine aminotransferase. This model achieved AUROCs of 0.86, 0.86, and 0.83 for predicting 5-, 10-, and 15-year risk, respectively. The HCC-4 risk score and the HBV risk predictive model require external validation. External validation is critically important for predictive models because the models can show overfitting to the cohort used to develop the model. Consequently, the model performance may drop substantially when the model is applied to other populations.

**Biomarkers for Surveillance for Early HCC Detection**

To date, AFP is the only HCC biomarker that has been studied through to phase 5 of biomarker development. A prospective cluster randomized trial conducted in Shanghai, China showed that a surveillance program using AFP and liver ultrasound performed every 6 months resulted in a 37% reduction in HCC mortality. Another population-based cohort study of HBV-infected Alaska natives showed that surveillance using AFP alone every 6 months improved survival. As compared with historical control subjects who were not screened, over 16 years follow-up, prospectively screened patients had significantly longer survival. Because both studies were conducted in chronically infected patients with HBV, the benefit of AFP may not be directly extrapolated to patients with other chronic liver diseases. Despite these findings, imaging currently remains the backbone of HCC surveillance and the AFP is used to complement ultrasound. Current evidence is conflicting as to whether AFP provides additional value when added to ultrasound.

**Biomarkers for HCC Diagnosis**

There is variation in the performance of AFP and *Lens culinaris* agglutinin-reactive glycoform of AFP (AFP-L3) in early HCC diagnosis in different at-risk populations. Generally, AFP is more sensitive than AFP-L3%, which measures the AFP isoform that binds to *Lens culinaris* agglutinin. The sensitivity and specificity of AFP at a cutoff of 10.9 ng/mL and AFP-L3% at a cutoff of 1.7% were 65% and 82% for AFP and 37% and 94% for AFP-L3%, respectively. The combination of AFP with AFP-L3% or des-γ-carboxyprothrombin (DCP) only slightly improved the AFP performance for early HCC diagnosis (Figure 2).
DCP, or protein induced by vitamin K absence/antagonist-II, is an abnormal prothrombin resulting from defective posttranslational carboxylation of the prothrombin precursor. The serum DCP performance for HCC diagnosis varies among studies. Marrero and co-workers\textsuperscript{14} reported DCP at a cutoff of 125 mAU/mL better distinguished HCC from chronic liver diseases and cirrhosis than AFP at a cutoff of 11 ng/mL (sensitivity of 89\% vs. 77\% and specificity of 95\% vs. 73\%), whereas Nakamura and coworkers\textsuperscript{15} reported AFP outperformed DCP for the diagnosis of HCC <3 cm but DCP had better performance than AFP for the diagnosis of HCC >5 cm. These differences may be caused by population differences in patient and tumor characteristics. DCP is more likely to be elevated in patients with more advanced HCCs (eg, larger tumors, vascular invasion, or metastasis).\textsuperscript{15,16} Additionally, nonspecific elevation of DCP can occur in vitamin K deficiency from impaired liver function or administration of vitamin K antagonists. Although the diagnostic performance of AFP-L3 and DCP has been studied in the United States, it is important to note that the Food and Drug Administration–approved indications for AFP-L3 and DCP are for risk-stratification of patients with chronic liver disease, rather than in screening for HCC.

Glypican-3 (GPC3), a plasma membrane bound heparan sulfate proteoglycan, regulates cell growth by modulating activities of several tyrosine kinases and the Wnt signaling pathway. Serum GPC3 had comparable performance with AFP for HCC diagnosis with AUROC of 0.81.\textsuperscript{17} When used in combination with AFP, the sensitivity increased from 52\% (AFP) or 57\% (GPC3) to 77\% (AFP plus GPC3) without a significant decrease in specificity.\textsuperscript{17}

**Prognostic Biomarkers for Predicting Patient Survival**

Single or serial AFP measurements can be used as predictors of survival and outcome of patients with HCC after therapy. A preoperative AFP >100 ng/mL was associated with a higher risk of recurrence after surgical resection.\textsuperscript{18} However, the AFP level was not prognostic for survival of patients with Child A cirrhosis with a single HCC ≤3 cm treated with curative intent.\textsuperscript{19} Patients who had HCCs within the Milan criteria with an AFP >15 ng/mL at the time of transplantation had worse survival than patients who underwent transplantation for non-HCC indications.\textsuperscript{20} Radiologic response and survival after locoregional therapy or systemic chemotherapy can also be predicted by the change of AFP after treatment. Three months after chemoembolization or radioembolization, a >50\% decrease in AFP from a baseline value of >200 ng/mL was associated with tumor response; similarly, a decrease in AFP >20\% from a baseline value of >20 ng/mL was significantly associated with better survival after 2 cycles of systemic chemotherapy.\textsuperscript{21,22}

The performance of AFP in predicting survival of patients with HCC improved when the AFP was combined with patient and tumor characteristics, including the Model For End-Stage Liver Disease score, albumin, tumor size and number, vascular invasion, and metastasis in the Model for Estimating Survival in Ambulatory Hepatocellular Carcinoma patients (score calculator available at http://www.mayoclinic.org/meld/mayomodel10.html).\textsuperscript{23} The Model for Estimating Survival in Ambulatory Hepatocellular Carcinoma model outperforms the Barcelona Clinic Liver Cancer, Cancer of the Liver Italian Program, and Japan Integrated Staging scores in predicting patient survival and has been already validated externally.\textsuperscript{23}

**Predictive Biomarkers for Response to Treatment**

It is controversial whether the AFP is useful for predicting response to sorafenib treatment.\textsuperscript{24,25} Nakazawa and coworkers\textsuperscript{25} reported that an increase in AFP >20\% within 4 weeks after sorafenib treatment was associated with shorter survival, whereas Llovet and coworkers\textsuperscript{24} found a change in AFP from baseline to Week 12 after sorafenib initiation did not predict survival or time to progression. It is important to note that AFP is not considered a standard means of assessing sorafenib response; imaging is the currently accepted standard for assessing sorafenib response.

**Maximizing the Use of Currently Available HCC Biomarkers**

Recent efforts have focused on strategies for maximizing the utility of AFP, AFP-L3\%, DCP, and GPC3,
which have been investigated for years but have not reached the stage of widespread clinical utility in Europe or North America. The AFP has been used for many years worldwide as a HCC biomarker, and the AFP-L3% and DCP have been used for several years in Asia, particularly in Japan, as an adjunct to ultrasound and AFP in HCC surveillance. However, studies in Western countries have demonstrated that the AFP, AFP-L3%, and DCP have relatively low sensitivity for early HCC detection. Consequently, their use has become controversial, with current US and European guidelines recommending against their use in HCC surveillance. Because few clinicians act on the result of a single biomarker test in isolation, recent efforts have focused on integrating the time trends in biomarker levels into clinical decision-making. For example, the variability of AFP over time, determined by the standard deviation of the AFP, has been shown to be more useful than the value of AFP itself for early HCC detection.5

Advances in techniques for detecting and measuring biomarkers have been used to improve the initial limited performance of conventional DCP and GPC3 assays. A major problem of the conventional DCP assay is its low specificity; this has been addressed by the use of a novel DCP assay. The new DCP assay uses a sandwich electrochemiluminescence immunoassay with 2 monoclonal antibodies, the P-11 and P-16 antibody (designated NX-DCP), to detect nonspecifically elevations of DCP caused by vitamin K deficiency. In contrast, the conventional assay uses the MU-3 antibody, which detects both DCP elevations induced by HCC and DCP elevations induced by vitamin K deficiency. Because NX-DCP is not elevated in HCC, patients with HCC have a higher DCP/NX-DCP ratio than individuals without HCC. The new DCP assay uses a sandwich enzyme-linked immunosorbent assay improved the specificity from 55% to 97% when compared with the traditional DCP.31

Although several potential HCC biomarkers have been reported, AFP will likely continue to be the most commonly used HCC biomarker in practice over the next few years, despite its limited sensitivity and specificity. Additional approaches being explored to improve the AFP performance in diagnosis and prognostic prediction include the use of AFP in combination with other biomarkers, laboratory, or clinical parameters. For example, as compared with AFP alone at a cutoff of 19.8 ng/mL, when AFP was used in combination with vascular endothelial growth factor, the sensitivity for HCC detection increased from 68% to 96% and the specificity increased from 75% to 85.26

### Newer Biomarkers for HCC

#### Biomarkers of HCC Risk Prediction

Genome-wide association studies have identified a number of single nucleotide polymorphisms (SNPs) associated with HCC susceptibility (Table 1).29–34 Although these results are encouraging, there are practical challenges to the use of SNPs for predicting HCC risk. First, the results of genome-wide association studies vary across different ethnic populations. More importantly, the odds ratio of each susceptibility SNP identified thus far has been less than 1.5, the cutoff odds ratio considered to be clinically meaningful. This reflects the immense genetic heterogeneity of populations and suggests that using a single SNP as a genetic test for HCC risk prediction may not be feasible. The development of a panel of genetic variants would likely be more informative and potentially more useful for predicting HCC risk. Additionally, risk prediction models including genetic and nongenetic risk factors for HCC could address the complexity of the interaction between host genetic risk factors, viral factors, and other environmental risk factors in contributing to HCC development. This approach was taken by Guyot and coworkers,35 who developed a risk predictive model for HCC development in patients with HCC.

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Minor allele</th>
<th>Risk for HCC</th>
<th>Patient settings</th>
<th>Odds ratio (95% CI)</th>
<th>P value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intron 1 of TPTE2</td>
<td>rs2880301</td>
<td>C</td>
<td>Decrease</td>
<td>HBV/HCV, Korea</td>
<td>0.27 (0.19–0.39)</td>
<td>1.74 × 10^{-12}</td>
<td>Cliftord et al</td>
</tr>
<tr>
<td>KIF1B</td>
<td>rs17401966</td>
<td>G</td>
<td>Decrease</td>
<td>HBV, China</td>
<td>0.61 (0.55–0.67)</td>
<td>1.70 × 10^{-18}</td>
<td>Zhang et al</td>
</tr>
<tr>
<td>GRIK1</td>
<td>rs455804</td>
<td>A</td>
<td>Decrease</td>
<td>HBV, China</td>
<td>0.84 (0.80–0.89)</td>
<td>5.24 × 10^{-10}</td>
<td>Li et al</td>
</tr>
<tr>
<td>STAT4</td>
<td>rs7574865</td>
<td>G</td>
<td>Increase</td>
<td>HBV, China</td>
<td>1.22 (1.15–1.29)</td>
<td>1.66 × 10^{-11}</td>
<td>Jiang et al</td>
</tr>
<tr>
<td>HLA-DQA1/DRB1</td>
<td>rs9272015</td>
<td>A</td>
<td>Increase</td>
<td>HBV, China</td>
<td>1.28 (1.22–1.35)</td>
<td>5.24 × 10^{-12}</td>
<td>Li et al</td>
</tr>
<tr>
<td>MICA</td>
<td>rs2596542</td>
<td>A</td>
<td>Increase</td>
<td>HCV, Japan</td>
<td>1.39 (1.27–1.52)</td>
<td>4.21 × 10^{-13}</td>
<td>Kumar et al</td>
</tr>
<tr>
<td>HLA-DQ</td>
<td>rs9273519</td>
<td>A</td>
<td>Increase</td>
<td>HBV, China</td>
<td>1.51 (1.39–1.66)</td>
<td>8.65 × 10^{-19}</td>
<td>Jiang et al</td>
</tr>
<tr>
<td>DEPDC5</td>
<td>rs1012068</td>
<td>G</td>
<td>Increase</td>
<td>HCV, Japan</td>
<td>1.75 (1.51–2.03)</td>
<td>1.27 × 10^{-13}</td>
<td>Miki et al</td>
</tr>
<tr>
<td>Upstream of DDX18</td>
<td>rs2551677</td>
<td>A</td>
<td>Increase</td>
<td>HBV/HCV, Korea</td>
<td>3.38 (2.07–5.53)</td>
<td>1.41 × 10^{-10}</td>
<td>Cliftord et al</td>
</tr>
</tbody>
</table>

DEPDC5, DEP domain containing 5; DDX18, DEAD (Asp-Glu-Ala-Asp) box polypeptide 18; GRIK1, glutamate receptor, ionotropic, kainate 1; HCV, hepatitis C virus; HLA-DQA1/DRB1, major histocompatibility complex, class II, DQ alpha 1, DR beta 1; KIF1B, kinesin-like factor 1 B; MICA, MHC class I polypeptide-related sequence A; STAT4, signal transducer and activator of transcription 4; TPTE2, transmembrane phosphoinositide 3-phosphatase and tensin homolog 2.
with alcoholic cirrhosis. The model including the rs738409 GG genotype in the PNPLA3 gene, age, gender, and body mass index predicted HCC risk at 3 and 6 years.

**Biomarkers for Early HCC Detection**

**Osteopontin.** Osteopontin is a glycoprotein produced by several different cell types, particularly bone and epithelial cells, and highly expressed in various cancers, including HCC. For differentiating early HCC from cirrhosis, the plasma osteopontin at an optimal cutoff of 91 ng/mL was found to be superior to AFP at a cutoff of 20 ng/mL, with an AUROC (95% CI) of 0.73 (0.62–0.85) versus 0.68 (0.54–0.82), a sensitivity of 75% versus 46%, and a specificity of 62% versus 93%, respectively. The combination of osteopontin and AFP at these cutoffs had better performance than either test alone, with an AUROC (95% CI) of 0.81 (0.70–0.91), a sensitivity of 83%, and a specificity of 63%. Further studies are needed to validate the benefit of using osteopontin as a complementary test to AFP for early HCC diagnosis.

**Golgi protein 73 (GP73).** GP73 is a transmembrane protein localized in the Golgi complex. Its expression is upregulated in chronic liver diseases, and substantially elevated in HCC. Whether GP73 is a better biomarker than AFP for HCC diagnosis remains controversial, with the available results varying with the GP73 assay method. Serum GP73 measured by immunoblot was superior to AFP for differentiating early HCC from cirrhosis. However, serum GP73 measured by enzyme-linked immunosorbent assay was not better than AFP for early HCC diagnosis.

**Circulating miRNAs.** Given that a large number of miRNAs are dysregulated in HCC, miRNA levels in serum, plasma, or urine have been explored for their utility as HCC biomarkers. Of the miRNAs assessed thus far, increased serum or plasma miR-21 of patients with HCC has been the most frequently reported. Compared with AFP at a cutoff of 19.0 ng/mL, plasma miR-21 had an only slightly better performance for distinguishing HCC from chronic hepatitis, with an AUROC (95% CI) of 0.77 (0.69–0.86) versus 0.74 (0.66–0.82), resulting in a sensitivity of 61% versus 60% and a similar specificity of 83%. However, when AFP and miR-21 were combined, the assay performance improved to an AUROC (95% CI) of 0.82 (0.74–0.90), sensitivity of 81%, and specificity of 77%.

**Blood and urine metabolomics.** Advanced chromatography and mass spectrometry technologies are enabling the detection of small-molecule metabolites produced by dysregulated metabolic pathways during hepatocarcinogenesis. Canavaninosuccinate, an organic acid metabolite produced in the liver, is a promising serum biomarker for early HCC diagnosis in patients with cirrhosis. Compared with the AFP at a cutoff of 20 ng/mL, the serum canavaninosuccinate outperformed AFP, with an AUROC of 0.90 versus 0.61, a sensitivity of 79% versus 74%, and a specificity of 100% versus 38%. Combining AFP with canavaninosuccinate increased the sensitivity to 96% while maintaining a specificity of 100%. Urine metabolomics in HCC is under active investigation. Recently, glycocholic acid, a secondary bile acid, was identified in urine of patients with HCC. Further studies are needed to determine the performance of glycocholic acid in distinguishing between HCC and cirrhosis.

**Biomarkers for Prognostication and Stratification for Therapy**

**Five-gene score.** A score for predicting survival and outcome of patients with HCC treated with curative surgical resection was created from a panel of 5 genes involved in different dysregulated pathways in hepatocarcinogenesis (TAF9, RAMP3, HN1, KRT19, and RAN). The score has been shown to accurately predict disease-specific survival and early recurrence. Patients with a score ≥0 had worse survival than those with the score <0, with a hazard ratio (95% CI) of 3.5 (1.9–6.6). The 5-gene score was more accurate than the previously described G3 transcriptome classification, which primarily reflected cell-cycle activation. The hazard ratio (95% CI) of the 5-genes score versus the G3 signature was 4.7 (2.7–8.2) versus 1.0 (0.5–1.7) with Wald test P value of 7.7 × 10⁻⁸ versus 0.9. The 5-gene score has also been validated in different ethnic populations in Europe, United States, and Asia. Despite the impressive performance, the generalizability of the score may be limited because it was developed using surgically resected specimens and has only been proved to reliably predict the outcomes of patients treated with curative resection. Whether it can be applied to patients with unresectable disease is currently unknown.

**Fibroblast growth factors 3 and 4 (FGF3/FGF4) amplification.** FGF3 and FGF4, ligands for the FGF receptors, serve as proto-oncogenes that are amplified in HCCs. FGF3 and FGF4 amplification has been shown to predict response to sorafenib (Figure 3). Of a group of 48 patients studied, 30% (3 of 10) of the partial or complete responders to sorafenib, versus 0% (0 of 38) of non-responders had FGF3/FGF4 amplification. Although this result suggests that FGF3/FGF4 amplification can be used as a biomarker to guide targeted therapy for HCC, prospective validation in a larger cohort is required. No other biomarkers have been proved to predict response to sorafenib.

**High Met expression.** The expression level of c-Met, a proto-oncogene encoding the hepatocyte growth factor receptor, can predict patient outcome and guide treatment. High expression of c-Met in resected HCC tumors was associated with a higher rate of recurrence after surgery. Furthermore, in a randomized phase 2 study of the Met inhibitor tivantinib, patients with high Met expression in the HCC tumor tissue had a significantly
better response to tivantinib than those with low Met expression (median time to progression, 2.7 vs. 1.4 months; hazard ratio [95% CI], 0.43 [0.19–0.97]; P = .03). This result suggests that Met expression can be used as a biomarker for selecting patients for treatment with Met inhibitors.

**miRNA signature.** Analysis of tumor miRNA expression profiles has been used to classify HCC subtype, to assess associations with tumor characteristics or prognostic clinical features, and to identify miRNAs that modulate response to systemic chemotherapy or targeted therapy. Jiang and coworkers have reported a 19-miRNA signature that predicts overall survival. In the context of predicting recurrence after curative resection, Budhu and coworkers reported a 20-miRNA signature associated with venous metastasis that predicted survival and recurrence, whereas Sato and coworkers reported a different set of 13 miRNAs that predicted early recurrence and another set of 20 miRNAs that predicted late recurrence. The differences in these findings may be caused by

![Figure 3](image-url)  
**Figure 3.** Arterial-phase computed tomography image of a large HCC tumor shown to have FGF3/FGF4 amplification before sorafenib treatment (A). After treatment with sorafenib for 2 months, the tumor size was substantially decreased as shown in the arterial phase (B) and hepatocyte phase (C) magnetic resonance images (Panel A reproduced with permission from Arao T et al. FGF3/FGF4 amplification and multiple lung metastases in responders to sorafenib in hepatocellular carcinoma. Hepatology 2013;57:1407-1415; Panels B and C courtesy of Dr Masatoshi Kudo, used with permission).

### Table 2. Summary of Known and Newer HCC Biomarkers

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Source</th>
<th>Current phase for surveillance</th>
<th>Application of biomarker</th>
<th>Proved or potential clinical utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFP</td>
<td>Serum</td>
<td>5</td>
<td>Risk stratification/surveillance/diagnosis/prognosis</td>
<td>The only HCC biomarker for which a prospective study has proved a reduction in mortality from HCC in the surveillance setting9</td>
</tr>
<tr>
<td>AFP-L3%</td>
<td>Serum</td>
<td>2</td>
<td>Risk stratification/diagnosis/prognosis</td>
<td>Changes in AFP-L3% over time may be more sensitive than changes in AFP for predicting HCC progression after transarterial chemoembolization (personal communication from L. Roberts)</td>
</tr>
<tr>
<td>DCP</td>
<td>Serum</td>
<td>2</td>
<td>Risk stratification/diagnosis/prognosis</td>
<td>Combination of DCP and AFP was superior to either AFP or DCP alone for HCC diagnosis12</td>
</tr>
<tr>
<td>Glypican 3</td>
<td>Serum</td>
<td>2</td>
<td>Diagnosis</td>
<td>Comparable performance with AFP for HCC diagnosis15</td>
</tr>
<tr>
<td>Osteopontin</td>
<td>Plasma</td>
<td>2</td>
<td>Early detection</td>
<td>Slightly better than AFP in early HCC detection. Performance was improved when used in combination with AFP6</td>
</tr>
<tr>
<td>GP73</td>
<td>Serum</td>
<td>2</td>
<td>Early detection</td>
<td>Diagnostic performance depends on the method of measurement34–36</td>
</tr>
<tr>
<td>Canavaninosuccinate</td>
<td>Serum</td>
<td>2</td>
<td>Early detection</td>
<td>Outperforms AFP for HCC diagnosis39 Predicts survival and early recurrence after curative resection11</td>
</tr>
<tr>
<td>Five-gene score</td>
<td>Tissue</td>
<td>1</td>
<td>Prognosis</td>
<td>Predicts response to sorafenib; may potentially be used to personalize therapy44</td>
</tr>
<tr>
<td>FGF3/FGF4 amplification</td>
<td>Tissue</td>
<td>1</td>
<td>Prediction</td>
<td>Associated with better response to tivantinib; may potentially be used as a biomarker for selecting patients for treatment with MET inhibitors46</td>
</tr>
<tr>
<td>High Met expression</td>
<td>Tissue</td>
<td>1</td>
<td>Prediction</td>
<td></td>
</tr>
</tbody>
</table>
differences in the populations studied, the outcomes assessed, the sources of tissues used for comparison (healthy liver tissue vs. peritumoral tissue), the techniques used for profiling, and tumor heterogeneity.

EpCAM-positive circulating tumor cells. Stem cell–like, epithelial adhesion molecule–positive circulating tumor cells (EpCAM+ CTC) are a subpopulation of tumor-initiating cells. The presence of EpCAM+ CTC in blood reflects more advanced disease and may be used as a biomarker for stratifying patients for curative or systemic therapy. Detection of EpCAM+ CTC was associated with tumor aggressiveness (ie, higher Barcelona Clinic Liver Cancer stage), vascular invasion, and worse survival (1.3 vs. 2.8 years). Detection of ≥2 EpCAM+ CTC cells in 7.5 mL before hepatectomy was the strongest independent predictor of HCC recurrence postresection.

Summary

Despite its performance limitations, AFP remains the most widely used HCC biomarker. Many novel HCC biomarkers have been discovered during the past few decades (Table 2). However, none have achieved broad acceptance in clinical practice as yet. Several biomarkers are currently under development to improve assay performance and to demonstrate proof of efficacy in clinical practice. Because of the heterogeneous nature of both tumors and humans, it is unlikely that a single ideal biomarker with excellent performance will be identified. Future studies should focus on efforts to combine biomarkers to achieve maximum diagnostic and predictive ability.

References

25. Nakazawa T, Hidaka H, Takada J, et al. Early increase in alpha-fetoprotein for predicting unfavorable clinical outcomes in


