A multi-analyte cell-free DNA–based blood test for early detection of hepatocellular carcinoma


1The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, China
2The First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China
3Helio Health, Inc., Irvine, California, USA
4Laboratory for Advanced Medicine, Inc., Irvine, California, USA
5Guangzhou Youze Biological Pharmaceutical Technology Company Ltd, Guangzhou, Guangdong Province, China
6Department of Internal Medicine, Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center, Dallas, Texas, USA
7Division of Computational Biomedicine, Department of Biological Chemistry, School of Medicine, University of California, Irvine, California, USA
8Texas Gastro Research, El Paso, Texas, USA
9University Health Network, Toronto, Ontario, Canada
10Hepatitis B Foundation, Doylestown, Pennsylvania, USA
11Memorial Sloan Kettering Cancer Center, New York, New York, USA
12Division of Gastroenterology and Hepatology and Department of Epidemiology and Population Health, Stanford University Medical Center, Stanford, California, USA
13Department of Medicine, Chao Family Comprehensive Cancer Center, University of California, Irvine, California, USA

Correspondence
David J. Taggart, Helio Health, Inc., 9950 Research Dr., Irvine, CA, USA.
Email: davidt@heliohealth.com
Yujin Hoshida, Department of Internal Medicine, Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas TX, 75390, USA.
Email: yujin.hoshida@utsouthwestern.edu
Wei Li, Division of Computational Biomedicine, Department of Biological Chemistry, School of Medicine, University of California, Irvine, 5141 California Ave., Room 243, Irvine CA, 92617, USA.
Email: wei.li@uci.edu

Funding information
Helio Health and the Laboratory for Advanced Medicine

Abstract
The limited performance of guideline-recommended abdominal ultrasound and serum alpha-fetoprotein (AFP) highlights the urgent, unmet need for new biomarkers for more accurate detection of early hepatocellular carcinoma (HCC). To this end, we have conducted a prospective clinical validation study to evaluate the performance of the HelioLiver Test, a multi-analyte blood test combining cell-free DNA methylation patterns, clinical variables, and protein tumor markers. A blinded, multicenter validation study was performed with 247 subjects, including 122 subjects with HCC and 125 control subjects with chronic liver disease. The performance of the HelioLiver Test was compared with AFP and the GALAD score as established HCC surveillance blood tests. The performance of the HelioLiver Test (area under the receiver operating characteristic curve [AUROC] = 0.944) was superior to both AFP (AUROC
INTRODUCTION

Mortality of liver cancer, with hepatocellular carcinoma (HCC) as the major histological type, has substantially increased over the past two to three decades in the United States and worldwide, and its incidence is projected to continue to increase over the next decade.\(^1\) HCC prognosis remains dismal (5-year survival <15%) due to frequent diagnoses at late, noncurable stages.\(^4\,5\)  

To increase HCC tumor detection at earlier stages that are more amenable to curative treatment, practice guidelines recommend semi-annual HCC surveillance using ultrasound with or without serum alpha-fetoprotein (AFP) for high-risk populations.\(^6\)\(^–\)\(^8\)  

However, the suboptimal performance of current HCC surveillance tests hinders effective early tumor detection. The sensitivity of ultrasound for detecting early-stage HCC is only 45%, and a systematic review and meta-analysis study has suggested that the improvement to ultrasound by adding AFP is limited to only 63%.\(^9\) Additionally, the performance of ultrasound is reduced by various factors such as obesity, which is sharply increasing globally.\(^10\,11\) Thus, HCC surveillance tests with superior performance are urgently needed. In an effort to improve HCC detection, the GALAD score was developed by combining age, sex, AFP, Lens culinaris agglutinin-reactive AFP (AFP-L3%), and des-gamma-carboxy prothrombin (DCP).\(^12\) The GALAD score has shown superior HCC detection performance compared with AFP, but there is room for improvement\(^13\)\(^,\)\(^14\) indicating the still unmet need for HCC surveillance tests with substantially improved performance characteristics.  

“Liquid biopsies” assaying the methylation of circulating cell-free DNA (cfDNA) released from cancer cells have been actively explored as a promising noninvasive biomarker to sensitively detect various cancer types, including HCC, at early stages.\(^15\) Previously, from a comprehensive methylome profiling of HCC tissue/plasma samples combined with machine-learning analysis, we identified cfDNA methylation markers associated with the presence of HCC in patients with chronic liver diseases as a potential HCC detection biomarker.\(^16\)\(^,\)\(^17\) Based on this foundational work, we further optimized the candidate cfDNA methylation panel and developed the HelioLiver Test, a multi-analyte blood test that combines cfDNA methylation markers with patient demographic information and clinically available HCC tumor markers (components of the GALAD score) to enable robust and accurate HCC detection. Here, we validated the performance characteristics of the HelioLiver Test in a prospective, blinded, multi-center phase 2 biomarker study (the ENCORE study; NCT05059665).  

PATIENTS AND METHODS

Study design

The HelioLiver Test was prospectively validated in an independent, multicenter study (the HelioLiver Test Validation Set 1) for the blinded evaluation of test performance as an HCC early detection marker (The Specimen Collection for Multi-analyte Blood Test for Hepatocellular Carcinoma [ENCORE] study; NCT05059665). The full study details can be accessed on clinicaltrials.gov.  

Study subjects

Subjects recruited in this study were patients newly diagnosed with HCC or patients with a benign liver disease that were recommended for HCC surveillance and were found to be without HCC (control subjects). Subjects with HCC were diagnosed by histopathologic examination or by specific radiologic characteristics...
according to current practice guidelines in China\(^{[18]}\). HCC stage (i.e., extent of tumor spread) was determined for subjects according to the American Joint Commission on Cancer (AJCC) 8th Edition. The control subjects were patients who were recommended to HCC surveillance in China\(^{[18]}\) due to underlying chronic liver disease, including chronic fibrotic liver diseases from any cause, chronic hepatitis B virus (HBV) infection, chronic hepatitis C virus infection, fatty liver disease, and nonalcoholic fatty liver disease.

The presence of cirrhosis was defined by histology or clinical evidence of portal hypertension in subjects with chronic liver disease. All clinical information, including patient demographics and clinical characteristics, were prospectively obtained from medical records.

All subjects were prospectively and consecutively enrolled at the Third Affiliated Hospital of Sun Yat-sen University (Guangzhou, China) and the First Affiliate Hospital of Guangzhou Medical University (Guangzhou, China) between 2020 and 2021 with written informed consent. The study was approved by their respective ethical review boards.

In total, the study included 140 patients with HCC and 150 patients diagnosed with a benign liver disease without HCC (control subjects). A total of 93 subjects were enrolled at the Third Affiliated Hospital of Sun Yat-sen University, and 210 subjects were enrolled at the First Affiliate Hospital of Guangzhou Medical University. Subsequently, 5 subjects were excluded for incomplete health and/or demographic information, and 44 subjects were excluded for failing to meet quality control criteria for the HelioLiver Test, specifying an average sequencing coverage of ≥50 times among all target sites. The final study population analyzed consisted of 122 patients with HCC and 125 control subjects (Table 1).

### Serum protein HCC tumor markers

Serum concentrations of AFP, AFP-L3%, and DCP were measured by using commercially available assays (Hotgen Biotech, Beijing, China) on a HotGen MQ60 instrument according to the manufacturer’s instructions.

### cfDNA methylation detection with targeted capture assay

The Helios Eclipse platform was used to evaluate methylation patterns of cfDNA at target sites (Table S1). To this end, total cfDNA was isolated from specimens by using the EliteHealth cfDNA Extraction Kit (EliteHealth, Guangzhou Youze, China). Isolated cfDNA was eluted into nuclease-free low-bind 1.5-mL microcentrifuge tubes and stored at −80°C. DNA concentration was measured using the Qubit dsDNA High Sensitivity Assay Kit (Thermo Fisher Scientific, USA) as per manufacturer’s instructions. A total of 5 ng cfDNA per sample was used to prepare the barcoded next-generation sequencing (NGS) libraries by using the NEB Next Enzymatic Methyl-seq Kit (New England Biolabs, USA) according to the manufacturer’s instructions. The libraries were then pooled in groups of 24 barcoded libraries at 100 ng each and hybridized with a custom set of HelioLiver capture probes (Twist Bioscience, USA) to capture the target library sequences using the Twist Fast Hybridization and Wash Kit, along with the Twist Universal Blocker. The captured libraries were then supplemented with 20% PhiX genomic DNA library to increase base calling diversity and submitted for NGS on either a HiSeq X or a NovaSeq 6000 platform (Illumina, USA).

### cfDNA methylation data analysis

Raw sequencing data were first trimmed by TrimGalore (ver. 0.6.5) to remove low-quality (Phred score < 20) sequences and potential adapter contamination. To remove M-bias, 5 bp and 10 bp of sequence was trimmed from the 5’ end of Read 1 and Read 2, respectively. Cleaned

<table>
<thead>
<tr>
<th>TABLE 1: Clinical characteristics of study participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects (n)</td>
</tr>
<tr>
<td>Age, median, years</td>
</tr>
<tr>
<td>Sex</td>
</tr>
<tr>
<td>Male (n)</td>
</tr>
<tr>
<td>Female (n)</td>
</tr>
<tr>
<td>Liver disease</td>
</tr>
<tr>
<td>Cirrhosis, n (%)</td>
</tr>
<tr>
<td>Chronic HBV, n (%)</td>
</tr>
<tr>
<td>Chronic HCV, n (%)</td>
</tr>
<tr>
<td>Fatty liver disease, n (%)</td>
</tr>
<tr>
<td>Other(a), n (%)</td>
</tr>
<tr>
<td>Protein tumor markers</td>
</tr>
<tr>
<td>AFP, median (ng/mL)</td>
</tr>
<tr>
<td>AFP-L3%, median (%)</td>
</tr>
<tr>
<td>DCP, median (ng/mL)</td>
</tr>
<tr>
<td>GALAD</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
</tr>
<tr>
<td>II</td>
</tr>
<tr>
<td>III</td>
</tr>
<tr>
<td>IV</td>
</tr>
<tr>
<td>Unstaged or unknown</td>
</tr>
</tbody>
</table>

Abbreviations: HBV, hepatitis B virus; HCV, hepatitis C virus.
\(a\)Other liver conditions include liver cysts and benign liver tumors.
sequencing reads were then aligned to the hg19 human reference genome by using BSMAP (ver. 2.90).[19] The aligned reads were further processed by Samtools (ver. 1.13)[20] and Bedtools (ver. 2.29.1)[21] to select only primarily mapped reads with fragment size between 80 bp and 200 bp. Methratio.py (BSMAP) was finally used to extract the methylation ratio from aligned bam files. Samples with insufficient sequencing depth (<50 times) were excluded from the downstream analysis.

**HelioLiver Test**

The HelioLiver Test (Figure 1) was developed to discriminate between patients with HCC from high-risk patients without HCC. Figure 2 summarizes the overall HelioLiver Test development. Briefly, DNA methylation sites that were shown to be differentially methylated between HCC and control subjects for tissue and blood specimens were first identified from The Cancer Genome Atlas and evaluated in previous studies.[16,17] These target DNA methylation sites were included in a Discovery methylated cell-free DNA (m-cfDNA) panel. DNA methylation sites with undesirable characteristics (e.g., repeated elements, poor sequence capture) in the NGS-based cfDNA methylation assay were then removed to generate a preliminary NGS m-cfDNA panel. Finally, an optimized subset of m-cfDNA markers, clinically available serum protein markers (AFP, AFP-L3%, and DCP), and patient demographics (age and sex) were combined to generate the HelioLiver Test. To this end, we first selected cytosine-guanine dinucleotide (CpG) sites showing significant methylation alteration in HCC samples compared to non-HCC control samples. Subsequently, the feature selection R package “Boruta” was used to identify the optimal cfDNA methylation markers within the Integrative Training Set. This approach identified 77 CpG sites in 28 genes (Table S1) as being significantly and consistently differentially methylated for HCC and was used to construct the cfDNA methylation model. For model training, we assessed different off-the-shelf machine learning models and chose the random forest model (implemented by R package “Ranger”) that showed the best performance. The hyper parameters of the random forest model were fine-tuned by the grid-search method. The cfDNA methylation component, protein tumor marker component, and demographic component were combined by using a decision tree model to generate the HelioLiver Test diagnostic algorithm. The threshold of the HelioLiver Test diagnostic algorithm was fixed based on the out-of-bag predictions in the Training Set to achieve approximately 90% specificity. The HelioLiver diagnostic algorithm was then locked before the initiation of this validation study (ENCORE).

**Statistical analysis**

For the independent clinical validation of the HelioLiver Test (in Validation Set 1), the primary endpoint was to compare the area under receiver operating characteristic (AUROC) curve of the HelioLiver Test to both AFP alone and the GALAD score. The co-secondary end-points were to compare the sensitivity and specificity of the HelioLiver Test (using a prespecified diagnostic algorithm and cutoffs) to AFP at the most commonly reported clinical cutoff of 20 ng/mL,[22] at a lower cutoff of 10 ng/mL, and to the GALAD score at a proposed cutoff of −0.63.[12] As an exploratory endpoint, the sensitivity of the HelioLiver Test was compared with AFP and the GALAD score at standardized specificities. As a post hoc analysis, the performance characteristics

**Figure 1** The HelioLiver Test Workflow. Blood specimens were collected from patients and then assayed for alpha-fetoprotein (AFP), Lens culinaris agglutinin-reactive AFP (AFP-L3%), and des-gamma-carboxy prothrombin (DCP) by using immunoassays, and for cell-free DNA (cfDNA) methylation patterns of 28 genomic regions by using a cfDNA Methylation Assay. The data for each assay were then input into a diagnostic algorithm with the patient’s age and sex, to generate a qualitative (positive/negative) test result.
of AFP-L3% alone, DCP alone, and the combination of AFP and DCP were also calculated for comparison. Due to the relatively high prevalence of chronic HBV within the study population, a post hoc subgroup analysis was additionally performed in a subpopulation of subjects without chronic HBV infection, to compare the AUROC curve, sensitivity, and specificity of the HelioLiver Test, AFP alone, and the GALAD score.

The comparison of the AUROCs for both all subjects with HCC and only early (stage I and II) HCC were performed by sample permutation–based Wilcoxon signed-rank test (10,000 permutations) with Bonferroni correction. The comparisons of the sensitivity and specificity of the HelioLiver Test to AFP and GALAD score were performed using McNemar’s test for paired proportions. A two-tailed \( p \) value less than 0.05 was regarded as statistically significant. All statistical analyses were performed by using Prism software version 8.0 (GraphPad, La Jolla, CA). To assess confounding, the logit function from the python statsmodels module (statsmodels.formula.api.logit) was used to perform logistic regression, with the cancer status as the response variable, and the HelioLiver Test result along with age, gender, and several benign liver conditions as explanatory variables. For each variable, the exponential of the coefficient was calculated to determine the odds ratio.

**RESULTS**

**Development of the HelioLiver Test and algorithm**

The HelioLiver Test algorithm was developed as outlined in Figure 2. Interestingly, 10 of the 28 genes in our cfDNA panel are involved in molecular pathways...
implicated in HCC pathogenesis\textsuperscript{[24,25]} whereas of the 497 unselected genes, only MLK1\textsuperscript{[26]} has been associated in molecular pathways implicated in HCC pathogenesis (Figure S1).

**Independent validation of the HelioLiver Test (the ENCORE study)**

We prospectively enrolled 247 evaluable subjects (Figure 3), including 122 subjects diagnosed with HCC and 125 subjects with a chronic liver disease, who were found to be without HCC after undergoing HCC surveillance (control subjects). The demographic and clinical characteristics of all eligible subjects are described in Table 1. The subjects with HCC were older (median age = 55 years) compared with the control subjects (median age = 47 years). The major disease etiology was HBV infection similarly among both subjects with HCC (72%) and control subjects (58%), in part due to the high rate of HBV infections in China. As expected, AFP, AFP-L3%, DCP, and the GALAD score were higher in the subjects with HCC compared to the control subjects. The individual performance of the 525 genes included in the Preliminary NGS cfDNA panel used to evaluate ENCORE subjects and the performance of the 28 genes included in the HelioLiver Test are described in Table S1.

**Performance of the HelioLiver Test for HCC detection compared with AFP and GALAD**

As the primary endpoint of the study, AUROC curves were used to compare the performance characteristics of the HelioLiver Test to both AFP alone and the GALAD score for the detection of HCC (Figure 4). The HelioLiver Test demonstrated a significantly higher AUROC of 0.944 (95% CI 0.917–0.975) compared with AFP (AUROC 0.851; 95% CI 0.777–0.903; \( p < 0.0001 \)), AFP-L3% (AUROC 0.801; 95% CI 0.755–0.847; \( p < 0.0001 \)), DCP (AUROC 0.780; 95% CI 0.719–0.842; \( p < 0.0001 \)), and the GALAD score (AUROC 0.899; 95% CI 0.833–0.941; \( p < 0.0001 \)) for the detection of HCC overall (Figure 4A). The HelioLiver Test (AUROC 0.924; 95% CI 0.846–0.986) also outperformed both AFP (AUROC 0.806; 95% CI 0.653–0.902; \( p < 0.0001 \)), AFP-L3% (AUROC 0.769; 95% CI 0.686–0.852; \( p < 0.0001 \)), DCP (AUROC 0.742; 95% CI 0.632–0.852; \( p < 0.0001 \)), and the GALAD score (AUROC 0.842; 95% CI 0.693–0.926; \( p = 0.0003 \)) for the detection of early-stage (AJCC stage I and II) HCC (Figure 4B). As anticipated, the performance of GALAD was superior to AFP, AFP-L3%, and DCP alone for detection of both HCC overall and early HCC (Figure 4).

To investigate whether confounding variables influenced the HelioLiver Test results, we used logistic regression analysis to assess the relationship between the patient group (subjects with HCC or control subjects) and the HelioLiver Test result in the presence of potential confounding variables including age, gender, and underlying liver disease. We then calculated an odds ratio for the HelioLiver Test result, adjusted for these potential confounders (Table S2). The coefficient associated with the HelioLiver Test prediction was calculated to be 3.9 with a \( p \) value < 2.2e-16 and an odds ratio = 50. This suggests that patients that have a positive HelioLiver Test are approximately 50 times more likely to actually have HCC than patients with a negative test result. This odds ratio was adjusted for patient demographic data (age and gender) and the underlying liver disease of the subjects.

To further confirm that the underlying etiology of liver disease for ENCORE subjects did not influence the performance characteristics of the HelioLiver Test, a subset of 100 subjects diagnosed with HCC and 100 control subjects with matched liver disease etiologies.
HEPATOLOGY COMMUNICATIONS

Within the etiology-matched subgroup of subjects, the HelioLiver Test demonstrated superior performance characteristics (Table S4) for HCC overall (AUROC 0.933; 95% CI 0.905–0.964) compared with AFP (AUROC 0.844; 95% CI 0.789–0.898), AFP-L3% (AUROC 0.797; 95% CI 0.745–0.848; \( p < 0.0001 \)), DCP (AUROC 0.750; 95% CI 0.678–0.821; \( p < 0.0001 \)), and the GALAD score (AUROC 0.881; 95% CI 0.832–0.930) (Figure S2). The HelioLiver Test (AUROC 0.917; 95% CI 0.866–0.968) similarly outperformed the AFP (AUROC 0.803; 95% CI 0.708–0.898), AFP-L3% (AUROC 0.765; 95% CI 0.682–0.849; \( p < 0.0001 \)), DCP (AUROC 0.733; 95% CI 0.622–0.844; \( p < 0.0001 \)), and GALAD score (AUROC 0.834; 95% CI 0.743–0.924) for the detection of early (Stage I and II) HCC within the etiology matched subgroup of subjects.

As co-secondary endpoints, the sensitivity and specificity of the HelioLiver Test (using a prespecified diagnostic algorithm and cutoffs) were compared with GALAD and the individual protein tumor markers at standard clinical cutoffs. These test performance characteristics are summarized in Table 2. The HelioLiver Test (85.2%; 95% CI 77.8%–90.4%) demonstrated a superior overall sensitivity for the detection of all-stage HCC compared with the commonly used cutoff 20 ng/mL (62.3%; 95% CI 53.5%–70.4%) and a lower cutoff of 10 ng/mL (68.0%; 95% CI 59.3%–75.6%).

### Table 2: Comparison of test performance characteristics for detection of HCC

<table>
<thead>
<tr>
<th>Test Description</th>
<th>Early-stage (I + II) Sensitivity, % (95% CI) (n = 37)</th>
<th>Late-stage (III + IV) Sensitivity, % (95% CI) (n = 71)</th>
<th>Overall Sensitivity, % (95% CI) (n = 122)</th>
<th>Specificity, % (95% CI) (n = 125)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HelioLiver Test</td>
<td>75.7 (59.9, 86.7)</td>
<td>91.5 (85.2, 95.3)</td>
<td>85.2 (77.8, 90.4)</td>
<td>91.2 (84.9, 95.0)</td>
</tr>
<tr>
<td>GALAD (≥−0.63)</td>
<td>64.9 (48.8, 82.8)</td>
<td>80.3 (72.4, 86.4)</td>
<td>75.4 (67.1, 82.2)</td>
<td>93.6 (87.9, 96.7)</td>
</tr>
<tr>
<td>GALAD (≥−1.2)</td>
<td>70.3 (54.3, 82.5)</td>
<td>81.7 (73.9, 87.6)</td>
<td>77.9 (69.8, 84.4)</td>
<td>91.2 (84.9, 95.0)</td>
</tr>
<tr>
<td>AFP (≥10 ng/mL)</td>
<td>62.2 (46.1, 76.0)</td>
<td>69.0 (57.5, 78.6)</td>
<td>68.0 (59.3, 75.6)</td>
<td>90.4 (84.0, 94.4)</td>
</tr>
<tr>
<td>AFP (≥12.1 ng/mL)</td>
<td>59.5 (43.5, 73.4)</td>
<td>67.6 (58.9, 75.3)</td>
<td>66.4 (57.6, 74.2)</td>
<td>91.2 (84.9, 95.0)</td>
</tr>
<tr>
<td>AFP (≥20 ng/mL)</td>
<td>56.8 (40.1, 71.4)</td>
<td>63.4 (54.6, 71.4)</td>
<td>62.3 (53.5, 70.4)</td>
<td>96.8 (92.1, 98.8)</td>
</tr>
<tr>
<td>AFP-L3% (≥10%)</td>
<td>51.4 (35.9, 66.6)</td>
<td>60.6 (49.0, 71.1)</td>
<td>59.8 (50.9, 68.1)</td>
<td>97.6 (93.2, 99.2)</td>
</tr>
<tr>
<td>DCP (≥7.5 ng/mL)</td>
<td>40.5 (26.3, 56.5)</td>
<td>62.0 (50.4, 72.4)</td>
<td>52.5 (43.7, 61.2)</td>
<td>93.6 (87.9, 96.7)</td>
</tr>
<tr>
<td>AFP (≥20 ng/mL) + DCP (≥7.5 ng/mL)</td>
<td>67.6 (51.5, 80.4)</td>
<td>80.3 (72.4, 86.4)</td>
<td>76.2 (67.9, 82.3)</td>
<td>91.2 (84.9, 95.0)</td>
</tr>
</tbody>
</table>

*a*Cutoff value corresponds to HelioLiver Test specificity of 91%.

**Figure 4** Receiver operating characteristic (ROC) curves for HCC blood tests. (A) ROC curves for analysis of all subjects diagnosed with HCC and control (benign liver disease) subjects. (B) Subjects diagnosed with early-stage (American Joint Commission on Cancer [AJCC] stage I and II) HCC and control (benign liver disease) subjects. AUROC, area under receiver operating characteristic.
an established cutoff of $-0.63$ (75.4%; 95% CI 67.1%–82.2%). The HelioLiver Test demonstrated a superior sensitivity (75.7%; 95% CI 59.9%–86.7%) for early-stage (I and II) HCC when compared with AFP at both the 20-ng/mL cutoff (56.8%; 95% CI 40.1%–71.4%) and the 10-ng/mL cutoff (62.2%; 95% CI 46.1%–76.0%), and the GALAD score (64.9%; 95% CI 48.8%–78.2%) at the cutoff of $-0.63$. The specificity of the HelioLiver Test (91.2%; 95% CI 84.9%–95.0%) was comparable to AFP at the 10-ng/mL cutoff (90.4%; 95% CI 84.0–94.4%) and the GALAD score (93.6%; 95% CI 87.9%–96.7%). The sensitivity of both the HelioLiver Test and the GALAD score (at both the $-0.63$ and $-1.2$ cutoffs) was found to be superior to AFP-L3% ($\geq 10%$ cutoff), DCP (7.5 ng/mL cutoff), and the combination of AFP ($\geq 20$ ng/mL cutoff) and DCP ($\geq 7.5$ ng/mL cutoff)$^{[23]}$ for the detection of both HCC overall and early-stage HCC (Table 2).

As an exploratory endpoint, the sensitivity of the HelioLiver Test was compared with AFP and the GALAD score at the specificity determined for the HelioLiver Test (91.2%) (Table 2). At this standardized specificity, the sensitivity of the HelioLiver Test for HCC detection overall was 85.2% (95% CI 77.8%–90.4%), which was higher than AFP (cutoff = 12.1 ng/mL; 66.4%; 95% CI 57.6%–74.2%) and the GALAD score (cutoff = $-1.2$; 77.9%; 95% CI 69.8%–84.4%). The sensitivity of early-stage HCC detection for the HelioLiver Test was 75.7% (95% CI 59.9%–86.7%), which remained higher than AFP (cutoff = 12.1 ng/mL; 59.5%; 95% CI 43.5%–73.4%) and the GALAD score (cutoff = $-1.2$; 70.3%; 95% CI 54.3%–82.5%). The sensitivity of the HelioLiver Test also remained higher than both AFP and the GALAD score at the remaining standardized specificities between 85% and 95% (Figure S3, Table S5).

The major underlying liver disease etiology in the ENCORE study was HBV, which is more prevalent in China compared with many other areas of the world. To gain insight surrounding this issue, a post hoc exploratory subgroup analysis was performed in subjects with non-HBV etiologies (Table S6). AUROC of the HelioLiver Test (0.93; 95% CI 0.863–0.983) remained higher than AFP (0.913; 95% CI 0.852–0.974) and the GALAD score (0.901; 95% CI 0.825–0.977) within this non-HBV subgroup (Figure S4). Additionally, the sensitivity of the HelioLiver Test (86.7%; 95% CI 70.4%–94.7%) was also higher than AFP (66.7%; 95% CI 48.8%–80.8%) and the GALAD score (80.0%; 95% CI 62.7%–90.5%) within this subgroup for all HCC (Table S7). These results suggest that the performance of the HelioLiver Test is etiology agnostic.

**DISCUSSION**

In this phase 2 cancer detection biomarker study as defined by the NCI Early Detection Research Network,$^{[27]}$ we confirmed that the HelioLiver Test demonstrated superior performance in detecting early-stage (AJCC stage I and II) HCC compared with clinically available tests (AFP, AFP-L3%, and DCP), the combination of AFP and DCP,$^{[23]}$ and the GALAD score.

The current standard of care for HCC surveillance is ultrasound either with or without AFP. A recent meta-analysis of 32 separate studies (consisting of 13,367 patients) indicated ultrasound alone has a sensitivity of 84% (95% CI 76%–92%) for HCC of any stage and 47% (95% CI 33%–61%) for early-stage HCC.$^{[9]}$ These relatively wide CIs for ultrasound sensitivity are indicative of the variability of ultrasound performance in different clinical settings. This variability is due in part to differences in the skill and training of the ultrasound operator, the quality of the instrument, and confounding factors that reduce ultrasound sensitivity, such as obesity and underlying liver disease.$^{[28]}$ Interestingly, many of these challenges posed to ultrasound as an HCC surveillance tool may be overcome by the development of a sensitive and easy-to-use blood test. Further subgroup analysis of studies comparing the performance of ultrasound with or without AFP indicated that including AFP increases the sensitivity for early HCC lesions from 45% for ultrasound alone to 63% for the combination of ultrasound and AFP.$^{[9]}$ However, the specificity is reduced from 92% to 84%.

Various meta-analyses have calculated the pooled sensitivity and specificity for the individual protein tumor markers evaluated in this study to be 61% and 86% for AFP,$^{[29]}$ 56% and 90% for AFP-L3%,$^{[30]}$ and 69% and 88% for DCP,$^{[31]}$ respectively. The combination of these three markers as part of the GALAD score consistently outperforms any of these markers individually. Within separate, multicenter studies of subjects with HCC with different etiologies of liver disease, the GALAD score was calculated to have sensitivities from 68% to 92% and specificities from 88% to 95%.$^{[13,14,32]}$ However, the combination of GALAD and ultrasound was found to possess superior performance characteristics than either test alone, with a sensitivity of 95% and specificity of 91%.$^{[32]}$ In this study, the sensitivity of the HelioLiver Test was found to be superior to AFP, AFP-L3%, and DCP individually, and to the combination of these markers as part of the GALAD score.

Of note, the HelioLiver Test, which incorporates components of the GALAD score, achieved improved performance over the GALAD score itself by additionally measuring cfDNA methylation patterns from blood. Similar attempts also showed promising improvement of the GALAD score’s performance by combining additional tests and examinations such as ultrasound and tumor glycomics biomarkers,$^{[19,20]}$ suggesting that this is a valid strategy to incorporate such biomarkers and/or other HCC surveillance modalities for clinical translation.

Various types of biomolecules have been evaluated as potential HCC biomarkers. These include (1) a proteomics approach that identified a panel of seven serum protein markers$^{[33]}$; (2) the DNA methylation marker Sept9,$^{[34,35]}$ which has previously been evaluated for
characteristics when compared with ultrasound-based HCC surveillance, then HCC surveillance using a blood test may prove to be a more accessible option for patients. Reducing logistical barriers to semi-annual surveillance is anticipated to significantly improve adherence or even inspire more frequent monitoring of at-risk patients.

Despite the promising performance of our test, there are several limitations inherent to our study design. First, all subjects with HCC were diagnosed before enrollment. HCC diagnosis of all study subjects occurred after suspicion of HCC by ultrasound and/or serum AFP within a surveillance setting, or in subjects previously presenting with symptoms of HCC, which are not generally assumed in a surveillance setting. Thus, it is not feasible to directly compare the performance of the experimental tests assessed in this study to the current standard-care surveillance method of ultrasound with or without AFP for the detection of HCC, as this population is biased for subjects in whom HCC was detected by the current surveillance techniques. Second, some control subjects with chronic liver diseases were diagnosed as HCC-free based only on ultrasound and AFP rather than with diagnostic imaging such as multiphasic magnetic resonance imaging (MRI) or computed tomography. Therefore, although unlikely, it is possible that some of the control subjects may have undiagnosed HCC. Third, surveillance tests such as the HelioLiver Test are expected to have the most significant impact on health outcomes when HCC is detected at an early stage, when curative treatment options are more likely. The enrollment of a greater number of subjects with early-stage HCC in follow-up case-control studies or prospective studies will allow for a more robust analysis of test performance for subjects with early HCC. Finally, the underlying HCC etiology was biased toward HBV infection, which is representative of the patient population in China. The limitations in the design of our study are common among recently published studies to evaluate the performance of novel blood tests for HCC.[33–38,45] To address these limitations, the HelioLiver Test is being further evaluated as part of a currently ongoing phase 3 biomarker study, in which the performance of the HelioLiver Test will be directly compared with standard-of-care ultrasound using multiphasic MRI as the gold standard for HCC diagnosis (CLiMB trial; NCT03694600). Prospective trials that directly compare the performance characteristics of the current standard-of-care surveillance methods to these novel blood tests for HCC detection are essential for evaluating the true contribution of these potential surveillance tools.

CONCLUSIONS

The HelioLiver Test was found to have a superior sensitivity for HCC and a similar specificity when compared to both AFP alone and the GALAD score. Most importantly, the HelioLiver Test demonstrated a superior
sensitivity for early-stage (AJCC I and II) HCC when compared with either AFP testing alone or the GALAD score. The implementation of a blood test such as the HelioLiver Test will enable easy, flexible, noninvasive, and accurate HCC detection at early stages, and significantly improve treatment outcomes for a transformative reduction of HCC mortality.

**ACKNOWLEDGMENT**

The authors thank the patients who generously participated in the clinical trial and the principal investigators and institutions who oversaw their enrollments. They would also like to acknowledge Dr. Shivani Mahajan for performing the logistic regression analysis, to assess confounding variables for the HelioLiver Test.

**CONFLICT OF INTEREST**

Jianfeng Xu, Maxime A. Gallant, Dhruvajyoti Roy, and David Taggart are employees of Helio Health. Dan Liu is an employee of the Laboratory for Advanced Medicine. Diange Li is employed by Guangzhou Youze Biological Pharmaceutical Technology Company. Robert Gish consults for, advises for, is on the speakers' bureau for, and received grants from Gilead. He consults for, advises for, and is on the speakers' bureau for AbbVie, Bayer, Bristol-Myers Squibb, Eisai, Intercept, and Salix. He consults for, advises for, and is on the data safety monitoring board for Arrowhead. He consults for, advises for, and owns stock in Eiger, Genentantis, and HepQuant. He consults for and advises for Abbott, Access Biologicals, Antios, Dova, Dynavax, Enyo, Forty-Seven, Fujifilm/Wako, eStudySite, Genentech, Gerson Lehrman, HepaTX, Janssen, Helios, Lilly, Merck, Shionogi, and Viking. He consults for ADMA, AEC, Aligos, Arena, Arterys, Alexion, Alimmune, AprosTx, Cirina, Consumer Health Products Association, DiaSorin, DRG Abacus, DURECT, Echosens, Exelixis, IDLogiq, Intellia, Inotek, Iqvia, KannaLife, Laboratory for Advanced Medicine, Labyrinth Holdings, MedImmune, New Enterprise Associates, Ogivy, Organovo, Patient Connect, ProdigY Biotech, Prometheus, Refuah Solutions, Regulus, Spring Bank, and Trimaran. He advises for Biocollections, Prodigy, and Quest. He owns stock in RiboSciences, CoCrystal, and AngioCrine. Ghassan Abou-Alfa serves as a consultant for and serves on scientific advisory boards for Agios, AstraZeneca, Autem Medical, Bayer, Beigene, Berry Genomics, Celgene, Eisai, Flatiron Health, Genoscience, Gilead, Incyte, Ipsen, LAM, Lilly, Merck Serono, Minapharm, QED, Red-Hill Biopharma, Roche/Genentech, SillaJen, Therabionic, twoXAR, and Vector Health. He receives travel/accommodation funding from Polaris. He receives research funding (to institution) from Agios, AstraZeneca, Bayer, Bristol Myers Squibb, CASI, Exelixis, Incyte, Polaris, Puma Biotechnology, and QED outside the submitted work. Mindie H. Nguyen was previously a consultant for the Laboratory for Advanced Medicine and has received research support from Pfizer, Enanta, Gilead, Exact Sciences, Vir Biotech, Helio Health, National Cancer Institute, Glycotest, and B.K. Kee Foundation. Mindie Nguyen also received personal fees as a consultant and/or advisory board member for Intercept, Exact Sciences, Gilead, GSK, Eli Lilly, and Janssen. Richard Van Etten serves as an advisory board member for Helio Health. Yujin Hoshida serves as an advisory board member for Helio Health and founding shareholder for Alentis Therapeutics and received a research funding from Morphic Therapeutics. Wei Li is a consultant for Helio Health and ChosenMed.

**TRIAL REGISTRATION**

ClinicalTrials.gov identifier: NCT05059665.

**ORCID**

Robert G. Gish 🌐 https://orcid.org/0000-0001-6306-3189
Mindie H. Nguyen 🌐 https://orcid.org/0000-0002-6275-4989
Wei Li 🌐 https://orcid.org/0000-0001-9931-5990

**REFERENCES**


