Screening Tests for Hepatocellular Carcinoma in Patients With Chronic Hepatitis C: A Systematic Review

Kelly A. Gebo,[1,2] Geetanjali Chander,[1] Mollie W. Jenckes,[1] Khalil G. Ghanem,[1] H. Franklin Herlong,[1] Michael S. Torbenson,[3] Samer S. El-Kamary,[4] and Eric B. Bass[1,5]

for the Johns Hopkins University Evidence Based Practice Center Team

This systematic review addresses the following questions: (1) What is the efficacy of using screening tests for hepatocellular carcinoma (HCC) in improving outcomes in chronic hepatitis C, and (2) what are the sensitivity and specificity of screening tests for HCC in chronic hepatitis C? The search strategy involved searching Medline and other electronic databases between January 1985 and March 2002. Additional articles were identified by reviewing pertinent articles and journals and by querying experts. Articles were eligible for review if they reported original human data from studies of screening tests that used virological, histological, pathologic, or clinical outcome measures. Data collection involved paired reviewers who assessed the quality of each study and abstracted data. One nonrandomized prospective cohort study suggested that HCC was detected earlier and was more often resectable in patients who had twice yearly screening with serum alpha-fetoprotein (AFP) and hepatic ultrasound than in patients who had usual care. Twenty-four studies, which included patients with chronic hepatitis C or B or both, addressed the sensitivities and specificities of screening tests. They were relatively consistent in showing that the sensitivity of serum AFP for detecting HCC usually was moderately high at 45% to 100%, with a specificity of 70% to 95%, for a threshold of between 10 and 19 ng/mL. The few studies that evaluated screening with ultrasound reported high specificity, but variable sensitivity. In conclusion, screening of patients with chronic hepatitis C with AFP and ultrasound may improve detection of HCC, but studies are needed to determine whether screening improves clinical outcomes. (HEPATOLOGY 2002;36:S84-S92.)

Hepatocellular carcinoma (HCC) is one of the most common non-dermatologic cancers in the world. Incidence rates vary regionally with rates reported in Asia as high as 80 per 100,000.[1] The incidence of HCC in patients with hepatitis C ranges between 0% and 1.6% per year[2] and 0.46% per year in patients with hepatitis B.[3-6] In addition, the annual risk of developing HCC with cirrhosis is between 1% and 6%.[7-13] The mortality from HCC is substantial, with survival rates as low as 1% at 2 years in untreated patients.[14-16] Despite knowledge of the risk factors for HCC, screening for HCC is controversial, as there have been no randomized controlled trials demonstrating the efficacy of screening for HCC. The 1997 Consensus Development Conference on Management of Hepatitis C made no specific recommendations about screening for HCC.

Although numerous studies have examined screening for HCC in patients with chronic hepatitis B, the natural history of chronic hepatitis C is different from that of chronic hepatitis B.[17] For example, in contrast to hepatitis B, HCC rarely occurs in patients with hepatitis C virus (HCV) who do not have cirrhosis.[18] Thus, conclusions about the value of HCC screening in chronic hepatitis B are likely not to apply to patients with chronic hepatitis C.

We conducted a systematic review of the literature on 2 key questions: (1) What is the efficacy of using screening tests for HCC to improve clinical outcomes in patients with chronic hepatitis C and (2) what are the sensitivity and specificity of screening tests for HCC, especially resectable HCC, in patients with chronic hepatitis C? Sensitivity refers to the ability to detect HCC in those who have HCC, and specificity refers to the ability to rule out HCC when HCC is not present.

Patients and Methods

Identification of Target Population and Specific Questions. For this systematic review of the literature, we defined the targeted population as patients with chronic hepatitis C. We formulated the key questions after meeting with the Planning Committee responsible for the 2002 Consensus Development Conference on Management of Hepatitis C and then obtaining input from experts. Details about the identification of questions are available in our complete evidence report.[19] For the purposes of the current review, we define screening as the one time performance of a screening test to detect previously unrecognized disease, and we define surveillance as the repeated use of a screening test over a period of time.

Literature Search. We conducted a search of literature from January 1996 through March 2002 using DIALOG (Cary, NC), a commercial database vendor that has access to electronic databases including MEDLINE; Biological Abstracts - BIOSIS Previews; Science Citation Index - SciSearch; Manual, Alternative, and Natural Therapy - MANTIS; Allied and Complementary Medicine Database; PsycINFO; and Sociological Abstracts. To minimize the chance of missing an important study, we also searched MEDLINE back to 1985, reviewed reference lists in key studies, and reviewed the table of contents of recent issues of journals most likely to publish relevant studies.

We included the search for studies on screening for HCC in our overall search for articles on key questions in the management of hepatitis C, as described in the complete report.[19] The search strategy was designed by experts at the National Library of Medicine (Bethesda, MD) to maximize the likelihood of detecting all pertinent peer-reviewed studies. After developing a search strategy for MEDLINE, the search experts modified the strategy for the other electronic databases. Search results were entered into a reference manager database (ProCite, ISI Research Soft, Berkeley, CA).

Abstract Review. Paired reviewers independently reviewed the article titles identified by the search and ex-

cluded those that did not meet our eligibility criteria. The exclusion criteria were: language other than English, no human data, no original data, no information relevant to management of hepatitis C, only basic science, did not apply to one of our questions, meeting abstract without a full article for review, or other incomplete reports.

The paired reviewers then independently reviewed the abstracts of the remaining citations. Abstracts were excluded if both reviewers agreed that the abstracts did not meet the eligibility criteria. Disagreements between reviewers about eligibility were adjudicated at meetings.

Article Review and Data Extraction. To review and extract data from eligible articles, we used a standardized form for assessing study quality against pre-established criteria and also developed forms for extracting relevant information on characteristics and results of each type of study. To focus the review on articles that were most likely to give useful data on the questions, we excluded studies that (1) were not designed to address one of the questions; (2) addressed management of hepatitis C in liver transplant patients only; (3) had less than 30 study subjects; or (4) did not document the presence or absence of HCC with appropriate histological or pathologic evidence. For the question on whether screening tests improved clinical outcomes, we also required that studies have at least 6 months of follow-up. For the question on performance characteristics of screening tests, we included only studies that reported data on patients with hepatitis C, although studies were included if some patients had only hepatitis B or were co-infected with hepatitis B and C. Because the pathophysiology and natural history of hepatitis C differs substantially from that of hepatitis B,[6,17,18,20] we excluded studies that focused solely on hepatitis B.

The study quality assessment form included the following categories: Representativeness of study population (5 items); bias and confounding (1 to 3 items); description of management or test (1 to 2 items); outcomes and follow-up (4 to 5 items); and statistical quality (3 to 4 items). The items on this form were derived from forms used in previous systematic reviews[21] and can be found in our complete evidence report.[19] Not all items applied to both questions addressed in this review. Each applicable item could be assigned a score of zero (criteria not met), 1 (criteria partially met), or 2 (criteria fully met). The score for each category of study quality was the percentage of points available in each category for a given study and could range from 0% to 100%. The overall quality score was the average of the 5 categorical scores.

The data extraction form included items that described the type of study, geographical location, definition of

study groups, specific aims, eligibility criteria, screening test characteristics, characteristics of subjects, and results.

We reviewed each eligible article in pairs, including at least one reviewer with clinical training and one with training in epidemiology and research methods. One reviewer completed the quality assessment and data extraction forms, and the second reviewer checked the material abstracted; they resolved differences by consensus.

After entering data into an Access (Microsoft, Seattle, WA) database, we created evidence tables to display information across studies. The detailed evidence tables are in our comprehensive evidence report.[19]

Evidence Grades. We graded the strength of evidence using a scheme derived from previous projects.[22,23] For the question on sensitivity and specificity of screening tests, Grade A (strong evidence) was given when appropriate data were available, including at least one well done study, the study population was sufficiently large, there was an adequate reference standard, data was consistent, and the test definitively was or was not useful. Grade B (moderate evidence) was given when appropriate data were available, the study population was sufficiently large, there was an adequate reference standard, the data were reasonably consistent, and the data indicated the test was or was not likely to be useful, but there was insufficient evidence to conclude definitively. Grade C (weak evidence) was given when some data were available, the study population was reasonably large, the data indicated a trend supporting benefit (or no benefit) of one intervention compared with another, and there was insufficient evidence to conclude that a test was likely to be superior, equivalent, or inferior to another. Grade I (insufficient evidence) was given when appropriate data were not available or an insufficient number of patients was studied. We used a similar set of evidence grades for the question on the efficacy of screening.

When all the evidence from eligible studies was assembled, 5 members of our team independently submitted written grades on the strength of the evidence on each question. To resolve differences in grading, the final grade was based on the opinion of the majority.

Peer Review Process. An external group of experts reviewed a draft report detailing the methods and results of our review. This group included experts in hepatology and infectious diseases, experts in clinical epidemiology, representatives of selected professional organizations (e.g., the American Association for the Study of Liver Diseases and the American College of Physicians-American Society of Internal Medicine), and representatives of governmental agencies (e.g., the National Institutes of Health).

Results

Results of Literature Search. We found 1 surveillance study that answered Key Question 1[24] regarding outcomes of screening for HCC and 24 studies relevant to Key Question 2 regarding performance characteristics of the screening tests. Five of these 24 studies were in patients infected with hepatitis C alone,[25-29] and 19 were in patients with hepatitis B or C or both.[18,20,24,30-45] Ten of these studies were surveillance studies,[20,24,26,30-36] and 14 were screening studies.[18,25,27-29,37-45] No study was excluded solely for lack of appropriate histological or pathologic evidence of HCC.

Study on Outcomes of Screening for HCC. We were unable to identify any randomized controlled trials on screening for HCC, but we identified one prospective cohort surveillance study that evaluated outcomes for HCC in patients with chronic liver disease. A total of 360 patients who were followed in a hepatitis clinic received alpha-fetoprotein (AFP) and ultrasound screening twice a year. The control group was 2,170 patients who received usual care in other hepatology clinics.[24]

During a mean follow-up of 56 months, focal lesions that proved to be HCC were found in 24 (6.7%) of the patients in the screening group and 129 (5.5%) in the control group. Of the 24 malignancies noted in the screening group, 75% were unifocal and less than 3 cm, compared with 16% in the control group, a statistically significant difference.[24] In the screening group, at the time of diagnosis, serum AFP was normal (< 20 ng/mL) in 11 patients, between 20 and 200 ng/mL in 9 patients, and above 200 ng/mL in 4 patients. At these thresholds, sensitivities for detecting HCC were 46%, 38%, and 17%, respectively.

Overall, this study indicated that HCC was detected earlier (i.e., smaller tumors) and was more often resectable when the screening group was compared with patients who received usual care. However, the strength of the evidence was weak (Evidence Grade C: some data available, study population reasonably large, data indicated a trend supporting benefit [or no benefit] of 1 intervention compared with another, but insufficient evidence to conclude that the intervention was likely to be superior, equivalent, or inferior to another).

Sensitivity and Specificity of Screening Tests. To answer the question regarding performance characteristics of tests for screening for HCC, we identified 24 articles that evaluated serologic, urinary, or radiologic tests, including screening and surveillance studies. Nineteen studies evaluated use of serum AFP for detection of HCC. As shown in Table 1, 10 studies evaluated other serologic markers for HCC,[25,27-31,35,37-39] 1 study evaluated urinary

Table 1.

Author, Year	N	Study Quality Score (%)	Surveillance Study?	Screening Test	Sensitivity (%)	Specificity (%)
Serologic test						
Ishii, 2000	734	81	Yes	PIVKA-II > 60 mAU/mL	41	91
			No	AFP ≥ 40 ng/mL & PIVKA-II > 80 mAU/mL	66	85
Izzo, 1999	1,520	67	Yes	S-IL2R > 850 U/mL	99	96
				US	66	
				CT/MRI	100	
Kakumu, 1997	80	67	No	IL10 > 5 pg/mL	63	
				IL15 > 70 pg/mL	45	
Larcos, 1998	232	40	Yes	AFP > 81 U/mL	24	
				US	92	
Nagai, 2001	129	51	No	CK > 2.5 ng/mL	47	95
Nomura, 1996	128	35	No	Conventional DCP	17	
				Overnight DCP	29	
				Avidin biotin complex DCP	33	
Raedle, 1995	174	65	No	Positive anti-p53 auto Ab	43	100
				AFP > 20 ng/mL + positive anti-p53 auto Ab	86	86
				AFP > 100 ng/mL + positive anti-p53 auto Ab	71	99
Sassa, 1999	195	33	No	AFP > 200 ng/mL and AFP L3 > 10%	25	99
				AFP L3 > 10%	23	99
				H-DCP > 40 mAU/mL	45	99
				H-DCP and AFP > 200 ng/mL	48	99
				H-DCP and AFP L3 > 10%	54	98
Tsuzurahara, 1997	170	65	No	MAGE-4 > 1.04 ng/mL	47	77
Tsai, 1995	256	80	No	3% PEG CIC	65	
				AFP > 120 ng/mL + 3% PEG CIC	84	100
				AFP > 400 ng/mL + 3% PEG CIC	83	100
Urinary test						
Tsai, 1997	238	63	No	TGFβ1 ≥ 50 mg/g Cr	53	99
				TGFβ1 ≥ 50 mg/g Cr or AFP > 100 ng/mL	84	98
				TGFβ1 ≥ 50 mg/g Cr or AFP > 400 ng/mL	80	99
Radiologic test						
Colombo, 1991	447	69	Yes	US	49	
				US and CT	93	
Izzo, 1998	1,125	50	Yes	US	87	
				US or AFP > 10 ng/mL	100	
Kasahara, 1998	1,022	87	Yes	US and CT		96
Tong, 2001	1,204	56	Yes	US	100	98
Van Roey, 2000	140	63	No	US	51	

Abbreviations: Auto Ab, autoantibodies; CIC, circulating immune complexes; CK, cytokeratin; Cr, Creatinine; DCP, des-gamma-carboxy prothrombin; H-DCP, high sensitivity des-gamma-carboxy prothrombin; IL, interleukin; MAGE-4, melanoma antigen 4; mAU, milliarbitrary units; PEG, polyethylene glycol; PIVKA-II, protein induced by vitamin K absence; S-IL2R, soluble interleukin 2 receptor.

transforming growth factor β-1,[40] 6 studies evaluated hepatic ultrasound (US),[26,32-35,40] 1 study evaluated computerized tomography (CT) and magnetic resonance imaging (MRI),[30] and 2 studies evaluated CT and US.[26,34]

AFP. Of 19 studies that evaluated serum AFP for detection of HCC, 3 were in patients with HCV alone,[28,36,46] and 16 were in patients with HCV or hepatitis B virus (HBV) or both.[10,20,27,30-33,35,37,39,40,40-44] Nine of the 10 surveillance studies were AFP studies.[20,24,30-36] These studies varied significantly in study design, patient characteristics, and sample size. Most were conducted in Europe or Asia with 1 study from Australia[35] and 1 from the United States.[32] Almost all studies excluded patients with other forms of liver disease. In all of the studies, the majority of participants were men, with the mean age ranging between 31 and 66 years. Most had advanced liver disease, the group thought to be at highest risk of HCC. Duration of infection, if reported, was generally over 10 years. Genotypes varied according to the country in which the study was performed. The mean quality score of these studies was 64% and they tended to be weakest in the bias category.

Figure 1 includes surveillance and screening studies and demonstrates the increasing sensitivity of serum AFP with decreasing AFP threshold. With an AFP threshold of 400 ng/mL, the sensitivity ranged from 0% to 64%. With an AFP threshold of 10 to 19 ng/mL, sensitivity ranged from 45% to 100%. Figure 2 demonstrates how both the sensitivity and specificity of serum AFP varied according
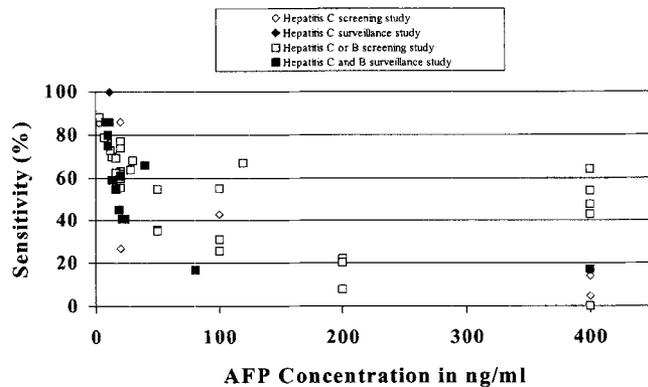
Fig. 1. Sensitivity of HCC detection by AFP threshold value.

to the value of the test threshold. As the AFP threshold increased above 20 ng/mL, the sensitivity began to decrease below 60%. A large case-control study evaluating different AFP thresholds found 16 ng/mL as the threshold that maximized sensitivity and specificity.[18] A receiver operating characteristic (ROC) curve in this study found an area under the curve of 0.819 indicating a high level of diagnostic performance. Overall, the studies were relatively consistent in suggesting that the sensitivity of serum AFP for detecting HCC in patients with hepatitis C increases from very low levels to moderately high levels of 60% to 80% as the threshold value decreases from 400 ng/mL to 10 ng/mL, with the corresponding specificity decreasing from 100% to about 70% to 90% (Evidence Grade B, moderate evidence).

Other Serologic Markers. Table 1 summarizes results of the 10 studies that evaluated other serologic markers for HCC.[25,27-31,37-40] These studies included 3 surveillance studies and 7 screening studies. They used different study designs, with only 2 using diagnostic test design permitting calculation of both sensitivity and specificity of a test.[25,28] One of these studies examined the performance of the tests in patients with chronic hepatitis C,[38] while the other studies all included patients with either hepatitis B or C.

The mean study quality score for these studies was 61%. Generally, these studies received high scores for providing good descriptions of the screening tests (mean score, 84%), and variable scores for representativeness of the study population (mean score, 57%), outcome assessment (i.e., documentation of presence or absence of HCC) (mean score, 61%) and statistical analysis (mean score, 70%).

These studies reported sensitivities that ranged from 17% to 99%. The best sensitivity (99%) was obtained in a study that evaluated serum interleukin 2 receptor in 1,520 patients with either hepatitis B or hepatitis C, or both. All patients had been infected for at least 5 years.

The mean duration of HCV infection was 10.6 years. In this study, the sensitivity of serum interleukin 2 receptor was 99%, and the specificity was 96%.[30]

Overall, we graded the evidence on other serologic markers as Evidence Grade I (insufficient evidence).

Urinary Tests. One case control screening study conducted in Taiwan evaluated the use of urinary transforming growth factor (TGF) β-1 in screening for HCC in patients with and without cirrhosis (see Table 1).[40] This study had a quality score of 63%. The 238 patients included 94 cirrhotics with HCC, 94 cirrhotics without HCC, and 50 healthy controls, had a median age between 55 and 58 years, and were nearly 80% male.

In this population, the sensitivity of urinary TGF β-1 for detecting HCC was 53% and the specificity was 99%, using a threshold of greater than 50 μg/g creatinine.[40] When urinary TGF β-1 was used in combination with serum AFP, the sensitivity for detecting HCC was 84% if the AFP threshold was 100 ng/mL and 80% if the AFP threshold was 400 ng/mL.[40] Specificities for these same thresholds were 98% and 99%, respectively. We graded the evidence for urinary TGF β-1 as Evidence Grade I (insufficient evidence).

US. We identified 7 studies that evaluated US[26,30,32-35,41] 6 through surveillance[26,30,32-35] and 1 through screening with hepatic US, 1 in patients with HCV alone,[26] and 6 in patients with chronic liver disease, including, but not limited to hepatitis C. These studies were limited in that some were designed to assess the incidence of HCC and not to assess the performance characteristics of US. Studies varied in the screening frequency and the extent of liver disease in the screened patients.

Overall, the sensitivity of US for detecting HCC varied from 11% to 99% with a high specificity of 95% to 100%. We graded the evidence on use of US as Evidence Grade C (weak evidence).
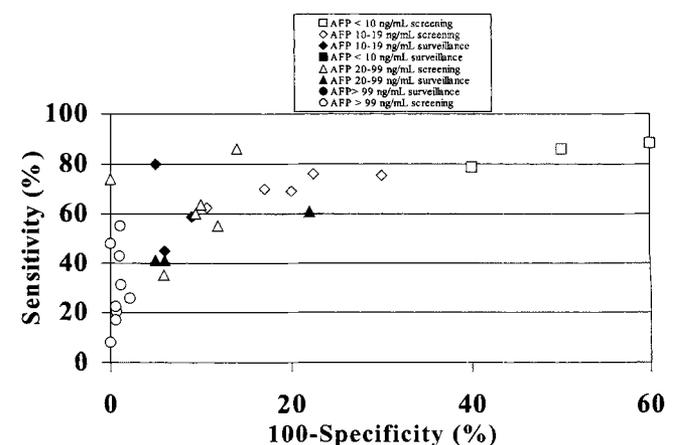

Fig. 2. Sensitivity and specificity of AFP for HCC screening by threshold value.

CT and MRI. One surveillance study evaluated the use of CT and/or US[26] to detect HCC in patients with hepatitis C. This study had limited data on the utility of the screening tests, as the study was designed primarily to evaluate the incidence of HCC in patients with hepatitis C. The study indicated a specificity of 96% for the combination of the tests. The study did not examine the sensitivity of the tests.

Another surveillance study reported a sensitivity of 93% for the combination of US and CT in cirrhosis patients,[34] but it did not give the specificity. A third surveillance study in patients with either hepatitis B or C or both reported a sensitivity of 100% for CT or MRI of the liver.[30] Overall, we graded the evidence on use of CT or MRI as Evidence Grade C (weak evidence).

AFP and US. Three surveillance studies[30,32,35] and one screening study[41] compared the sensitivities of US and serum AFP for detecting HCC, but they did not use the tests in combination. In contrast, one surveillance study evaluated the combination of AFP and US in 1,125 patients with HCV, HBV, or both, who had been infected for at least 5 years.[33] The overall quality score for this study was 50%. The mean age of the population was 56 years, and the mean duration of HCV infection was 10.2 years. This study reported a sensitivity of 100% when using a serum AFP greater than 10 ng/mL together with US, which compared with a sensitivity of 75% using only a serum AFP greater than 10 ng/mL and a sensitivity of 87% when using only US.[33] We graded the evidence on AFP and US as Evidence Grade C (weak evidence).

Discussion

Our literature search of surveillance and screening studies showed only one prospective cohort study and no randomized controlled trials evaluating the efficacy of screening for HCC in patients with chronic hepatitis C infection. The prospective cohort study suggested that HCC was detected earlier and was more often resectable in patients who underwent routine screening with AFP and hepatic US than in those who had usual care. However, this study had important limitations, especially the fact that it included patients with chronic liver disease, primarily due to hepatitis B or C, but also due to other causes, and thus may not be representative of the development of HCC in patients with hepatitis C. Because the study was not randomized and was susceptible to both selection bias and observer bias, we concluded that there was not enough evidence to draw any firm conclusions about the efficacy of screening using AFP or US for HCC in patients with chronic hepatitis C.

In the absence of good direct evidence about the efficacy of screening for HCC, clinicians and policy makers may want to consider the evidence on the performance characteristics of screening tests for HCC. The most frequently used tests have been serum AFP and hepatic US. One must keep in mind that the efficacy of screening for HCC will depend on more than just the performance characteristics of screening tests. It also will depend on the efficacy of treatment for HCC discovered by screening and on the prevalence of HCC in specific subsets of patients with chronic hepatitis C.

Numerous trials evaluated the performance characteristics of serum AFP in screening for HCC in patients with chronic hepatitis C. These studies, including surveillance and screening studies, were relatively consistent in suggesting that a serum AFP level of 10 to 19 ng/mL has a moderately high sensitivity of 45% to 100% and a specificity of 70% or 95%, and a serum AFP level of greater than 400 ng/mL has a low sensitivity with a specificity of nearly 100% in screening for HCC. These results indicate that an AFP threshold of 10 ng/mL may maximize sensitivity and retain relatively high specificity. One study has shown that a discriminating value of 16 ng/mL can be used, as this maximized the sensitivity and specificity in patients with liver disease.[18] Di Bisceglie and Hoofnagle[47] have shown, however, that AFP is not always specific for HCC, and titers can increase with flares of active hepatitis. Previous work has shown that elevated AFP can be associated with changes in HBV replication status.[47] AFP levels may increase transiently, intermittently, or permanently in patients with viral hepatitis without HCC. Increases are most often paralleled by an increase in aminotransferase levels; however, the diagnostic dilemma occurs in differentiating HCC from a viral illness when an increase in AFP levels does not correlate with an increase in aminotransferase levels or occurs in the presence of normal levels.[48] Further studies are needed, as currently there are no guidelines as to when an increase in AFP in the presence of a normal US should trigger a further evaluation.

Although many clinicians use US to screen for HCC, few studies evaluated the performance characteristics of US in screening patients with hepatitis C. A previous study by Pateron et al.[49] used US and AFP every 6 months in a Caucasian population with cirrhosis, with primarily alcoholic cirrhosis, but the screening methods used did not effectively identify potentially resectable tumors. The studies using US that we evaluated were relatively consistent in demonstrating high specificity, but had variable sensitivity depending on the population screened. Combination screening with AFP and US demonstrated an increase in sensitivity in at least 1 trial of patients with hepatitis B or C. Further study is needed, and trials comparing AFP and US would be very useful.

As previously shown by Collier and Sherman,[48] the surveillance intervals studied varied from 3 to 12 months, and reasons for choosing a given interval often were not reported. In 1 study of patients with hepatitis B, the AFP levels corresponded with tumor doubling time in 55% of the tumors; however no patients included in this study had hepatitis C.[50] In this study, the most rapidly growing tumor increased from 1 cm to 3 cm in 5 months. Although the investigators suggested a 4- to 5-month interval for screening[50] and others have suggested that a 6-month interval may be most appropriate,[48] the ideal time for re-screening has not been identified in patients with hepatitis C.

Several other serologic and urinary screening tests were evaluated, but none of these were evaluated in more than 2 studies. Few of these studies had a large enough population of patients with chronic hepatitis C to provide reliable estimates of the performance characteristics of the tests.

Three surveillance studies reported on the performance characteristics of CT and/or MRI. The studies were consistent in demonstrating both a high sensitivity and specificity in patients with hepatitis C, but MRI is significantly more expensive than US (personal communication, June 2002, Johns Hopkins University Radiology).

The limitations of the studies varied with the screening test used. Studies evaluating AFP varied widely in study design and patient eligibility criteria. Studies using US varied by screening frequency, experience of the ultrasonographer, and the extent of liver disease in screened patients. Finally, the studies using CT or MRI were not designed to assess the performance characteristics of these tests in screening, but to evaluate the incidence of HCC. In addition, the limitations in specificity varied by study. Even a false positive rate of 20% can lead to expensive tests that have morbidity. Finally, the literature did not allow us to estimate the rates of detection of resectable compared with non-resectable HCC, which is a key issue in determining the utility of screening tests for improving long-term outcomes.

The criteria for judging a screening test include the following: (1) The disease must be common and have substantial morbidity or mortality; (2) the target population must be easily identifiable; (3) the screening test must have low morbidity and high sensitivity and specificity; (4) there must be standardized recall procedures, including what constitutes a positive result; (5) the screening test must be acceptable to the population; and (6) there must be an acceptable and effective therapy.[51] Screening for HCC in patients with chronic hepatitis C seems to meet at least some of these criteria since: (1) HCC occurs at an annual incidence of 1% to 6% in patients with cirrhosis and has a high risk of death; (2) hepatitis C patients are at risk and can be easily identified; (3) serum AFP and hepatic US have relatively high sensitivity and specificity; (4) a variety of radiologic studies and hepatic angiography are available to confirm the diagnosis; (5) the morbidity and cost of these tests are low enough to expect that they would be acceptable to many patients in the western world; and (6) effective treatment, including resection and transplant, exists, although it is expensive and is limited by the availability of transplantable organs.[3] However, from our review of evidence, it is not clear whether screening for HCC in patients with chronic hepatitis C will identify enough resectable tumors to improve long-term outcomes and offset the risk of invasive tests on patients with false positive screening tests. Therefore, clinicians may want to carefully consider the varying sensitivities and specificities of different tests, as well as the potential complications and costs of screening when discussing the pros and cons of screening with patients. Screening strategies are most likely to be successful if they use tests that have the highest sensitivity and specificity.

Future Research Needs

Clearly, additional research is needed to determine whether routine screening and surveillance for HCC actually improve outcomes in patients with chronic hepatitis C. This would be best determined by prospective, randomized, controlled trials in patients who are at greatest risk for HCC, such as those with cirrhosis or advanced fibrosis. Both the types of screening tests (AFP, US, CT, or MRI) and the timing of repeat testing for surveillance require study in a prospectively, randomized, controlled fashion. A critical component in these studies is assessment of outcome, i.e., whether these test detect HCC at an early enough stage that therapy (resection, ablation, liver transplantation) is effective in prolonging life. These studies should be designed to assess both sensitivity and specificity of these test in detecting HCC at different sizes (such as 1 to 3 cm, 3 to 5 cm, or > 5 cm). These studies should also include estimates of cost and cost-effectiveness of different protocols for surveillance. In addition, new tests will need to be compared with, and possibly combined with, the currently most sensitive screening options. Given the lack of direct evidence on the benefit of screening for HCC, studies should seek to elucidate the ideal screening population, test(s), and surveillance protocols.

References

1. Parking DM, Muir CS, Whelan SL, Gao Y-T, Ferlay J, Powell J. Cancer incidence in five continents. 1994 World Health Statistics Annual (Vol 6). Lyon IARC Scientific Publications. 117-129: World Health Organization. Geneva WHO; 1995.

2. Horiike N, Fujisawa T, Michitaka K, Tada K, Masumoto T, Iuchi H, Kojima N, et al. The effectiveness of interferon therapy on occurrence of hepatocellular carcinoma in chronic hepatitis C. Oncol Rep 1998;5:1171-1174.

3. Bruix J. Treatment of hepatocellular carcinoma. HEPATOLOGY 1997;25:259-262.

4. Villeneuve JP, Desrochers M, Infante-Rivard C, Willems B, Raymond G, Bourcier M, Cote J, et al. A long-term follow-up study of asymptomatic hepatitis B surface antigen-positive carriers in Montreal. Gastroenterology 1994;106:1000-1005.

5. McMahon BJ, Alberts SR, Wainwright RB, Bulkow L, Lanier AP. Hepatitis B-related sequelae. Prospective study in 1400 hepatitis B surface antigen-positive Alaska native carriers. Arch Intern Med 1990;150:1051-1054.

6. Sherman M, Peltekian KM, Lee C. Screening for hepatocellular carcinoma in chronic carriers of hepatitis B virus: incidence and prevalence of hepatocellular carcinoma in a North American urban population. HEPATOLOGY 1995;22:432-438.

7. Kato Y, Nakata K, Omagari K, Furukawa R, Kusumoto Y, Mori I, Tajima H, et al. Risk of hepatocellular carcinoma in patients with cirrhosis in Japan. Analysis of infectious hepatitis viruses. Cancer 1994;74:2234-2238.

8. Ikeda K, Saitoh S, Koida I, Arase Y, Tsubota A, Chayama K, Kumada H, et al. A multivariate analysis of risk factors for hepatocellular carcinogenesis: a prospective observation of 795 patients with viral and alcoholic cirrhosis. HEPATOLOGY 1993;18:47-53.

9. Benvegnu L, Fattovich G, Noventa F, Tremolada F, Chemello L, Cecchetto A, Alberti A. Concurrent hepatitis B and C virus infection and risk of hepatocellular carcinoma in cirrhosis. A prospective study. Cancer 1994; 74:2442-2448.

10. Cottone M, Turri M, Caltagirone M, Parisi P, Orlando A, Fiorentino G, Virdone R, et al. Screening for hepatocellular carcinoma in patients with Child's A cirrhosis: an 8-year prospective study by ultrasound and alphafetoprotein. J Hepatol 1994;21:1029-1034.

11. Zoli M, Magalotti D, Bianchi G, Gueli C, Marchesini G, Pisi E. Efficacy of a surveillance program for early detection of hepatocellular carcinoma. Cancer 1996;78:977-985.

12. Oka H, Tamori A, Kuroki T, Kobayashi K, Yamamoto S. Prospective study of alpha-fetoprotein in cirrhotic patients monitored for development of hepatocellular carcinoma. HEPATOLOGY 1994;19:61-66.

13. Borzio M, Bruno S, Roncalli M, Mels GC, Ramella G, Borzio F, Leandro G, et al. Liver cell dysplasia is a major risk factor for hepatocellular carcinoma in cirrhosis: a prospective study. Gastroenterology 1995;108:812-817.

14. Okuda K, Ohtsuki T, Obata H, Tomimatsu M, Okazaki N, Hasegawa H, Nakajima Y, et al. Natural history of hepatocellular carcinoma and prognosis in relation to treatment. Study of 850 patients. Cancer 1985;56:918-928.

15. Calvet X, Bruix J, Gines P, Bru C, Sole M, Vilana R, Rodes J. Prognostic factors of hepatocellular carcinoma in the west: a multivariate analysis in 206 patients. HEPATOLOGY 1990;12:753-760.

16. Stuart KE, Anand AJ, Jenkins RL. Hepatocellular carcinoma in the United States. Prognostic features, treatment outcome, and survival. Cancer 1996; 77:2217-2222.

17. Keeffe EB. Gastroenterology: Chronic Hepatitis. In: Scientific American Medicine. Dale DC and Federman DD eds. New York. Healtheon/WebMD, 4.VIII(1-9).

18. Trevisani F, D'Intino PE, Morselli-Labate AM, Mazzella G, Accogli E, Caraceni P, Domenicali M, et al. Serum alpha-fetoprotein for diagnosis of hepatocellular carcinoma in patients with chronic liver disease: influence of HBsAg and anti-HCV status. J Hepatol 2001;34:570-575.

19. Gebo KA, Jenckes MW, Chander G, Torbenson MS, Ghanem KG, Herlong HF, Sulkowski MS, et al. Management of chronic hepatitis C. Evidence Report/Technology Assessment No. 60 (Prepared by the Johns Hopkins University Evidence-based Practice Center under Contract No. 290-97-0006). AHRQ Publication No. 02-E030. Rockville, MD: Agency for Healthcare Research and Quality. July 2002.

20. Sato Y, Nakata K, Kato Y, Shima M, Ishii N, Koji T, Taketa K, et al. Early recognition of hepatocellular carcinoma based on altered profiles of alpha-fetoprotein. N Engl J Med 1993;328:1802-1806.

21. McNamara RL, Bass EB, Miller MR, Segal JB, Goodman SN, Kim NL, Robinson KA, et al. Management of new onset atrial fibrillation No. 12 (Prepared by the Johns Hopkins University Evidence-based Practice Center under Contract No. 290-097-0006.) AHRQ Publication No. 01-E026. Rockville, MD: Agency for Healthcare Quality and Research; 2001.

22. Garbutt JC, West SL, Carey TS, Lohr KN, Crews FT. Pharmacological treatment of alcohol dependence: a review of the evidence. JAMA 1999; 281:1318-1325.

23. Schein OD, Friedman DS, Fleisher LA, Lubomski LH, Magaziner J, Spritz M, Kemper J, et al. Evidence Report/Technology Assessment No. 16 (Prepared by the Johns Hopkins University Evidence-based Practice Center under Contract No. 290-097-0006.) AHRQ Publication No. 01-E017. Rockville, MD: Agency for Healthcare Research and Quality; 2001.

24. Solmi L, Primerano AMM, Gandolfi L. Ultrasound follow-up of patients at risk for hepatocellular-carcinoma — results of a prospective study on 360 cases. Am J Gastroenterol 1996;91:1189-1194.

25. Tsuzurahara S, Sata M, Iwamoto O, Shichijo S, Kojiro M, Tanikawa K, Itoh K. Detection of MAGE-4 protein in the sera of patients with hepatitis-C virus-associated hepatocellular carcinoma and liver cirrhosis. Jpn J Cancer Res 1997;88:915-918.

26. Kasahara A, Hayashi N, Mochizuki K, Takayanagi M, Yoshioka K, Kakumu S, Iijima A, et al. Risk factors for hepatocellular carcinoma and its incidence after interferon treatment in patients with chronic hepatitis C. HEPATOLOGY 1998;27:1394-1402.

27. Sassa T, Kumada T, Nakano S, Uematsu T. Clinical utility of simultaneous measurement of serum high-sensitivity des-gamma-carboxy prothrombin and Lens culinaris agglutinin A-reactive alpha-fetoprotein in patients with small hepatocellular carcinoma. Eur J Gastroenterol Hepatol 1999;11:1387-1392.

28. Raedle J, Roth WK, Oremek G, Caspary WF, Zeuzem S. Alpha-fetoprotein and p53 autoantibodies in patients with chronic hepatitis C. Dig Dis Sci 1995;40:2587-2594.

29. Kakumu S, Okumura A, Ishikawa T, Yano M, Enomoto A, Nishimura H, Yoshioka K, et al. Serum levels of Il-10, Il-15 and soluble tumour necrosis factor-alpha (TNF-alpha) receptors in type C chronic liver disease. Clin Exp Immunol 1997;109:458-463.

30. Izzo F, Cremona F, Delrio P, Leonardi E, Castello G, Pignata S, Daniele B, et al. Soluble interleukin-2 receptor levels in hepatocellular cancer: a more sensitive marker than alfa fetoprotein. Ann Surg Oncol 1999;6:178-185.

31. Ishii M, Gama H, Chida N, Ueno Y, Shinzawa H, Takagi T, Toyota T, et al. Simultaneous measurements of serum alpha-fetoprotein and protein induced by vitamin K absence for detecting hepatocellular carcinoma.

South Tohoku District Study Group. Am J Gastroenterol 2000;95:1036-1040.

32. Tong MJ, Blatt LM, Kao VW. Surveillance for hepatocellular carcinoma in patients with chronic viral hepatitis in the United States of America. J Gastroenterol Hepatol 2001;16:553-559.

33. Izzo F, Cremona F, Ruffolo F, Palaia R, Parisi V, Curley SA. Outcome of 67 patients with hepatocellular cancer detected during screening of 1125 patients with chronic hepatitis. Ann Surg 1998;227:513-518.

34. Colombo M, Franchis R, Del Ninno E, Sangiovanni A, De Fazio C, Tommasini M, Donato MF, et al. Hepatocellular carcinoma in Italian patients with cirrhosis. N Engl J Med 1991;325:675.

35. Larcos G, Sorokopud H, Berry G, Farrell GC. Sonographic screening for hepatocellular carcinoma in patients with chronic hepatitis or cirrhosis: an evaluation. AJR Am J Roentgenol 1998;171:433-435.

36. Tradati F, Colombo M, Mannucci PM, Rumi MG, De Fazio C, Gamba G, Ciavarella N, et al. A prospective multicenter study of hepatocellular carcinoma in Italian hemophiliacs with chronic hepatitis C. The Study Group of the Association of Italian Hemophilia Centers. Blood 1998;91:1173-1177.

37. Nomura F, Ishijima M, Horikoshi A, Nakai T, Ohnishi K. Determination of serum des-gamma-carboxy prothrombin levels in patients with small-sized hepatocellular carcinoma: comparison of the conventional enzyme immunoassay and two modified methods. Am J Gastroenterol 1996;91:1380-1383.

38. Nagai T, Murota M, Nishioka M, Fujita J, Ohtsuki Y, Dohmoto K, Hojo S, et al. Elevation of cytokeratin 19 fragment in serum in patients with hepatoma: its clinical significance. Eur J Gastroenterol Hepatol 2001;13:157-161.

39. Tsai JF, Jeng JE, Ho MS, Chang WY, Lin ZY, Tsai JH. Clinical evaluation of serum alpha-fetoprotein and circulating immune complexes as tumour markers of hepatocellular carcinoma. Br J Cancer 1995;72:442-446.

40. Tsai JF, Jeng JE, Chuang LY, Chang WY, Tsai JH. Urinary transforming growth factor beta(1) levels in hepatitis C virus-related chronic liver disease: correlation between high levels and severity of disease. HEPATOLOGY 1997;25:1141-1146.

41. VanRoey G, Fevery J, VanSteenbergen W. Hepatocellular carcinoma in Belgium: clinical and virological characteristics of 154 consecutive cirrhotic and non-cirrhotic patients. Eur J Gastroenterol Hepatol 2000;12:61-66.

42. Trevisani F, D'Intino PE, Caraceni P, Pizzo M, Stefanini GF, Mazziotti A, Grazi GL, et al. Etiologic factors and clinical presentation of hepatocellular carcinoma. Cancer 1995;75:2220-2232.

43. Tsai J-F, Chang WY, Jeng JE, Ho MS, Lin ZY, Tsai JH. Frequency of raised alpha-fetoprotein level among Chinese patients with hepatocellular carcinoma related to hepatitis B and C. Br J Cancer 1994;69:1157-1159.

44. Cedrone A, Covino M, Caturelli E, Pompili M, Lorenzelli G, Villani MR, Valle D, et al. Utility of alpha-fetoprotein (AFP) in the screening of patients with virus-related chronic liver disease: does different viral etiology influence AFP levels in HCC? A study in 350 western patients. Hepatogastroenterology 2000;47:1654-1658.

45. Zaman SN, Melia WM, Johnson RD, Portmann BC, Johnson PJ, Williams R. Risk factors in development of hepatocellular carcinoma in cirrhosis: prospective study of 613 patients. Lancet 1985;1:1357-1360.

46. Bruno S, Silini E, Crosignani A, Borzio F, Leandro G, Bono F, Asti M, et al. Hepatitis C virus genotypes and risk of hepatocellular carcinoma in cirrhosis: A prospective study. HEPATOLOGY 1997;25:754-758.

47. Di Bisceglie AM, Hoofnagle JH. Elevations in serum alpha-fetoprotein levels in patients with chronic hepatitis B. Cancer 1989;64:2117-2120.

48. Collier J, Sherman M. Screening for hepatocellular carcinoma. HEPATOLOGY 1998;27:273-278.

49. Pateron D, Ganne N, Trinchet JC, Aurousseau MH, Mal F, Meicler C, Coderc E, et al. Prospective study of screening for hepatocellular carcinoma in Caucasian patients with cirrhosis. J Hepatol 1994;20:65-71.

50. Sheu JC, Sung JL, Chen DS, Yang PM, Lai MY, Lee CS, Hsu HC, et al. Growth rate of asymptomatic hepatocellular carcinoma and its clinical implications. Gastroenterology 1985;89:259-266.

51. Prorok PC. Epidemiologic approach for cancer screening. Problems in design and analysis of trials. Am J Pediatr Hematol Oncol 1992;14:117-128.