In whom, how and how often is surveillance for hepatocellular carcinoma cost-effective?

Ana Lilia Ruelas-Villavicencio; Florencia Vargas-Vorácková

Abstract

Background: Despite well known worldwide differences in hepatocellular carcinoma incidence, which reflect different risk profiles, current recommendation of surveillance with ultrasound and alpha-fetoprotein twice-a-year has been restricted to cirrhotic patients. To evaluate the generalizability of this recommendation, we reviewed the clinical charts of hepatocellular carcinoma cases in a Mexican scenario. To evaluate efficiency, we performed a literature based cost-effectiveness analysis. Methods: Charts pertaining to 174 consecutive patients with histologically proven hepatocellular carcinoma, seen at a tertiary health care centre were analysed. A decision tree, based on the surveillance and recall algorithm of the European Association for the Study of the Liver was constructed. Ultrasound and/or alpha-fetoprotein, performed every six or twelve months were the diagnostic alternatives, and accurate diagnoses, direct medical costs and cost-effectiveness ratios were the outcomes of interest. Results: Male:female ratio was 1.2:1, underlying liver disease was secondary to alcohol in 44% and to hepatitis C virus in 26%, documented cirrhosis was present in 42%. Cost-effectiveness ratios for twice-a-year ultrasound and alpha-fetoprotein ranged from $303.09 to $346.22 U.S. dollars per accurate diagnosis, and for annual ultrasound from $115.86 to $116.42 U.S. dollars. Conclusions: Male gender, hepatitis C and cirrhosis were not predominant characteristics in our series. If a hepatocellular carcinoma surveillance program were to be instituted in our setting, or where patient characteristics are similar to ours, it probably should not be restricted to cirrhotic patients. Recommended performance of ultrasound and alpha-fetoprotein every six months is the least cost-effective surveillance strategy. Instead, annual ultrasound optimises diagnoses and costs.

Key words: Hepatocellular carcinoma, cirrhosis, surveillance, alpha-fetoprotein, ultrasound, cost-effectiveness.

Introduction

Hepatocellular carcinoma (HCC) is the fifth most common neoplasm in the world. It accounts for more than 500,000 new cases every year, and is the third cause of mortality due to cancer. Known risk factors for HCC development are cirrhosis of any etiology, hepatitis B (HBV) and C virus (HCV) infection, hereditary liver diseases, and exposure to carcinogens (e.g. aflatoxins). Many HCC develop in cirrhotic livers, with an estimated annual incidence of 1% to 6%. Approximately 78% of HCC coincide with cirrhosis, but this percentage can be as low as 50% to 60%. In large surgical series of hepatic resections for HCC, cirrhosis was absent in 15% to 25%. In patients with chronic viral hepatitis, cirrhosis amplifies HCC risk, but is not a prerequisite for liver carcinogenesis. Recently, obesity and dyslipidemias have been proposed as independent risk factors for HCC development. Accepted means for HCC detection are imaging techniques such as ultrasound (US), computed tomography (CT) scan, magnetic resonance imaging, angiography, and Tecnetium-99 scintiscan. Biochemical markers such as alpha-fetoprotein (AFP) concentration are other diagnostic means. Definite diagnosis is established by histology.

Diagnosis of early HCC, when tumour size is ≤2 cm, allows therapeutic intervention and, in theory, improves survival. Treatment options are surgical resection, percutaneous alcohol injection, arterial embolisation, intraarterial chemotherapy or liver transplantation. Newer therapeutic developments are radio-frequency, microwave, cryotherapy or laser destruction.

To become detectable, i.e. about 2 cm size, a HCC needs between four to twelve months to grow. This growth rate sustains surveillance at three to twelve-
month intervals. Based on the assumption that an early detection will lead to better outcomes, surveillance with US and AFP at intervals of six months has become common practice in cirrhotic patients. This schema was recently ratified by the Panel of Experts on HCC of the European Association for the Study of the Liver (EASL) (Figure 1). Aside of this expert opinion, there is no evidence to support that this is the most efficient surveillance strategy. The restriction of HCC surveillance to cirrhotic patients may exclude, on the other hand, an important proportion of patients at risk, given that up to 40% HCC patients might not have cirrhosis, and twenty to 56% can present silent cirrhosis. Variations in HCC frequency and risk factors around the world, suggest that surveillance strategies should be adjusted according to local epidemiological situation.

This study was undertaken to assess the generalizability and efficiency of currently recommended HCC surveillance strategy, namely performance of US and AFP in cirrhotic patients at intervals of six months. For the former, clinical charts of HCC cases in a Mexican scenario were reviewed. For the latter, a literature based cost-effectiveness analysis was carried out.

Materials and methods

Clinical chart review

Patients

Clinical charts of patients with histologically proven HCC, seen in the last 16 years at the Instituto Nacional de Ciencias Medicas y Nutricion Salvador Zubiran, a Mexican third level health care centre, were reviewed. Information was gathered on age, gender, history of obesity (BMI >30 Kg/m²), diabetes mellitus, dyslipidemia, arterial hypertension, smoking, alcohol consumption, clinical manifestations, liver disease and etiology, liver function tests, cholesterol levels, triglyceride, AFP, HCC detection methods, tumour characteristics, treatment, and survival.

Diagnosis of cirrhosis

Presence of cirrhosis was supported by histology, or by clinical presentation and laboratory abnormalities associated to endoscopic and/or imaging findings (esophageal varices and/or irregular liver contour at US or CT scan, respectively).

Classification of underlying liver disease

Underlying liver disease was classified as i) alcohol-related when history of alcohol intake was present and both, HBsAg and anti-HCV were negative, ii) viral when HBsAg and/or anti-HCV were positive, and iii) cryptogenic when no viral, metabolic or immune causes were identified. When the cause of the liver disease was primary biliary cirrhosis, haemochromatosis, autoimmune hepatitis, or mixed it was classified as “other”. Under “unknown” were grouped all cases of HCC in whom no etiology was identified.

Data management

Data extracted from clinical charts were summarized as absolute and relative frequencies, or mean ± standard deviation and intervals. The Pearson correlation coefficient was used to estimate correlation between tumour size and AFP concentrations, and the Kaplan-Meier approach to estimate cumulative survival after HCC diagnosis. The Stata 7.0 statistical package was used.

Figure 1. EASL HCC surveillance and recall algorithm. Modified from G (2). US=ultrasound, AFP=alphafetoprotein, CT=computed tomography, MRI=magnetic resonance imaging, HCC=hepatocarcinoma.
Cost-effectiveness analysis

Costs

From the point-of-view of the health care provider, direct medical costs of HCC detection were considered. Charges were used as estimates of costs. Adopting an exchange rate of $10.00 Mexican pesos per US dollar, the cost of AFP determination was $12.20, for liver US $86.20, for US-guided liver biopsy $183.00, and for liver CT scan $260.10 US dollars.

Benefits

These were defined in terms of diagnostic effectiveness. Correct diagnoses (i.e., true positive or true negative) were weighted with a value of one, and incorrect diagnoses (i.e., false positive or false negative) were weighted with a value of zero.

Decision analysis

A decision tree, based on the surveillance and recall algorithm proposed by the EASL, was constructed. The population at risk included patients with chronic liver disease or Child A cirrhosis. In it, three branches parted from the starting decision node, namely AFP determination alone, US performance alone, and combined performance of US and AFP. The subsequent recall and diagnostic confirmation strategy included fine needle aspiration biopsy, and either CT scan, magnetic resonance imaging or angiography. The branch pertaining tumour size of <1 cm was pruned, due to its association to a very low prevalence (close to 0) (Figure 2). Two surveillance intervals were analyzed, namely six and twelve months.

Tree probabilities were extracted and averaged from published literature. Among all nodules with size =1 cm detected during surveillance, a prevalence of 60% was assumed for those >2 cm of diameter. HCC annual incidence was assumed constant over time and averaged at 3% in patients at risk.

The lowest and highest sensitivity and specificity values published in the literature for AFP, US, histology and CT scan were adopted, to build a worst and best scenario, respectively. Thus, AFP sensitivity ranged from 39% to 64%, and specificity from 76% to 91%. US sensitivity ranged from 71% to 78%, and specificity was 93%. Sensitivity of liver histology ranged from 69% to 96%, and specificity from 71% to 100%. Sensitivity of CT scan ranged from 87% to 97%, and specificity was 81%. Independence among diagnostic tests was assumed.

Starting with a cohort of 1000 patients, followed for a one year period, expected values were calculated for each decision branch by means of the Bayes theorem. Total and incremental costs, benefits and cost-benefit ratios were determined. A sensitivity analysis was carried out on the prevalence of tumour size above 2 cm.

The TreeAge Data 3.5.1 software was used to construct the decision tree and conduct the cost-effectiveness analysis.

Results

Characteristics of HCC patients

One hundred and seventy four patients with histologically proven HCC were analyzed in this series. Ninety four of them were males, with a male:female ratio of 1.2:1. The mean age at the time of HCC diagnosis was 58.9 ± 15.5 years. Esophageal varices were found in 58 of 129 patients who underwent upper gastrointestinal endoscopy. Prevalence of diabetes mellitus was 32%, hypertension 20%, obesity 12%, and hyperlipidemia 5%. Fifty eight percent of the patients had history of alcohol abuse, and 54% history of smoking.

In 114 (66%) patients the underlying liver disease was associated to alcohol (44%), to HCV infection (26%), or

![Figure 2. HCC surveillance and recall decision tree. Squares represent decision nodes, circles probability nodes, and triangles end nodes. For the analysis of surveillance on a six-month basis, filled probability nodes were turned into decision nodes before the whole strategy was repeated. AFP=alphafetoprotein, US=ultrasound, CT=computed tomography, Bx=biopsy, HCC=hepatocarcinoma.](image-url)
HBV infection (12%). In 10% it was associated to more than one cause, or to metabolic or autoimmune disturbances, and in 8% it was classified as cryptogenic. In 60 (34%) patients the underlying liver disease was “unknown”. Seventy-three (42%) patients fulfilled the cirrhosis criteria. Among 59 patients who underwent liver biopsy, 18 (30%) showed findings compatible with cirrhosis.

Anti-HCV and HBsAg were determined in 87 (50%) and 126 (72%) patients, respectively, and were positive in 39 (45%) and 18 (14%). Taken together, both viral markers were determined in 82 patients, and were positive in four.

At the time of HCC diagnosis, 141 (81%) patients were symptomatic. Most frequent clinical manifestations were abdominal pain (45%), weight loss (36%), abdominal mass (23%), hepatomegaly (18%), and jaundice (15%). Less frequent were fever, gastric fullness, vomiting, diarrhea, gastrointestinal bleeding, ascites, bulging abdomen, deep vein thrombosis, portal thrombosis, acute abdomen, hypoglycemia and thrombocytopenia.

Laboratory exams at the time of HCC diagnosis are summarized in table I.

HCC was detected by US and AFP in 31 (18%) patients. The mean AFP concentration was 218.8 ± 537.9 ng/mL, with a minimum of zero and a maximum of 5400 ng/mL. Eighty seven (54%) patients presented AFP levels above 20 ng/mL. Correlation between AFP concentrations and tumour size, as determined by the largest diameter in single lesions or the sum of the largest diameters of all detected lesions, was $r = -0.07$.

HCC was present as a single focal lesion in 77 (44%) patients, as multifocal lesions in 92 (53%), and as diffuse lesion in four (2%). Tumour size, as determined by the largest diameter of the largest lesion, was measured in 123 patients, with a mean of 8.8 ± 5.5 cm, and minimum and maximum of 1.3 and 30 cm, respectively. Only five (4%) patients presented a tumour size of ≤2 cm and, among them, none a tumour size of <1 cm. In the same patients, vascular invasion was documented in 22 (18%), positive nodes in 25 (20%), and metastases in 23 (19%). The site of metastases was lung (26%), mesenterum (26%), peritoneum (22%), retroperitoneum (13%), bone (9%), and pleura, diaphragm and colon (4%).

Liver ultrasound was performed in 110 patients, CT scan in 108 and magnetic resonance imaging in 50. The overall sensitivity to detect HCC was 90%, 97% and 98%, respectively, and for tumour size of ≤2 cm sensitivity was 60%, 100% and 100%.

HCC was well differentiated in 44 patients, intermediately differentiated in 25 and poorly differentiated in 8. The tumour was classified as compatible to HCC in 12 patients, as fibrolamellar in 8, and as mixed carcinoma (non-differentiated hepatocellular and adenocarcinoma) in one patient. In 76 (44%) patients, histological findings were not described.

Due to an advanced HCC stage and/or poor liver function, 53 (30%) patients received symptomatic treatment only. The remaining 121 (70%) received one or more of the following: tamoxifen (41%), intraarterial chemotherapy (36%), systemic chemotherapy (10%), percutaneous alcohol injection (8%), arterial embolisation (7%), and other therapeutic interventions (6%). Among the latter were radio-frequency, conventional radiotherapy, gastrojejunal bypass, and bile endoprosthesis.

Median survival time was 4.6 months. (Figure 3)

Cost-effectiveness analysis

Among 1000 patients with chronic liver disease, and either no- or Child A-B cirrhosis, 30 are expected to de-

<table>
<thead>
<tr>
<th>Test</th>
<th>n</th>
<th>mean ± SD (interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total bilirubin, mg/dL</td>
<td>174</td>
<td>1.4 ± 4.0 (0.1-37.5)</td>
</tr>
<tr>
<td>ALT, IU/dL</td>
<td>174</td>
<td>74.5 ± 73.6 (2-638)</td>
</tr>
<tr>
<td>AST, IU/dL</td>
<td>172</td>
<td>97.3 ± 93.3 (6-699)</td>
</tr>
<tr>
<td>Alkaline phosphatase, IU/dL</td>
<td>174</td>
<td>285.4 ± 331.2 (31-2491)</td>
</tr>
<tr>
<td>Albumin, g/L</td>
<td>173</td>
<td>2.1 ± 0.8 (1-5.3)</td>
</tr>
<tr>
<td>Platelet count, x10^9/mm³</td>
<td>171</td>
<td>217.3 ± 133.3 (3.6-687)</td>
</tr>
<tr>
<td>Cholesterol, mg/dL</td>
<td>158</td>
<td>177.4 ± 86.1 (78-663)</td>
</tr>
<tr>
<td>Triglyceride, mg/dL</td>
<td>78</td>
<td>120.1 ± 67.5 (17-416)</td>
</tr>
</tbody>
</table>

Figure 3. Cumulative survival of HCC patients. The filled diamond shows the median survival, namely 4.6 months.
velop HCC in one year. If these patients undergo surveillance with both AFP and liver US every six months, followed by the EASL recall algorithm, at the end of one year 879 and 945 correct diagnoses will be made under the worst and best scenarios, respectively. This will cost between $286 319 and $393 807 US dollars, or $303.09 and $346.22 US dollars per correct diagnosis. If surveillance relies on AFP determinations every six months, followed by liver US and by the EASL recall algorithm if concentrations are above normal, at the end of one year 903 to 961 correct diagnoses will be made, at a total cost of $92 489 to $189 118 US dollars, or $96.21 to $209.43 US dollars per correct diagnosis. If only liver US is carried out every six months, followed by AFP if normal or the EASL recall algorithm if a nodule is detected, 957 to 971 correct diagnoses will be obtained, at a total cost of $217 677 to $221 681 US dollars, or $227.50 to 228.32 US dollar per correct diagnosis. Taking the three alternatives together, surveillance with both US and AFP every six months is the least beneficial and the most expensive strategy. This is more noticeable if the same alternatives are performed on a yearly basis (Table II).

When the nature of correct diagnoses is analyzed, surveillance with both US and AFP allows the detection of most HCC, specifically 22 to 29 cases of the 30 expected. AFP alone, in turn, allows the detection of 11 to 23 cases, and US alone the detection of 19 to 26 cases. Performance of liver US allows also the attainment of both the highest number of true positives and the lowest number of false positives. When performed every six months, US leads to the detection of 23 to 26 HCC at a cost of 25 to 36 false positives. When performed every year, it yields 19 to 23 HCC at an expense of 12 to 20 false positives (Table III). Comparing performance of US at six and twelve months intervals, the three to four additional HCC cases detected by a more frequent search represent a total additional cost of $105 415 to $107 518 US dollars (Table II), or $26 353.75 to $35 839.33 US dollars per additional HCC detected.

Results were not sensitive to changes in the prevalence of tumours with diameter above 2 cm. This prevalence was increased from 0% up to 100%.

Discussion

Nowadays known risk factors for HCC development are male gender, increasing age and liver cirrhosis of any aetiology. In cirrhotic patients, alcohol and viral hepatitis are additional risk factors for HCC. This knowledge has lead to recommend HCC surveillance to cirrhotic patients only.

Male predominance was not observed in our series, where the male:female ratio was 1.2:1. The mean age of our patients was 59 years, which agrees with the 50 to 70 years age interval published in the literature, but applies to developed environments mainly. In developing countries,

<table>
<thead>
<tr>
<th>Table II. Effectiveness, costs and cost-effectiveness ratios of HCC surveillance strategies.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategy</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>6-month intervals</td>
</tr>
<tr>
<td>US + AFP</td>
</tr>
<tr>
<td>AFP</td>
</tr>
<tr>
<td>US</td>
</tr>
<tr>
<td>12-month interval</td>
</tr>
<tr>
<td>US + AFP</td>
</tr>
<tr>
<td>AFP</td>
</tr>
<tr>
<td>US</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table III. Diagnostic effectiveness of HCC surveillance strategies.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategy</td>
</tr>
<tr>
<td>6-month intervals</td>
</tr>
<tr>
<td>US + AFP</td>
</tr>
<tr>
<td>AFP</td>
</tr>
<tr>
<td>US</td>
</tr>
<tr>
<td>12-month interval</td>
</tr>
<tr>
<td>US + AFP</td>
</tr>
<tr>
<td>AFP</td>
</tr>
<tr>
<td>US</td>
</tr>
</tbody>
</table>
HCC appears at a younger age, which explains the lower proportion of underlying cirrhosis. In spite of being older, only 42% of our HCC patients showed signs of cirrhosis, which is not far from the 56% found by Mondragon et al in their Mexican series of HCC patients. Given the nature of our review, we should not discard, however, that our prevalence might be underestimated due to the lack of systematic search for underlying liver disease and/or missing information. In Western industrialized countries, cirrhosis is present in 80% to 90% of patients who develop HCC. These variations suggest that although regeneration inherent to cirrhosis increases susceptibility for HCC development, its presence is not a prerequisite for liver carcinogenesis, and that malignant transformation may be influenced by the same disorder that caused cirrhosis. In view of this evidence, current restriction of HCC surveillance to cirrhotic patients might exclude, in settings such as ours, a significant proportion of patients at risk. It might exclude, additionally, another group at risk, namely patients with silent or non-diagnosed cirrhosis.

Significant ethnic and regional variability in the pathogenesis of HCC has been observed. Anti-HCV is the most frequently found associated factor in both blacks and whites with HCC (53.7% and 51.7%), whereas HBsAg is frequently associated factor in both blacks and whites with HCC (53.7% and 51.7%), whereas HBsAg is mainly found in Asians (49.5%).

The most frequent cause of underlying liver injury in our series was alcohol abuse. The risk for HCC development in subjects with alcohol-related cirrhosis has been estimated to be 3% to 15% and has been attributed mainly to the underlying cirrhosis, and the coexistence of HCV infection, and an older age. Consistent with the latter, HCV infection was present in 45% and 33% of our alcohol- and non-alcohol-related cirrhotics, and mean age was 63.7 and 54.5 years, respectively.

Other factors associated to an increased risk of HCC are diabetes mellitus, obesity with hyperlipidemia, and fatty liver disease. Reported prevalence of diabetes mellitus in HCC patients varies from 19% to 23% (38-40), and determines a 1.4 to 3.3 fold increase in HCC risk. In our series, the prevalence of diabetes was somewhat higher, namely 32%, and was not affected by the presence of cirrhosis. This higher prevalence might reflect the increased susceptibility of Mexican people to present diabetes. Whether it determines an increased risk for HCC development remains to be established. Obesity has been reported in 34% of HCC patients, but was significantly lower in ours (12%). As for hyperlipidemia, its prevalence in HCC varies from 5% to 15% in the literature, and was similar in our series (5%).

More than half of our patients presented with multifocal lesions, and almost 70% had a tumour size above 5 cm. This advanced disease was out of any therapeutic possibility in 30%, and determined a very short median survival time (4.6 months). Consistency of our findings with those published by Mondragon et al shows that, in our setting, HCC is diagnosed at a late stage.

The increasing incidence of HCC, and its usually advanced stage at diagnosis underscores the need for effective surveillance strategies. Randomized controlled trials assessing this effectiveness are lacking, and are not expected to be conducted in the future due to ethical restrictions. In most countries, common resources such as US are accessible and widely used in patients thought at risk, making withdrawal of these procedures unacceptable for the patients. So, there is no evidence up to date that an early HCC diagnosis and treatment lead to an improved survival. This is why the usefulness of HCC surveillance remains controversial and raises an important issue of costs.

Assuming that surveillance might be useful, our analysis shows that current recommendation of US and AFP performance at intervals of six months is the least cost-effective, given that it renders the lowest number of correct diagnoses at the highest cost. This is because US and AFP together, in spite of detecting most HCC, determine a larger number of false positives. Given that therapeutic usefulness of early HCC remains unproven, and that it can have significant risks when applied to patients without HCC, emphasis should be given not only to HCC detection but also to patients who do not have HCC but test positive. US alone increases the number of correct diagnoses at a lower cost, but reduces the number of HCC detected. Specifically, under the worst scenario, US detects 63% to 77% of all HCC expected. Compared with the combined use of US and AFP, it misses one and five HCC cases, but reduces false positives by 79 (67%) and 95 (83%), respectively (Table III). As to its frequency, an US performed every six months instead of once a year implies an additional cost of $26,353.75 US dollars for each additional HCC case detected. This amount might be socially acceptable, at least in the developed world. However, the gain of four additional HCC cases detected is obscured by the cost and risk of 16 additional false positives. Taking all former considerations together, and given the current state of evidence, an annual US seems to be an acceptable and cost-effective HCC surveillance strategy. Whether its use should be restricted to cirrhotics or to other groups at risk is an issue that needs to be weighted with the disease profile present in different geographic regions.

In conclusion, male gender, hepatitis C and cirrhosis were not predominant characteristics in our series. For Mexican third level health care settings, and for settings that have similar HCC patient characteristics as ours, current restriction of HCC surveillance to cirrhotic patients might not be generalizable. As to the HCC surveillance strategy, the recommended practice of US and AFP at six months intervals is the least cost-effective. Annual US followed by the EASL recall algorithm is an alternative that optimizes both true positives and false positives, as well as costs.
References


