Review

JHEP|Reports

How to improve HCC surveillance outcomes

Morris Sherman[1,*]

Summary

Outside of expert centres, surveillance programmes for hepatocellular carcinoma (HCC) are not well executed. There are deficiencies in every stage of the process. Overcoming these obstacles is the most important method for improving surveillance. However, even if these obstacles were overcome, there would still be room for improvement. Assessing who is at risk of developing HCC remains incompletely validated. At present, risk scores have been developed for different causes of liver disease, but scores developed in different parts of the world for the same disease do not always agree. Furthermore, most scores stratify patients by risk but do not examine what level of risk should trigger surveillance. Which surveillance tools to use remains controversial – schemes have been proposed that use biomarkers alone, ultrasound alone, or a combination of both. However, the requisite level of test sensitivity that would be associated with high cure rates has not been defined, so at this point it is not clear whether surveillance requires both ultrasound and biomarkers, or whether the use of biomarkers alone is sufficient. Finally, surveillance should result in the identification of HCC at a very early stage. Diagnosing these lesions is difficult and optimal algorithms for lesions that are atypical on radiology have yet to be developed. Algorithms for the follow-up of abnormal biomarkers in the absence of ultrasound have also not been developed yet.

Keywords: hepatocellular carcinoma; Risk Score; Surveillance; LI-RADS.

Received 24 August 2019; received in revised form 21 October 2019; accepted 23 October 2019; available online 5 November 2019

Introduction

It is generally accepted, by hepatologists at least, that surveillance for hepatocellular carcinoma (HCC) should be standard practice in patients known to be at risk of this cancer. However, it has been difficult to demonstrate this conclusively and there remain those who believe that this should not be generally instituted until we have better evidence. In part, the problem is one common to many surveillance programmes, namely that the majority of individuals currently defined as being at risk will not get HCC, and therefore the potential for harm resulting from overdiagnosis, in the form of unnecessary downstream investigations and potentially unnecessary treatment, is significant. Nonetheless, as all hepatologists know, patients whose HCC can be diagnosed early have a significant likelihood of cure, whereas cure is uncommon for those whose cancer presents only when symptoms develop. Therefore, in order to justify a surveillance programme, it is incumbent on us to try to target surveillance to those most likely to benefit and to maximise the efficiency and minimise the costs of providing surveillance.

It is important to consider surveillance as a programme rather than the provision of surveillance tests alone. The components of this programme include identification of the at-risk population, determining the optimal tests(s) and surveillance interval, and establishing the optimal recall strategy. The recall strategy includes when and how to investigate abnormal surveillance test results and to come to an appropriate diagnosis. Improvements in any of these components could enhance the outcomes achieved with surveillance programmes.

A surveillance programme includes the identification of candidates for surveillance, as well as the provision of the appropriate tests and appropriate recall procedures in the case of a positive surveillance test (Fig. 1). The first step in a programmatic approach (as opposed to simply providing tests) is to identify that liver disease is present. Then the presence of cirrhosis must be established. Doctors have to be convinced that such patients need surveillance and patients have to be convinced of its necessity. Abnormal tests must be recognised and appropriate investigations initiated. There is evidence that many steps along this pathway are carried out poorly. In one study only about 2% of patients with cirrhosis underwent bi-annual surveillance.[1] Once liver disease has been identified, cirrhosis must be recognised, yet often this step is carried out poorly.[2] Cirrhosis is essentially silent until liver failure supervenes. Nonetheless, there are clues such as elevated alanine aminotransferase (ALT), a falling platelet count, coarse appearance of the liver on ultrasound and abnormal albumin

[1]*University Health Network, Toronto*

* Corresponding author. Address: 560 Briar Hill Avenue, Toronto, ON, M5N 1M9, Canada. Tel.: +1 416 804 1146. *E-mail address:* morris.sherman@uhn.on.ca (M. Sherman).

or international normalised ratio that should lead to further investigation. Today, any patient known to have liver disease should have one of the non-invasive tests of liver fibrosis. Although none of these have high accuracy, it is better than not doing any investigations and will increase the number of patients in whom cirrhosis is identified.

One of the concepts around cancer surveillance is that the benefits increase as the incidence rate increases. This is because the benefits accrue only to those with cancer, while the harms (but not all harms) accrue to those with and without cancer. Therefore, as cancer incidence increases more patients benefit, while the harms remain the same (or decrease if unnecessary tests are included as harms). This means that the higher the incidence of the cancer in question the more likely the programme is to be effective. Thus, better identification of the at-risk population is required for HCC.

Key points

Liver cancer is one of the most common cancers in the world. Early detection is key to improving survival.

For most patients who develop liver cancer there is an underlying liver disease that leads to the development of the cancer.

Several steps are required for the early detection of HCC.

The underlying liver disease must be recognised, the patient must then undergo regular check-ups most commonly using ultrasound, but also blood tests. Abnormal ultrasounds or blood tests must be recognised and investigated.

All steps in the surveillance process can be improved, hopefully leading to improved clinical outcomes in patients at risk of HCC.

Risk stratification

It is well recognised that within broadly defined risk groups only a minority will ever develop cancer. The risk groups are well known for HCC, but an attempt to target those who have the highest HCC incidence has led to the development of a number of risk scores that claim to do just that, and equally importantly, to identify those who have a low likelihood of developing HCC and who can possibly be excluded from surveillance programmes.

The decision to offer a patient surveillance depends on the degree of risk, but the level of risk that is necessary to trigger surveillance and how this should be determined remain unclear. Even patients with no liver disease, no diabetes and none of the usual risk factors still have a finite, albeit very small risk of developing HCC. The decision to offer surveillance thus depends on what point along the spectrum of risk, from negligible to more than 5% per year, the benefits exceed the harms, and whether this can be achieved at a reasonable cost. In the absence of large-scale randomised controlled trials, cost-efficacy analyses must answer this question.

A further consideration is whether the patient is a candidate for therapy. Patients with Child-Pugh C status should not undergo surveillance unless the finding of an HCC would be an indication for liver transplantation.

Risk scores in hepatitis B

An early attempt to target the at-risk population was described in the AASLD guidelines in 2005,[3] where a cut-off incidence of 0.2%/year for hepatitis B and 1.5%/year for hepatitis C and other forms of cirrhosis was established. These criteria were established on the basis of cost-efficacy analyses. More sophisticated models based on multivariable regression analysis are now available that stratify individuals into different risk categories. Presumably the highest risk cohort should undergo surveillance, but whether the incidence of HCC in the middle and lower risk strata warrant surveillance has not been established. One of the earliest models was the GAG-HCC score, derived by multivariate analysis of risk factors in untreated patients with chronic hepatitis B who developed HCC[4]; the score consists of age, gender, HBV DNA, presence of cirrhosis and presence of the core-promoter mutation. The CU-HCC score looked at a similar cohort of patients with chronic hepatitis B using similar methodology,[5] but identified different predictors of risk, consisting of age, albumin, bilirubin, HBV DNA and the presence or absence of cirrhosis. More recently, the authors substituted the categorical variable cirrhosis for liver stiffness, measured by FibroScan®, which was categorised into 3 strata.[6] This improved the predictive ability of the score. However, the fact that 2 different risk scores were derived from essentially identical populations (hepatitis B carriers in Hong Kong) indicates the difficulty of developing a universal scoring system.

The REACH-B score was derived from the REVEAL study population in Taiwan (also Asian patients with hepatitis B).[7,8] REVEAL was a large-scale community study in Taiwan performed in more than 4,000 individuals with hepatitis B who were followed for more than 10 years. Analysis of the potential risk factors in those who developed HCC and those who did not initially led to the development of several nomograms[8] to predict who would get HCC, but this was ultimately simplified into a score based on age, gender, HBV DNA concentration, ALT and HBeAg/anti-HBe status.[9] The score was externally validated in Hong Kong and South Korea.[10] Like the GAG-HCC and CU-HCC score this score is only applicable to hepatitis B. The study that led to the REACH-B score did not include cirrhosis as a variable, which has advantages and disadvantages. The advantage is that it makes the score easier to use since the non-invasive diagnosis of cirrhosis may

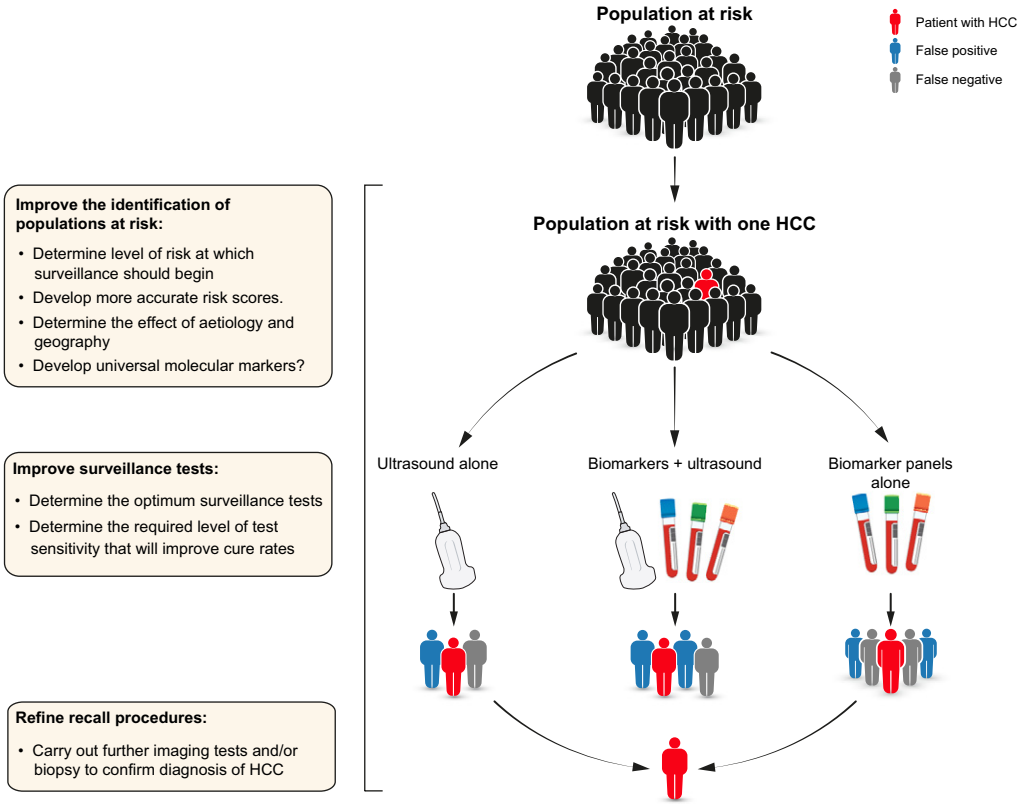

Fig. 1. Possible surveillance paradigms. Note that the use of ultrasound and biomarkers increases the false-positive rate. Using current biomarkers without ultrasound increases both false-positive and false-negative rates.

be inaccurate. Older age then becomes a surrogate for cirrhosis. The disadvantage is that it may not perform as well in younger patients with cirrhosis.

All the aforementioned models were developed in untreated cohorts of patients with chronic hepatitis B. There is now ample evidence that treatment of hepatitis B (and hepatitis C) reduces HCC incidence. Whether the risk scores remain applicable after treatment has recently started to be addressed. Wong *et al.*.[11] evaluated the REACH-B score, the GAG-HCC score and the CU-HCC score in patients with hepatitis B treated with entecavir. They found that at baseline (prior to treatment) each of the scores predicted the development of HCC with varying degrees of accuracy (based on ROC curve analysis). As expected, risk and risk scores decreased over time while on treatment. On-treatment risk scores were relatively good at predicting who would or would not develop HCC, although the 3 risk scores did not perform equally well. However, the scores were less reliable predictors in a more mixed European or North American population.[12,13] CU-HCC still showed high accuracy (area under the receiver operating curve [AUROC] 0.89) but the other risk scores did not perform as well. The PAGE-B score was developed in a European cohort of patients with hepatitis B.[14] This has been validated in another European cohort[15] and in Asia.[16]

Each of these scores is different in terms of the variables identified by regression analysis and the accuracy of risk prediction. This may be due to heterogeneity in the populations in which the scores were derived, but the differences indicate that each score requires considerably more validation before being suitable for general use.

Risk scores in hepatitis C

There have been fewer modelling studies in patients with chronic hepatitis C. The first risk score was derived from the HALT-C study.[17] The score included age, black race, alkaline phosphatase, oesophageal varices, ever smoked and platelets. This score clearly separated out those at high risk from those at minimal risk. Although the authors did not specifically address this, the incidence of HCC in the low risk group was so low as to make surveillance unlikely to be beneficial if the AASLD criteria were applied. Other more recent risk scores have been developed for hepatitis C,[18,19] but the same criticisms exist regarding the heterogeneity of the populations and hence the validity and usability of the scores themselves.

Risk scores in cirrhosis

Others have attempted to assess those at highest risk in an undifferentiated cirrhotic population. The ADRESS-HCC model was derived from patients

on the transplant waiting list.[20] Therefore, this scoring system may not be applicable to any other situation, since patients on the transplant waiting list have the most advanced disease and are thus at the highest risk of developing HCC. This score includes age, race, presence of diabetes, fibrosis stage (F3 or F4), and platelets. The value of this score is not clear, since all patients on the transplant waiting list would undergo regular surveillance, and it is not clear that a low risk score would result in surveillance not being offered. It is also not clear why there were patients with F3 fibrosis on the transplant waiting list unless HCC was the indication for transplant, in which case this would not be a predictive score, but a diagnostic score.

Effect of aetiology of cirrhosis on risk of HCC

Not all patients with cirrhosis will develop HCC, yet the recommendations suggest that all should undergo surveillance. Until recently it has not been possible to identify those aetiologies associated with a higher or lower likelihood of developing HCC. A Danish study was the first to evaluate whether the risk of HCC was high enough to warrant surveillance in a population of patients with alcohol-related cirrhosis.[21] This study showed that the risk of HCC was below the AASLD cut-off of 1.5% for non-hepatitis B cirrhosis. In contrast, a French study found an incidence of HCC of 2.9% in alcoholic cirrhosis, high enough to warrant surveillance.[22] The discrepancy between these 2 studies probably relates to the differences in populations and differences in how alcoholic cirrhosis was diagnosed, by biopsy in France, and from hospital records (biopsy not specified) in Denmark.

The effect of aetiology of cirrhosis was evaluated by Sharma *et al,.*[23] The incidence was highest in patients with cirrhosis due to viral hepatitis and lowest in those with autoimmune disease. These factors were included in their risk score. This is one of the few risk scores that takes aetiology of liver disease into account.

There is also a risk score for HCC in patients with diabetes.[24] The score includes age, gender, smoking, variation in haemoglobin, serum glutamic-pyruvic transaminase, liver cirrhosis, hepatitis B, hepatitis C, antidiabetes medications, and antihyperlipidemia medications and total/high-density lipoprotein cholesterol ratio. This is a complex score and requires additional validation.

HCC is one of the most common causes of cancer death in Taiwan and other places in Asia. There is, therefore, interest in screening whole populations. One approach that has been considered is to initially define risk using blood tests, before performing surveillance with ultrasound for those found to be at sufficiently high risk.[25] A different approach was taken by Hung *et al.,*[26] who assessed the HCC risk in the general population, showing the importance of elevated aminotransferases as risk factors. Adding hepatitis B or hepatitis C to the model improved predictability. This approach may be feasible in countries where the HCC incidence is high, but in Western countries this is not likely to be cost effective.

The study by Hung *et al.* also looked at how to assess what level of risk made surveillance worthwhile, using a technique called decision curve analysis. Essentially, this is a modelling study using real life data, in which the potential benefits and harms of surveillance are quantitated, usually in terms of quality-adjusted life years. A level of "net benefit" (benefit minus harm) is chosen based on societal and/or cost considerations. The threshold level of risk for which surveillance provides a net benefit is then calculated. In this particular study they showed that there was a category of patients, particularly younger patients, who would be excluded from surveillance by AASLD criteria but would nonetheless benefit from surveillance.

Others have attempted to define HCC risk by liver stiffness measurements[27,28] and by incorporating these measurements into models combining other variables.

There are now many models purporting to identify those at risk of HCC. The mere fact that there are so many indicates that none are ideal. All the studies described above are either retrospective collections of data or *post hoc* analyses of cohorts collected prospectively for other reasons. Furthermore, the size of HCC at diagnosis is seldom specified. This is important because an algorithm that identifies someone whose cancer is diagnosed late may be different from one that identifies someone whose cancer is diagnosed early. Furthermore, even though the various scores may clearly separate out risk by category, with a few exceptions, they do not indicate whether a risk category that is not the highest risk still warrants surveillance.

How to improve risk assessment

What is needed is a prospective, large-scale multinational study with stratification on the basis of disease aetiology, liver disease severity, and country. The target population should include patients with stage 3 or 4 liver fibrosis. The risk scores should be calculated at baseline and at intervals during follow-up. The study would need to be several years in duration, allowing enough time for a sufficient number of HCC events to occur in each of the different strata to enable analysis. However, only cost-efficacy or decision curve analysis can determine what level of risk warrants surveillance. Such a study is unlikely in an era when every expert seems to want to develop their own model and such multinational cooperation in liver disease research is uncommon.

One of the unknowns is the level of risk at which surveillance should begin. The AASLD

guidelines somewhat arbitrarily defined the threshold of 1.5–2%/year as the incidence above which surveillance is necessary in patients with non-hepatitis B cirrhosis, and 0.2%/year in those with hepatitis B.[3] However, these numbers were based on cost-efficacy analyses and represented the threshold at which surveillance became cost effective. The analyses on which these conclusions were based are old and should be repeated using modern thinking about the natural history of HCC, the efficacy of surveillance tests and the response to treatment – they should also be disease specific. It is likely that the threshold would be different for different diseases. For example, resection is more likely to be possible in patients with hepatitis B because liver function is less likely to be impaired than in patients with hepatitis C. In patients with diabetes and/or non-alcoholic fatty liver disease, competing causes of death from cardiovascular events will likely alter the cost-efficacy ratios.

In the future, molecular markers might improve prediction. Studies of gene expression in the non-tumorous liver of patients with HCC have identified several 'signatures' that are associated with the development of HCC.[29] As yet, there are no studies that incorporate these factors into risk assessment, and these markers require liver tissue, which is a drawback in clinical practice. Attempts have been made to identify markers from circulating tissue (liquid biopsy), but at present this is still in the early stages of development.

Given that no score can, at present, be universally applied and that the vast majority of risk scores merely identify risk along a scale rather than establishing a cut-off below which surveillance is unnecessary, risk scores are interesting but of little use. In this era of "personalised medicine" is it possible that risk scores could be used to "personalise" surveillance programmes? At present there are too many unknowns to even suggest how this could be achieved, and second, the concept of personalised medicine is contrary to one of the principles of surveillance, *i.e.*, that it should be simple to implement and uniform in design.

Risk scores also need to be accurate with few false-positives and few false-negatives. Experimentally, the accuracy of prediction is assessed using the AUROC or the c-statistic. No model will provide 100% accuracy, but there is no consensus as to how high the AUROC or c-statistic needs to be. Usually anything below 0.8 is considered inadequate but, given the poor outcomes associated with the late diagnosis of HCC, we should perhaps aim higher.

Surveillance tests

Recommended surveillance tests include ultrasound and biomarkers, alpha-fetoprotein (AFP), des-gamma-carboxy prothrombin (DCP, also known as PIVKA-II), and the L3 fraction of AFP (AFP-L3). There is clear evidence that addition of biomarkers to ultrasound increases the sensitivity of the surveillance protocol.[30] However, it has never been shown that this is necessary. It is commonly assumed that if one surveillance test has a higher sensitivity than another it must be a better test. However, this is mostly a factor of tumour size. The larger the lesion the more likely it is to be identified by biomarkers or on ultrasound. However, this does not necessarily translate into improved mortality because mortality might be determined by factors other than tumour size, such as invasiveness or ability to metastasise. Furthermore, the most effective treatment may still be available even with the less sensitive test. Translating this into HCC surveillance means that it may not be necessary to find the smallest detectable lesion, and that cure rates may be similar if the lesion is detected at 1.5 cm rather than 0.5 cm, for example.

Some small HCCs will be detected by biomarkers while not visible on ultrasound. However, if these biomarkers were not used, a small lesion would grow and become detectable on ultrasound at a later point, assuming surveillance was conducted on schedule, at which point it would be treated. Whether the delay in diagnosis would affect the likelihood of cure is unknown. However, as long as the lesion is smaller than about 2.5 cm at diagnosis cure remains highly likely.[31]

The biomarkers currently in use are somewhat non-specific, being elevated in patients with active hepatitis and in some patients with cirrhosis. Only a minority of small HCCs will secrete AFP or other biomarkers. Even in larger HCCs only 40–60% will secrete AFP. Finally, the commonly used biomarkers, AFP, AFP-L3 and DCP are all also markers of advanced HCC. This may be because it requires a certain tumour size to produce enough biomarker to be detectable, but it may also be because they are markers of aggressive cancers, that are less likely to be cured. Thus, adding biomarkers to ultrasound surveillance may increase the sensitivity of HCC detection without improving the cure rate.

There is a single randomised controlled study comparing surveillance with ultrasound alone to ultrasound plus the panel of biomarkers listed earlier. The study is still underway and should demonstrate whether or not it is necessary to use biomarkers if ultrasound is also being used. This study is being undertaken at a centre of excellence in ultrasound and the results might be different in centres with less experience.

A number of ultrasound-free surveillance models have been constructed, in an attempt to overcome the problem of false-positive results on ultrasound in patients with nodular cirrhotic livers and false-negative results in obese people with fatty livers.[32,33] These algorithms have varying

degrees of accuracy, with AUROCs of about 0.7–0.8 or even 0.9. However, the characteristics of the lesion that the model purports to find are frequently not provided, so we do not know whether these were potentially curable lesions or not.

Some have suggested that surveillance should be with CT scan or MRI, particularly because of the difficulty of obtaining quality ultrasound in cirrhosis and in the obese and in NAFLD. However, these modalities are expensive and not widely available and cannot be recommended for routine use. A recent cost-efficacy analysis from Korea found MRI to be cost-effective[34], but whether this holds true in the US and Europe where MRI is more expensive than elsewhere is debatable. Even if cost-effective there aren't enough MRI's available to provide surveillance for all who need it.

Others have stressed that static measurements (AFP at a single time point) are less informative than sequential measurements, for example an AFP value that is higher than the measurement 6 months prior is a significant indicator of HCC.[30,35] One validated model incorporates age, ALT and platelet count.[32,36] This was a good predictor of HCC at 6 months, but it had an even better AUROC when serial measurements of AFP were used.[36] This model takes advantage of the fact that patients with hepatitis C or hepatitis B who have achieved either a sustained virologic response or adequate viral suppression have normalised ALT and under these circumstances are less likely to a have a false-positive elevation in AFP. Many studies on biomarkers for surveillance have looked at homogeneous populations, such as hepatitis B in Asia, hepatitis C in North America and hepatitis B in Greece. Whether the result obtained in these homogeneous populations will hold true in other, more heterogeneous populations has yet to be determined.

How to improve surveillance

If only 2% of candidates for surveillance get guideline appropriate surveillance, then clearly the biggest problem is not the performance of the surveillance test. Even if the number were 20–30% this would still be the biggest barrier to effective surveillance. Therefore, the first step is to improve the awareness of primary care providers through better education. How this should be done is beyond the scope of this article. Of course, surveillance tests can be improved, but each new test or combination of tests must be assessed in a prospective manner. It is not possible to perform controlled trials against no surveillance, no matter how desirable this may be, but prospective collection of data in well characterised cohorts can be used to compare different surveillance tests in the same cohort. If one method is demonstrated to be superior to another it is also likely to be superior to no surveillance. One difficulty is in choosing an endpoint. Sensitivity of detection is not adequate as an endpoint for the reasons given. Tumour size at the time of treatment (not diagnosis) may be a surrogate of cure. Hard outcomes such as mortality are impractical, because if surveillance is successful the majority of cancers will be cured in either arm. Other classical endpoints in oncology such as progression-free survival (although a very indirect measure) may also be useful.

Recall procedures

The thresholds that are considered abnormal for biomarkers have not been rigorously defined. Different investigators have used different cut-offs that would trigger further investigation. For example, an AFP >20 ng/L used to be an accepted cut-off, but more recent studies suggest that an increase of AFP over baseline is a better indication, although the magnitude of increase that should trigger cross-sectional imaging has not been well-defined.[30,35] If surveillance was with biomarkers alone, the next step should be an ultrasound rather than immediately jumping to CT or MRI. CT or MRI may show vascular lesions that may not be HCC, such as transient hyperattenuation lesions or intrahepatic shunts. An ultrasound is necessary to show that there is a nodule to be further investigated. Without a nodule, further imaging may not be helpful. However, this does require a certain level of ultrasonographer skill. In the USA it is not uncommon to go straight from an abnormal biomarker measurement to cross-sectional imaging. However, this raises the possibility of false-positive results and over-diagnosis.

If surveillance is performed with ultrasound, a new nodule not seen in prior ultrasounds should be a warning that an HCC might be present. Similarly, any nodule that has grown significantly is likely to be HCC. However, given the vagaries of measurement on an ultrasound image, a minor increase of 1 mm or so does not necessarily count as growth. Even if this does turn out to be HCC, a lesion that increases by no more than 1 mm over the surveillance interval (usually 6 months) is not going to be incurable 6 months later.

Cross-sectional imaging is required to complete the recall procedure after a nodule has been demonstrated on ultrasound. AASLD guidelines from 2011 include an algorithm that has been validated. This has not been improved on since. Perhaps the only change that might improve the recall would be to jump to cross-sectional imaging for lesions smaller than 1 cm, but this cannot be recommended at this stage due to a lack of evidence that this is effective in either identifying more HCCs, or at identifying HCCs at a more curable stage than the current algorithm.

Even cross-sectional imaging and biopsy have their drawbacks. In order to improve the reporting of imaging of liver nodules, the American Radiological Society devised a set of reporting guidelines that

classified the risk of a lesion being HCC from Li-RADS® (Liver Imaging Data and Reporting System) 1 (benign) to Li-RADS 5 (definitely HCC).[37] Li-RADS 2 is reported as probably benign, Li-RADS 3 as indeterminate, and Li-RADS 4 as probably HCC. However subsequent studies have shown that, with the exception of Li-RADS 1, HCC occurs with appreciable frequency in patients classified in the other categories. Therefore, this recall aspect of the surveillance programme needs to be further refined or, at this stage, ignored in favour of the AASLD 2011 algorithm.

Finally, the interpretation of liver biopsy must be improved. Since the objective of surveillance is to find early stage disease and since the earliest stages of HCC more closely resemble normal liver tissue than typical later stage HCC, diagnosis is difficult even with a biopsy. Criteria for the diagnosis of the earliest stage of HCC have been published and expounded upon by experts.[38] However, the extent to which these features are recognised in practice is not clear. Recognition of the so-called very early HCC requires expert interpretation, which is often limited to academic centres with a special interest in HCC. Therefore, algorithms must be developed that minimise the possibility of a false-negative biopsy report. These will likely include additional biopsies and/or changes to follow-up that include more frequent imaging rather than simply returning to 6-month surveillance. These have yet to be developed.

Conclusion

In summary, although surveillance strategies from identification of risk to diagnosis of HCC have been developed, they have not been widely implemented. Risk scores still require additional validation and are probably going to be population specific. Surveillance tools also require refinement. Currently, surveillance programmes pick up more than 80% of HCCs <2 cm when conducted in expert centres, but the diagnosis rate is much lower in non-expert centres. One or more non-invasive tests are needed that can detect small HCCs in addition to being able to detect a higher proportion of all HCCs than current biomarkers, by reflecting a more universal abnormality in HCC, *e.g.* DNA methylation patterns.[39] Finally, the recommendations that radiologists make need not assess the likelihood of HCC, but should suggest appropriate next steps according to the appearance of lesions; except for obviously benign lesions (haemangioma, cysts), further follow-up should always be recommend, at the very least, until the benign nature of the lesion is assured.

Abbreviations

AFP, alpha-fetoprotein; ALT, alanine aminotransferase; AUROC, area under the receiver operating curve; DCP, des-gamma-carboxy prothrombin; HCC, hepatocellular carcinoma; Li-RADS, Liver Imaging Data and Reporting System.

Conflict of interest

The authors declare no conflicts of interest that pertain to this work.

Please refer to the accompanying ICMJE disclosure forms for further details.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhepr.2019.10.007.

References

Author names in bold designate shared co-first authorship

[1] Singal AG, Tiro J, Li X, Adams-Huet B, Chubak J. Hepatocellular Carcinoma Surveillance among Patients with Cirrhosis in a Population-Based Integrated Healthcare Delivery System. J Clin Gastroenterol 2017;51:650–655.

[2] Walker M, El-Serag HB, Sada Y, Mittal S, Ying J, Duan Z, et al. Cirrhosis is under-recognised in patients subsequently diagnosed with hepatocellular cancer. Aliment Pharmacol Ther 2016;43:621–630.

[3] **Bruix J**, **Sherman M**, Management of hepatocellular carcinoma. Practice Guidelines Committee, American Association for the Study of Liver Diseases. Hepatology 2005;42:1208–1236.

[4] Yuen MF, Tanaka Y, Fong DY, Fung J, Wong DK, Yuen JC, et al. Independent risk factors and predictive score for the development of hepatocellular carcinoma in chronic hepatitis B. J Hepatol 2009;50:80–88.

[5] Wong VWS, Chan SL, Mo F, Chan TC, Loong HHF, Wong GLH, et al. Clinical scoring system to predict hepatocellular carcinoma in chronic hepatitis B carriers. J Clin Oncol 2010;28:1660–1665.

[6] Wong GL, Chan HL, Wong CK, Leung C, Chan CY, Ho PP, et al. Liver stiffness-based optimization of hepatocellular carcinoma risk score in patients with chronic hepatitis B. J Hepatol 2014;60:339–345.

[7] Chen CJ, Yang HI, Su J, Jen CL, You SL, Lu SN, et al. Risk of hepatocellular carcinoma across a biological gradient of serum hepatitis B virus DNA level. JAMA 2006;295:65–73.

[8] Yang HI, Sherman M, Su J, Chen PJ, Liaw YF, Iloeje UH, et al. Nomograms for risk of hepatocellular carcinoma in patients with chronic hepatitis B virus infection. J Clin Oncol 2010;28:2437–2444.

[9] Yang HI, Tseng TC, Liu J, Lee MH, Liu CJ, Su TH, et al. Incorporating Serum Level of Hepatitis B Surface Antigen or Omitting Level of Hepatitis B Virus DNA Does not Affect Calculation of Risk for Hepatocellular Carcinoma in Patients Without Cirrhosis. Clin Gastroenterol Hepatol 2016;14:461–468.

[10] Yang HI, Yuen MF, Chan HL, Han KH, Chen PJ, Kim DY, et al. Risk estimation for hepatocellular carcinoma in chronic hepatitis B (REACH-B): development and validation of a predictive score. Lancet Oncol 2011;12:568–574.

[11] Wong GL, Chan HL, Chan HY, Tse PC, Tse YK, Mak CW, et al. Accuracy of risk scores for patients with chronic hepatitis B receiving entecavir treatment. Gastroenterology 2013;144:933–944.

[12] Arends P, Sonneveld MJ, Zoutendijk R, Carey I, Brown A, Fasano M, et al. Entecavir treatment does not eliminate the risk of hepatocellular carcinoma in chronic hepatitis B: limited role for risk scores in Caucasians. Gut 2015;64:1289–1295.

[13] Abu-Amara M, Cerocchi O, Malhi G, Sharma S, Yim C, Shah H, et al. The applicability of hepatocellular carcinoma risk prediction scores in a North American patient population with chronic hepatitis B infection. Gut 2016;65:1347–1358.

[14] Papatheodoridis G, Dalekos G, Sypsa V, Yurdaydin C, Buti M, Goulis J, et al. PAGE-B predicts the risk of developing hepatocellular carcinoma in Caucasians with chronic hepatitis B on 5-year antiviral therapy. J Hepatol 2016;64:800–806.

[15] Brouwer WP, van der Meer AJP, Boonstra A, Plompen EPC, Pas SD, de Knegt RJ, et al. Prediction of long-term clinical outcome in a diverse chronic hepatitis B population: Role of the PAGE-B score. J Viral Hepat 2017;24:1023–1031.

[16] Lee HW, Kim SU, Park JY, Kim DY, Ahn SH, Han KH, et al. External validation of the modified PAGE-B score in Asian chronic hepatitis B patients receiving antiviral therapy. Liver Int 2019 May 3, https://doi.org/10.1111/liv.14129 [Epub ahead of print].
[17] Lok AS, Seeff LB, Morgan TR, di Bisceglie AM, Sterling RK, Curto TM, et al. Incidence of hepatocellular carcinoma and associated risk factors in hepatitis C-related advanced liver disease. Gastroenterology 2009;136:138–148.
[18] Ganne-Carrié N, Layese R, Bourcier V, Cagnot C, Marcellin P, Guyader D, et al. Nomogram for individualized prediction of hepatocellular carcinoma occurrence in hepatitis C virus cirrhosis (ANRS CO12 CirVir). Hepatology 2016;64:1136–4117.
[19] Chang KC, Wu YY, Hung CH, Lu SN, Lee CM, Chiu KW, et al. Clinical-guide risk prediction of hepatocellular carcinoma development in chronic hepatitis C patients after interferon-based therapy. Br J Cancer 2013;109:2481–2488.
[20] **Flemming JA**, Yang JD, Vittinghoff E, **Kim WR**, **Terrault NA**. Risk prediction of hepatocellular carcinoma in patients with cirrhosis: the ADRESS-HCC risk model. Cancer 2014;120:3485–3493.
[21] Jepsen P, Ott P, Andersen PK, Sørensen HT, Vilstrup H. Risk for hepatocellular carcinoma in patients with alcoholic cirrhosis: a Danish nationwide cohort study. Ann Intern Med 2012;156:841–847.
[22] Ganne-Carrie N, Chaffaut C, Bourcier V, Archambeaud I, Pernaud JM, Oberti F, et al. for CIRRAL Group. Estimate of heaptcellular caricnoma incidence in patients with lcoholic cirrhosis. J Hepatol 2018 Dec;69(6):1274–1283.
[23] Sharma SA, Kowigier M, Hansen B, Brouwer WP, Wong D, Shah H, et al. Toronto HCC risk index: A validated scoring system to predict 10-year risk of HCC in patients with cirrhosis. J Hepatol 2018;68:92–99.
[24] Li TC, Li CI, Liu CS, Lin WY, Lin CH, Yang SY, et al. Risk score system for the prediction of hepatocellular carcinoma in patients with type 2 diabetes: Taiwan Diabetes Study. Semin Oncol 2018;45:264–274.
[25] Wen CP, Lin J, Yang YC, Tsai MK, Tsao CK, Etzel C, et al. Hepatocellular carcinoma risk prediction model for the general population: the predictive power of transaminases. J Natl Cancer Inst 2012;104:1599–1611.
[26] Hung YC, Lin CL, Liu CJ, Hung H, Lin SM, Lee SD, et al. Development of risk scoring system for stratifying population for hepatocellular carcinoma screening. Hepatology 2015;61:1934–1944.
[27] Jung KS, Kim SU, Ahn SH, Park YN, Kim DY, Park JY, et al. Risk assessment of hepatitis B virus-related hepatocellular carcinoma development using liver stiffness measurement (FibroScan). Hepatology 2011;53:885–894.
[28] Masuzaki R, Tateishi R, Yoshida H, Goto E, Sato T, Ohki T, et al. Prospective risk assessment for hepatocellular carcinoma development in patients with chronic hepatitis C by transient elastography. Hepatology 2009;49:1954–1961.
[29] Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, Camargo A, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. N Engl J Med 2008;359:1995–2004.
[30] Lee E, Edward S, Singal AG, Lavieri MS, Volk M. Improving screening for hepatocellular carcinoma by incorporating data on levels of α-fetoprotein, over time. Clin Gastroenterol Hepatol 2013;11:437–440.
[31] Sala M, Llovet JM, Vilana R, Bianchi L, Solé M, Ayuso C, et al. Initial response to percutaneous ablation predicts survival in patients with hepatocellular carcinoma. Hepatology 2004;40:1352–1360.
[32] El-Serag HB, Kanwal F, Davila JA, Kramer J, Richardson P. A new laboratory-based algorithm to predict development of hepatocellular carcinoma in patients with hepatitis C and cirrhosis. Gastroenterology 2014;146:1249–1255.
[33] Wang M, Devarajan K, Singal AG, Marrero JA, Dai J, Feng Z, et al. The Doylestown Algorithm: A Test to Improve the Performance of AFP in the Detection of Hepatocellular Carcinoma. Cancer Prev Res (Phila) 2016;9:172–179.
[34] Wang M, Devarajan K, Singal AG, Marrero JA, Dai J, Feng Z, et al. The Doylestowmn Algorithm: A Test to Improve the Performance of AFP in the Detection of Hepatocellular Carcinoma. Cancer Prev Pes (Phila) 2016;(9):172–179.
[35] Tayob N, Lok AS, Do KA, Feng Z. Improved Detection of Hepatocellular Carcinoma by Using a Longitudinal Alpha-Fetoprotein Screening Algorithm. Clin Gastroenterol Hepatol 2016;14:469–475.
[36] White DL, Richardson P, Tayoub N, Davila JA, Kanwal F, El-Serag HB. The Updated Model: An Adjusted Serum Alpha-Fetoprotein-Based Algorithm for Hepatocellular Carcinoma Detection With Hepatitis C Virus-Related Cirrhosis. Gastroenterology 2015;149:1986–1987.
[37] Mitchell DG, Bruix J, Sherman M, Sirlin CB. LI-RADS (Liver Imaging Reporting and Data System): summary, discussion, and consensus of the LI-RADS Management Working Group and future directions. Hepatology 2015;61:1056–1065.
[38] Pathologic diagnosis of early hepatocellular carcinoma: a report of the international consensus group for hepatocellular neoplasia. The International Consensus Group for Hepatocellular NeoplasiaHepatology 2009;49:658–664.
[39] Wu HC, Yang HI, Wang Q, Chen CJ, Santella RM. Plasma DNA methylation marker and hepatocellular carcinoma risk prediction model for the general population. Carcinogenesis 2017;38:1021–1028.