Please use this PDF proof to check the layout of your article. If you would like any changes to be made to the layout, you can leave instructions in the online proofing interface. First, return to the online proofing interface by clicking "Edit" at the top page, then insert a Comment in the relevant location. Making your changes directly in the online proofing interface is the quickest, easiest way to correct and submit your proof.

Please note that changes made to the article in the online proofing interface will be added to the article before publication, but are not reflected in this PDF proof.

Editorial

JOURNAL OF HEPATOLOGY

In search of an evidence base for HCC surveillance: purity or pragmatism?

Philip J. Johnson*

Department of Molecular and Clinical Cancer Medicine, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, United Kingdom

For decades, clinicians faced by a patient with advanced hepatocellular carcinoma (HCC) have felt helpless. The cancer had developed 'silently', causing no symptoms until it had spread widely or overwhelmed liver function and was no longer amenable to potentially curative therapy. Despite the recent 'revolution' in systemic therapies, leading investigators opined in 2019 that 'although [*new systemic agents*] have the potential to improve outcomes, a survival increase of 2–5 months remains poor'.[1] To clinicians, the obvious answer was to detect the cancer at a treatable stage by undertaking surveillance of patients perceived to be at highest risk, namely those with chronic liver disease at the stage of cirrhosis.

Although the gold standard for evidence is widely agreed to be the prospective randomised clinical trial (RCT),[2] many aspects of HCC management such as resection or liver transplantation have been adopted into practice without an RCT because they are perceived to be of undoubted benefit. RCTs are reserved for interventions such as systemic therapy where there is doubt. But the definition of 'doubt' is not always clear-cut. A great furore surrounded the usually authoritative Cochrane Review that gave a downbeat assessment of the importance of direct-acting antiviral-induced sustained virologic response (SVR) in patients with chronic HCV infection, largely because the necessary RCTs had not been undertaken.[3] The clinical community were already convinced of the benefit and vehemently rejected the Cochrane Review.[4] One authority sees doubt where another does not and HCC surveillance is another example thereof. Thus, all practice guidelines[5] call for surveillance (6 monthly ultrasound examination with or without the biomarker alpha-fetoprotein) implying that, at least in the eyes of the clinical community, an RCT is not required before it is recommended or, in the case of Japan, nationally implemented. However, in this issue of the *Journal*, Jensen and West,[6] like the US National Cancer Institute ('based on fair evidence, screening of persons at elevated risk does not result in a decrease in mortality') make a cogent case that there *is* serious doubt and suggest that surveillance should only be implemented if there is evidence of benefit based on an RCT. This editorial concurs with the sentiment but highlights the multiple practical obstacles that would surround any such trial and considers alternative more pragmatic strategies.

The very fact that clinical guidelines all recommend surveillance, immediately poses an obstacle to an RCT. Can clinicians (upon whose opinion most patients rely) really recommend to their patients, contrary to their national guidelines, entry into a controlled trial, in which they might not receive surveillance? In Japan, clinicians have lost court cases for having failed to effectively implement surveillance.[7] Furthermore, whilst RCTs are widely acknowledged to be one of the great methodological advances of the last century, their limitations are increasingly recognised.

In the world of 'precision medicine' apparently homogenous cancers are being extensively sub-categorised, usually on the basis of molecular characterisation. Nowadays, 'the gold standard of randomised confirmatory phase III trials is not always ethical or feasible when developing drugs for treatment of small cancer populations'.[8] Each molecular variant may require its own trial. Even at a clinical level the heterogeneity of HCC will result in multiple smaller sub-groups who may need their own evidence base. Those who remain at risk of HCC after achievement of SVR come to mind.

Such limitations are compounded by the fact that RCTs report a 'central tendency' which is difficult to translate into advice to the individual patient and those within trials may be unrepresentative of the population in whom the trial results will be implemented. A single trial is unlikely to be generalizable to different regions. All these limitations would impact on any individual's decision or their physician's recommendation as to whether to enter a surveillance program.

Can a methodology (*i.e.* an RCT) that will likely take a decade to move from conception to practice-changing implementation really be fit for purpose in the 21st century? A trial conceived today would emerge into a different world. Chronic viral hepatitis and the insensitive ultrasound may be history, cirrhosis may be reversible and effective systemic therapy may render surgical resection (and early diagnosis) irrelevant. If NAFLD becomes the most common liver disease associated with HCC what would a trial's entry criteria be? Two randomised surveillance trials were run in China in the 1980s.[9] In the West they were dismissed on the basis of methodological shortcomings and lack of relevance to the West. They were of course simply 'of their time and place' and one suspects that the same fate would befall any trial started today for the reasons listed above.

Furthermore, the 'goal posts' will likely have moved from aiming to decrease *disease* specific mortality to *overall*

Received 17 February 2021; accepted 17 February 2021; available online xxx
DOI of original article: https://doi.org/10.1016/j.jhep.2020.12.029.
* Corresponding author. Address: Department of Molecular and Clinical Cancer Medicine, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, 2nd floor Sherrington Building, Ashton Street, Liverpool L69 3GE, United Kingdom; Tel.: +44 7850 3737 44.
E-mail address: Philip.Johnson@liverpool.ac.uk (P.J. Johnson).
https://doi.org/10.1016/j.jhep.2021.02.018

mortality.[10] Over the last 10 years mortality in a surveillance population is attributable to HCC in about a quarter. Yes, resection is *potentially* curative but that cure (however defined) only occurs in around 20% of cases[11] so there would have to be a dramatic improvement in survival consequent on surveillance to demonstrate any improvement in overall mortality.

In Japan, a government-funded and well-designed surveillance program which was started in the 1980s, without any preparatory trial, probably represents the upper limit of what can be expected from current surveillance strategies.[12] Median survival has increased, decade on decade, from less than 4 months (as it was then in other countries) to more than 4 years. Yes, lead time bias and new treatments may account for some of this improvement but surely, not all.[13] A, or perhaps *the*, key contributor to the success of the program (in stark contrast to the situation in the West) has been an enthusiastic 'buy-in' from the population. The extent of 'buy-in' may be influenced more by culture than science, an observation consistent with the vaccination program for COVID-19; well-run RCTs that clearly demonstrate the benefit of an intervention do not necessarily convince the target population.

As in many areas of screening, interest is currently focussed on refining 'risk stratification' such that some patients under surveillance, according to current guidelines, may be exempted on the basis of very low risk, whereas high-risk groups may have surveillance intensified using more sensitive approaches.[14] Resources saved from the former will allow roll out of the latter, although the theoretical underpinning of such approaches deserves careful attention.[15]

The major refinements to risk assessment have been developed for patients with HBV-related HCC who represent, we should remember, the majority of cases worldwide. Early examples included CU-HCC[16] and REACH-B.[17] These combined clinical features and HBV-DNA levels, gave clear risk estimates and were extensively validated.[18] However, HBV-DNA loses discriminatory function after treatment with nucleoside analogues and may be replaced by quantitative HBsAg and hepatitis B core-related antigen.[19] These original HBV-related surveillance models had a real link with the biology of the disease they were being applied in. However, the latest iterations of risk stratification scores are not restricted by aetiology. These are statistical models that comprise a common backbone of age and gender (which presumably identifies a population at risk)[20] to which are added various clinical parameters as exemplified by 'PAGE-B'[21] [adding platelet count] and aMAP [adding PALBI - platelets, albumin and bilirubin].[22] The latter used a wealth of data collected from multiple sources and was extensively validated. Subsets which are at minimal risk of HCC and others (a much smaller fraction) at high risk, are readily identifiable.

As such they will give traditional translational researchers considerable intellectual indigestion. Whilst in other cancers precision medicine is being introduced to risk stratification and screening,[14] none of the above scores were developed on the basis of any understanding of, or research upon, the biology of HCC or cirrhosis and all variables in the models have been available for at least 50 years. There is still much scope for mechanistically based and biologically plausible approaches, but perhaps these will be enhanced by a second wave of translational research based on the vast quantities of newly accessible data and new analytical (biostatistical and artificial intelligence) approaches. These will be technologically rather than scientifically driven, with the aim of answering very specific questions rather than testing hypotheses.

Many of the issues facing surveillance for HCC are but a reflection of the issues facing Health Services at the beginning of the 21st century, brought into sharper focus by the COVID-19 pandemic. What *level* of evidence is required such that clinicians and the public are comfortable to change practice, who decides where 'doubt' and 'certainty' lie? At present, the RCT remains the gold standard. However, as already detailed, several factors conspire to challenge this position. The ethos of evidence-based medicine (EBM) is being questioned with suggestions that it is 'in crisis' and should be 'refocused on providing useable evidence that can be combined with context and professional expertise so that individual patients get optimal treatment'.[23] The colossal amounts of data that are now available for the first time, in conjunction with the requisite methodology and the processing power, are likely to change the balance between RCTs and more pragmatic approaches.[24]

The author remains to be convinced that any single RCT could provide large scale generalizable practice-changing results. A more practical approach may be to develop consensus around which patients might be exempted from current surveillance strategies and accept this approach's limitations in exchange for quicker, and more readily available, answers.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2021.02.018.

References

[1] Yarchoan M, Agarwal P, Villanueva A, Rao S, Dawson LA, Llovet JM, et al. Recent developments and therapeutic strategies against hepatocellular carcinoma. Canc Res 2019;79(17):4326–4330.
[2] Hariton E, Locascio JJ. Randomised controlled trials–the gold standard for effectiveness research: study design: randomised controlled trials. BJOG 2018;125(13):1716.
[3] Jakobsen JC, Nielsen EE, Feinberg J, Katakam KK, Fobian K, Hauser G, et al. Direct-acting antivirals for chronic hepatitis C. Cochrane Database Syst Rev 2017;9:CD012143.
[4] Powderly WG, Naggie S, Kim AY, Vargas HE, Chung RT, Lok AS. IDSA/AASLD response to Cochrane Review on direct-acting antivirals for hepatitis C. Clin Infect Dis 2017;65(11):1773–1775.
[5] Frenette CT, Isaacson AJ, Bargellini I, Saab S, Singal AG. A practical guideline for hepatocellular carcinoma screening in patients at risk. Mayo Clin Proc Innov Qual Outcomes 2019;3(3):302–310.
[6] Jepsen P, West J. We need stronger evidence for (or against) hepatocellular carcinoma surveillance. J Hepatol 2021.
[7] Kudo M. Japan's successful model of nationwide hepatocellular carcinoma surveillance highlighting the urgent need for global surveillance. Liver Cancer 2012;1(3-4):141–143.
[8] Verweij J, Hendriks HR, Zwierzina H, Cancer Drug Development F. Innovation in oncology clinical trial design. Canc Treat Rev 2019;74:15–20.
[9] Zhang BH, Yang BH, Tang ZY. Randomized controlled trial of screening for hepatocellular carcinoma. J Canc Res Clin Oncol 2004;130(7):417–422.
[10] Prasad V, Lenzer J, Newman DH. Why cancer screening has never been shown to "save lives"–and what we can do about it. BMJ 2016;352:h6080.
[11] Cucchetti A, Zhong J, Berhane S, Toyoda H, Shi K, Tada T, et al. The chances of hepatic resection curing hepatocellular carcinoma. J Hepatol 2020;72(4):711–717.
[12] Kudo M. Surveillance, diagnosis, treatment, and outcome of liver cancer in Japan. Liver Cancer 2015;4(1):39–50.
[13] Johnson P, Berhane S, Kagebayashi C, Satomura S, Teng M, Fox R, et al. Impact of disease stage and aetiology on survival in hepatocellular carcinoma: implications for surveillance. Br J Canc 2017;116(4):441–447.
[14] Roberts MC. Implementation challenges for risk-stratified screening in the era of precision medicine. JAMA Oncol 2018;4(11):1484–1485.

[15] Duffy SW. Risk stratification in breast screening: a word of caution. J Med Screen 2019;26(2):57–58.
[16] Wong VW, Chan SL, Mo F, Chan TC, Loong HH, Wong GL, et al. Clinical scoring system to predict hepatocellular carcinoma in chronic hepatitis B carriers. J Clin Oncol 2010;28(10):1660–1665.
[17] 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(6):568–574.
[18] Jung KS, Kim SU, Song K, Park JY, Kim DY, Ahn SH, et al. Validation of hepatitis B virus-related hepatocellular carcinoma prediction models in the era of antiviral therapy. Hepatology 2015;62(6):1757–1766.
[19] Liang LY, Wong VW, Toyoda H, Tse YK, Yip TC, Yuen BW, et al. Serum hepatitis B core-related antigen predicts hepatocellular carcinoma in hepatitis B e antigen-negative patients. J Gastroenterol 2020;55(9):899–908.
[20] Johnson PJ, Krasner N, Portmann B, Eddleston AL, Williams R. Hepatocellular carcinoma in Great Britain: influence of age, sex, HBsAg status, and aetiology of underlying cirrhosis. Gut 1978;19(11):1022–1026.
[21] 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(4):800–806.
[22] Fan R, Papatheodoridis G, Sun J, Innes H, Toyoda H, Xie Q, et al. aMAP risk score predicts hepatocellular carcinoma development in patients with chronic hepatitis. J Hepatol 2020.
[23] Greenhalgh T, Howick J, Maskrey N. Evidence Based Medicine Renaissance G. Evidence based medicine: a movement in crisis? BMJ 2014;348:g3725.
[24] Thorlund K, Dron L, Park JJH, Mills EJ. Synthetic and external controls in clinical trials - a primer for researchers. Clin Epidemiol 2020;12:457–467.