Editorial

JOURNAL
OF HEPATOLOGY

Radiological endpoints as surrogates for survival benefit in hepatocellular carcinoma trials: All that glitters is not gold

Giuseppe Cabibbo[1,*], Jordi Bruix[2,*]

DOI of original article: https://doi.org/10.1016/j.jhep.2022.09.006

When evaluating efficacy in oncology trials, overall survival (OS) provides an unequivocal measure of true clinical benefit. OS, defined as the time from random assignment to the date of death from any cause, is a precise, objectively measured, and easy-to-interpret clinical outcome. It is considered the most reliable and clinically meaningful endpoint to evaluate drug efficacy in oncology trials.[1,2] However, the interpretation of OS can be affected by post-progression survival and/or subsequent cancer treatments, potentially confounding survival analysis (treatment crossover or second-line therapies), which has led to the use of several intermediate parameters based on radiology findings that are supposed to be surrogates of survival benefit.[3] These include objective response rate (ORR), progression-free survival (PFS) and time-to-progression, which have been extensively used in oncology trials despite their controversial value in predicting survival benefit, irrespective of cancer type. Indeed, a major rate of response with long term duration has led to accelerated approvals in the field of hepatocellular carcinoma (HCC). However, their relevance remains debated,[4] and aggregate-data meta-analyses showed a modest correlation with OS, with substantial variability according to cancer type, the class of drug(s) administered and line of treatment.[5,6] The most commonly evaluated surrogate endpoints in the field of HCC are generally related to radiological tumour response, and they might be classified as time-independent (*e.g.* ORR) or time-dependent (*e.g.* PFS) endpoints, whether they are able to capture the time to the event or not.

The study by Kudo *et al.*[7] is a retrospective analysis of the REFLECT study[8] suggesting that, at the individual level, a patient treated with sorafenib or lenvatinib whose tumour responds or whose time to disease progression is longer lives longer. Median OS was 21.6 months for responders compared to 11.9 months for non-responders (hazard ratio 0.61; 95% CI 0.49-0.76; p <0.001). OS benefit significantly persisted in responders at 2-month and 6-month landmarks. Interestingly, when using modified RECIST criteria (both for investigator or independent radiologic review) but not RECIST criteria, non-responder patients with stable disease had a median survival closer to responders. One is thus prompted to ask, what is the true driver of survival outcome?

The study by Kudo *et al.*[7] is relevant, particularly considering that, for a long time, tumour "shrinkage" (objective response) was postulated to be clinically relevant and to translate into enhanced survival, but adequate data were absent. Accordingly, the current data represent a major piece of information. However, the study focuses on lenvatinib and sorafenib in a single trial data set. Would it also be the case under other agents or in a separate population at a different evolutionary time point? Furthermore, should the decision to transition from early phase evaluation to phase III trials be based on evaluation of response?

It is important to note that tumour burden measurements by any system were intended simply to describe what happens to tumours during therapy—not to infer a meaningful benefit from those changes. Reduction in tumour burden is a marker of activity, but activity is not to be ruled out in the absence of response as previously exemplified by sorafenib.[9] Moreover, a significantly higher ORR does not always automatically translate into a proportional survival benefit. For example, the ORR of durvalumab plus tremelimumab was 4 times higher than that of sorafenib,[10] reaching a rate of 20% for the combination, but this significant ORR difference translated into a median OS difference of only 3 months between the two treatments. The same discrepancy was observed in the trial comparing the efficacy of lenvatinib *vs.* lenvatinib plus pembrolizumab. ORR increased from 17.5 to 26.1 but OS was not significantly different.[11]

In the context of sequential treatments, time-to-event outcomes such as PFS are proposed to have great clinical utility owing to the fact that PFS better captures timing of transition to second-line

* Corresponding authors. Addresses: Section of Gastroenterology and Hepatology, Department of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties PROMISE, University of Palermo, Piazza delle Cliniche n 2, 90127, Palermo, Italy (G. Cabibbo), or BCLC Group, Liver Unit, Hospital Clinic Barcelona, IDIBAPS, CIBEREHD, University of Barcelona, Villarroel 170, Barcelona, 08036, Spain (J. Bruix).
E-mail addresses: giuseppe.cabibbo@unipa.it (G. Cabibbo), jbruix@clinical.cat (J. Bruix).
Received 1 October 2022; received in revised form 12 October 2022; accepted 16 October 2022; available online xxx
https://doi.org/10.1016/j.jhep.2022.10.018

therapy after progression.[12] Its interpretation cannot be confounded by post-progression survival and treatment crossover but the surrogacy of PFS has not been validated. Also, in this case, it is possible to observe differences between the radiological endpoint and the survival benefit, as for COSMIC-312 trial[13] in which, despite achieving significantly longer median PFS *vs.* sorafenib (6.8 *vs.* 4.2 months), cabozantinib plus atezolizumab failed to improve median OS. Similarly, in the REFLECT trial,[8] the significant advantage of lenvatinib over sorafenib in PFS (7.4 *vs.* 3.7 months) did not translate into an equivalent survival benefit. In the case of sunitinib,[14] on the other hand, the high toxicity led to a significantly reduced survival compared to sorafenib while maintaining the PFS substantially unchanged. Therefore, the importance of validating PFS at an individual level with respect to OS is clear.

Kudo and colleagues[7] explored the surrogate value of PFS and found a moderate to strong individual-level association between PFS and OS, that does not suffice to endorse PFS as a robust endpoint for future trials. Indeed, it is important to consider that the clinical value of PFS could be weakened by the fact that it does not consider the pattern of progression,[15,16] which has been proven to be an important (and extensively validated) survival predictor. Moreover, different classes of drugs could have different impacts on the pattern of progression and consequently on survival. Finally, the so-called pseudo-progression phenomenon, associated with immune checkpoint inhibitors (ICIs), may further reduce the value of PFS, as response may be registered later during follow-up. Therefore, since the evaluation of the benefit associated with ICIs is increasingly complex, traditional radiological response criteria, such as modified RECIST or RECIST, may no longer be appropriate. The accuracy of specifically designed radiological response criteria for ICIs, such as immune-RECIST, remains to be extensively explored in the field of HCC[17] as ICIs are now the backbone of systemic HCC treatment. Moreover, from a methodological point of view, the late pattern of response associated with ICIs, with a proportion of patients obtaining a long-term survival benefit, may violate the assumption of proportional hazards, making median OS and hazard ratio a potentially flawed measure to assess treatment benefit. In this line, milestone analysis or restricted mean survival time could overcome the issue of non-proportional hazard ratios and they should be routinely reported in ICI trials for HCC.

Other fundamental issues affecting evaluation of radiological response to treatment are reliability due to interobserver variability and use of several radiologic criteria across trials. Particularly in the setting of systemic therapies, the advantage of one system over the other is still to be validated and it might be more useful to use the one that requires the fewest measurements.

These comments highlight that caution should be exercised when evaluating benefit in terms of radiology-based endpoints which do not translate into clear OS benefits; this is particularly relevant if there is no adequate information on the withdrawal rates for adverse events (AEs), duration of treatment, safety measures, pattern of progression, and further treatment lines after progression.[4,18]

AEs are usually presented as a frequency at the end of follow-up (as time-independent variables), without describing their occurrence over time. Thus, treatments that improve survival may appear to cause a disproportionate rate of AEs, as data accumulate over a longer follow-up time than for alternative treatments that confer less survival benefit. Moreover, treatment toxicity may overestimate PFS if a substantial proportion of

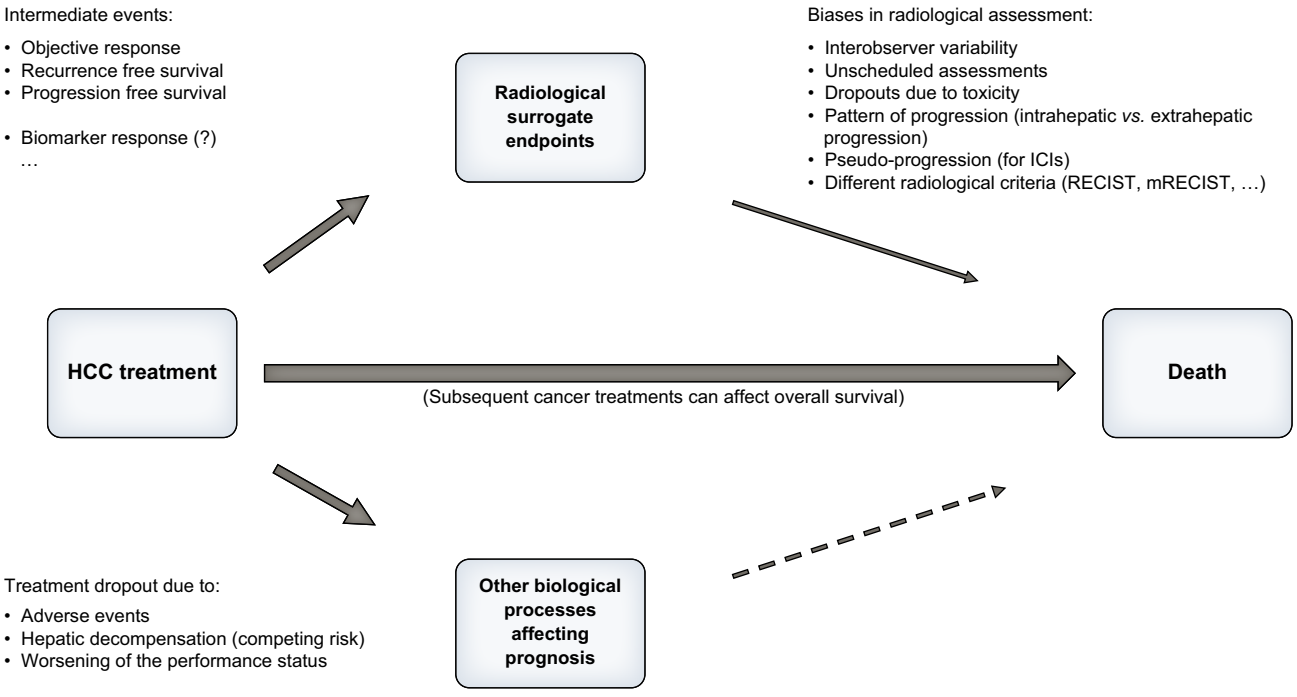

Fig. 1. Pitfalls of surrogate endpoints. HCC, hepatocellular carcinoma; ICIs, immune checkpoint inhibitors; mRECIST, modified RECIST.

patients dropout for AEs before radiological progression could be documented, introducing an informative censoring bias.

All these issues should be considered in future HCC clinical trials. Lack of inclusion of hepatic decompensation as an outcome (which could play a role as a competing risk)[19–21] in trials of HCC superimposed on cirrhosis, could partially explain the wide heterogeneity in OS surrogacy (Fig. 1).

Obviously, an earlier initial response may be desirable. However, analysis should consider both dropout and progression as "events". Kudo and colleagues[7] partially address this area of need, evaluating the fine balance between antitumor drug efficacy and hepatic toxicity, considering death due to hepatic failure as a competing risk. However, hepatic decompensation, type of decompensation during follow-up, and decompensation-free survival, are not clearly assessed. A careful evaluation of the treatment failure (due to severe AEs that also consider hepatic decompensation) would avoid misleading conclusions resulting from uneven dropout rates, and failure to tolerate a drug is as much a sign of ineffective treatment as tumour progression.

In summary, demonstrating associations between the surrogate endpoints (ORR and PFS) and OS at the individual level is a necessary but not sufficient condition to conclude that ORR and PFS can replace OS for the assessment of clinical benefit in oncology trials.[22] While use of intermediate surrogate endpoints provides rapid assessments of experimental interventions, timeliness should not be achieved at the expense of a misleading risk-benefit profile. Moreover, the highest level of validation of surrogacy requires an analysis of correlation in terms of treatment effects between arms based on data from randomised-controlled trials (trial-level association). Also, surrogacy must be validated for all pharmacological approaches and lines of treatment. In the end, reliable blood- or tissue-based biomarkers that can accurately predict clinical benefit (*i.e.* improved survival) are an urgent unmet need. While waiting for such treatment-specific tools, the prediction of benefit in an individual patient or the assessment of activity in early phase trials will remain an educated guess far from precision oncology.

Affiliations

[1]Section of Gastroenterology and Hepatology, Department of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties PROMISE, University of Palermo, Palermo, Italy; [2]BCLC Group, Liver Unit, Hospital Clinic Barcelona, IDIBAPS, CIBEREHD, University of Barcelona, Barcelona, Spain

Financial support

The authors received no financial support to produce this manuscript.

Conflict of interest

Giuseppe Cabibbo has served as a consultant or on advisory boards for Bayer, Eisai, Ipsen, AstraZeneca, Roche. Jordi Bruix has consulted for Arqule, Bayer-Shering Pharma, Novartis, BMS, BTG- Biocompatibles, Eisai, Kowa, Terumo, Gilead, Bio-Alliance, Roche, AbbVie, MSD, Sirtex, Ipsen, Astra-Medimmune, Incyte, Quirem, Adaptimmune, Lilly, Basilea, Nerviano, Sanofi; and received research/educational grants from Bayer, and lecture fees from Bayer-Shering Pharma, BTG-Biocompatibles, Eisai, Terumo, Sirtex, Ipsen.

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

Authors' contributions

Interpretation of data and drafting of the manuscript (all authors); critical revision of the manuscript for important intellectual content (all authors). All authors approve final version of the manuscript.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2022.10.018.

References

Author names in bold designate shared co-first authorship

[1] Ladanie A, Schmitt AM, Speich B, Naudet F, Agarwal A, Pereira TV, et al. Clinical trial evidence supporting US food and drug administration approval of novel cancer therapies between 2000 and 2016. JAMA Netw Open 2020 Nov 2;3(11):e2024406.

[2] Gyawali B, Hey SP, Kesselheim AS. Assessment of the clinical benefit of cancer drugs receiving accelerated approval. JAMA Intern Med 2019 Jul 1;179(7):906–913.

[3] Zhao F. Surrogate end points and their validation in oncology clinical trials. J Clin Oncol 2016 May 1;34(13):1436–1437.

[4] Bruix J. Endpoints in clinical trials for liver cancer and their value in evidence-based clinical decision making: an unresolved Gordian knot. J Hepatol 2021 Jun;74(6):1483–1488.

[5] Cooper K, Tappenden P, Cantrell A, Ennis K. A systematic review of meta-analyses assessing the validity of tumour response endpoints as surrogates for progression-free or overall survival in cancer. Br J Cancer 2020 Nov;123(11):1686–1696.

[6] **Cabibbo G**, **Celsa C**, Enea M, Battaglia S, Rizzo GEM, Busacca A, et al. Progression-free survival early assessment is a robust surrogate endpoint of overall survival in immunotherapy trials of hepatocellular carcinoma. Cancers (Basel) 2020 Dec 30;13(1):90.

[7] Kudo M, Finn RS, Qin S, Han KH, Ikeda K, Cheng AL, et al. Overall survival and objective response in advanced unresectable hepatocellular carcinoma: a subanalysis of the REFLECT study. J Hep 2022. https://doi.org/10.1016/j.jhep.2022.09.006.

[8] Kudo M, Finn RS, Qin S, Han KH, Ikeda K, Piscaglia F, et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. Lancet 2018 Mar 24;391(10126):1163–1173.

[9] Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, et al., Study Group. Sorafenib in advanced hepatocellular carcinoma. N Engl J Med 2008 Jul 24;359(4):378–390.

[10] Abou-Alfa GK, Lau G, Kudo M, Chan SL, Kelley RK, Furuse J, et al. Tremelimumab plus durvalumab in unresectable hepatocellular carcinoma. NEJM Evid 2022;1(8). https://doi.org/10.1056/EVIDoa2100070.

[11] Finn RS, Kudo M, Merle P, Meyer T, Qin S, Ikeda M, et al. Primary results from the phase III LEAP-002 study: lenvatinib plus pembrolizumab versus lenvatinib as first-line (1L) therapy for advanced hepatocellular carcinoma (aHCC). Ann Oncol 2022;33(suppl_7):S808–S869. https://doi.org/10.1016/annonc/annonc1089.

[12] **Cabibbo G**, **Reig M**, Celsa C, Torres F, Battaglia S, Enea M, et al. First-line immune checkpoint inhibitor-based sequential therapies for advanced hepatocellular carcinoma: rationale for future trials. Liver Cancer 2021 Nov 23;11(1):75–84.

[13] Kelley RK, Rimassa L, Cheng AL, Kaseb A, Qin S, Zhu AX, et al. Cabozantinib plus atezolizumab versus sorafenib for advanced hepatocellular carcinoma (COSMIC-312): a multicentre, open-label, randomised, phase 3 trial. Lancet Oncol 2022;23:995–1008.

[14] Cheng AL, Kang YK, Lin DY, Park JW, Kudo M, Qin S, et al. Sunitinib versus sorafenib in advanced hepatocellular cancer: results of a randomized phase III trial. J Clin Oncol 2013 Nov 10;31(32):4067–4075.

[15] Reig M, Rimola J, Torres F, Darnell A, Rodriguez-Lope C, Forner A, et al. Post-progression survival of patients with advanced hepatocellular carcinoma. Rationale for second line trial design. HEPATOLOGY 2013;58:2023–2031.
[16] Bruix J, da Fonseca LG, Reig M. Insights into the success and failure of systemic therapy for hepatocellular carcinoma. Nat Rev Gastroenterol Hepatol 2019 Oct;16(10):617–630.
[17] Rimola J, Da Fonseca LG, Sapena V, Perelló C, Guerrero A, Simó MT, et al. Radiological response to nivolumab in patients with hepatocellular carcinoma: a multicenter analysis of real-life practice. Eur J Radiol 2021 Feb;135:109484.
[18] Tannock IF, Pond GR, Booth CM. Biased evaluation in cancer drug trials-how use of progression-free survival as the primary end point can mislead. JAMA Oncol 2022 May 1;8(5):679–680.
[19] Cabibbo G, Petta S, Barbara M, Attardo S, Bucci L, Farinati F, et al. Hepatic decompensation is the major driver of death in HCV-infected cirrhotic patients with successfully treated early hepatocellular carcinoma. J Hepatol 2017 Jul;67(1):65–71.
[20] Reig M, Cabibbo G. Antiviral therapy in the palliative setting of HCC (BCLC-B and -C). J Hepatol 2021 May;74(5):1225–1233.
[21] Cabibbo G, Aghemo A, Lai Q, Masarone M, Montagnese S, Ponziani FR. Italian Association for the Study of the Liver (AISF). Optimizing systemic therapy for advanced hepatocellular carcinoma: the key role of liver function. Dig Liver Dis 2022 Apr;54(4):452–460.
[22] Bruix J, Reig M, Sangro B. Assessment of treatment efficacy in hepatocellular carcinoma: response rate, delay in progression or none of them. J Hepatol 2017 Jun;66(6):1114–1117.