Review

JOURNAL OF HEPATOLOGY

Randomized trials and endpoints in advanced HCC: Role of PFS as a surrogate of survival

Josep M. Llovet[1,2,3,*], Robert Montal[1], Augusto Villanueva[2,4]

See Editorials, pages; 1060–1061; 1062–1064

Keywords: Trial design; Liver cancer; Systemic therapies; FDA approval.

Received 24 July 2018; received in revised form 21 December 2018; accepted 29 January 2019

Summary

Hepatocellular carcinoma (HCC) is a major cause of cancer-related mortality worldwide. Around half of patients with HCC will receive systemic therapies during their life span. The pivotal positive sorafenib trial and regulatory approval in 2007 was followed by a decade of negative studies with drugs leading to marginal antitumoral efficacy, toxicity, or trials with a lack of enrichment strategies. This trend has changed over the last 2 years with several compounds, such as lenvatinib (in first-line) and regorafenib, cabozantinib, ramucirumab and nivolumab (in second-line), showing clinical benefit. These successes came at a cost of increasing the complexity of decision-making, and ultimately, impacting the design of future clinical trials. Nowadays, life expectancy with single active agents has surpassed the threshold of 1 year and sequential strategies have provided encouraging outcomes. Overall survival (OS) remains the main endpoint in phase III investigations, but as in other solid tumours, there is a clear need to define surrogate endpoints that both reliably recapitulate survival benefits and can be assessed before additional efficacious drugs are administered. A thorough analysis of 21 phase III trials published in advanced HCC demonstrated a moderate correlation between progression-free survival (PFS) or time to progression (TTP) and OS (R = 0.84 and R = 0.83, respectively). Nonetheless, the significant differences in PFS identified in 7 phase III studies only correlated with differences in OS in 3 cases. In these cases, the hazard ratio (HR) for PFS was ≤0.6. Thus, this threshold is herein proposed as a potential surrogate endpoint of OS in advanced HCC. Conversely, PFS with an HR between 0.6–0.7, despite significance, was not associated with better survival, and thus these magnitudes are considered uncertain surrogates. In the current review, we discuss the reasons for positive or negative phase III trials in advanced HCC, and the strengths and limitations of surrogate endpoints (PFS, TTP and objective response rate [ORR]) to predict survival.

[1]*Translational Research in Hepatic Oncology, Liver Unit, IDIBAPS, Hospital Clinic Barcelona, University of Barcelona, Barcelona, Catalonia, Spain*; [2]*Mount Sinai Liver Cancer Program, Division of Liver Diseases, Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA*; [3]*Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Catalonia, Spain*; [4]*Division of Hematology and Medical Oncology, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA*

Introduction

The incidence of hepatocellular carcinoma (HCC) is increasing and will soon surpass one million annual cases worldwide.[1] Up to 80% of patients with HCC have concomitant liver cirrhosis, mainly as result of hepatitis B and C virus infection, alcohol abuse or non-alcoholic steatohepatitis in the context of metabolic syndrome.[2] Coexistence of cancer and cirrhosis in HCC is an essential hallmark that has shaped clinical trial design in HCC, as encapsulated in the Barcelona Clinic Liver Cancer (BCLC) algorithm.[3,4] Only 40% of patients with HCC are diagnosed at early stages, when potentially curative treatments (*i.e.* resection, liver transplantation and local ablation) are applicable.[4] As disease progresses, transarterial chemoembolization[5] (for intermediate HCC) and systemic targeted therapies[6] (for advanced HCC) have shown survival benefits. Since sorafenib was shown to have a positive impact on survival,[7] at least 10 trials have shown negative results in the frontline setting. In the last 2 years, however, numerous systemic agents have demonstrated clinical benefit in the context of phase III trials. The fact that 5 effective drugs and 2 immune checkpoint inhibitors with signs of efficacy have been approved by the Food and Drug Administration (FDA) for the management of advanced HCC poses a challenge in terms of trial design in this arena. Overall survival (OS) is an unquestionable, unbiased primary endpoint in oncology and in all randomized studies testing systemic first- and second-line therapies in advanced HCC.[7–27] However, studies in other solid tumours have identified surrogate endpoints of survival that led to accelerated regulatory approval, notably objective response rate (ORR) and progression-free survival (PFS).[28,29] The aim of these endpoints is to reflect survival benefit, with the advantage that they can be assessed prior to a patient receiving additional efficacious drugs. By thoroughly analysing the clinical experience following sorafenib approval, we have assessed the correlation between surrogate endpoints, such as PFS, time to progression (TTP) and ORR, with clinically meaningful improvements in OS in 21 reported phase III studies.[7–27] Based on this analysis and a conservative approach to define surrogates of OS in HCC, we propose PFS with a threshold of hazard

ratio (HR) ≤0.6 as a reliable surrogate with solid positive predictive value, whereas the threshold of HR = 0.6–0.7, despite leading to positive statistical results, is defined as clinically uncertain in terms of capturing true advantages in OS. In addition, we revisit the correlation between ORR by modified RECIST (mRECIST) and OS. We confirm that ORR is an independent predictor of OS at early, intermediate and advanced stages, meaning that responders survive significantly longer. Nonetheless, ORR is still a suboptimal surrogate marker because of its low sensitivity in capturing those patients that benefit from a given drug. Ultimately, we aim to provide a historical perspective on HCC trial design, focussing on the lessons that can be learned and the ways to maximize clinical trial success in the near future.

Overview of phase III and practice-changing phase II trials reported during the last 10 years

Current estimates suggest that around 50% of patients with HCC will receive systemic therapies at one time point or another during their lifespan.[2,4,30] Several trials have tried to show survival benefits of systemic agents in advanced disease, a traditionally challenging setting due to the limited efficacy and high toxicity of conventional systemic chemotherapy.[16–18,31] Randomized studies also failed to prove any clinical efficacy for anti-oestrogen therapies.[32] In 2007, the landmark SHARP trial assessing the multi-tyrosine kinase inhibitor sorafenib (VEGFRs, PDGFRs, RAF and KIT) was the first to significantly expand survival (HR of 0.69) with manageable adverse events.[7] Similar efficacy was demonstrated in the phase III trial testing sorafenib in Asian patients.[8] These successful results helped establish contemporary concepts in trial design that have been implemented in phase III trials over the succeeding years (Table 1 and Fig. 1).[33] The main concepts implemented in these trials are: a) selection of patients with well-preserved liver function (*i.e.*, Child-Pugh class A) to minimize the competing risk of liver failure and death as a result of the natural history of cirrhosis; b) restriction of the investigational niches to those stages with unmet medical needs such as advanced stage (BCLC C), intermediate stage (BCLC B) progressing after transarterial chemoembolization, or adjuvant setting after resection/local ablation; c) use of OS as the cornerstone primary endpoint to assess efficacy in advanced stages, and d) use of critical prognostic factors as tools for stratification prior to randomization based upon Eastern Cooperative Oncology Group (ECOG) grade 0 *vs.* 1, macrovascular invasion, extrahepatic spread, and alpha-fetoprotein (AFP) levels (>400 ng/ml). Aetiology is not considered a prognostic factor, but needs to be incorporated into stratification when testing sorafenib, since it has been demonstrated to be a predictor of response, like absence of extrahepatic spread and low neutrophil-to-lymphocyte ratio.[34]

In this context, new treatment modalities emerged to challenge sorafenib in first-line or placebo in second-line. These included brivanib (VEGFRs and FGFRs),[10,19] sunitinib (VEGFRs, PDGFRs and KIT),[9] linifanib (VEGFRs and PDGFRs),[11] erlotinib (EGFR) in combination with sorafenib,[12] everolimus (mTOR),[20] tivantinib (MET),[23] doxorubicin loaded nanoparticles[27] and ADI-PEG 20 (arginine deiminase enzyme)[26] (Table 1). All of these drugs led to disappointing results and it was not until 2016 that the RESORCE study led to the first positive phase III trial in advanced HCC for nearly a decade. Regorafenib (VEGFRs, PDGFRs, KIT and Tie2) improved OS compared to placebo from 7.8 to 10.6 months in patients who progressed and were tolerant to sorafenib.[22] Notably, in patients with advanced HCC, the OS from the start of sorafenib was 26 months in those that received subsequent regorafenib treatment compared to 19 months for those receiving placebo after sorafenib.[35] Besides regorafenib, other phase III clinical trials have recently improved OS in second-line when compared to placebo: The CELESTIAL study, showing median OS of 10.2 months with cabozantinib (VEGFRs, MET and AXL) *vs.* 8 months with placebo;[24] and the REACH-2 study, where ramucirumab (VEGFR2 monoclonal antibody) provided a median OS of 8.5 months in patients with AFP equal or higher than 400 ng/ml *vs.* 7.3 months with placebo.[21,25] AFP is well-known for its independent prognostic capacity in HCC.[36] As such, REACH-2 was the first positive phase III trial in a biomarker-driven population of patients with HCC. In parallel, lenvatinib (VEGFRs, FGFRs, RET, KIT and PDGFRA) has become an option in front-line treatment, after the positive result of the non-inferiority REFLECT study.[13] In contrast, 3 phase III trials testing internal radiation with Y-90 for advanced HCC, either as single treatment (SARAH[14] and SIRveNIB[15]) or in combination Y-90 with sorafenib[37] did not meet the primary endpoint of improved OS compared to sorafenib. As a result, Y-90 was discouraged for the management of advanced HCC in the recent European Association for the Study of the Liver (EASL) guidelines.[4]

Finally, the FDA has granted accelerated approval to the immune checkpoint inhibitor nivolumab (monoclonal antibody against PD1) in second-line after a large phase II single-arm trial showing promising ORR of 14% by RECIST (responses lasting more than 12 months in 55% of cases).[38,39] Pembrolizumab has recently shown an ORR of 17% and median OS of 12.9 months in the second-line setting. However, results of the phase III RCT testing pembrolizumab *vs.* placebo in second-line were negative,[40,41] thus highlighting the importance of biomarkers form detection of responders. The revolution of immune therapies

Key point

In the last years 4 systemic agents (*i.e.*, lenvatinib, regorafenib, cabozantinib and ramucirumab) have demonstrated clinical benefit in phase III trials and 2 (*i.e.*, nivolumab, pembrolizumab) have been granted accelerated approval based on a phase II trial. Thus, these new agents are expanding the pipeline of effective drugs in advanced HCC.

* Corresponding author.
Address: Translational Research in Hepatic Oncology, IDIBAPS-Hospital Clinic, Rosselló 153, 08039 Barcelona, Catalonia, Spain. Tel. +34-932.279.155.
E-mail address: jmllovet@clinic.cat (J.M. Llovet).

Key point

The improvement in the number of effective agents comes at a cost of increased complexity of clinical decision-making, and thus, in the design of future clinical trials.

Table 1. Phase III trials in advanced hepatocellular carcinoma conducted in the last decade.

	Trial	Arms	N	ORR	TTP		PFS		OS	
					Median	HR	Median	HR	Median	HR
First-line	SHARP[7]	Sorafenib	299	2.3	5.5	***0.58 (0.45–0.74)***		NR	10.7	***0.69 (0.55–0.87)***
		Placebo	303	0.7	2.8				7.9	
	Asian-Pacific[8]	Sorafenib	150	3.3	2.8	***0.57 (0.42–0.79)***		NR	6.5	***0.68 (0.50–0.93)***
		Placebo	76	1.3	1.4				4.2	
	SUN1170[9]	Sunitinib	530	6.6	4.1	1.13 (0.98–1.31)	3.6	1.13 (0.99–1.30)	7.9	1.30 (1.13–1.50)
		Sorafenib	544	6.1	3.8		3		10.2	
	BRISK-FL*[10]	Brivanib	577	12.0	4.2	1.01 (0.88–1.16)		NR	9.5	1.07 (0.94–1.23)
		Sorafenib	578	8.8	4.1				9.9	
	LIGHT[11]	Linifanib	514	10.1	5.4	***0.76 (0.64–0.90)***	4.2	***0.81 (0.70–0.95)***	9.1	1.05 (0.90–1.22)
		Sorafenib	521	6.1	4		2.9		9.8	
	SEARCH[12]	Sorafenib + erlotinib	362	6.6	3.2	1.14 (0.94–1.37)	NR	1.11 (0.94–1.31)	9.5	0.93 (0.78–1.11)
		Sorafenib	358	3.9	4				8.5	
	REFLECT*[13]	Lenvatinib	478	24.1	8.9	***0.63 (0.53–0.73)***	7.4	***0.66 (0.57–0.77)***	13.6	**0.92 (0.79–1.06)**
		Sorafenib	476	9.2	3.7		3.7		12.3	
	SARAH[14]	Y90	237	15.2		NR	4.1	1.03 (0.85–1.25)	8	1.15 (0.94–1.41)
		Sorafenib	222	10.4			3.7		9.9	
	SIRveNIB[15]	Y90	182	16.5	6.1	0.88 (0.7–1.1)	5.8	0.89 (0.70–1.10)	8.8	1.10 (0.90–1.40)
		Sorafenib	178	1.7	5.4		5.1		10	
	EACH[16]	Folfox4	184	8.2		NR	2.93	***0.62 (0.49–0.79)***	6.4	0.80 (0.63–1.02)
		Doxorubicin	187	2.7			1.77		4.97	
	CALGB80802[17]	Sorafenib + doxorubicin	173	NR		NR	3.6	0.90 (0.72–1.20)	9.3	1.06 (0.80–1.40)
		Sorafenib	173	NR			3.2		10.5	
	SILIUS*[18]	Sorafenib + HAIC	103	36.3	5.3	***0.65 (0.48–0.87)***	4.8	0.75 (0.57–1.00)	11.8	1.01 (0.74–1.37)
		Sorafenib	103	17.5	3.5		3.5		11.5	
Second-line	BRISK-PS*[19]	Brivanib	263	9.9	4.2	***0.56 (0.42–0.76)***		NR	9.4	0.89 (0.69–1.15)
		Placebo	132	1.5	2.7				8.2	
	EVOLVE-1[20]	Everolimus	362	2.2	3	0.93 (0.75–1.15)		NR	7.6	1.05 (0.86–1.27)
		Placebo	184	1.6	2.6				7.3	
	REACH[21]	Ramucirumab	283	7.1	3.5	***0.59 (0.49–0.72)***	2.8	***0.63 (0.52–0.75)***	9.2	0.87 (0.72–1.05)
		Placebo	282	0.7	2.6		2.1		7.6	
	RESORCE*[22]	Regorafenib	379	10.6	3.2	***0.44 (0.36–0.55)***	3.1	***0.46 (0.37–0.56)***	10.6	***0.63 (0.50–0.79)***
		Placebo	194	4.1	1.5		1.5		7.8	
	METIV-HCC[23]	Tivantinib	226	0.0	2.4	0.96 (0.74–1.25)	2.1	0.96 (0.75–1.22)	8.4	0.97 (0.75–1.25)
		Placebo	114	0.0	3		2		9.1	
	CELESTIAL[24]	Cabozantinib	470	3.8	5.4	***0.41 (0.34–0.49)***	5.2	***0.44 (0.36–0.52)***	10.2	***0.76 (0.63–0.92)***
		Placebo	237	0.4	1.9		1.9		8	
	REACH-2[25]	Ramucirumab	197	4.6	3.02	***0.43 (0.31–0.58)***	2.8	***0.45 (0.34–0.60)***	8.5	***0.71 (0.53–0.95)***
		Placebo	95	1.1	1.61		1.6		7.3	
	ADI-PEG 20[26]	Adi-peg 20	424	NR		NR	2.6	1.18 (0.96–1.43)	7.8	1.02 (0.85–1.23)
		Placebo	211	NR			2.6		7.4	
	ReLive[27]	Doxorubicin transdrug	263	0.8		NR	2.3	0.95 (0.74–1.22)	9.1	1.00 (0.78–1.28)
		Placebo	134	0.7			2.3		9.0	

*Radiological evaluation by mRECIST.
Bold-italic = Positive for superiority (p <0.05).
Bold = Positive for non-inferiority (upper 95% CI <1.08).
Regular = Negative for superiority or non-inferiority.
NR, not reported.

that has changed the paradigm of treatment in oncology is now finding its way in HCC, with ongoing phase III trials targeting key mediators of the anti-cancer immune response (*e.g.*, PD1, PDL1, CTLA4, LAG3), in both first- (NCT02576509, NCT03298451, NCT03434379, NCT03713593) and second-line (NCT02702401) settings. Overall, these successful results have amplified the number of effective drugs available to clinicians for the management of advanced HCC (Fig. 2). New studies will be crucial to ascertain the most efficient way to utilize these drugs and maximize clinical benefit.

Reasons for positive/negative results in phase III investigations

Negative phase III clinical trials

Until 2016, sorafenib was the only systemic agent able to significantly increase survival in patients with advanced HCC.[7] This was despite numerous attempts to improve, or parallel (*i.e.*, non-inferiority trials), its efficacy and develop new second-line therapies in the context of phase III trials (Table 1). Until then none of the 8 randomized clinical trials (RCTs) testing systemic treatments (*vs.* sorafenib in frontline[9–12,17] or placebo in second-line[19–21]) were able to achieve positive results. Nowadays, 6 of 21 (29%) trials have been able to meet the primary endpoint and potentially change the standard of care. This success rate is lower than in other tumour types, with a reported success rate of 37%,[42] and resonates with the difficulties of developing effective drugs in HCC. Negative HCC trials enrolled a total of 8,604 patients and consumed a significant amount of resources.

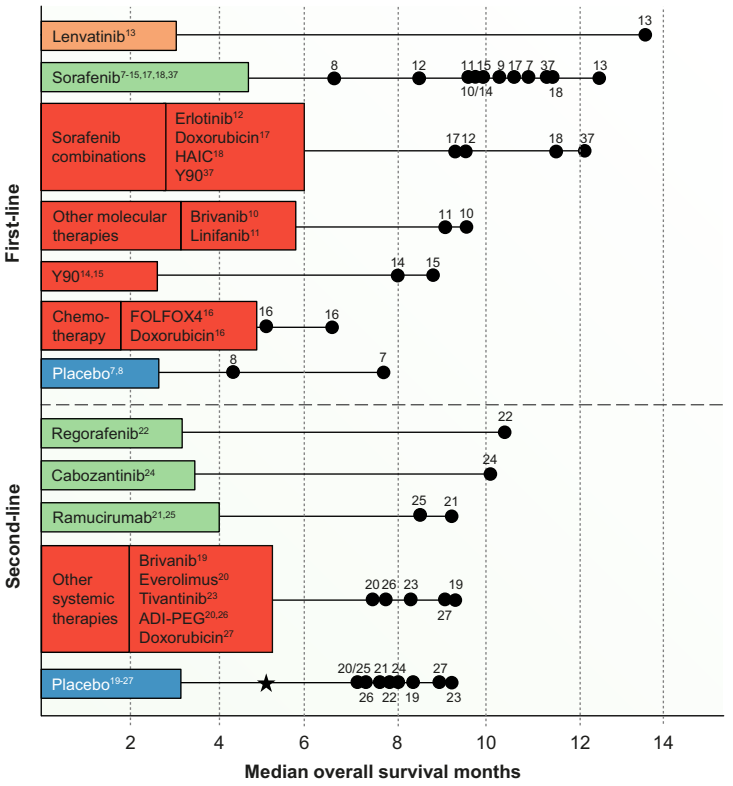

Fig. 1. Median overall survival of treatment modalities assessed in phase III trials for advanced hepatocellular carcinoma. Treatments with more than one dot represent all the results obtained from different clinical trials testing the same compound. Trials are coloured based on whether the final result was positive for superiority (green), negative (red) or positive for non-inferiority (orange) for the primary endpoint (OS). Placebo appears in blue. Relevant inclusion/exclusion criteria that may impact on median OS are: no portal vein invasion,[13] no pulmonary metastases,[37] sorafenib tolerant,[22] MET high[23] and AFP >400 ng/ml.[25] Star represents pooled individual patient data from REACH and REACH-2[130] in patients with AFP >400 ng/mL. AFP, alpha-fetoprotein; HAIC, hepatic infusion arterial chemotherapy; OS, overall survival.

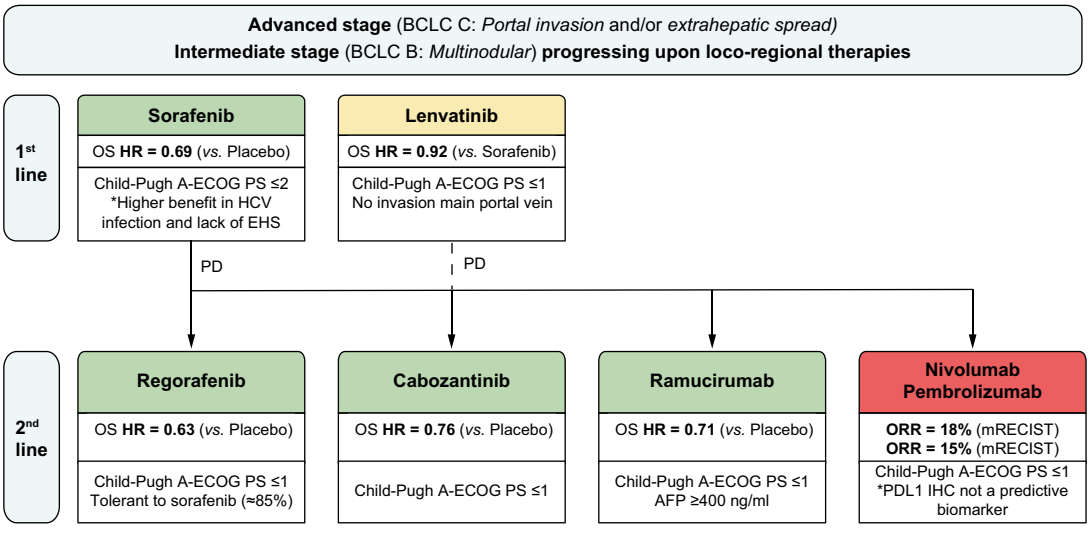

Fig. 2. Treatment strategy for advanced hepatocellular carcinoma. Adapted from Llovet *et al.* Nat Rev Clin Oncol 2018.[6] Drugs in green have positive results from phase III trials with a superiority design (sorafenib in the first-line setting and regorafenib, cabozantinib and ramucirumab in the second-line setting). Drugs in orange have positive results from phase III trials with a non-inferiority design (lenvatinib in the first-line setting). Drugs in red have received accelerated approval from the FDA on the basis of promising efficacy results in phase II trials (nivolumab and pembrolizumab in the second-line setting). Key details of the patient populations are provided. AFP, Alpha-fetoprotein; BCLC, Barcelona Clinic Liver Cancer (classification); ECOG PS, Eastern Cooperative Oncology Group performance status; EHS: extrahepatic spread; HCV, hepatitis C virus; HR, hazard ratio; mRECIST, modified Response Evaluation Criteria In Solid Tumors; ORR, objective response rate; OS, overall survival.

Failed drugs include linifanib,[11] erlotinib,[12] brivanib,[10,19] sunitinib,[9] doxorubicin,[17,27] everolimus,[20] tivantinib,[23] ADI-PEG 20[26] and radioembolization with Y-90[14,15] (Table 1). These drugs have different molecular targets, mechanisms of action, and include various treatment modalities (pharmacological *vs.* radiation-based). Thus, it is likely that multiple factors contributed to their failure, but we will dissect 3 key factors, reviewed in 43: a) limited antitumoral efficacy (or biological activity), b) significant toxicity, and c) lack of effective enrichment strategies for patient enrolment.

Limited antitumoral efficacy (or biological activity)

The first factor relates to limited antitumoral activity of the drug as per its main molecular targets. This implies that either they have a marginal role in HCC progression, or their selective inhibition is insufficient to induce a significant clinical benefit. For instance, evidence from murine models demonstrated that aberrant activation of MTOR signalling promotes liver cancer,[44] and that its selective abrogation has antitumoral effects in xenografts.[45] In human HCC, the MTOR pathway is deregulated in up to 45% of samples,[46,47] and yet, in a phase III trial, everolimus was unable to improve survival compared to placebo in second-line treatment, with an HR for OS of 1.05.[20] Data from phase II already suggested a modest median survival of 8.4 months[47] associated with marginal response rates. Moreover, the companion biomarker study for the phase III trial failed to find any robust predictive biomarker of response to everolimus.[48] Altogether, these data suggest that MTOR inhibition has no antitumoral activity in advanced HCC. Other drugs potentially falling under this category include the EGFR inhibitor erlotinib,[12] and the FGFR2/VEGFR inhibitor brivanib.[10,19] It has also been shown that Y-90 resin microsphere treatment is not superior to sorafenib. The published phase III trials[14,15] were negative, with an HR for OS of 1.1. It is important to note that failure to demonstrate superiority does not mean similar efficacy, which requires an *ad hoc* trial design for non-inferiority or equivalence, a concept that will be further discussed in this review.[49]

Drug toxicity

The second reason is significant drug toxicity, which is relevant in cirrhotic patients since liver dysfunction decreases the threshold for severe adverse effects. The best example is sunitinib,[50] which was unable to improve survival when compared to sorafenib in frontline.[9] Despite having a molecular target profile similar to sorafenib, the trial was prematurely terminated due to futility and safety concerns relating to sunitinib. The median OS with sunitinib was 7.9 months, compared to the 10.2 months with sorafenib (HR of 1.3). Treatment-related deaths occurred in 3.2% and 0.4% of patients receiving sunitinib and sorafenib, respectively. When this trial was conducted, sunitinib was already FDA approved for advanced kidney cancer and gastrointestinal stromal tumours, where toxicity was not a major clinical issue. However, in patients with underlying liver disease the toxicity of sunitinib was severe enough to obscure any beneficial antitumoral efficacy. Sunitinib has a higher inhibitory potency than sorafenib, particularly regarding its anti-angiogenic activity via VEGFR and PDGFR inhibition.[51] Angiogenesis is critical during liver fibrogenesis,[52] so the strong and sustained anti-angiogenic effect achieved with sunitinib seems detrimental, favouring liver failure. Previous phase II trials testing sunitinib in HCC offer additional insights into the hepatic toxicity of this drug,[53–55] including up to 4/37 (11%) treatment-related deaths.[55] An adequate identification of toxicity signals at this stage could help mitigate this problem. Another example is the VEGFR/PDGFR inhibitor linifanib, tested against sorafenib in a frontline setting.[11] This trial was terminated early based on futility (median OS for linifanib and sorafenib were 9.1 and 9.8 months), but grade 3–4 adverse events were significantly more frequent in linifanib than in sorafenib, including hypertension (21% *vs.* 11%) and hepatic encephalopathy (7% *vs.* 3%). Besides the negative effect of toxicity in clinical outcomes, there is a subtler effect of non-lethal toxicity as it associates with dose reductions, which could also decrease antitumor potency.

Lack of trial enrichment strategies

A third reason for clinical trial failure is the lack of effective enrichment strategies for patient enrolment based on predicted biomarkers of response. Trial enrichment in oncology is closely linked to the concept of oncogene addiction. This term describes those molecular alterations, generally DNA mutations or chromosomal aberrations, required for cancer cell proliferation and survival.[56] There are numerous examples in oncology of survival benefits after a clinical trial testing a drug only in those patients with mutations in its target[57] (*e.g.*, ALK rearrangements in lung cancer and response to crizotinib[58]). Only 2/21 (10%) phase III trials in advanced HCC incorporated patient enrichment, likely due to: a) limited access to tumour tissue in patients already diagnosed by non-invasive criteria; and b) few druggable targets among the most common genetic alterations in HCC.[59,60] In fact, the most common mutations in HCC (*TERT* promoter, *CTNNB1*, *TP53*, *AXIN1*, *ARID1A* and *ARID1B*) are untargetable.[2,61] One trial evaluated tivantinib *vs.* placebo in second-line treatment for patients with high expression of MET assessed by immunohistochemistry.[23] This trial was based on a *post hoc* analysis of 37 patients from a previous phase II trial[62] and failed to meet its primary endpoint with an HR for OS of 0.97. Arguably, the signal in the phase II trial was weak,

but most importantly, recent data questions the specificity of tivantinib as a MET inhibitor.[63] It was also thought that MET was a prognostic factor, but the median survival of 9.1 months for the placebo arm in MET-high patients in a second-line setting challenges this concept.[23] The second trial tested the VEGFR2 monoclonal antibody ramucirumab *vs.* placebo in second-line (*i.e.* REACH-2) in patients with AFP higher than 400 ng/ml, and showed a significant improvement in OS *vs.* placebo (HR of 0.71[25]). A difference with the tivantinib case is that the rationale for REACH-2 came from a *post hoc* analysis of the negative phase III trial in all-comers (*i.e.*, REACH[21]) which enrolled 565 patients. This showed a robust improvement in OS with ramucirumab in patients with high AFP ($p = 0.02$). AFP is a well-known poor prognostic marker,[36,47] highly expressed in tumours with a supposed progenitor cell origin,[64] but it does not provide a neat link between any specific driver oncogenic event (*i.e.*, structural DNA alterations or signalling pathways) and ramucirumab's main molecular target. Experimental evidence identifies VEGFR2 as a marker of hepatic progenitors,[65] which could hypothetically explain the efficacy of ramucirumab in patients with high AFP.

Positive phase III clinical trials

Successful drugs in frontline include sorafenib and lenvatinib, whereas regorafenib, cabozantinib, and ramucirumab demonstrated efficacy in second-line treatment for patients with high AFP levels (Table 1 and Fig. 2). The PD1 inhibitor nivolumab has shown promising results in phase II with an ORR of 14% by RECIST (18% by mRECIST) and a median OS of 15.6 months,[38,66] which led to its accelerated approval by the FDA. A similar ORR (17%) but lower median OS (12.9 months) was reported for pembrolizumab, another immune-based therapy.[41] Recent data failed to show survival benefit from pembrolizumab compared to placebo in phase III in the second-line setting.[40] Data from phase III trial in first-line (nivolumab *vs.* sorafenib) will be critical to determine if immune-based therapies should be recommended in clinical practice guidelines.[4] Since the strength of evidence so far comes from phase II data, current EASL guidelines posited a weak recommendation for nivolumab.[4]

The reasons for trial failure provide the best clues regarding the qualities required for a drug to be successful in HCC, which essentially are: a) adequate clinical trial design with an emphasis on selection criteria and robust endpoints; b) a fine balance between drug efficacy and toxicity; and c) a proper interpretation of efficacy and toxicity signals in phase II trials. Sorafenib epitomizes these qualities, and to certain extent, the design principles implemented in the pivotal SHARP trial[7] were adopted as best-practices for design in subsequent studies.[33] The target population must include patients with well-preserved liver function (*i.e.*, Child-Pugh A with compensated liver disease) to avoid competing risks from deaths due to progression of the liver disease, and to minimize drug toxicity. Also, patients need to be fit enough to tolerate the drug, with a life expectancy of at least 3 months, which can be reasonably guaranteed by enrolling patients with an ECOG performance status test grade of 0–1. It is paramount to enrol patients at the same clinical stage as per the BCLC classification.[4,67] The SHARP trial was instrumental in eradicating the misleading concept of 'unresectable' HCC when conducting HCC trials. This concept included a heterogeneous population of patients at intermediate (BCLC B) and advanced (BCLC B) stages,[68] which imposed significant bias when interpreting trial results. In addition to the same clinical stage, patients need to be adequately stratified for known HCC prognostic factors and geographic region.

Regarding patient selection, the success of the REACH-2 trial underscores the importance of properly interpreting *post hoc* analyses. The pooled analysis of REACH and REACH-2 assessing ramucirumab in those patients with AFP >400 ng/ml further confirms a significant and clinically meaningful benefit of ramucirumab *vs.* placebo in second-line (median survival 8.1 *vs.* 5.0 months; HR = 0.694).[25] The rationale to enrich trials based on predicted oncogene addiction is twofold: first, to maximize antitumoral response by perturbing the cancer drivers active in a given patient, and second, to spare unnecessary toxicity in those patients without the oncogene addiction. HCC has few druggable targets among the most frequent driver mutations, but a recent proof-of-concept trial reinforces the validity of this approach to explore treatment response.[69] In a phase II trial testing the efficacy of the MEK inhibitor refametinib, 1,318 patients with HCC were screened for RAS mutations, of whom 59 (4.4%) had RAS mutations, detected using circulating tumour DNA (ctDNA), and were therefore eligible for treatment.[69] Mutation analysis of ctDNA is feasible in HCC[70] and facilitates screening of large populations. Other potential druggable oncogenic alterations in HCC include high-level DNA amplifications of *FGF19*[71,72] or *VEGFA*.[71,73] Phase II clinical trials are currently exploring selective inhibition of these candidate oncogene addiction loops.[74] There is also increasing interest in developing biomarkers to identify the 20% of patients who respond to immune-based therapies, who show outstanding OS. The use of PD-L1 staining seems irrelevant in HCC,[38] and other potential biomarkers such as tumour mutational burden[75] or the HCC immune class[76] are under investigation. To facilitate the implementation of biomarker-based clinical trials in HCC, it is essential to enforce mandatory tissue collection in all clinical trials testing new compounds.[4] In this regard, the impact of intratumor heterogeneity in single-biopsy predictions is still

debated despite recent studies showing that driver gene mutations are common between different regions of the tumour.[77,78]

Traditionally, new therapies were compared with standard of care or placebo to demonstrate greater efficacy of the new drug. Despite this being the recommended trial design in HCC,[4,33] some studies after SHARP used non-inferiority designs to challenge sorafenib in first-line. The hypothesis in non-inferiority trials is that the new compound is not substantially worse than the current standard, as opposed to equivalence trials, which are designed to demonstrate that the experimental treatment is neither worse nor better than the standard therapy.[79] Non-inferiority trials are required to claim similar efficacy, as opposed to assuming it from a negative superiority trial, as previously explained for the Y-90 trials. The non-inferiority trial scenario in HCC has been described extensively elsewhere,[43] and caveats include the need for larger sample sizes and a very small window of opportunity, as defined by the tight non-inferiority margins. For instance, the BRISK-FL trial was designed to demonstrate non-inferiority of brivanib compared to sorafenib in first-line.[10] The trial assumptions set the upper limit of the 95% confidence interval of HR for OS to 1.08. To confirm non-inferiority, the HR could cross 1, but the upper boundary needed to fall between 1 and 1.08. This threshold is very stringent and can be interpreted as the requirement to demonstrate a robust non-significant trend towards superiority for the new drug. The value proposed by the FDA has been calculated based upon capturing at least >60% of the survival benefit obtained with sorafenib.[13] The BRISK-FL trial did not meet this endpoint since the HR confidence interval limits for OS were 0.94 and 1.23. The concept of non-inferiority trials introduces other considerations in treatment recommendations such as toxicity or cost, which will surely contribute to frame the landscape of systemic therapies in HCC.

Hard and surrogate endpoints: Implications in clinical trial design

The overreaching goal of oncological treatments is to allow patients to live longer and better lives than they would do without treatment.[80] Thus, clinical research needs to unequivocally demonstrate statistically and clinically meaningful improvements of the experimental arm over the standard of care. Three types of endpoints have been defined: i) Hard endpoints, such as OS and cancer-specific survival; ii) Surrogate endpoints such as PFS, TTP and ORR, and iii) Patient-reported endpoints, such as quality of life (QoL).

Key point

OS is the most robust endpoint in advanced HCC but the increasing number of effective therapies after progression underscore the need for surrogate endpoints.

Overall survival

OS is a hard endpoint that quantifies the time between random trial allocation and death, whatever the cause. Since is not subject to investigator bias, OS has been traditionally recommended by international HCC guidelines as the primary endpoint for randomized phase III trials testing new therapies.[33] In fact, all regular FDA drug approvals in advanced HCC were based upon improvements in OS.[7,22] Cancer-specific survival, where only deaths due to cancer are considered and non-cancer-related deaths are censored, is more difficult to assess in conventional trial settings. Deaths due to competing risks, such as liver failure, require a subjective interpretation by the investigator, and thus are more prone to bias.[33] OS is still the most robust endpoint in advanced HCC but the increasing number of treatments after progression underscore the need for surrogate endpoints.

What magnitude of benefit can be defined as clinically meaningful? This is a controversial topic, highly dependent on the expected outcome for the target population, with conflicting opinions between patients, providers, regulatory agencies and health insurers.[81,82] In HCC, there is no consensus on what absolute survival benefit (or the magnitude of benefit in OS according to HR) can be defined as clinically relevant. Reported thresholds of OS with HR <0.8 are sound for capturing the benefit of patients in advanced HCC trials.[83]

Survival has some limitations as a sole endpoint in cancer research. First, it might require a long follow-up time to capture enough events, due to significant improvement in median OS in the experimental arm.[35] This negatively impacts feasibility and delays patients' access to highly effective drugs. Second, it can be affected by sequential therapies received after tumour progression (post-progression survival), for example, regorafenib after a first-line therapy. This might involve one-third of patients in recent phase III trials.[13] In this context, validation of surrogate endpoints of OS is paramount to facilitate trial execution and rapid deployment of effective drugs in routine clinical conditions.

Surrogate endpoints: PFS, TTP and ORR

Ideally, significant improvement in OS is preferred, but many drugs have been approved based on their ability to improve other less robust endpoints, termed surrogates (*i.e.*, TTP, PFS and ORR). These are outcomes not inherently meaningful from the clinical standpoint, but thought to accurately predict hard outcomes such as OS.[84] The development of surrogate endpoints became a necessity in clinical trials in cardiology, where the long time to accumulate enough events for a hard endpoint made most studies unfeasible. Surrogate endpoints are becoming increasingly important in oncology, as more effective post-progression therapies are available.

Accelerated approval based upon surrogate endpoints is becoming the most relevant path for regulatory approval of cancer drugs in the US.

Between 2009 and 2014; the FDA approved 83 drugs in oncology, 66% of them on the basis of surrogate endpoints.[85] The FDA's Accelerated Approval Program was introduced in 1992 as a social compromise during the worse years of the HIV epidemic to expedite access to agents for life-threating conditions based on surrogate endpoints. The programme included a "safety net" that required the manufacturer to conduct post-marketing studies and confirm the efficacy of the drug using hard endpoints.[86] A recent analysis of approved drugs during the period 1992–2017 led to the following conclusions:[28] a) Accelerated approval was granted for 93 indications, ORR being the most common surrogate endpoint used (87% of cases), b) Among drugs approved through this path, 55% were ultimately confirmed for regular approval, 5% of indications were withdrawn (*e.g.*, bevacizumab in metastatic breast cancer[87]), whereas in others the process has not been concluded.

Despite the increasing importance of surrogate endpoints in oncology, they have 2 main limitations. Firstly, since they usually rely on the radiological definition of tumour progression or response, they are vulnerable to interpretation bias. This can be minimized by using central radiology reviews and a designated adjudicator of response. Secondly, and more importantly, in order to be reliable, they require validation as credible predictors of OS.[84] Validation of surrogate endpoints can be conducted at the individual- or trial-levels.[88] While validation of individual-level surrogacy requires individual patient data from at least one clinical trial, trial-level surrogacy uses assembled data from multiple trials. The Institute for Quality and Efficiency in Health Care has proposed a set of criteria to quantify the association between a surrogate and hard endpoint, which includes low ($R < 0.7$), moderate ($R > 0.7$ to $R < 0.85$) and high correlation ($R > 0.85$).[89] R refers to the weighted Pearson coefficient between HR of OS and HR of the surrogate endpoint. Alternative methods to study this correlation have been reported.[90] A systematic review and meta-analysis of trial-level surrogate endpoints (PFS, TTP and ORR) for OS in oncology, including 36 articles and 352 clinical trials, found low, moderate and high correlation with OS in 52%, 25% and 23% of surrogate endpoints, respectively.[84]

Analysis of surrogate end-points in HCC

In order to explore the concept of surrogate endpoints recapitulating OS in advanced HCC we have identified 21 RCTs assessing systemic therapies with or without loco-regional therapies in advanced HCC (12 in first-line and 9 in second-line) (Table 1) published between 2008 and 2018 through a MEDLINE search via PubMed using the keywords "advanced hepatocellular carcinoma". Results were limited to "clinical trial, phase III". Trials recently presented at international meetings (2016–2018) were also included despite the full manuscripts not yet being available. For each trial, data on sample size, radiological response, TTP, PFS and OS were collected. TTP, PFS and OS were determined in terms of HR using published data (values less than 1 denote a favourable result in the experimental group). In addition, ORR was established with the odds ratio calculated from the published radiological response (values greater than 1 denote a favourable result in the experimental group). For the purpose of the trial-level analysis, we first assessed the overall correlation between PFS and OS ($R = 0.84$; $R^2 = 0.71$) (Fig. 3A), and then the correlation of TTP and OS ($R = 0.83$; $R^2 = 0.69$) (Fig. 3B). Afterwards, we established a conservative threshold of positive predictive value for PFS since this is the most documented surrogate time-to-event endpoint in oncology, and the one showing a higher correlation with OS.[84] Finally, we explored the correlation between ORR assessed by mRECIST and survival in early, intermediate and advanced HCC.

Progression-free survival

PFS is a composite endpoint of 2 variables: death and evidence of radiological progression, usually defined by standard criteria as RECIST[91] or mRECIST.[92] International guidelines initially discouraged this endpoint in HCC due to the competing risk of dying due to progressed liver dysfunction despite a relevant antitumoral benefit.[33] However, this limitation has been mitigated since most trials in HCC have adopted restrictive inclusion criteria in terms of liver function (*i.e.*, Child-Pugh A without decompensation). In this scenario, the likelihood of death as a result of liver decompensation (*i.e.*, gastrointestinal bleeding, encephalopathy or ascites and spontaneous peritonitis) is 5% at 1 year.[93] When we evaluate the association between PFS and OS in HCC phase III trials, we observe a Pearson correlation (R) of 0.84 (Fig. 3A). This figure falls in the upper boundary of a moderate correlation (R between 0.7 and 0.85). When specifically analysing the positive predictive value of theoretical thresholds of PFS correlating with OS, only 3/7 PFS reported an HR ≤0.6 that was significantly associated with a positive survival benefit (in all cases with an HR for OS <0.8). Conversely, those 4 studies reporting a positive PFS with an HR between 0.6–0.8 were associated with no significant survival benefits (HR for OS between 0.87 and 1.05) (Fig. 3A). In our study, according to the linear regression equation obtained [$\log HR_{OS} = 0.072 + 0.487 \times \log HR_{PFS}$], a threshold PFS HR of 0.6 (representing a 40% risk reduction) will decrease the risk of OS by ~17% (OS HR = 0.83) (see Fig. 3A). In summary, a moderate correlation has been established between PFS and OS in 21 RCTs in advanced HCC. A value of HR ≤0.6 is proposed as surrogate threshold effect,[94] and is likely to predict a clinically meaningful improvement in OS. Is worth

Key point

PFS has a significant correlation with OS at trial level (R = 0.84). A conservative minimum surrogate threshold effect of HR ≤0.6 is highly predictive of a significant improvement in OS, whereas HR ranging from 0.6–0.7 are uncertain surrogates.

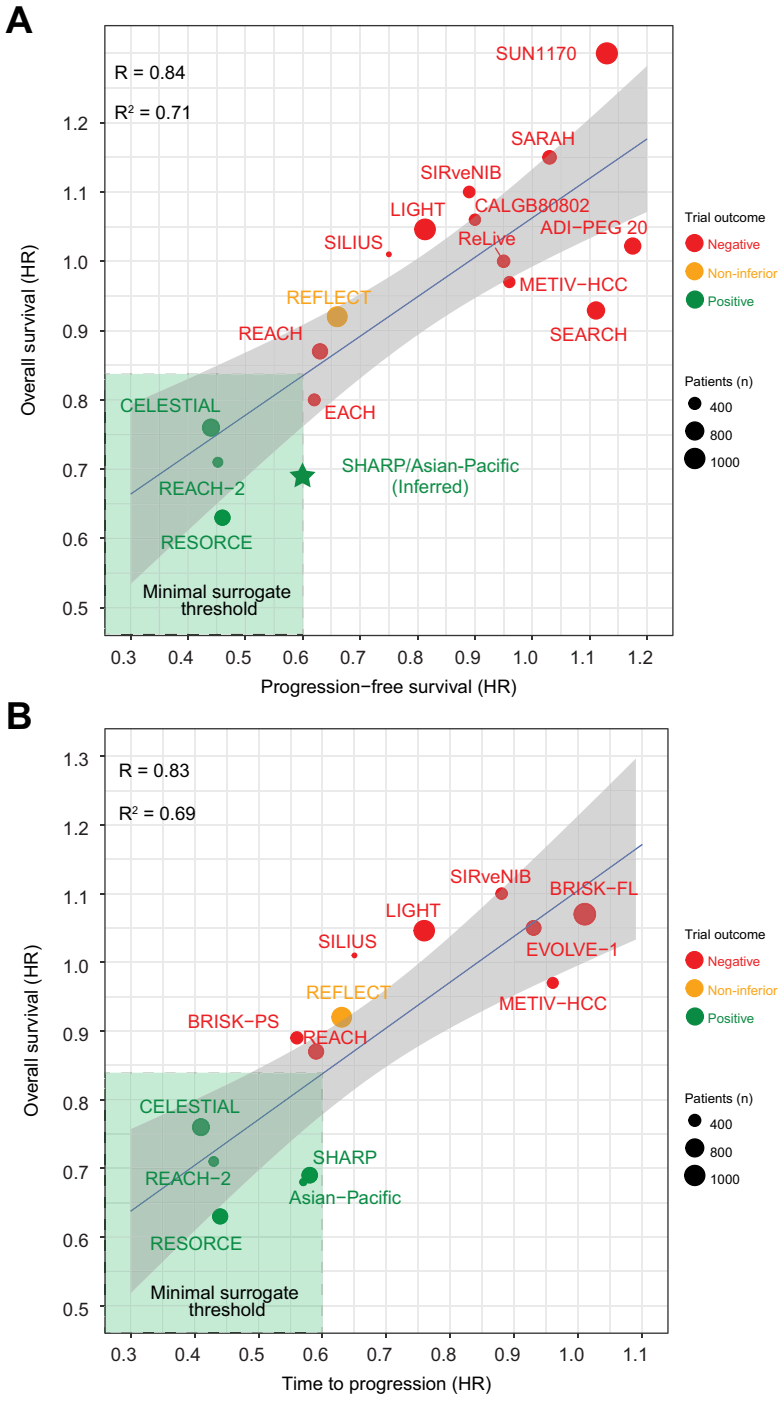

Fig. 3. Correlation between surrogate endpoints and hard endpoint (overall survival). (A) Correlation between progression-free survival and overall survival. (B) Correlation between time to progression and overall survival. Trial-level correlation between endpoints using criteria from the Institute for Quality and Efficiency in Health Care (IQWIG). R and R2 refers to the weighted Pearson coefficient between the HR of OS and the HR of the surrogate endpoint. IQWIG categorizes the strength of the correlation based on the value of R as low (R <0.7), moderate (R >0.7 to R <0.85) and high (R >0.85).[89] Each dot represents one of the phase III clinical trials conducted on advanced HCC. Size of the dot is proportional to the total number of patients enrolled in the trial. Trials are coloured based on whether the final result was positive, negative or non-inferior for the primary endpoint (OS). X and Y axis depict the value of the HR for the surrogate and the hard endpoint, respectively. Gray shaded areas represent the upper and lower limits of the 95% confidence intervals for the regression. HCC, hepatocellular carcinoma; HR, hazard ratio; OS, overall survival.

mentioning that this rule is supported mainly by positive trials in the second-line setting comparing active drugs *vs.* placebo. We assume that such an association is retained in a frontline setting, where 2 active drugs are being compared, but any recommendation in that setting should be tempered due to the lack of confirmatory data. In conclusion, PFS has a moderate correlation at trial level with OS (R = 0.84). A conservative minimum surrogate threshold effect of HR ≤0.6 is highly predictive of a significant improvement in OS, whereas HR ranging from 0.6–0.7 is an uncertain surrogate.

Time to progression

This endpoint quantifies the time between trial allocation and radiological progression, usually defined by standard criteria as RECIST[91] or mRECIST.[92] Deaths are censored as non-radiological progressions at the time of death or at an earlier visit, with a cause-specific hazard, representing informative censoring. Symmetric repeated radiological measurements every 6–8 weeks are required to avoid missing moderate differences between treatment groups.[33] This recommendation was not followed in the SIRveNIB,[15] SARAH[14] and ADI-PEG 20[26] trials.

To delineate the adequacy of TTP as a surrogate of OS in HCC we also conducted a trial-level meta-analysis to evaluate the correlation between TTP and OS in 21 RCTs (Fig. 3B). The Pearson correlation (R) was 0.83, which indicates a moderate association according to the IQWIG guidelines.[84] In 10 phase III trials there was a significant difference in TTP in favour of the investigational arm. However, these positive results in TTP were not reflected by superiority in OS in 5 (50%) trials. Brivanib[19] and ramucirumab[21] in second-line showed efficacy according to TTP (HR = 0.56 and 0.59, respectively), while not significantly improving the hard endpoint of OS (HR = 0.89 and 0.87, respectively). A trial comparing first-line lenvatinib *vs.* sorafenib showed a significant positive difference in TTP with levantinib (HR 0.63), which was not reflected by superiority in OS (HR = 0.92).[13] This trial was positive since it was designed for non-inferiority (upper 95% CI lower than 1.08). Finally, linifanib[11] and hepatic infusion arterial chemotherapy (HAIC)[18] in first-line failed to show any benefit in terms of OS (HR = 1.05 and 1.01, respectively) even though there was a clear benefit when measuring TTP (HR = 0.76 and 0.65, respectively). These results do not support the initial recommendation after the SHARP trial of using TTP as the optimal surrogate endpoint[33] in phase II trials, and reinforce the need for accurate evaluation of surrogacy in clinical trials. Based on the linear regression model obtained [log $HR_{OS} = 0.083 + 0.491 \times \log HR_{TTP}$], we can extrapolate that a therapy producing a 40% risk reduction in TTP will yield an estimated ~16% risk

reduction in OS (HR = 0.84) (see Fig. 3B). Moreover, in order to directly compare the performance of PFS and TTP, we analysed the correlation between both surrogate endpoints, obtaining a Pearson correlation (R) of 0.99 (Fig. 4). Thus, in the modern era of HCC trial design, with minimal cirrhosis-related deaths (due to the inclusion of Child-Pugh A), there is a strong correlation between both endpoints. In fact, when we inferred the non-reported PFS HR of SHARP and the Asia Pacific trials according to the linear equation obtained by comparing both surrogate endpoints [log HR_{PFS} = 0.014 + 0.927 × log HR_{TTP}], the HRs values are close to 0.60, just at the previously proposed minimum threshold.

There are 2 other considerations regarding this endpoint. First, not all types of tumour progression necessarily have the same clinical meaning. Recent data also suggest that TTP may capture heterogeneous features, with essentially 2 types of progression at advanced stages.[95,96] In particular, survival after progression is significantly worse for patients who develop a new extrahepatic lesion and/or vascular invasion (median OS = 7.1 months) compared to those who progress due to the growth of existing intrahepatic/extrahepatic lesions or the development of a new intrahepatic lesion (median OS = 14.9 months). Second, factors including evaluation bias, trial attrition or informative censoring may weaken the association between the TTP and OS.[29] Finally, prolonged exposure to a given therapy might lead to a phenotypic change in tumours, thus, offsetting any initial advantage from the treatment captured by the surrogate endpoint.[97]

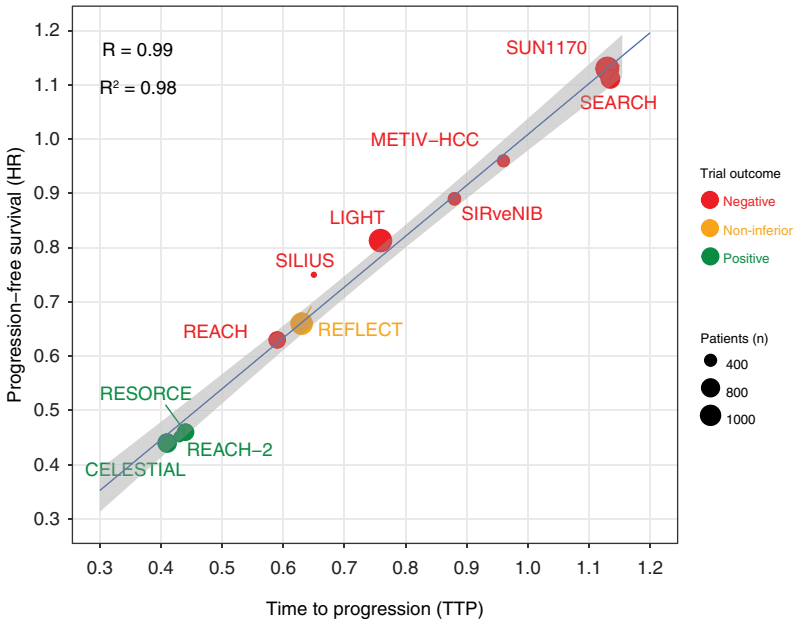

Fig. 4. Correlation between surrogate endpoints (progression-free survival and time to progression). Trial-level correlation between endpoints. R and R2 refers to the weighted Pearson coefficient between the HR of PFS and the HR of TTP. Each dot represents one of the phase III clinical trials conducted on advanced HCC. Size of the dot is proportional to the total number of patients enrolled in the trial. Trials are coloured based on whether the final result was positive, negative or non-inferior for the primary endpoint (OS). X and Y axis depict the value of the HR for the surrogate TTP and PFS, respectively. Gray shaded areas represent the upper and lower limits of the 95% confidence intervals for the regression. HCC, hepatocellular carcinoma; HR, hazard ratio; OS, overall survival; PFS, progression-free survival; TTP, time to progression.

Objective response rate

Tumour response in oncology trials is typically measured using RECIST (Response Evaluation Criteria in Solid Tumors).[91] These criteria standardize methods for converting radiological observations into a quantitative and statistically tractable framework to define tumour response (*i.e.*, a 30% decrease in the diameter of target lesion). ORR is the percentage of patients who achieve an objective tumour response. ORR by sensitive criteria in single-arm phase II trials could be a useful tool to prioritize treatments for testing in phase III trials. Disease control rate is the combination of ORR and stable disease, but it has 2 disadvantages that limit its adoption for regulatory approval: a) the definition of duration of stable disease varies between studies; and b) stable disease can reflect inherent characteristics of the tumour rather than treatment efficacy.

The RECIST criteria were originally developed to evaluate cytotoxic agents. The generalization of targeted therapies has challenged this simplistic approach that relies on tumour shrinkage to indicate clinical efficacy. Sorafenib was associated with an ORR of only 2–3%, despite providing clear survival benefits.[7,8] Given the poor correlation between tumour response assessed with conventional tools and OS, a group of experts convened by the American Association for the Study of Liver Diseases (AASLD) proposed specific amendments to standard RECIST.[33] Further description of response and progression resulted in the criteria named modified RECIST (mRECIST), which ultimately incorporates the concept of viable tumour defined as the portions of tumour showing arterial enhancement.[92] The mRECIST criteria in HCC have improved the sensitivity to quantify tumour response with targeted therapies: ORR of 9–17% with sorafenib,[10,13,18] 10–12% with brivanib,[10,19] 11% with regorafenib[22] and 24% with lenvatinib.[13] Retrospective studies have consistently demonstrated that patients who achieved an objective response on sorafenib had a longer survival than non-responders.[98–100] Recently, data from double-blind randomized trials assessing brivanib, nintedanib and lenvatinib further validated this association.[101,102] Thus, the association between tumour response and improved OS in patients with HCC at advanced stages complement what was already known in patients at early and intermediate stages treated with loco-regional therapies[103–109] (Table 2).

When we evaluate the trial-level correlation between ORR and OS (Fig. S1), the R weighted Pearson coefficient obtained is 0.54. This is significantly lower than the correlation obtained with PFS/TTP and OS as depicted in Fig. 3. There are 2

Key point

ORR, assessed by sensitive criteria, in single arm phase II trials could be a useful tool to prioritize treatments for testing in phase III trials.

Table 2. Studies analysing associations between radiological response and survival in hepatocellular carcinoma.

	Treatment	Study	N	Radiological response	Median OS (HR/OR OS) [R *vs.* Non-R]	Prediction of OS
Local ablation	RFA	Cabibbo G *et al.* (retrospective RFA)* [103]	151	CRR = 78%	59.4 m *vs.* 26 m (HR = 0.3)	Yes (UV)
	RFA/PEI	Sala *et al.* (retrospective)* [105]	282	CRR = 68%	43 m *vs.* 28 m (OR = 0.58)	Yes (MV)
Chemo-embolization	TAE/TACE	Gillmore R *et al.* (retrospective)[106]	83	ORR = 57%	20.7 m *vs.* 13.3 m (HR = 0.58)	Yes (MV)
	TACE	Kim BK *et al.* (retrospective)[107]	292	ORR = 71.9%	33.8 m *vs.* 17.1 m (HR = 0.48)	Yes (MV)
	TACE	Jung ES *et al.* (retrospective)[108]	114	ORR = 63.3%	41.1 m *vs.* 20.7 m (HR = 0.31)	Yes (MV)
	DEB TACE	Prajapati HJ *et al.* (retrospective)[109]	120	ORR = 52.5%	28 m *vs.* 9.1 m (HR = 0.4)	Yes (MV)
Sorafenib	Sorafenib	Edeline J *et al.* (retrospective)[98]	53	ORR = 23%	18.2 m *vs.* 7.7 m (NR)	Yes (UV)
	Sorafenib	Ronot M (retrospective)** [99]	82	ORR = 28%	25.5 m *vs.* 5.7 m (HR = 0.19)	Yes (MV)
	Sorafenib	Takada J (retrospective)[100]	191	ORR = 13.1%	~21 m *vs.* ~10 m (NR)	Yes (UV)
Other systemic therapies	Brivanib	Lencioni R *et al.* (phase III trial)[101]	226	ORR = 11.5%	14.3 m *vs.* 9.4 m (HR = 0.48)	Yes (MV)
	Nintedanib	Meyer T *et al.* (2 phase II trials)[102]	180	ORR = 15.6%	16.7 m *vs.* 10.9 m (HR = 0.62)	Yes (MV)
	Nivolumab	El-Khoueiry A *et al.* (phase II trial)*** [66]	145	ORR = 14%	Non-reached *vs.* 13.4 m^ (NR)	Yes (UV)
	Lenvatinib	Kudo M *et al.* (phase III trial)[13] (https://meetinglibrary.asco.org/record/169342/abstract)	478	ORR = 24.1%	22.4 m *vs.* 11.4 m (HR = 0.61)	Yes (MV)

CRR, complete response rate; HR, hazard ratio; m, months; MV, multivariate; NR, not reported; OR, odds ratio; ORR, objective response rate; R, response (Objective or complete response); UV, univariate.
* EASL criteria.
** Non-R does not include stable disease.
*** RECIST 1.1 criteria, ^0–25% reduction (17.7 m); 0–25% increase (11.7 m); ≥25% increase (8.9 m).

reasons for this: one is inherent to the use of odds ratios instead of HRs to compare differences in ORR. The accuracy of an odds ratio decreases for low values of ORRs. The second reason is that only a small proportion of patients within these trials achieved ORR (~10–20%), which is, in fact, the event that correlates with better survival.[101] A direct comparison between RECIST and mRECIST for OS surrogacy through an independent meta-analysis of trials using either criteria would be ideal to define the role of ORR to predict OS in advanced HCC. This will also help determine the best tool for evaluating tumour response to systemic therapies. However, since only 5 RCTs reported response data using mRECIST, we did not sub-analyse this endpoint according to the tool used to evaluate response.

Some other questions remain unanswered. As observed for other solid tumours treated with efficacious targeted therapies,[110–112] the reported rates of responders are still suboptimal to estimate the maximum number of patients who would benefit from the treatment. In addition, the duration of response might be more clinically relevant than the extent of tumour reduction. Finally, the strategy to evaluate response might require a thoughtful revision when assessing immunotherapies. As shown in melanoma patients treated with checkpoint inhibitors, standard RECIST may not provide a reliable assessment of antitumor efficacy.[113] In fact, response to immunotherapy may take longer compared to other agents and can even falsely mirror criteria for progression (*i.e.*, pseudo-progression).[114] Immune-related response criteria have been developed,[115,116] including the concept of "confirmation of progression" by a second scan obtained at least 4 weeks after progressive disease has been registered.

Despite all the challenges that evaluation of tumour response raises in oncology, and particularly in HCC, the importance of ORR as a surrogate endpoint is recognized by regulatory agencies and frequently used for accelerated drug approval. This was the case for nivolumab, approved in second-line based on an ORR of 18% by mRECIST and 14% by RECIST.[38,39] Remarkably, objective response to nivolumab has been associated with prolonged OS.[66] Overall, the fact that a high ORR in phase II trials was considered a robust criterion for drug approval,[117] and further success in phase III trials,[118] indicates that ORR should be considered as a primary endpoint for single-arm phase II studies. Related to this, early clinical trials are showing promising results with combinations of checkpoint inhibitors and targeted therapies, as measured by ORR. In front-line advanced HCC, combination treatments with lenvatinib plus pembrolizumab[119] achieved an ORR of 46% by mRECIST and atezolizumab plus bevacizumab (n = 73)[120,121] 34% by mRECIST. As a result, the later combination was granted breakthrough therapy designation by the

FDA.[122] Of note, most of the drugs approved under the accelerated programme reported ORRs exceeding 30%.[123]

Patient-reported endpoints: Quality of life

Health-related QoL measures the effect of the disease on an individual's physical, psychological and social functioning and well-being.[124] Regulatory agencies recognize symptomatic improvement as a direct clinical benefit to patients and an important consideration in drug approval.[125] However, unlike OS, the interpretation of QoL is subjective. In HCC, 2 tools have been proposed to measure QoL: the European Organization for Research and Treatment of Cancer Quality of Live Questionnaire (EORTC QLQ-HCC18)[126] and the Functional Assessment of Cancer Therapy-Hepatobiliary (FACT-Hep) questionnaire.[127] They can be used to evaluate time to symptomatic progression (*i.e.* time between trial allocation and the occurrence of disease-related symptoms according to preestablished scores). QoL was measured in the SHARP trial[7] according to the FHSI-8 questionnaire,[128] a reduced version of FACT-Hep, with results that collided with OS. Conversely, the SARAH trial[14] reported better global health status with Y-90 than sorafenib based on QLQ-HCC18, which was inconsistent with the primary endpoint of OS. Evaluation of QoL is contingent on when it is assessed during disease progression. Also, significant changes in QoL have been observed across different cultures.[129] Defining and evaluating reliable QoL assessment tools has been established as one of the unmet needs in HCC research by international guidelines.[4] In summary, health-related QoL measures are not ready to support, as single tools, regulatory approval for drugs in HCC.

Conclusions

The current period of drug development in HCC is providing major advances in the management of this devastating disease. Currently, 5 drugs have shown activity in phase III clinical trials and 2 immune checkpoint inhibitors have been approved based on a phase II trial, which represents an unprecedented revolution in the context of the last 50 years. Novel drugs and combination strategies are emerging in the field. Thus, new tools will be required for the proper assessment of clinical benefits. OS is still the most robust endpoint but the increasing number of treatments available in advanced HCC necessitate the use of surrogate endpoints, which are less vulnerable to subsequent treatments after progression. In this scenario, PFS has shown a moderate correlation with OS (R = 0.84); a threshold of HR ≤0.6 defines a conservative approach which enables capturing survival differences in a superiority trial, using this surrogate endpoint, with a high positive predictive value. Two recent studies have been released supporting our threshold of HR ≤0.6 for PFS. The first, a meta-analysis of individual patient data from 2 RCTs (REACH[21] and REACH-2[25]), showed a significant OS HR with a PFS HR of 0.57.[130] The second, a phase III RCT comparing sorafenib with or without conventional transarterial chemoembolization in advanced HCC, was negative for its primary endpoint (OS), despite a PFS HR of 0.73.[131] Thus, a PFS HR of ≤0.6 could be considered a candidate endpoint in phase II and phase III RCTs when subsequent therapies are expected to impact overall outcome. ORR by sensitive criteria (mRECIST) may be useful, particularly in single-arm phase II trials with proof-of-concept drugs or in combination studies targeting accelerated approval with a threshold >30%. Finally, the current development of RCTs assessing immune therapies or drug combinations in HCC may probably change the paradigm of drug development and trial design. Whether the statements proposed in the present review are confirmed in trials designed with composite primary endpoints, such as OS-PFS for lenvatinib + pembrolizumab *vs.* lenvatinib (NCT03713593) or OS-ORR for atezolizumab + bevacizumab *vs.* sorafenib (NCT03434379) is of particular interest.

Financial support

JML is supported by grants from European Commission Horizon 2020 Program (HEPCAR, proposal number 667273-2), US Department of Defense (CA150272P3), National Cancer Institute (P30 CA196521), Spanish National Health Institute (SAF 2016-76390), Cancer Research UK – Asociación Española Contra el Cáncer (Accelerator Award-HUNTER-A26813), Samuel Waxman Cancer Research Foundation, Generalitat de Catalunya (AGAUR, SGR-1358). AV is supported by U.S. Department of Defense (CA150272P3) and Tisch Cancer Institute (Cancer Center Grant P30 CA196521). RM is supported by a FSEOM – Boehringer Ingelheim grant.

Conflict of interest

JML and AV are inventors of a patent for the immune molecular subclass in HCC. Dr Llovet is receiving research support from Bayer HealthCare Pharmaceuticals, Eisai Inc, Bristol-Myers Squibb and Ipsen, and consulting fees from Eli Lilly, Bayer HealthCare Pharmaceuticals, Bristol-Myers Squibb, Eisai Inc, Celsion Corporation, Exelixis, Merck, Ipsen, Glycotest, Navigant, Leerink Swann LLC, Midatech Ltd, Fortress Biotech, Sprink Pharmaceuticals, Nucleix and Cat-Fite pharmaceuticals. Dr. Villanueva has received consulting fees from Guidepoint and Fujifilm; advisory board fees from Exact Sciences, Nucleix and NGM; and lecture fees from Exelixis.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2019.01.028.

References

Author names in bold designate shared co-first authorship

[1] Torre L, Bray F, Siegel R, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA Cancer J Clin 2015;65:87–108.
[2] Llovet JM, Zucman-Rossi J, Pikarsky E, Sangro B, Schwartz M, Sherman M, et al. Hepatocellular carcinoma. Nat Rev Dis Prim 2016;2:16018.
[3] Llovet JM, Brú C, Bruix J. Prognosis of hepatocellular carcinoma: the BCLC staging classification. Semin Liver Dis 1999;19:329–338.
[4] Galle PR, Forner A, Llovet JM, Mazzaferro V, Piscaglia F, Raoul J, et al. EASL clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol 2018;69:182–236.
[5] Llovet JM, Bruix J. Systematic review of randomized trials for unresectable hepatocellular carcinoma: chemoembolization improves survival. Hepatology 2003;37:429–442.
[6] Llovet JM, Montal R, Sia D, Finn RS. Molecular therapies and precision medicine for hepatocellular carcinoma. Nat Rev Clin Oncol 2018, in press.
[7] Llovet J, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc J, et al. Sorafenib in advanced hepatocellular carcinoma. N Engl J Med 2008;359:378–390.
[8] Cheng A-L, Kang Y-K, Chen Z, Tsao C-J, Qin S, Kim JS, et al. Efficacy and safety of sorafenib in patients in the Asia-Pacific region with advanced hepatocellular carcinoma: a phase III randomised, double-blind, placebo-controlled trial. Lancet Oncol 2009;10:25–34.
[9] Cheng A-L, Kang Y-K, Lin D-Y, Park J-W, Kudo M, Qin S, et al. Sunitinib versus sorafenib in advanced hepatocellular cancer: results of a randomized phase III trial. J Clin Oncol 2013;31:4067–4075.
[10] Johnson PJ, Qin S, Park J-W, Poon RTP, Raoul J-L, Philip PA, et al. Brivanib versus sorafenib as first-line therapy in patients with unresectable, advanced hepatocellular carcinoma: results from the randomized phase III BRISK-FL study. J Clin Oncol 2013;31:3517–3524.
[11] Cainap C, Qin S, Huang W-T, Chung IJ, Pan H, Cheng Y, et al. Linifanib versus Sorafenib in patients with advanced hepatocellular carcinoma: results of a randomized phase III trial. J Clin Oncol 2015;33:172–179.
[12] Zhu AX, Rosmorduc O, Evans TRJ, Ross PJ, Santoro A, Carrilho FJ, et al. SEARCH: a Phase III, randomized, double-blind, placebo-controlled trial of sorafenib plus erlotinib in patients with advanced hepatocellular carcinoma. J Clin Oncol 2014;33:559–566.
[13] Kudo M, Finn R, Qin S, Han SS, Ikeda K, Piscaglia F, et al. Randomised phase 3 trial of lenvatinib vs. sorafenib in firstline treatment of patients with unresectable hepatocellular carcinoma. Lancet 2018;391:1163–1173.
[14] Vilgrain V, Pereira H, Assenat E, Guiu B, Ilonca AD, Pageaux GP, et al. Efficacy and safety of selective internal radiotherapy with yttrium-90 resin microspheres compared with sorafenib in locally advanced and inoperable hepatocellular carcinoma (SARAH): an open-label randomised controlled phase 3 trial. Lancet Oncol 2017;18:1624–1636.
[15] Chow PKH, Gandhi M, Tan S-B, Khin MW, Khasbazar A, Ong J, et al. SIRveNIB: selective internal radiation therapy versus sorafenib in asia-pacific patients with hepatocellular carcinoma. J Clin Oncol 2018;36:1913–1921.
[16] Qin S, Bai Y, Lim HY, Thongprasert S, Chao Y, Fan J, et al. Randomized, multicenter, open-label study of oxaliplatin plus fluorouracil/leucovorin versus doxorubicin as palliative chemotherapy in patients with advanced hepatocellular carcinoma from Asia. J Clin Oncol 2013;31:3501–3508.
[17] Abou-Alfa GK, Niedzwieski D, Knox JJ, Kaubisch A, Posey J, Tan BR, et al. Phase III randomized study of sorafenib plus doxorubicin versus sorafenib in patients with advanced hepatocellular carcinoma (HCC): CALGB 80802 (Alliance). J Clin Oncol 2016;34(4_suppl).
[18] Kudo M, Ueshima K, Yokosuka O, Ogasawara S, Obi S, Izumi N, et al. Sorafenib plus low-dose cisplatin and fluorouracil hepatic arterial infusion chemotherapy versus sorafenib alone in patients with advanced hepatocellular carcinoma (SILIUS): a randomised, open label, phase 3 trial. Lancet Gastroenterol Hepatol 2018;3:424–432.
[19] Llovet JM, Decaens T, Raoul JL, Boucher E, Kudo M, Chang C, et al. Brivanib in patients with advanced hepatocellular carcinoma who were intolerant to sorafenib or for whom sorafenib failed: results from the randomized phase III BRISK-PS study. J Clin Oncol 2013;31:3509–3516.
[20] Zhu AX, Kudo M, Assenat E, Cattan S, Kang Y-K, Lim HY, et al. Effect of everolimus on survival in advanced hepatocellular carcinoma after failure of sorafenib: the EVOLVE-1 randomized clinical trial. JAMA 2014;312:57–67.
[21] Zhu AX, Park JO, Ryoo B-Y, Yen C-J, Poon R, Pastorelli D, et al. Ramucirumab versus placebo as second-line treatment in patients with advanced hepatocellular carcinoma following first-line therapy with sorafenib (REACH): a randomised, double-blind, multicentre, phase 3 trial. Lancet Oncol 2015:859–870.
[22] Bruix J, Qin S, Merle P, Granito A, Huang Y-H, Bodoky G, et al. Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet 2017;389:56–66.
[23] Rimassa L, Assenat E, Peck-Radosavljevic M, Zagonel V, Pracht M, Rota Caremolli E. Second-line tivantinib (ARQ 197) vs placebo in patients (Pts) with MET-high hepatocellular carcinoma (HCC): results of the METIV-HCC phase III trial. J Clin Oncol 2017;35:15_supp.
[24] Abou-Alfa GK, Meyer T, Cheng A-L, El-Khoueiry AB, Rimassa L, Ryoo B-Y, et al. Cabozantinib in patients with advanced and progressing hepatocellular carcinoma. N Engl J Med 2018;379:54–63.
[25] Zhu AX, Kang Y, Yen C, Finn RS, Galle PR, Llovet JM, et al. Ramucirumab after sorafenib in patients with advanced hepatocellular carcinoma and increased α-fetoprotein concentrations (REACH-2): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol 2019;2045:1–15.
[26] Abou-Alfa GK, Qin S, Lu S-N, Yen C-J, Feng Y-H, Lim Y, et al. Phase III randomized study of second line ADI-peg 20 (A) plus best supportive care versus placebo (P) plus best supportive care in patients (pts) with advanced hepatocellular carcinoma (HCC). Ann Oncol 2018;29:1402–1408.
[27] Merle P, Bodoky G, López-López C, Saad AS, Casadei Gardini A, Borbath I, et al. Safety and efficacy results from the phase 3 relive study of doxorubicin-loaded nanoparticles versus best standard of care in patients with advanced hepatocellular carcinoma after failure or intolerance to previous treatment including sorafenib. Int Liver Cancer Assoc 11th Annu Conf 2017.
[28] Beaver JA, Howie LJ, Pelosof L, Kim T, Liu J, Goldberg KB, et al. A 25-year experience of US Food and Drug Administration accelerated approval of malignant hematology and oncology drugs and biologics. JAMA Oncol 2018;4:849–856.
[29] Kemp R, Prasad V. Surrogate endpoints in oncology: when are they acceptable for regulatory and clinical decisions, and are they currently overused? BMC Med 2017;15:1–7.
[30] Bruix J, Sherman MAA for the S of LD. Management of hepatocellular carcinoma: an update. Hepatology 2011;53:1020–1022.
[31] Yeo W, Mok TS, Zee B, Leung TWT, Lai PBS, Lau WY, et al. A randomized phase III study of doxorubicin versus cisplatin/interferon alpha-2b/doxorubicin/fluorouracil (PIAF) combination chemotherapy for unresectable hepatocellular carcinoma. J Natl Cancer Inst 2005;97:1532–1538.
[32] Chow P, Tai B-C, Tan C-K, Machin D, Win KM, Johnson PJ, et al. High-dose tamoxifen in the treatment of inoperable hepatocellular carcinoma: a multicenter randomized controlled trial. Hepatology 2002;36:1221–1226.
[33] Llovet JM, Di Bisceglie AM, Bruix J, Kramer BS, Lencioni R, Zhu AX, et al. Design and endpoints of clinical trials in hepatocellular carcinoma. J Natl Cancer Inst 2008;100:698–711.
[34] Bruix J, Cheng A-L, Meinhardt G, Nakajima K, De Sanctis Y, Llovet J. Prognostic factors and predictors of sorafenib benefit in patients with hepatocellular carcinoma: analysis of two phase 3 studies. J Hepatol 2017;67:999–1008.
[35] Finn RS, Merle P, Granito A, Huang Y-H, Bodoky G, Pracht M, et al. Outcomes with sorafenib followed by regorafenib or placebo for HCC: additional analyses from the phase 3 RESORCE trial. J Hepatol 2018;69:353–358.
[36] Llovet JM, Pena CEA, Lathia CD, Shan M, Meinhardt G, Bruix J. Plasma biomarkers as predictors of outcome in patients with advanced hepatocellular carcinoma. Clin Cancer Res 2012;18:2290–2300.
[37] Ricke J, Sangro B, Amthauer H, Bargellini I, Bartenstein P, De Toni E, et al. The impact of combining Selective Internal Radiation Therapy (SIRT) with sorafenib on overall survival in patients with advanced hepatocellular carcinoma: the SORAMIC trial palliative cohort. J Hepatol 2018;68:S65–S102.
[38] El-Khoueiry AB, Sangro B, Yau T, Crocenzi TS, Kudo M, Hsu C, et al. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. Lancet 2017;389:2492–2502.

[39] FDA grants accelerated approval to nivolumab for HCC previously treated with sorafenib n.d. https://www.fda.gov/Drugs/InformationOnDrugs/ApprovedDrugs/ucm577166.htm.
[40] Merck. Merck Provides Update on KEYNOTE-240, a Phase 3 Study of KEYTRUDA® (pembrolizumab) in Previously Treated Patients with Advanced Hepatocellular Carcinoma. February 19, 2019. Accessed at: https://investors.merck.com/news/press-release-details/2019/Merck-Provides-Update-on-KEYNOTE-240-a-Phase-3-Study-of-KEYTRUDA-pembrolizumab-in-Previously-Treated-Patients-with-Advanced-Hepatocellular-Carcinoma/default.aspx.
[41] Zhu AX, Finn RS, Edeline J, Cattan S, Ogasawara S, Palmer D, et al. Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): a non-randomised, open-label phase 2 trial. Lancet Oncol 2018;19.
[42] Gan HK, You B, Pond GR, Chen EX. Assumptions of expected benefits in randomized phase III trials evaluating systemic treatments for cancer. J Natl Cancer Inst 2012;104:590–598.
[43] Llovet JM, Hernandez-Gea V. Hepatocellular carcinoma: reasons for phase III failure and novel perspectives on trial design. Clin Cancer Res 2014;20:2072–2079.
[44] Guri Y, Colombi M, Dazert E, Hindupur SK, Roszik J, Moes S, et al. mTORC2 promotes tumorigenesis via lipid synthesis. Cancer Cell 2017;32(807–823) e12.
[45] Villanueva A, Chiang DY, Newell P, Peix J, Thung S, Alsinet C, et al. Pivotal role of mTOR signaling in hepatocellular carcinoma. Gastroenterology 2008;135:1972–1983.
[46] Sahin F, Kannangai R, Adegbola O, Wang J, Su G, Torbenson M. mTOR and P70 S6 kinase expression in primary liver neoplasms. Clin Cancer Res 2004;10:8421–8425.
[47] Zhu AX, Abrams TA, Miksad R, Blaszkowsky LS, Meyerhardt JA, Zheng H, et al. Phase 1/2 study of everolimus in advanced hepatocellular carcinoma. Cancer 2011;117:5094–5102.
[48] Zhu AX, Chen D, He W, Kanai M, Voi M, Chen LT, et al. Integrative biomarker analyses indicate etiological variations in hepatocellular carcinoma. J Hepatol 2016;65:296–304.
[49] Llovet JM, Finn RS. Negative phase 3 study of 90Y microspheres versus sorafenib in HCC. Lancet Oncol 2018;19 e69.
[50] Faivre S, Delbaldo C, Vera K, Robert C, Lozahic S, Lassau N, et al. Safety, pharmacokinetic, and antitumor activity of SU11248, a novel oral multitarget tyrosine kinase inhibitor, in patients with cancer. J Clin Oncol 2006;24:25–35.
[51] Stein MN, Flaherty KT. Sorafenib and sunitinib in renal cell carcinoma. Clin Cancer Res 2007;13:3765–3770.
[52] Poisson J, Lemoinne S, Boulanger C, Durand F, Moreau R, Valla D, et al. Liver sinusoidal endothelial cells: physiology and role in liver diseases. J Hepatol 2017;66:212–227.
[53] Koeberle D, Montemurro M, Samaras P, Majno P, Simcock M, Limacher A, et al. Continuous Sunitinib Treatment in Patients with Advanced Hepatocellular Carcinoma: A Swiss Group for Clinical Cancer Research (SAKK) and Swiss Association for the Study of the Liver (SASL) Multicenter Phase II Trial (SAKK 77/06). Oncologist 2010;15:285–292.
[54] Zhu AX, Sahani DV, Duda DG, Di Tomaso E, Ancukiewicz M, Catalano OA, et al. Efficacy, safety, and potential biomarkers of sunitinib monotherapy in advanced hepatocellular carcinoma: a phase II study. J Clin Oncol 2009;27:3027–3035.
[55] Faivre S, Raymond E, Boucher E, Douillard J, Lim HY, Kim JS, et al. Safety and efficacy of sunitinib in patients with advanced hepatocellular carcinoma: an open-label, multicentre, phase II study. Lancet Oncol 2009;10:794–800.
[56] Luo J, Solimini NL, Elledge SJ. Principles of cancer therapy: oncogene and non-oncogene addiction. Cell 2009;136:823–837.
[57] Villanueva A, Llovet JM. Targeted therapies for hepatocellular carcinoma. Gastroenterology 2011;140:1410–1426.
[58] Shaw AT, Kim D-W, Nakagawa K, Seto T, Crinó L, Ahn M-J, et al. Crizotinib versus chemotherapy in advanced ALK-positive lung cancer. N Engl J Med 2013;368:2385–2394.
[59] Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat Med 2017;23:703–713.
[60] Schulze K, Nault J-C, Villanueva A. Genetic profiling of hepatocellular carcinoma using next-generation sequencing. J Hepatol 2016;65:1031–1042.
[61] Zucman-Rossi J, Villanueva A, Nault JC, Llovet JM. Genetic landscape and biomarkers of hepatocellular carcinoma. Gastroenterology 2015;149:1226–1239.
[62] Santoro A, Rimassa L, Borbath I, Daniele B, Salvagni S, Van Laethem JL, et al. Tivantinib for second-line treatment of advanced hepatocellular carcinoma: a randomised, placebo-controlled phase 2 study. Lancet Oncol 2013;14:55–63.
[63] Rebouissou S, La Bella T, Rekik S, Imbeaud S, Calatayud AL, Rohr-Udilova N, et al. Proliferation markers are associated with MET expression in hepatocellular carcinoma and predict tivantinib sensitivity in vitro. Clin Cancer Res 2017;23:4364–4375.
[64] Hoshida Y, Nijman SMB, Kobayashi M, Chan JA, Brunet J-P, Chiang DY, et al. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. Cancer Res 2009;69:7385–7392.
[65] Goldman O, Han S, Sourrisseau M, Dziedzic N, Hamou W, Corneo B, et al. KDR identifies a conserved human and murine hepatic progenitor and instructs early liver development. Cell Stem Cell 2013;12:748–760.
[66] El-Khoueiry AB, Melero I, Yau TC, Crocenzi TS, Kudo M, Hsu C, et al. Impact of antitumor activity on survival outcomes, and nonconventional benefit, with nivolumab (NIVO) in patients with advanced hepatocellular carcinoma (aHCC): Subanalyses of CheckMate-040. J Clin Oncol 2018;36, suppl.
[67] Forner A, Reig M, Bruix J. Hepatocellular carcinoma. Lancet 2018;391:1301–1314.
[68] Llovet JM, Bustamante J, Castells A, Vilana R, Ayuso MDC, Sala M, et al. Natural history of untreated nonsurgical hepatocellular carcinoma: rationale for the design and evaluation of therapeutic trials. Hepatology 1999;29:62–67.
[69] Lim HY, Merle P, Weiss KH, Yau T, Ross PJ, Mazzaferro V, et al. Studies with refametinib or refametinib plus sorafenib in patients with RAS-mutated hepatocellular carcinoma. Clin Cancer Res 2018:27.
[70] Labgaa I, Villacorta-Martin C, D'Avola D, Craig AJ, von Felden J, Martins-Filho SN, et al. A pilot study of ultra-deep targeted sequencing of plasma DNA identifies driver mutations in hepatocellular carcinoma. Oncogene 2018:1–13.
[71] Chiang DY, Villanueva A, Hoshida Y, Peix J, Newell P, Minguez B, et al. Focal gains of VEGFA and molecular classification of hepatocellular carcinoma. Cancer Res 2008;68:6779–6788.
[72] Sawey ET, Chanrion M, Cai C, Wu G, Zhang J, Zender L, et al. Identification of a therapeutic strategy targeting amplified FGF19 in liver cancer by Oncogenomic screening. Cancer Cell 2011;19:347–358.
[73] Horwitz E, Stein I, Andreozzi M, Nemeth J, Shoham A, Pappo O, et al. Human and mouse VEGFA-amplified hepatocellular carcinomas are highly sensitive to Sorafenib treatment. Cancer Discov 2014;4:730–743.
[74] Kang Y-K, Macarulla T, Yau T, Sarker D, Choo SP, Meyer T, et al. Activity of Blu-554, a Potent, Highly-Selective FGFR4 Inhibitor in Advanced Hepatocellular Carcinoma (HCC) with FGFR4 Pathway Activation. Int Liver Cancer Assoc 11th Annu Conf 2017.
[75] Hellmann MD, Ciuleanu T-E, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, et al. Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. N Engl J Med 2018;378:2093–2104.
[76] Sia D, Jiao Y, Martinez-Quetglas I, Kuchuk O, Villacorta-Martin C, Castro De Moura M. Identification of an immune-specific class of hepatocellular carcinoma, based on molecular features. Gastroenterology 2017;153:812–826.
[77] Torrecilla S, Sia D, Harrington AN, Zhang Z, Cabellos L, Cornella H, et al. Trunk mutational events present minimal intra- and inter-tumoral heterogeneity in hepatocellular carcinoma. J Hepatol 2017;67:1222–1231.
[78] Reiter JG, Makohon-Moore AP, Gerold JM, Heyde A, Attiyeh MA, Kohutek ZA, et al. Minimal functional driver gene heterogeneity among untreated metastases. Science 2018;361:1033–1037.
[79] Kaji AH, Lewis RJ. Noninferiority trials: is a new treatment almost as effective as another? JAMA 2015;313:2371–2372.
[80] Wilson MK, Karakasis K, Oza AM. Outcomes and endpoints in trials of cancer treatment: the past, present, and future. Lancet Oncol 2015;16: e32–e42.
[81] Booth CM, Tannock I. Reflections on medical oncology: 25 years of clinical trials–where have we come and where are we going? J Clin Oncol 2008;26:6–8.
[82] Cherny NI, Dafni U, Bogaerts J, Latino NJ, Pentheroudakis G, Douillard JY, et al. ESMO-magnitude of clinical benefit scale version 1.1. Ann Oncol 2017;28:2340–2366.
[83] Ellis LM, Bernstein DS, Voest EE, Berlin JD, Sargent D, Cortazar P, et al. American society of clinical oncology perspective: raising the bar for

clinical trials by defining clinically meaningful outcomes. J Clin Oncol 2014;32:1277–1280.
[84] Prasad V, Kim C, Burotto M, Vandross A. The strength of association between surrogate end points and survival in oncology: a systematic review of trial-level meta-analyses. JAMA Intern Med 2015;175:1389–1398.
[85] Kim C, Prasad V. Strength of validation for surrogate end points used in the us food and drug administration's approval of oncology drugs. Mayo Clin Proc 2016;91:713–725.
[86] Gyawali B, Kesselheim AS. Reinforcing the social compromise of accelerated approval. Nat Rev Clin Oncol 2018.
[87] Carpenter D, Kesselheim AS, Joffe S. Reputation and precedent in the bevacizumab decision. N Engl J Med 2011;365 e3.
[88] Zhao F. Surrogate end points and their validation in oncology clinical trials. J Clin Oncol 2016;34:1436–1437.
[89] Institute for Quality and Efficiency in Health Care [Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen] (IQWiG). Validity of surrogate endpoints in oncology: executive summary. n.d. http://www.iqwig.de/download/A10-05_Executive_Summary_v1-1_Surrogate_end-points_in_oncology.pdf.
[90] Ciani O, Buyse M, Garside R, Peters J, Saad ED, Stein K, et al. Meta-analyses of randomized controlled trials show suboptimal validity of surrogate outcomes for overall survival in advanced colorectal cancer. J Clin Epidemiol 2015;68:833–842.
[91] Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 2009;45(2):228–247.
[92] Lencioni R, Llovet J. Modified RECIST (mRECIST) Assessment for Hepatocellular Carcinoma. Semin Liver Dis 2010;30:052–60.
[93] D'Amico G, Garcia-Tsao G, Pagliaro L. Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. J Hepatol 2006;44:217–231.
[94] Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006;5:173–186.
[95] Reig M, Rimola J, Torres F, Darnell A, Rodriguez-Lope C, Forner A, et al. Postprogression survival of patients with advanced hepatocellular carcinoma: rationale for second-line trial design. Hepatology 2013;58:2023–2031.
[96] Iavarone M, Cabibbo G, Biolato M, Della Corte C, Maida M, Barbara M, et al. Predictors of survival in patients with advanced hepatocellular carcinoma who permanently discontinued sorafenib. Hepatology 2015;62:784–791.
[97] Booth CM, Eisenhauer EAGroup NCT. Progression-free survival: meanigful or simply measurable? J Clin Oncol 2012;30:1030–1033.
[98] Edeline J, Boucher E, Rolland Y, Vauléon E, Pracht M, Perrin C, et al. Comparison of tumor response by Response Evaluation Criteria in Solid Tumors (RECIST) and modified RECIST in patients treated with sorafenib for hepatocellular carcinoma. Cancer 2012;118:147–156.
[99] Ronot M, Bouattour M, Wassermann J, Bruno O, Dreyer C, Larroque B, et al. Alternative Response Criteria (Choi, European association for the study of the liver, and modified Response Evaluation Criteria in Solid Tumors [RECIST]) Versus RECIST 1.1 in patients with advanced hepatocellular carcinoma treated with sorafenib. Oncologist 2014;19:394–402.
[100] Takada J, Hidaka H, Nakazawa T, Kondo M, Numata K, Tanaka K, et al. Modified response evaluation criteria in solid tumors is superior to response evaluation criteria in solid tumors for assessment of responses to sorafenib in patients with advanced hepatocellular carcinoma. BMC Res Notes 2015;8:609.
[101] **Lencioni R**, **Montal R**, Torres F, Park J-W, Decaens T, Raoul J-L, et al. Objective response by mRECIST as a predictor and potential surrogate end-point of overall survival in advanced HCC. J Hepatol 2017;66:1166–1172.
[102] Meyer T, Palmer DH, Cheng A-L, Hocke J, Loembé A-B, Yen C-J. mRECIST to predict survival in advanced hepatocellular carcinoma: analysis of two randomised phase II trials comparing nintedanib vs sorafenib. Liver Int 2017:1–9.
[103] Cabibbo G, Maida M, Genco C, Alessi N, Peralta M, Butera G, et al. Survival of Patients with Hepatocellular Carcinoma (HCC) Treated by Percutaneous Radio-Frequency Ablation (RFA) Is Affected by Complete Radiological Response. PLoS ONE 2013;8.
[104] Vincenzi B, Di Maio M, Silletta M, D'Onofrio L, Spoto C, Piccirillo MC, et al. Prognostic Relevance of Objective Response According to EASL Criteria and mRECIST Criteria in Hepatocellular Carcinoma Patients Treated with Loco-Regional Therapies: A Literature-Based Meta-Analysis. PLoS ONE 2015;10 e0133488.
[105] 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.
[106] Gillmore R, Stuart S, Kirkwood A, Hameeduddin A, Woodward N, Burroughs AK, et al. EASL and mRECIST responses are independent prognostic factors for survival in hepatocellular cancer patients treated with transarterial embolization. J Hepatol 2011;55:1309–1316.
[107] Kim BK, Kim KA, Park JY, Ahn SH, Chon CY, Han KH, et al. Prospective comparison of prognostic values of modified Response Evaluation Criteria in Solid Tumours with European Association for the Study of the Liver criteria in hepatocellular carcinoma following chemoembolisation. Eur J Cancer 2013;49:826–834.
[108] Jung ES, Kim JHJS, Yoon EL, Lee HJ, Lee SJ, Suh SJ, et al. Comparison of the methods for tumor response assessment in patients with hepatocellular carcinoma undergoing transarterial chemoembolization. J Hepatol 2013;58:1181–1187.
[109] Prajapati HJ, Spivey JR, Hanish SI, El-rayes BF, Kauh JS, Chen Z, et al. mRECIST and EASL responses at early time point by contrast-enhanced dynamic MRI predict survival in patients with unresectable hepatocellular carcinoma (HCC) treated by doxorubicin drug-eluting beads transarterial chemoembolization (DEB TACE). Ann Oncol 2013;24:965–973.
[110] Hudes G, Carducci M, Tomczak P, Dutcher J, Figlin R, Kapoor A, et al. Temsirolimus, interferon alpha or both for advanced renal cell carcinoma. N Engl J Med 2007;356:2271–2281.
[111] Karapetis CS, Khambata-Ford S, Jonker DJ, O'Callaghan CJ, Tu D, Tebbutt NC, et al. K-ras mutations and benefit from cetuximab in advanced colorectal cancer. N Engl J Med 2008;359:1757–1765.
[112] Hurwitz H, Fehrenbacher L, Novotny W, Cartwright T, Hainsworth J, Heim W, et al. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. N Engl J Med 2004;350:2335–2342.
[113] Hodi FS, Hwu WJ, Kefford R, Weber JS, Daud A, Hamid O, et al. Evaluation of immune-related response criteria and RECIST v1.1 in patients with advanced melanoma treated with Pembrolizumab. J Clin Oncol 2016;34:1510–1517.
[114] Wolchok JD, Hoos A, O'Day S, Weber JS, Hamid O, Lebbé C, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. Clin Cancer Res 2009;15:7412–7420.
[115] Seymour L, Bogaerts J, Perrone A, Ford R, Schwartz LH, Mandrekar S, et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. Lancet Oncol 2017;18:e143–e152.
[116] Stephen Hodi F, Ballinger M, Lyons B, Soria JC, Nishino M, Tabernero J, et al. Immune-modified response evaluation criteria in solid tumors (imrecist): Refining guidelines to assess the clinical benefit of cancer immunotherapy. J Clin Oncol 2018;36:850–858.
[117] Tsimberidou AM, Braiteh F, Stewart DJ, Kurzrock R. Ultimate fate of oncology drugs approved by the US food and drug administration without a randomized trial. J Clin Oncol 2009;27:6243–6250.
[118] El-Maraghi RH, Eisenhauer EA. Review of phase II trial designs used in studies of molecular targeted agents: outcomes and predictors of success in phase III. J Clin Oncol 2008;26:1346–1354.
[119] Ikeda M, Sung MW, Kudo M, Kobayashi M, Baron AD, Finn RS, et al. A phase 1b trial of lenvatinib (LEN) plus pembrolizumab (PEM) in patients (pts) with unresectable hepatocellular carcinoma (uHCC). J Clin Oncol 2018;36, suppl.
[120] Hsu CH, Lee MS, Ryoo BY, Stein S, Lee KH, Verret W, et al. APASL: Safety and clinical activity results from atezolizumab (atezo) + bevacizumab (bev) in hepatocellular carcinoma (HCC): Updates from a Phase Ib study. Feb 22, 2019. Accessed at: https://medically.roche.com/en/search/pdfviewer.cfd683a3-ef40-4777-99f5-402cdc61c904.html.
[121] Stein S, Pishvaian MJ, Lee MS, Lee K-H, Hernandez S, Kwan A, et al. Safety and clinical activity of 1L atezolizumab + bevacizumab in a phase Ib study in hepatocellular carcinoma (HCC). J Clin Oncol 2018;36, suppl.
[122] FDA Grants Breakthrough Therapy Designation for Genentech's TECENTRIQ in Combination With Avastin as First-Line Treatment for Advanced or Metastatic Hepatocellular Carcinoma (HCC) 2018. https://www.gene.com/media/press-releases/14736/2018-07-17/fda-grants-breakthrough-therapy-designat?utm_source=F&utm_medium=P&utm_term=15538&utm_content=TecentriqHCCBTD&utm_campaign=TecentriqHCCBTD.

[123] Oxnard GR, Wilcox KH, Gonen M, Polotsky M, Hirsch BR, Schwartz LH. Response rate as a regulatory end point in single-arm studies of advanced solid tumors. JAMA Oncol 2016;2:772–779.
[124] Revicki DA, Osoba D, Fairclough D, Barofsky I, Berzon R, Leidy NK, et al. Recommendations on health-related quality of life research to support labeling and promotional claims in the United States. Qual Life Res 2000;9:887–900.
[125] Wilson MK, Collyar D, Chingos DT, Friedlander M, Ho TW, Karakasis K, et al. Outcomes and endpoints in cancer trials: bridging the divide. Lancet Oncol 2015;16:e43–e52.
[126] Chie W-C, Blazeby JM, Hsiao C-F, Chiu H-C, Poon RT, Mikoshiba N, et al. International cross-cultural field validation of an European Organization for Research and Treatment of Cancer questionnaire module for patients with primary liver cancer, the European Organization for Research and Treatment of Cancer quality-of-life ques. Hepatology 2012;55:1122–1129.
[127] Heffernan N, Cella D, Webster K, Odom L, Martone M, Passik S, et al. Measuring health-related quality of life in patients with hepatobiliary cancers: the functional assessment of Cancer Therapy-Hepatobiliary Questionnaire. J Clin Oncol 2002;20: 2229–2239.
[128] Yount S, Cella D, Webster K, Heffernan N, Chang CH, Odom L, et al. Assessment of patient-reported clinical outcome in pancreatic and other hepatobiliary cancers: The FACT Hepatobiliary Symptom Index. J Pain Symptom Manage 2002;24:32–44.
[129] Chie W-C, Blazeby JM, Hsiao C-F, Chiu H-C, Poon RT, Mikoshiba N, et al. Differences in health-related quality of life between European and Asian patients with hepatocellular carcinoma. Asia Pac J Clin Oncol 2017:e304–e311.
[130] Llovet JM, Kudo M, Finn R, Galle PR, Blanc J, Okusaka T, et al. Ramucirumab as second-line treatment in patients with hepatocellular carcinoma (HCC) and elevated alpha-fetoprotein (AFP) following sorafenib: pooled results from two global phase 3 studies (REACH-2 and REACH). Hepatology 2018;68, suppl.
[131] Park J-W, Kim YJ, Kim DY, Bae SH, Paik SW, Lee Y-J. Sorafenib with or without concurrent transarterial chemoembolization in patients with advanced hepatocellular carcinoma: a phase III STAH trial. J Hepatol 2018, in press.