mRECIST to Predict Survival in Advanced HCC: Analysis of Two Randomised Phase II Trials Comparing Nintedanib versus Sorafenib

Tim Meyer[1,*], Daniel H. Palmer[2], Ann-Lii Cheng[3], Julia Hocke[4], Arsène-Bienvenu Loembé[5], Chia-Jui Yen[6]

[1]UCL Cancer Institute, University College London, London, UK; [2]Department of Molecular and Clinical Cancer Medicine, University of Liverpool, UK; and Clatterbridge Cancer Centre, Bebington, UK; [3]National Taiwan University Hospital, Taipei City, Taiwan; [4]Boehringer Ingelheim, Pharma GmbH & Co. KG, Biberach, Germany; [5]Boehringer Ingelheim B.V., Alkmaar, The Netherlands; [6]National Cheng Kung University Hospital, Tainan City, Taiwan

Email addresses:

Daniel H. Palmer: daniel.palmer@liverpool.ac.uk

Ann-Lii Cheng: alcheng@ntu.edu.tw

Julia Hocke: julia.hocke@boehringer-ingelheim.com

Arsène-Bienvenu Loembé: arsene-bienvenu.loembe@boehringer-ingelheim.com

Chia-Jui Yen: yencj@mail.ncku.edu.tw

***Corresponding author:** Professor Tim Meyer, Department of Oncology, UCL Cancer Institute, 72 Huntley Street, London WC1E 6DD, UK

Email: t.meyer@ucl.ac.uk

Tel: +44 (020)7679 6731

Running header: mRECIST to predict survival in HCC

Word count: ~4363 (max 5000, excluding figures and/or tables)

Tables/figures: 3/3 (max 8)

Refs: 22 (50 max)

Abbreviations: RECIST, Response Evaluation Criteria in Solid Tumors; HCC, hepatocellular carcinoma; mRECIST, modified Response Evaluation Criteria in Solid Tumors; OS, overall survival; HR, hazard ratio; CI, confidence interval; TTP, time to progression; TAE, transcatheter embolisation; TACE, transarterial chemoembolisation; MTD, maximum tolerated dose; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ULN, upper limit of normal; PD, progressive disease; EHS, extrahepatic spread; MVI, macrovascular invasion; ECOG PS, Eastern Cooperative Oncology Group performance score; PFS, progression-free survival; OR, objective response; CR, complete response; PR, partial response; AFP, alpha-fetoprotein; SD, stable disease.

Financial support

The studies on which this analysis was based were sponsored by Boehringer Ingelheim. Analyses were conducted by Boehringer Ingelheim. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Conflicts of interest

Tim Meyer received consulting fees from Boehringer Ingelheim during the conduct of the study; Daniel H. Palmer received travel expenses from Boehringer Ingelheim and a grant from Bayer; Julia Hocke is an employee of Boehringer Ingelheim, Pharma GmbH & Co. KG, Biberach, Germany; Arsène-Bienvenu Loembé is an employee of Boehringer Ingelheim B.V., Alkmaar, The Netherlands. Ann-Lii Cheng and Chia-Jui Yen have nothing to disclose.

Authors' contributions

TM, JH and A-BL contributed to the conception and design of these analyses. TM, DHP, A-LC and C-JY were involved in the provision of study material, patients and data acquisition. All authors were involved in data analysis and interpretation. All authors contributed to the drafting of the manuscript and all approved the final version for publication.

Abstract

Background & Aims: Response Evaluation Criteria in Solid Tumors (RECIST) has been shown to be a poor surrogate for survival benefit with targeted therapy in advanced hepatocellular carcinoma (HCC). *Methods:* We investigated whether response evaluated using modified RECIST (mRECIST) predicted overall survival (OS) using data from two Phase II clinical trials. Analyses were conducted on pooled data from 188 patients with advanced HCC treated with nintedanib or sorafenib, of whom 180 were evaluable for response. Cox regression and Kaplan–Meier survival analyses were used to explore differences in OS between the responders and non-responders according to RECIST 1.0 and mRECIST criteria. Multivariate Cox proportional hazards models, including factors known to influence survival, were used to compare survival according to RECIST and mRECIST response. *Results:* Discordance between RECIST and mRECIST evaluation was most common for assessment of partial response (12.2%) and stable disease (13.3%). OS was significantly longer in patients with response compared to patients without response – RECIST: hazard ratio (HR) 0.325 (95% confidence interval [CI] 0.130–0.815); p=0.0122; mRECIST: HR 0.544 (95% CI 0.335–0.881); p=0.0122. HRs from the multivariate models used to evaluate response by RECIST or by mRECIST as predictors of OS approached significance for RECIST (0.40 [95% CI 0.16–1.01]; p=0.053) and for mRECIST (0.62 [95% CI 0.38–1.01]; p=0.053). *Conclusions:* Response according to RECIST or mRECIST is associated with improved survival and should be considered as a valid endpoint for use in HCC clinical trials.

Word count: 237 (250 max)

Keywords: angiogenesis, hepatocellular carcinoma, overall survival

Key points

- mRECIST has been proposed as a more appropriate method to assess response in HCC, although evidence supporting use for systemic therapy is limited and conflicting; to our knowledge prospective studies have not been published
- Data from two randomized Phase II trials in patients with HCC in which mRECIST and RECIST were prospectively compared, demonstrated a higher objective response rate for mRECIST vs RECIST 1.0 (15.6 vs 4.4%)
- Response according to both mRECIST and RECIST 1.0 were independent predictors of survival
- Both mRECIST and RECIST 1.0 can be used as validated response assessments in trials of systemic therapy for HCC

Introduction

Overall survival (OS) has long been deemed to be the most clinically relevant endpoint for assessing treatment efficacy in late-phase oncology trials. As the traditional ‘gold standard’, OS constitutes a clear and unambiguous endpoint and, as such, is recommended for use as the primary endpoint in Phase III studies evaluating primary treatments in hepatocellular carcinoma (HCC) (1). One of the challenges for drug development in HCC has been the identification of a reliable surrogate for survival that can be evaluated in Phase II trials and used to justify transition to Phase III clinical development. The value of response based on Response Evaluation Criteria in Solid Tumors (RECIST) (2) has been questioned after two Phase III trials of sorafenib demonstrated a survival benefit despite a very low rate of response (3, 4). According to conventional RECIST criteria, the response rate to sorafenib was 2–3%, yet the drug improved OS, apparently by delaying progression. This observation led to the proposal that time to progression (TTP) should be used in Phase II trials (1); however, recent trials have demonstrated that TTP is also not a reliable predictor of OS (5).

More recently, modified RECIST (mRECIST) has been proposed as a more appropriate method for assessing response in HCC (6). Although conventional RECIST evaluates unidimensional measurement of target lesions, mRECIST evaluates target lesion dimensions according to the diameter of viable tumour, as defined by contrast enhancement in the arterial phase. It has been suggested that mRECIST is more reflective of the mechanism of action of targeted agents such as sorafenib, which can induce tumour necrosis without changing the overall size of the tumour. Since the publication of the mRECIST guidelines, many studies have

provided validation of use of mRECIST in the assessment of locoregional therapies, such as transcatheter embolisation (TAE)/transarterial chemoembolisation (TACE) (7-12). A recent meta-analysis of seven reports including 1357 patients demonstrated a hazard ratio (HR) for survival of 0.39 (95% confidence interval [CI]: 0.26–0.61; p<0.0001) for mRECIST responders versus non-responders (13). However, as TAE/TACE causes acute devascularisation and necrosis, mRECIST response rates are high and correlation of mRECIST with outcome is unsurprising.

In contrast, evidence to support the use of mRECIST for systemic therapy is much more limited and confined to four relatively small retrospective studies, three of which were single centre and one of which was multicentre (14-17). Two studies showed that mRECIST correlated with survival (14, 15), and in another study patients classified as responders by mRECIST had significantly better OS than patients classified as non-responders (17). However, the fourth study failed to show a significant relationship between the two (16). These conflicting findings, combined with the small, retrospective nature of these studies, warrant further investigation of the value of mRECIST in determining prognosis in HCC. Furthermore, for mRECIST to be accepted as a valid endpoint in Phase II clinical trials investigating the use of targeted therapy for patients with HCC, it must be validated prospectively in larger multicentre trials conducted in well-characterised patient populations. Of note, currently only RECIST (and not mRECIST) is recognised by the European Medicines Agency and the Food and Drug Administration as validated criteria for use in patients with advanced HCC.

In this analysis, we used data from two Phase II clinical trials in HCC, in which responses by RECIST and mRECIST were prospectively collected. Both studies

were done to evaluate the efficacy of nintedanib and sorafenib, and found that these agents had similar efficacy in this patient population in terms of time to progression and OS (18, 19). In order to provide more robust evidence to support the use of mRECIST as a surrogate for survival, we investigated whether response evaluated using mRECIST predicted OS in these studies.

Patients and methods

Study design and patients

Two multicentre, open-label, Phase I/randomised Phase II studies were conducted to evaluate the efficacy and safety of nintedanib versus sorafenib as first-line treatment of predominantly Caucasian patients (study 1199.37; NCT01004003) or Asian patients (study 1199.39; NCT00987935) with advanced HCC. The dose-escalation Phase I part of both studies was designed to establish the maximum tolerated dose (MTD) in two different groups according to liver function; patients with mild hepatic impairment (alanine/aspartate aminotransferase [ALT/AST] ≤2 times upper limit of normal (ULN) and Child–Pugh score 5–6) and patients with moderate hepatic impairment (ALT or AST >2 to ≤5 times ULN or Child–Pugh score 7). In both trials, the MTD of nintedanib was determined to be 200 mg bid in both liver function groups investigated. Data from the Phase I part of these studies were not included in the analysis reported here as complete efficacy data were not collected. In Phase II, patients were randomised 2:1 to receive nintedanib 200 mg bid or sorafenib 400 mg bid continuously in 28-day cycles until intolerable adverse events or progressive disease (PD). Treatment beyond PD was allowed at the discretion of the investigator.

Patient randomisation was stratified by presence of extrahepatic spread (EHS) and/or macrovascular invasion (MVI) (MVI and/or EHS present vs both absent).

To be eligible for inclusion in either study, adult (≥18 years) patients with mild hepatic impairment (Child–Pugh score 5–6, ALT/AST levels ≤2 times ULN) were required to have advanced HCC not amenable to curative/locoregional therapy, ≥1 measurable lesion by RECIST 1.0, Eastern Cooperative Oncology Group performance score (ECOG PS) ≤2, >4 weeks since most recent local therapy and no prior systemic therapy for HCC; further details of the inclusion/exclusion criteria are reported elsewhere (18, 19). Both trials were approved by the following health authorities and independent ethics committees or institutional review boards at each country/centre based on local regulations: Medicines and Medical Devices Agency (Austria), Agence Nationale de Sécurité du Médicament et des Produits de Santé (France), Federal Institute for Drugs and Medical Devices (Germany), National Institute of Pharmacy (Hungary), Central Committee Research Involving Human Subjects (The Netherlands), Registration of Medicinal Products, Medical Devices and Biocidal Products (Poland), National Medicines Agency (Romania), Ministry of Food and Drug Safety (South Korea), Department of Health (Taiwan), Medicines and Healthcare Products Regulatory Agency (UK) and Food and Drug Administration (US). They were conducted in accordance with the guiding principles of the Declaration of Helsinki, the International Council for Harmonisation Harmonised Tripartite Guideline for Good Clinical Practice, and local legislation. All patients provided written informed consent.

In both trials, the primary endpoint was TTP by central independent review using RECIST 1.0. The secondary endpoints were centrally reviewed objective

tumour response and progression-free survival (PFS) using RECIST 1.0, and OS – defined as the duration from date of randomisation to the date of death. Further endpoints included TTP, objective tumour response and PFS by central review using mRECIST.

Radiological assessment

Tumour assessment was performed at screening and every 4 weeks for the first 16 weeks after the start of treatment and every 8 weeks thereafter. Computed tomography/magnetic resonance imaging of the chest, abdomen and pelvis was mandatory, whereas assessment of other body parts was performed as clinically indicated; scans were to be performed within 5 days prior to the scheduled visit. Assessment of tumour response was done by blinded central radiological review using both RECIST 1.0 (2) and mRECIST criteria for HCC (6). One to 10 target lesions (not exceeding five lesions per organ) were identified at screening. Objective response (OR) was defined as a best response of complete response (CR) or partial response (PR); this categorisation was done according to RECIST and according to mRECIST. Patients with unknown response due to missing data were not evaluable and were excluded from our analyses.

Statistical analysis

To increase the number of responses, both treatment arms were pooled for this analysis, as study results suggest that nintedanib and sorafenib treatment are comparable (18, 19). Survival data were analysed using the Kaplan–Meier method. Differences in OS between responders and non-responders were evaluated using the

Cox proportional hazards model stratified by MVI, EHS, or both present versus both absent, and study, where a HR of less than one favours response. The Score Test with Breslow method for tied observation times was used to determine p-values.

Multivariate Cox proportional hazards models were used to compare survival according to RECIST and mRECIST response. Analyses were conducted, including factors known to influence survival. Baseline characteristics that have consistently shown to be prognostic indicators for OS in patients with advanced HCC treated with sorafenib include ECOG PS, extent of tumour burden (defined as presence or absence of MVI, EHS, or both), and baseline levels of alpha-fetoprotein (AFP), albumin and total bilirubin (4). Two multivariate selection analyses, stratified by study, were conducted to identify which individual baseline variables predict survival at the 0.2 level of significance using the Likelihood Ratio Chi-Square Statistic with Breslow method for tied observation times to determine p-values. The selection models included RECIST or mRECIST response as a time-dependent covariate and other baseline variables (ECOG PS [0 vs >0], MVI [absent vs present], EHS [absent vs present], baseline albumin [<36 g/L vs ≥36 g/L], baseline bilirubin [<17 µmol/L vs ≥17 µmol/L] and age [continuous variable]) as time-independent covariates. Baseline AFP levels were not included due to the large amount of missing data in Study 1.

Results

Patient disposition and demographics

A total of 188 patients were randomised to treatment with nintedanib (1199.37: n=62; 1199.39: n=63) or sorafenib (n=31; n=32). At the time of data analysis, the majority of

patients had discontinued treatment with nintedanib (97.6%) or sorafenib (98.4%); three nintedanib patients and one sorafenib patient remained on treatment at this point. Full details of baseline demographics and disease characteristics by treatment group in each trial have been reported previously and were generally balanced between treatment groups, supporting our approach to pool treatment arms. Baseline demographics and disease characteristics for the pooled population are shown in Table 1. The majority of patients were male (84%) and the median age was 63 years.

Response assessment

Of the 188 patients treated in these trials, 180 patients (95.7%) were evaluable for RECIST and mRECIST response. Six patients in the nintedanib group and two patients in the sorafenib group had unknown response or were not evaluable – these patients were excluded from the analyses. Best overall response by RECIST and mRECIST is reported in Table 2, and concordance/discordance between RECIST and mRECIST response is shown in Table 3. Discordance between RECIST and mRECIST evaluation was most common for assessment of PR (n=22; 12.2%) and stable disease (SD) (n=24; 13.3%). Of note, among the 141 patients (nintedanib n=89; sorafenib n=52) who were classified as having SD by RECIST, 21 (14.9%; nintedanib n=11; sorafenib n=10) were reclassified as having PR by mRECIST. Overall, there was good concordance between RECIST and mRECIST when used to assess PD, and there were only two patients in whom there was discordance, giving rise to a PD rate of 17.2% by RECIST and 16.1% by mRECIST. Variation in best percentage changes from baseline in target lesions is shown in Fig. 1. Of note, five patients had a 100% reduction from baseline in the sum of target lesion dimensions

according to mRECIST criteria (Fig. 1B); these patients did not qualify as having a CR in the OR assessment due to non-target lesions and new lesion assessments.

Survival analysis according to radiological response

At the time of analysis, 140 patients (77.8%) had died and the median OS for the total population was 11.4 months (interquartile range [IQR] 6.6–20.5) (Fig. 2A). Kaplan–Meier survival curves by RECIST and mRECIST response are shown in Fig. 2B and 2C, respectively. OS was significantly longer in patients with RECIST response (n=8) compared to patients without RECIST response (23.6 months [IQR 12.7–n.c.] vs 11.2 months [IQR 6.5–19.8]; HR 0.325 [95% CI 0.130–0.815]; p=0.0122). Similar results were found in patients with and without mRECIST response, in which the median OS was 16.7 months (IQR 10.7–28.4) in those with mRECIST response versus 10.9 months (IQR 6.2–18.2) in those without mRECIST response; HR 0.544 (95% CI 0.335–0.881); p=0.0122. The 1-year survival rate was 75.0% in patients with RECIST response and 64.3% in patients with mRECIST response, with 2-year survival rates of 37.5% and 31.3%, respectively. The median OS in patients with both a RECIST and mRECIST response (n=7) was 23.9 months (IQR 17.3–n.c.).

Patients who experienced PD as best OR by RECIST had a median OS of 4.3 months and patients who experienced PD as best OR by mRECIST had a median

OS of 4.3 months (see supplementary Fig. 1). Median OS in patients with RECIST and mRECIST SD was 12.4 months and 11.9 months, respectively (see supplementary Fig. 1). Survival by concordance (i.e. in patients who exhibited a response according to RECIST and mRECIST, or according to only one of RECIST or mRECIST) is shown in Supplementary Table 1.

Analyses of RECIST and mRECIST as predictors of survival

HRs for OS from the multivariate model including RECIST or mRECIST and baseline variables are shown in Fig. 3. In both models, the presence of MVI and EHS at baseline were associated with worse OS. Neither RECIST response nor mRECIST response was a statistically significant predictor of survival, although both approached significance.

Analysis of survival models

Survival models were evaluated for relative quality. Two criteria for model selection were evaluated: the Akaike information Criterion (AIC) and the Schwarz Bayesian Information Criterion (SBC). For the final model selecting MVI, EHS and time-dependent RECIST response, the AIC was 1056.883 and the SBC was 1065.708. For the final model selecting MVI, EHS and time-dependent mRECIST response, the AIC was 1057.546 and the SBC was 1066.370. In summary, both the AIC and SBC were similar between the final survival models; this indicates that it cannot be concluded that either one of RECIST or mRECIST is a better predictor of survival.

Discussion

Previous research has suggested that RECIST, the standard method of response assessment for solid tumours, does not always represent the most appropriate tool for evaluation of response in patients with HCC, as it does not account for changes in lesion density that occur with targeted treatments. This is supported by observations from studies conducted with sorafenib, in which it was observed that significant improvements in survival compared to placebo were not accompanied by the expected differences in RECIST response rates; reported response rates with sorafenib by RECIST were only 2–3% (3, 4).

Here, we have analysed data from two randomised Phase II clinical trials in patients with HCC to determine whether response assessed using mRECIST can predict OS. Our findings show that both RECIST and mRECIST are of value in predicting long-term survival in patients with HCC treated with antiangiogenic agents, and that those patients with an OR by RECIST or mRECIST had significantly better survival compared to patients who only achieve SD or PD as best treatment response. The AIC and SBC were similar between the survival model that included RECIST as a variable and the survival model that included mRECIST as a variable; this indicates that it cannot be concluded that either RECIST or mRECIST response predicts survival better than the other.

These observations are in agreement with previously reported retrospective studies in patients treated with sorafenib that have also shown RECIST 1.1 and mRECIST response to successfully predict survival advantages compared to those without (14, 15). In a third study, patients classified as responders by mRECIST had significantly better OS than patients classified as non-responders (17). A fourth study conducted in 156 patients with HCC treated with sorafenib failed to show OS benefits

for RECIST 1.1 and mRECIST, although classification of response by mRECIST was found to be more strongly associated with OS than RECIST 1.1 (16).

Our analysis shows that mRECIST criteria identified more patients with a response to treatment than RECIST (15.6% vs 4.4%). This would be expected based on differences in the determination of response between RECIST and mRECIST, and higher response rates have also been consistently reported in retrospective evaluations (23% mRECIST vs 2% RECIST 1.1 (14); 28% mRECIST vs 3% RECIST 1.1 (15); 23% mRECIST vs 10% RECIST (16); 13.1% mRECIST vs 7.8% RECIST 1.1 (17). This finding has important implications, when combined with the observation that mRECIST response can reliably identify a subgroup of patients with a significantly better prognosis than those without response. Our results suggest that mRECIST may offer a suitable alternative to RECIST in Phase II clinical trials, in which detection of an efficacy signal is paramount. We also demonstrate that patients with PD measured by RECIST or mRECIST have a poor outcome, with a median survival of 4.3 months by either criteria. This observation suggests that treatment beyond radiological progression is not warranted and that patients should be actively monitored for radiological progression rather than waiting for symptomatic progression.

The limitation of the relatively small dataset used in this analysis should be taken into account when interpreting these findings and considering their clinical utility. Nonetheless, our analyses have a number of strengths. First, analyses were based on prospectively collected data from two well-designed trials complying with the European Association for the Study of the Liver recommendations for the conduct of trials in patients with HCC (20). Second, all radiological assessments of response

were performed by a central imaging unit, improving the consistency of response assessment. Third, analysis showed that mRECIST response remains an independent surrogate for survival when other known prognostic factors are considered. However, it should be noted that AFP levels, of known independent prognostic significance in HCC, were not available from the dataset for inclusion in our analysis. Finally, it should be considered that the RECIST criteria used in these trials (RECIST 1.0) are not the most up-to-date RECIST criteria, although differences between versions 1.0 and 1.1 are minor and are not expected to have affected the overall conclusions of this study. Additionally, other methods of response assessment have been proposed, such as the Response Evaluation Criteria in Cancer of the Liver (RECICL) (21) and the Choi criteria (22), although these methods have yet to be prospectively validated in HCC. Both methods are more complex than RECIST; RECICL requires bidimensional measurements of tumour size, whereas the Choi criteria requires evaluation of tumour density by selecting a region of interest and, as such, is less easy to standardise for routine use.

In conclusion, our findings show that objective mRECIST response is an independent marker for OS and support its use as a valid endpoint for use in HCC clinical trials of systemic therapy in HCC. To our knowledge, this is the first prospective validation of mRECIST and provides robust evidence to support management guidelines.

Acknowledgements

We thank all the patients who participated in both studies and their families, all the study centre staff, and the worldwide teams at Boehringer Ingelheim, Parexel and Perceptive. Suzanne Patel of inVentiv Medical Communications provided editorial assistance during the preparation of this manuscript, with funding from Boehringer Ingelheim.

References

1. Llovet J M, Di Bisceglie A M, Bruix J, et al. Design and endpoints of clinical trials in hepatocellular carcinoma. *J Natl Cancer Inst* 2008; **100**: 698-711.

2. Therasse P, Arbuck S G, Eisenhauer E A, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst* 2000; **92**: 205-16.

3. Cheng A L, Kang Y K, Chen Z, 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.

4. Llovet J M, Ricci S, Mazzaferro V, et al. Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med* 2008; **359**: 378-90.

5. Llovet J M, Decaens T, Raoul J L, 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**: 3509-16.

6. Lencioni R, Llovet J M. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. *Semin Liver Dis* 2010; **30**: 52-60.

7. Kim C J, Kim H J, Park J H, et al. Radiologic response to transcatheter hepatic arterial chemoembolization and clinical outcomes in patients with hepatocellular carcinoma. *Liver Int* 2014; **34**: 305-12.

8. Li H, Guo Z, Si T, Wang H. EASL and mRECIST responses are independent predictors of survival in hepatocellular carcinoma patients treated with cryoablation. *Eur J Gastroenterol Hepatol* 2013; **25**: 620-27.

9. Prajapati H J, Spivey J R, Hanish S I, 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**: 965-73.

10. Gillmore R, Stuart S, Kirkwood A, et al. EASL and mRECIST responses are independent prognostic factors for survival in hepatocellular cancer patients treated with transarterial embolization. *J Hepatol* 2011; **55**: 1309-16.

11. Jung E S, Kim J H, Yoon E L, et al. Comparison of the methods for tumor response assessment in patients with hepatocellular carcinoma undergoing transarterial chemoembolization. *J Hepatol* 2013; **58**: 1181-7.

12. Kim B K, Kim K A, Park J Y, 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**: 826-34.

13. Vincenzi B, Di Maio M, Silletta M, 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.

14. Edeline J, Boucher E, Rolland Y, 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**: 147-56.

15. Ronot M, Bouattour M, Wassermann J, 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.

16. Arizumi T, Ueshima K, Takeda H, et al. Comparison of systems for assessment of post-therapeutic response to sorafenib for hepatocellular carcinoma. *J Gastroenterol* 2014; **49**: 1578-87.

17. Takada J, Hidaka H, Nakazawa T, 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. *BMS Res Notes* 2015; **8**: 609.

18. Cheng A L, Yen C-J, Kim T-Y, et al. Efficacy and safety of nintedanib versus sorafenib in Asian patients with advanced hepatocellular carcinoma (HCC): A randomized phase II trial. *J Clin Oncol* 2015; **33 (Suppl. 3)**: Abstract 339.

19. Palmer D H, Ma T Y, Peck-Radosavljevic M, et al. Randomized phase II trial comparing the efficacy and safety of nintedanib versus sorafenib in patients with advanced hepatocellular carcinoma (HCC). *J Clin Oncol* 2015; **33 (Suppl. 3)**: Abstract 238.

20. European Association for the Study of the Liver, European Organisation for Research and Treatment of Cancer. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol* 2012; **56**: 908-43.

21. Kudo M, Kubo S, Takayasu K, et al. Response Evaluation Criteria in Cancer of the Liver (RECICL) proposed by the Liver Cancer Study Group of Japan (2009 Revised Version). *Hepatol Res* 2010; **40**: 686-92.

22. Choi H, Charnsangavej C, Faria S C, et al. Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate:

proposal of new computed tomography response criteria. *J Clin Oncol* 2007; **25**: 1753-9.

Tables and figures

Table 1. Patient baseline demographics and clinical characteristics (all randomised patients)

Characteristic		Study 1 (n=93)	Study 2 (n=95)	Combined (n=188)
Age, years	Median (range)	66.0 (28–86)	59 (32–84)	63 (28–86)
Male sex; n (%)		74 (79.6)	83 (87.4)	157 (83.5)
Race; n (%)	Indian	3 (3.2)	–	3 (1.6)
	Taiwanese or Chinese	1 (1.1)	64 (67.4)	65 (34.6)
	Korean	–	31 (32.6)	31 (16.5)
	Black	1 (1.1)	–	1 (0.5)
	Caucasian	81 (87.1)	–	81 (43.1)
	Missing	7 (7.5)	–	7 (3.7)
Time since diagnosis, months; median (range)		2.53 (0–101.4)	7.13 (0.1–131.3)	4.6 (0–131.3)
ECOG PS; n (%)	0	50 (53.8)	53 (55.8)	103 (54.8)
	1	38 (40.9)	41 (43.2)	79 (42.0)
	2	5 (5.4)	1 (1.1)	6 (3.2)
Child–Pugh score; n (%)	5	65 (69.9)	62 (65.3)	127 (67.6)
	6	27 (29.0)	32 (33.7)	59 (31.4)
	7[a]	1 (1.1)	1 (1.1)	2 (1.1)
BCLC stage; n (%)	0	1 (1.1)	–	1 (0.5)
	A	1 (1.1)	2 (2.1)	3 (1.6)
	B	22 (23.7)	10 (10.5)	32 (17.0)
	C	68 (73.1)	83 (87.4)	151 (80.3)
	D	1 (1.1)	–	1 (0.5)
MVI; n (%)		31 (33.3)	40 (42.1)	71 (37.8)
EHS; n (%)		61 (65.6)	65 (68.4)	126 (67.0)
Location of EHS; n (%)	Bone	11 (11.8)	9 (9.5)	20 (10.6)
	Lung	22 (23.7)	35 (36.8)	57 (30.3)
	Lymph	35 (37.6)	33 (34.7)	68 (36.2)
	Other	18 (19.4)	25 (26.3)	43 (22.9)
Aetiology of parenchymal liver disease; n (%)	Alcohol related	13 (14.0)	4 (4.2)	17 (9.0)
	HBV related	11 (11.8)	60 (63.2)	71 (37.8)
	HCV related	21 (22.6)	15 (15.8)	36 (19.1)
	HBV + HCV related	0	3 (3.2)	3 (1.6)

	Unknown	31 (33.3)	11 (11.6)	42 (22.3)
	Other	17 (18.3)	2 (2.1)	19 (10.1)
Parenchymal liver disease; n (%)	Chronic hepatitis	13 (14.0)	26 (27.4)	39 (20.7)
	Steatofibrosis	5 (5.4)	0	5 (2.7)
	Cirrhosis	49 (52.7)	61 (64.2)	110 (58.5)
	No evidence	16 (17.2)	6 (6.3)	22 (11.7)
	Unknown	9 (9.7)	1 (1.1)	10 (5.3)
	Other	1 (1.1)	1 (1.1)	2 (1.1)
Type of local therapy; n (%)	Complete surgical resection	12 (12.9)	9 (9.5)	21 (11.2)
	RFA	2 (2.2)	4 (4.2)	6 (3.2)
	TACE	29 (31.2)	48 (50.5)	77 (41.0)
	RT	1 (1.1)	2 (2.1)	3 (1.6)
	Other	6 (6.5)	11 (11.6)	17 (9.0)
Stratification group; n (%)	I: EHS and/or MVI present	72 (77.4)	82 (86.3)	154 (81.9)
	II: EHS and MVI both absent	21 (22.6)	13 (13.7)	34 (18.1)
AFP group at baseline; n (%)	≤20 µg/L	12 (12.9)	26 (27.4)	38 (20.2)
	>20 µg/L	18 (19.4)	69 (72.6)	87 (46.3)
	Missing	63 (67.7)	0	63 (33.5)

ECOG PS, Eastern Cooperative Oncology Group performance status; BCLC, Barcelona Clinic Liver Cancer; MVI, macrovascular invasion; EHS, extrahepatic spread; HBV, hepatitis B virus; HCV, hepatitis C virus; RFA, radiofrequency ablation; TACE, transarterial chemoembolisation; RT, radiotherapy; AFP, alpha-fetoprotein.

[a]Patients with a Child–Pugh score of 7 were protocol violations.

Table 2. Best overall response according to RECIST and mRECIST

Best response, n (%)	RECIST (n=180)	mRECIST (n=180)
Objective response	8 (4.4)	28 (15.6)
Complete response	2 (1.1)	2 (1.1)
Partial response	6 (3.3)	26 (14.4)
Stable disease	141 (78.3)	123 (68.3)
Progressive disease	31 (17.2)	29 (16.1)

RECIST, Response Evaluation Criteria in Solid Tumors; mRECIST, modified RECIST.

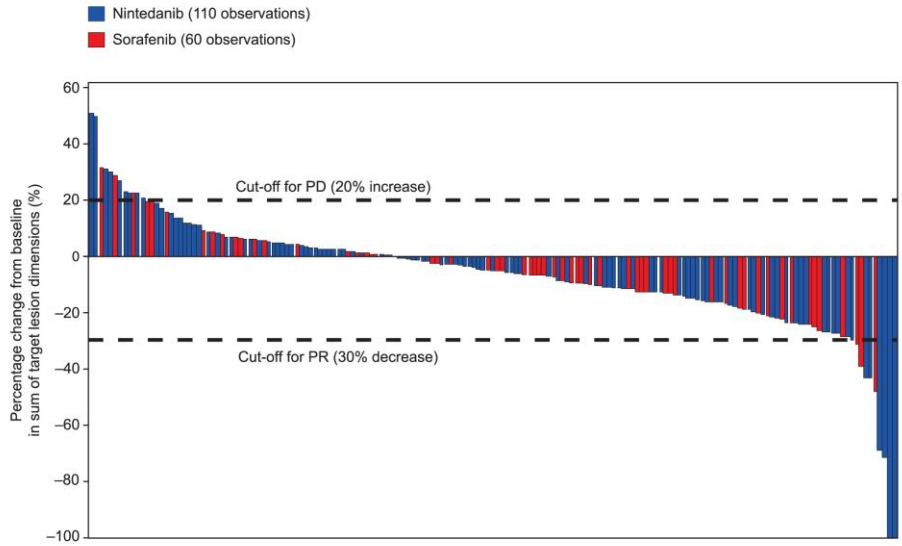
Table 3. Concordance and discordance between RECIST and mRECIST response

	Type of response (N=180)				
	Objective response	Complete response	Partial response	Stable disease	Progressive disease
Concordance between RECIST and mRECIST response, n (%)	7 (3.9)	2 (1.1)	5 (2.8)	120 (66.7)	29 (16.1)
Concordance between RECIST non-response and mRECIST non-response, n (%)	151 (83.9)	178 (98.9)	153 (85.0)	36 (20.0)	149 (82.8)
Discordance; RECIST response/mRECIST non-response, n (%)	1 (0.6)	0 (0.0)	1 (0.6)	21 (11.7)	2 (1.1)
Discordance; RECIST non-response/mRECIST response, n (%)	21 (11.7)	0 (0.0)	21 (11.7)	3 (1.7)	0 (0.0)

RECIST, Response Evaluation Criteria in Solid Tumors; mRECIST, modified RECIST.

Fig. 1. Waterfall plots showing best percentage change from baseline in sum of target lesion dimensions. (A) According to RECIST (n=170) criteria; (B) according to mRECIST (n=166) criteria. The total number of patients is less than the number of evaluable patients, as some patients had no evaluable target lesions according to central independent review. PD, disease progression; PR, partial response; RECIST, Response Evaluation Criteria in Solid Tumors; mRECIST, modified RECIST.

(A)



(B)

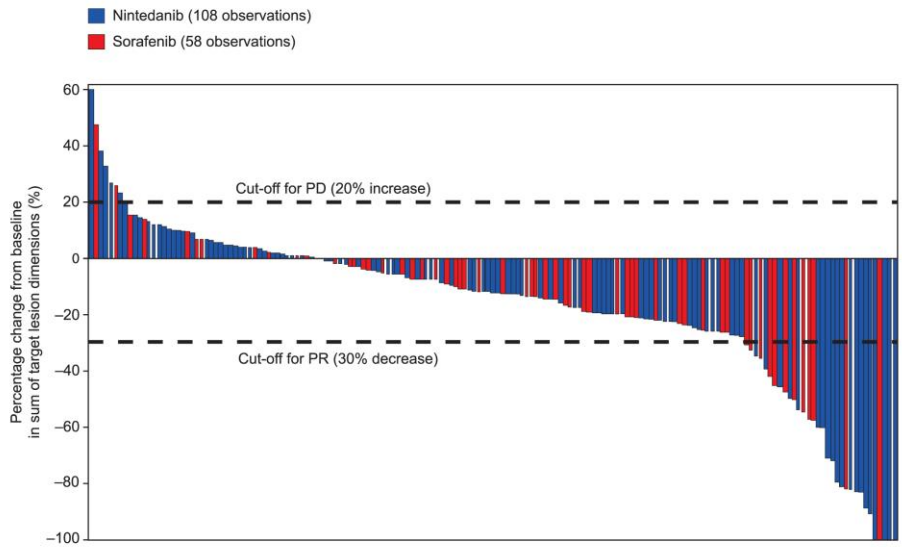
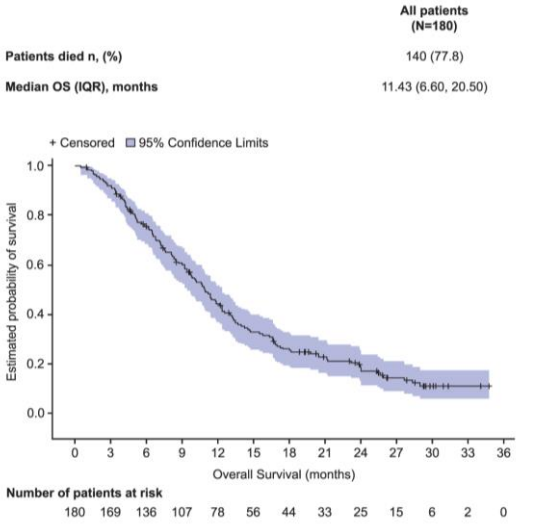
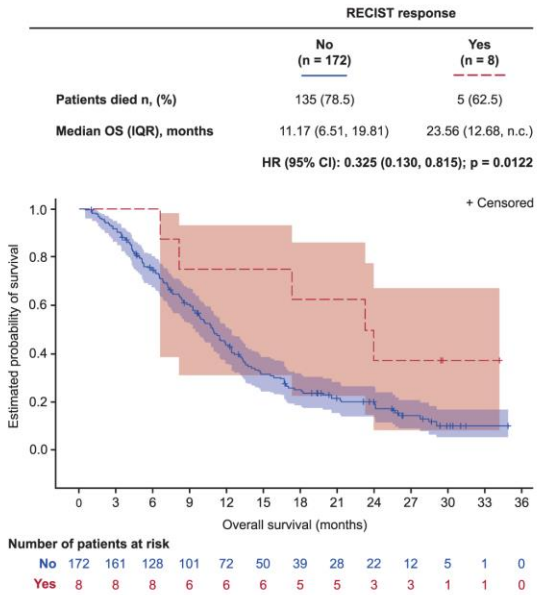


Fig. 2. Probability of overall survival in all patients and by RECIST/mRECIST response. (A) All patients; (B) By RECIST response; (C) By mRECIST response. Patients who had not died or who were lost to follow-up were censored on the last date on which they were known to have been alive. Shaded areas represent the 95% confidence limits. Hazard ratios represent patients with response compared to patients without response (yes vs no). RECIST, Response Evaluation Criteria in Solid Tumors; mRECIST, modified RECIST; OS, overall survival; CI, confidence interval; n.c., not calculable; HR, hazard ratio.

(A)



(B)



(C)

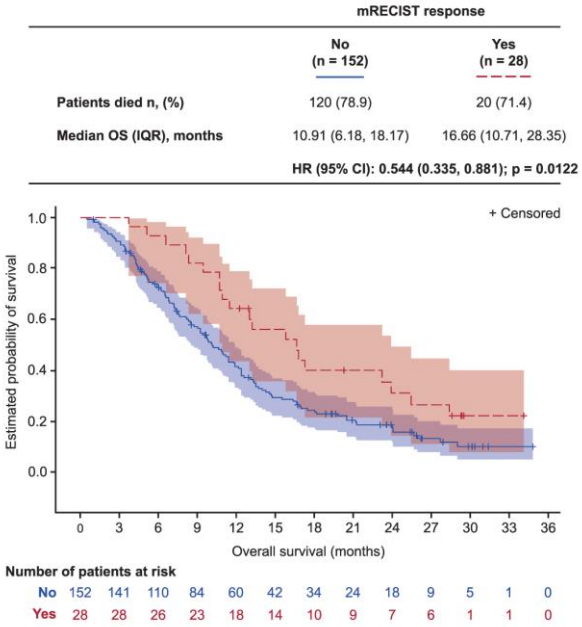


Fig. 3. Hazard ratios for overall survival from multivariate analyses. RECIST, Response Evaluation Criteria in Solid Tumors; modified RECIST; CI, confidence interval.