Lenvatinib versus sorafenib for first-line treatment of unresectable hepatocellular carcinoma: patient-reported outcomes from a randomised, open-label, non-inferiority, phase 3 trial

Arndt Vogel, Shukui Qin, Masatoshi Kudo, Yun Su, Stacie Hudgens, Tatsuya Yamashita, Jung-Hwan Yoon, Laetitia Fartoux, Krzysztof Simon, Carlos López, Max Sung, Kalgi Mody, Tatsuroh Ohtsuka, Toshiyuki Tamai, Lee Bennett, Genevieve Meier, Valery Breder

Summary

Background Hepatocellular carcinoma is the third-leading cause of cancer-related death worldwide. Preservation of health-related quality of life (HRQOL) during treatment is an important therapeutic goal. The aim of this study was to evaluate the effect of treatment with lenvatinib versus sorafenib on HRQOL.

Methods REFLECT was a previously published multicentre, randomised, open-label, non-inferiority phase 3 study comparing the efficacy and safety of lenvatinib versus sorafenib as a first-line systemic treatment for unresectable hepatocellular carcinoma. Eligible patients were aged 18 years or older with unresectable hepatocellular carcinoma and one or more measurable target lesion per modified Response Evaluation Criteria in Solid Tumors criteria, Barcelona Clinic Liver Cancer stage B or C categorisation, Child-Pugh class A, Eastern Cooperative Oncology Group (ECOG) performance status of 1 or lower, and adequate organ function. Patients were randomly assigned (1:1) via an interactive voice–web response system; stratification factors for treatment allocation included region; macroscopic portal vein invasion, extrahepatic spread, or both; ECOG performance status; and bodyweight. Patient-reported outcomes (PROs), collected at baseline, on day 1 of each subsequent cycle, and at the end of treatment, were evaluated in post-hoc analyses of secondary and exploratory endpoints in the analysis population, which was the subpopulation of patients with a PRO assessment at baseline. A linear mixed-effects model evaluated change from baseline in PROs, including European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire Core 30 (QLQ-C30) and hepatocellular carcinoma-specific QLQ-HCC18 scales (both secondary endpoints of the REFLECT trial). Time-to-definitive-deterioration analyses were done based on established thresholds for minimum differences for worsening in PROs. Responder analyses explored associations between HRQOL and clinical response. This study is registered with ClinicalTrials.gov, NCT01761266.

Findings Of 954 eligible patients randomly assigned to lenvatinib (n=478) or sorafenib (n=476) between March 14, 2013, and July 30, 2015, 931 patients (n=468 for lenvatinib; n=463 for sorafenib) were included in this analysis. Baseline PRO scores reflected impaired HRQOL and functioning and considerable symptom burden relative to full HRQOL. Differences in overall mean change from baseline estimates in most PRO scales generally favoured the lenvatinib over the sorafenib group, although the differences were not nominally statistically or clinically significant. Patients treated with lenvatinib experienced nominally statistically significant delays in definitive, meaningful deterioration on the QLQ-C30 fatigue (hazard ratio [HR] 0·83, 95% CI 0·69–0·99), pain (0·80, 0·66–0·96), and diarrhoea (0·52, 0·42–0·65) domains versus patients treated with sorafenib. Significant differences in time to definitive deterioration were not observed for other QLQ-C30 domains, and there was no difference in time to definitive deterioration on the global health status/QOL score (0·89, 0·73–1·09). For most PRO scales, differences in overall mean change from baseline estimates favoured responders versus non-responders. Across all scales, HRs for time to definitive deterioration were in favour of responders; median time to definitive deterioration for responders exceeded those for non-responders by a range of 4·8 to 14·6 months.

Interpretation HRQOL for patients undergoing treatment for unresectable hepatocellular carcinoma is an important therapeutic consideration. The evidence of HRQOL benefits in clinically relevant domains support the use of lenvatinib compared with sorafenib to delay functional deterioration in advanced hepatocellular carcinoma.

Funding Eisai and Merck Sharp & Dohme.

Lancet Gastroenterol Hepatol 2021

Published **Online** June 1, 2021 https://doi.org/10.1016/S2468-1253(21)00110-2

See **Online/Comment** https://doi.org/10.1016/S2468-1253(21)00139-4

Hannover Medical School, Hannover, Germany (Prof A Vogel MD); **Nanjing Bayi Hospital, Nanjing, Jiangsu, China** (Prof S Qin MD); **Kindai University Faculty of Medicine, Osaka, Japan** (Prof M Kudo MD); **Eisai, Woodcliff Lake, NJ, USA** (Y Su MD, K Mody MD, T Ohtsuka MS, T Tamai MS, G Meier Dphar); **Clinical Outcome Solutions, Tucson, AZ, USA** (S Hudgens MSc); **Kanazawa University Hospital, Kanazawa, Japan** (T Yamashita MD); **Seoul National University Hospital, Seoul, South Korea** (J-H Yoon MD); **Hôpital Saint Joseph, Paris, France** (L Fartoux MD); **Department of Infectious Diseases and Hepatology, Wroclaw Medical University, Wroclaw, Poland** (K Simon MD); **Hospital Universitario Marques de Valdecilla, IDIVAL, Santander, Spain** (C López PhD); **Tisch Cancer Institute at Mount Sinai, New York, NY, USA** (M Sung MD); **RTI Health Solutions, Durham, NC, USA** (L Bennett MS); **FSBSI N N Blokhin Russian Cancer Research Center, Moscow, Russia** (V Breder DrMedSci)

Correspondence to: Prof Arndt Vogel, Hannover Medical School, 30625 Hannover, Germany **vogel.arndt@mh-hannover.de**

Research in context

Evidence before this study

Hepatocellular carcinoma is the most common primary liver malignancy worldwide and the third most common cause of cancer-related death. Recent therapeutic advances have incrementally increased overall survival of patients with unresectable hepatocellular carcinoma. In the multicentre, randomised, open-label, non-inferiority phase 3 REFLECT study (NCT01761266) comparing the efficacy and safety of lenvatinib versus sorafenib as a first-line systemic treatment for unresectable hepatocellular carcinoma, lenvatinib showed non-inferiority in overall survival and statistically significant and clinically meaningful improvement in progression-free survival, time to progression, and objective response rate relative to sorafenib. In addition to understanding the clinical benefits of therapy, evaluation of patients' health-related quality of life (HRQOL) before and during treatment using valid and reliable measures is also important. Literature reviews for evidence related to the development of the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire Core 30 (QLQ-C30) and hepatocellular carcinoma-specific quality of life questionnaire (QLQ-HCC18) to identify potentially relevant HRQOL issues in hepatocellular carcinoma were done in Embase (1980–June, 2002; Jan 1, 1996–Aug 1, 2016), MEDLINE (1966–June, 2002), CINAHL (1982–April, 2002); and PsycINFO (Jan 1, 1996–Aug 1, 2016) databases. The searches were restricted to English-language publications. We used the search terms "hepatocellular carcinoma" and "surgery", "QoL", "questionnaire", "chemotherapy", "alcohol ablation", "chemoembolization", "physical distress", "psychological distress", "psychosocial distress", and "physical symptoms". From the 2055 articles identified in these searches, 33 potentially relevant HRQOL issues in hepatocellular carcinoma were identified. Across iterative qualitative research with clinical experts and patients, the items in the EORTC QLQ-HCC18 and the associated EORTC Quality of Life Questionnaire Core 30 (QLQ-C30) were shown to be reflective of the content and variation in symptoms and issues raised by patients with hepatocellular carcinoma. Despite the availability of valid HRQOL measures in hepatocellular carcinoma, the effect of first-line treatment on HRQOL for patients with unresectable hepatocellular carcinoma is unclear, as only scarce HRQOL data from prospective clinical trials in this setting have been previously reported.

Added value of this study

This post-hoc analysis of secondary and exploratory endoints from REFLECT was done to evaluate the effects of lenvatinib compared with sorafenib on HRQOL, a key therapeutic consideration. Analyses of patient-reported outcomes revealed clinically meaningful differences between lenvatinib and sorafenib arms in several domains. Differences in overall least-squares mean estimates for change from baseline in EORTC QLQ-C30 and QLQ-HCC18 scale scores generally favoured lenvatinib over sorafenib, and treatment with lenvatinib resulted in nominally statistically significant delays in definitive, meaningful deterioration in QLQ-C30 fatigue, pain, and diarrhoea compared with sorafenib.

Implications of all the available evidence

As the therapeutic landscape for unresectable hepatocellular carcinoma evolves, it is important to understand the effect of the available treatment options on HRQOL. The results of this analysis contribute to the evidence base regarding patients' experiences with systemic therapy for unresectable hepatocellular carcinoma. Additional clinical trials are ongoing to evaluate further the HRQOL effect of lenvatinib, in addition to its efficacy and safety, in patients with hepatocellular carcinoma to inform treatment decisions and treatment sequencing considerations.

Introduction

Hepatocellular carcinoma is the most common primary liver malignancy worldwide.[1] The incidence of hepatocellular carcinoma has more than doubled over the last 20 years and is expected to increase in the coming decades.[2] It represents a leading cause of cancer mortality: nearly 781 600 deaths related to liver cancer occur annually, making it the third most common cause of cancer-related death.[1,3,4] Hepatocellular carcinoma is often associated with chronic liver disease and is frequently diagnosed when potentially curative surgical resection is no longer a treatment option.[5] In 2007, the tyrosine kinase inhibitor (TKI) sorafenib was the first systemic therapy approved for unresectable hepatocellular carcinoma. Only recently have systemic treatment options for unresectable hepatocellular carcinoma begun to expand; the TKI lenvatinib and the immunotherapy–targeted therapy combination of atezolizumab plus bevacizumab are additional systemic therapies to have been approved in the USA for the first-line treatment of this disease.[3,6] Even as the treatment landscape for unresectable hepatocellular carcinoma evolves and novel therapies undergo investigation, TKIs will hold a place in sequential treatment of advanced disease. In addition, lenvatinib in combination with the PD-1 inhibitor pembrolizumab has shown promising activity in solid tumours including hepatocellular carcinoma and is being evaluated in several ongoing trials (NCT03713593, NCT04425226, NCT03006926, NCT04246177) in intermediate and advanced stages of the disease.

Although recent therapeutic advances have incrementally increased overall survival of patients with unresectable hepatocellular carcinoma, the evaluation of patients' health-related quality of life (HRQOL) before and during treatment using valid and reliable measures is also key in therapeutic decision-making.[7–9] Treatments might negatively affect HRQOL because of common adverse events.[3] On the other hand, if these treatments achieve

tumour response, symptoms could decrease, thereby improving HRQOL. The impact of first-line treatment on HRQOL for patients with unresectable hepatocellular carcinoma is unclear, as only limited HRQOL data from prospective clinical trials in this setting have been previously reported.[10–12]

The REFLECT trial demonstrated that the TKI lenvatinib showed non-inferiority versus sorafenib in terms of overall survival and statistically significant and clinically meaningful improvement in progression-free survival, time to progression, and objective response rate.[3] Median overall survival was 13·6 months (95% CI 12·1–14·9) for lenvatinib and 12·3 months (10·4–13·9) for sorafenib (hazard ratio [HR] 0·92, 95% CI 0·79–1·06). Median progression-free survival (as per independent review according to modified Response Evaluation Criteria in Solid Tumors [RECIST]) for lenvatinib exceeded that for sorafenib (7·3 [5·6–7·5] months *vs* 3·6 [3·6–3·7] months), as did time to progression (7·4 [7·2–9·1] months *vs* 3·7 [3·6–3·9] months) and objective response rate (40·6% [36·2–45·0] *vs* 12·4% [9·4–15·4]). Discontinuations due to adverse events were similar between treatment groups (63 [13%] of 478 patients receiving lenvatinib and 43 [9%] of 476 patients receiving sorafenib), as were discontinuations attributable to radiological or clinical progression (343 [72%] patients *vs* 380 [80%] patients), patient's choice or withdrawal of consent (37 [8%] patients *vs* 19 [4%] patients), loss to follow-up (three [1%] patients *vs* one [<1%] patient), and other reasons (three [1%] patients *vs* seven [2%] patients).[3] Treatment efficacy, as well as patients' experiences with toxicities during treatment, can influence their HRQOL.

A secondary objective of REFLECT was to compare the effects of treatment on HRQOL using the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire Core (QLQ-C30) and hepatocellular carcinoma-specific (QLQ-HCC18) modules. An exploratory objective of the study was to assess the effect of treatments on health states using the EuroQoL Five-Dimension Scale (EQ-5D).

Methods

Study design and participants

REFLECT was a multicentre, randomised, open-label, non-inferiority phase 3 study to compare the efficacy and safety of lenvatinib to sorafenib as a first-line systemic treatment in patients with unresectable hepatocellular carcinoma. The study design and primary results have been previously reported in detail.[3] All patients provided written informed consent. The study was approved by the relevant institutional review boards and was done in accordance with the Declaration of Helsinki.

Briefly, REFLECT was done in 154 sites in 20 countries throughout the Asia-Pacific, European, and North American regions. Eligible patients were aged 18 years or older with unresectable hepatocellular carcinoma with diagnoses confirmed histologically or cytologically, or confirmed clinically in accordance with American Association for the Study of Liver Diseases criteria, one or more measurable target lesion per modified RECIST criteria, Barcelona Clinic Liver Cancer (BCLC) stage B or C categorisation, Child-Pugh class A, Eastern Cooperative Oncology Group (ECOG) performance status of 1 or less, and adequate organ function. Patients with 50% or greater liver occupation, clear bile duct invasion, or portal vein invasion at the main portal vein were excluded, as were patients who had undergone previous systemic therapy for unresectable hepatocellular carcinoma.

Randomisation and masking

Patients were randomly assigned (1:1) to receive oral lenvatinib (12 mg/day for bodyweight ≥60 kg or 8 mg/day for bodyweight <60 kg) or sorafenib (400 mg twice daily) in 28-day cycles. An interactive voice–web response system was used for treatment allocation and functioned as the allocation concealment method. Randomisation was stratified by geographical region (Asia-Pacific or Western); macroscopic portal vein invasion, extrahepatic spread, or both; ECOG performance status (0 or 1); and bodyweight (<60 kg or ≥60 kg). Assessment of tumours was done in accordance with modified RECIST criteria on examination with CT or MRI using a triphasic scanning technique.

Procedures

The study comprised three phases: prerandomisation (screening and baseline), randomisation and treatment, and extension, with a total study duration of approximately 45 months. The treatment period for each patient began at randomisation and consisted of 28-day cycles of treatment until completion of the off-treatment visit within 30 days after the final treatment cycle. Participants discontinued treatment upon objectively documented disease progression, development of unacceptable toxicity, patient request, or withdrawal of consent.

During the randomisation and treatment phase, data for patient-reported outcomes (PRO) were collected at baseline; on day 1 of each postbaseline cycle, starting with cycle 2; and at the end-of-treatment visit. During the extension phase, PRO data were collected for each instrument on day 1 of each treatment cycle and at the end-of-treatment visit. Participants completed all PRO questionnaires. If a participant was unable to complete a questionnaire, site personnel administered the questionnaires via interview and completed the forms for the participant. Scores for assessments with missing data were computed using methods provided by the licensing organisation for each instrument.

Outcomes

The EORTC QLQ-C30 and QLQ-HCC18 questionnaires were used to evaluate HRQOL and hepatocellular

carcinoma-specific symptoms as secondary endpoints in REFLECT. HRQOL was evaluated using the EORTC QLQ-C30, a 30-item self-reported measure of health status, functioning, and symptoms in individuals with cancer participating in clinical trials. The EORTC QLQ-C30 includes five functional scales (physical functioning, role functioning, emotional functioning, cognitive functioning, and social functioning), three symptom scales (fatigue, nausea and vomiting, and pain), six single items (dyspnoea, insomnia, appetite loss, constipation, diarrhoea, and financial difficulties), and a global health status/QOL scale. Scores range from 0 to 100; higher scores for the summary, global, and functional domains indicate better HRQOL, whereas higher scores on symptom scales represent higher levels of symptomatology.

Hepatocellular carcinoma-specific symptoms were assessed using the EORTC QLQ-HCC18, an 18-item self-reported measure of the functional and symptom impact of hepatocellular carcinoma.[13] The EORTC QLQ-HCC18 includes five multi-item symptom scales (fatigue, jaundice, nutrition, pain, and fever), two single-item symptom scales (abdominal swelling and sexual interest), and one multi-item functional scale (body image). Unweighted scale scores range from 0 to 100, and higher scores represent more severe symptoms or problems.

As an exploratory endpoint, the EQ-5D was used to evaluate participants' current health state in five domains: mobility, self-care, usual activities, pain or discomfort, and anxiety or depression.[14] From the five domain scores, an EQ-5D index score can be calculated, with scores ranging from –1 ("worst imaginable health state") to 1 ("best imaginable health state"). Health states were mapped using US time trade-off values for US patients, Japan time trade-off values for Japanese patients, Chinese time trade-off values for Chinese and all other Asia-Pacific patients, and UK time trade-off values for all other non-Asian countries.[15,16] The visual analogue scale (VAS) component of the EQ-5D measures current health status from 0 to 100, where 0 represents "worst imaginable health state" and 100 represents "best imaginable health state."

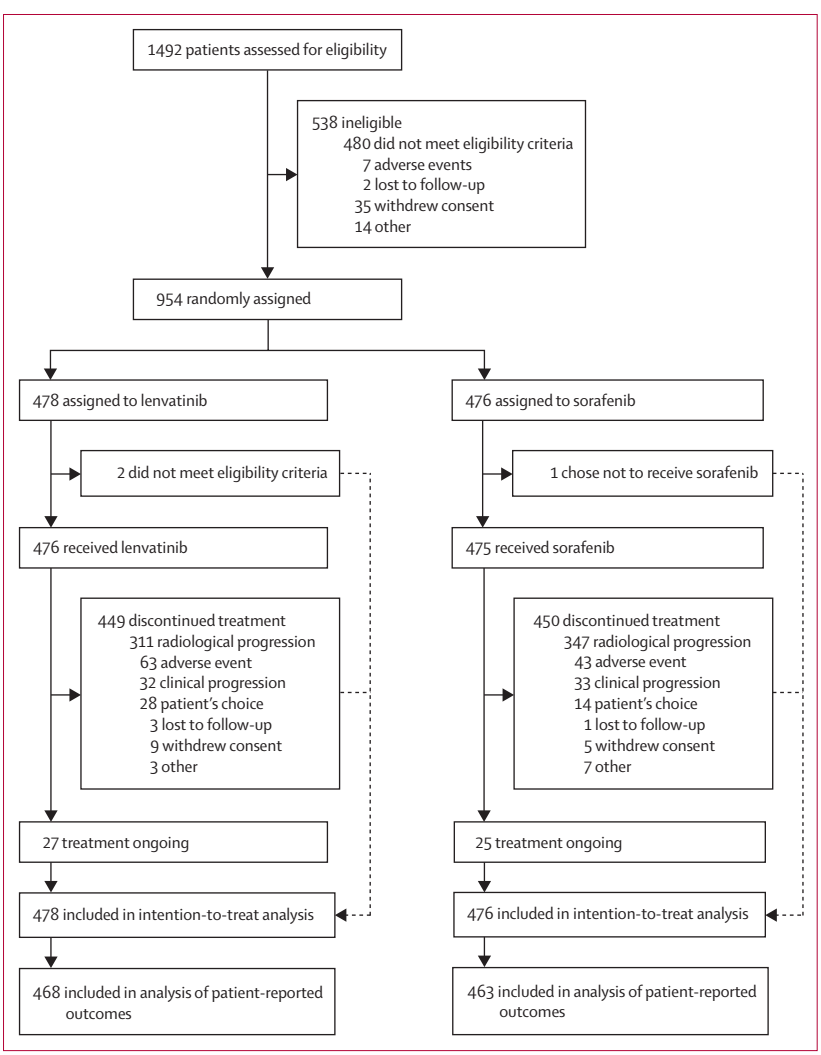

Figure 1: **Trial profile**

Statistical analysis

Post-hoc analyses of prespecified secondary and exploratory endpoints were done. The analysis population was defined as the subpopulation of patients with a PRO assessment at baseline. No formal sample size powering was done for PRO endpoints.

Descriptive summaries of each PRO score at baseline are shown by treatment group, along with compliance rates (ie, the proportion of patients who completed PRO measures according to the assessment schedule) for PROs over time. Missing data were evaluated at the score level for each patient at each cycle to assess the patterns of missing data within and between treatment groups. Based on this evaluation, missing data were assumed to be missing at random, and no further analyses to account for non-random missingness (eg, pattern-mixture models or selection models) were done. No adjustment was made for multiple comparisons. All analyses were based on available data only.

The change from baseline in PRO scale scores at mean postbaseline follow-up time was analysed using a linear mixed-effects model with random intercepts and slopes for the change in score over time. Treatment group, time (as the continuous value of postbaseline analysis day), and the interaction between treatment group and time were included in the model along with adjustment covariates for baseline score, as well as the four randomisation stratification factors (geographical region; macroscopic portal vein invasion, extrahepatic spread, or both; ECOG performance status; and bodyweight). Scores from the end-of-treatment visit were included in the model at the analysis day corresponding to that visit. All available PRO scores were used for fitting the model, and the assumption

of approximate linearity of the response over time was checked visually using longitudinal graphs of the scores. Overall least-squares (LS) means of postbaseline PRO scores for each treatment group and the differences between treatment groups, along with 95% CIs and nominal two-sided p values for the differences, were estimated at the mean postbaseline follow-up time. For each scale, the model-adjusted estimates of the mean score (LS means) were computed for each treatment group at the average postbaseline follow-up time (mean 31·6 weeks [range 31·4–31·7 weeks] after baseline; near the start of cycle 8). The difference between these estimates is the overall LS mean difference.

To capture sustained meaningful changes with treatment, analyses were done to evaluate time to definitive deterioration, accepted as a more relevant time-to-deterioration endpoint in advanced cancer than time to first deterioration.[17] Definitive deterioration was defined as the median time to earliest postbaseline time at which a score crossed the threshold for decline with no subsequent recovery. Individuals with subsequent recovery were considered to have no definitive deterioration. Time-to-definitive-deterioration analyses were done for each PRO scale using established thresholds for minimally important (ie, clinically significant) differences for worsening (10 points for EORTC QLQ-C30 and QLQ-HCC18 scores;[18] 7 points for EQ-VAS, 0·06 for EQ-5D index based on the US time trade-off, 0·08 for EQ-5D index based on all other time trade-offs[19]). Kaplan-Meier methods were used to estimate the distributions of time to definitive deterioration, and treatment groups were compared using an unstratified log-rank test. Cox models with treatment group as the only predictor were used to estimate HRs, along with 95% CIs, for each PRO scale. The randomisation stratification factors were not included as stratification factors in the time-to-event PRO analyses, as an informal sensitivity analysis showed that adding the randomisation stratification variables did not change the statistical inferences.

The association between clinical response and HRQOL was explored by analysing PRO scales, most notably symptoms and functional status, within groups of patients based on the observed best overall response (using investigator-assessed modified RECIST criteria) during the trial. Specifically, two responder analyses were done. In the first responder analysis, patients with complete or partial best response were classified as clinical responders, and all other patients (ie, those with stable or progressive disease) were classified as non-responders. In the second responder analysis, patients with complete or partial best response were classified as clinical responders and were compared with patients with stable disease; those with progressive disease were excluded from this second responder analysis. Using the same methods as for the main longitudinal analysis, descriptive analyses, longitudinal analysis using mixed models, and time-to-definitive-deterioration analyses were done to compare HRQOL outcomes for the two response categories in each of the responder analyses.

This study is registered with ClinicalTrials.gov, NCT01761266. All analyses were done using SAS, version 9.4.

Role of the funding source

Medical writing support was funded by Eisai and Merck Sharp & Dohme. Authors affiliated with Eisai participated in the design of this study; in data collection, analysis, and interpretation; and in development of this Article.

Results

A total of 954 patients enrolled from March 14, 2013, to July 30, 2015, were randomly allocated to receive either lenvatinib (n=478) or sorafenib (n=476) and were included in the intention-to-treat (ITT) population (figure 1). PRO analyses were done with the PRO population, defined as the subpopulation of patients with a PRO assessment at baseline (n=931 patients [n=468 for lenvatinib; n=463 for sorafenib]; 98% of the ITT population). Of the 954 patients in the ITT population, 23 (2%) were excluded from the PRO analyses because they did not complete a PRO assessment at baseline.

At baseline, demographic and health information characteristics were not significantly different between treatment groups for the PRO population (appendix p 2). The majority of patients had an ECOG-performance status of 0 at baseline (592 [64%] of 931), were Child-Pugh class A (923 [99%] of 931), and were Barcelona Clinic Liver Cancer stage C (737 [79%] of 931). The number of patients who remained on study at day 1 of every third cycle and who had PRO data for that cycle are shown in

See Online for appendix

	Lenvatinib	Sorafenib	Overall	Missing all data*		
				EORTC QLQ-C30	EORTC QLQ-HCC18	EQ-5D
Number of ITT† patients	478	476	954	..	..	..
Proportion of participants with PRO data‡						
Baseline	468 (98%)	463 (97%)	931 (98%)	3 (0%)	9 (1%)	6 (1%)
Cycle 3, day 1	389 (81%)	343 (72%)	732 (77%)	1 (0%)	4 (1%)	2 (0%)
Cycle 6, day 1	278 (58%)	186 (39%)	464 (49%)	1 (0%)	5 (1%)	0
Cycle 9, day 1	206 (43%)	125 (26%)	331 (35%)	2 (1%)	5 (2%)	0
Cycle 12, day 1	141 (30%)	80 (17%)	221 (23%)	2 (1%)	2 (1%)	1 (1%)
Cycle 15, day 1	101 (21%)	55 (12%)	156 (16%)	1 (1%)	1 (1%)	0
Cycle 18, day 1	76 (16%)	44 (9%)	120 (13%)	1 (1%)	2 (2%)	0
Off-treatment§	267 (56%)	301 (63%)	568 (60%)	0	3 (1%)	1 (0%)
Cycle 12, day 1	141 (30%)	80 (17%)	221 (23%)	2 (1%)	2 (1%)	1 (1%)

Data are n (%). EORTC=European Organisation for the Research and Treatment of Cancer. QLQ-C30=Quality of Life Questionnaire-Core 30. QLQ-HCC18=Quality of Life Questionnaire-Hepatocellular Carcinoma domain. EQ-5D=EuroQoL Five-Dimension Scale. ITT=intention to treat. PRO=patient-reported outcome. *Percentages computed based on patients who had survived, had continued in the study on day 1 of each cycle, and had available PRO data for that cycle. †The ITT population consisted of all patients who were randomised. ‡Proportion of patients with at least one valid PRO score at each cycle shown. §Visit occurred within 30 days following final dose of study treatment.

Table 1: **Summary of ITT population by treatment and compliance**

	Lenvatinib (n=468)	Sorafenib (n=463)
EORTC QLQ-C30 global health status/QOL	67·06 (20·47)	67·57 (20·52)
EORTC QLQ-C30 functional scales		
Physical functioning	84·80 (16·29)	85·63 (15·62)
Role functioning	85·41 (21·23)	85·79 (22·03)
Emotional functioning	83·69 (16·94)	83·41 (16·80)
Cognitive functioning	89·48 (14·26)	87·36 (17·14)
Social functioning	82·65 (23·06)	82·75 (22·19)
EORTC QLQ-C30 symptom scales or items		
Fatigue	26·35 (22·12)	25·26 (21·19)
Nausea and vomiting	4·85 (12·70)	5·03 (12·41)
Pain	18·45 (23·38)	16·52 (21·92)
Dyspnoea	13·76 (21·37)	13·99 (21·56)
Insomnia	19·27 (25·56)	20·68 (25·47)
Appetite loss	14·42 (23·75)	13·52 (22·90)
Constipation	11·80 (22·33)	12·61 (21·68)
Diarrhoea	7·74 (16·59)	5·87 (13·81)
Financial difficulties	27·80 (31·61)	26·65 (31·37)
EORTC QLQ-HCC18 scales		
Fatigue	23·33 (19·72)	21·80 (19·52)
Jaundice	8·44 (13·57)	7·84 (12·57)
Body image	15·12 (18·82)	14·16 (18·54)
Nutrition	14·74 (15·21)	12·29 (14·26)
Pain	13·93 (17·74)	14·78 (17·62)
Fever	5·72 (12·12)	5·45 (12·58)
Sex life	22·61 (31·60)	21·87 (31·26)
Abdominal swelling	13·53 (21·69)	12·56 (20·06)
EQ-VAS	73·46 (18·43)	74·59 (17·25)
EQ-5D index	0·84 (0·18)	0·85 (0·17)

Data are mean (SD). EORTC=European Organisation for the Research and Treatment of Cancer. QLQ-C30=Quality of Life Questionnaire-Core 30. QOL=quality of life. QLQ-HCC18=Quality of Life Questionnaire-Hepatocellular Carcinoma domain. EQ-VAS=EuroQol Visual Analogue Scale. EQ-5D=EuroQoL Five-Dimension Scale.

Table 2: **Patient-reported outcome population's scale scores at baseline**

table 1 to the end of cycle 18. Sample attrition was significant in both treatment groups, and by day 1 of cycle 12, less than 20% of patients in the sorafenib group and less than 30% of patients in the lenvatinib group remained in the study. Sample attrition was greater in the sorafenib group than in the lenvatinib group at all cycle time points (table 1; additional data not shown). The median time from randomisation to completion of the last PRO assessment was 4·6 months (IQR 2·0–9·4). Among the PRO population, study compliance for completion of all PRO measures was high, with less than 2% of patients missing all data at any particular assessment throughout the course of the study (table 1), even as attrition resulted in deteriorating sample sizes in later cycles.

Baseline scores for all scales in the EORTC QLQ-HCC18, EORTC QLQ-C30, and EQ-5D measures were nominally statistically similar between the lenvatinib and sorafenib treatment groups (table 2). Baseline scores on the QLQ-C30 and QLQ-HCC18 reflected impaired HRQOL relative to full HRQOL, as indicated by mean global health status/QOL scores, whereas functional and symptom scales indicated functional impairments and considerable symptom burden, particularly in the domains of fatigue, insomnia, and financial difficulties. Within each treatment group, patients who discontinued before 6 months of treatment had worse baseline scores than those who continued on treatment (appendix p 3). Observed scores at the off-treatment visit (ie, at the point of permanent treatment cessation), stratified by reason for treatment discontinuation (disease progression *vs* other reasons), are shown in the appendix (p 4). In general, observed mean scores at the off-treatment visit were better in both treatment groups for patients having disease progression than for those discontinuing treatment for other reasons. No clear treatment differences were observed, and patients who progressed on lenvatinib had mean scores similar to those who progressed on sorafenib.

During treatment, functional scores decreased while symptom scores increased, both reflecting deteriorating HRQOL, for both the lenvatinib and sorafenib treatment groups. These decreases were generally less pronounced for the lenvatinib group than the sorafenib group or similar between the two groups. Overall LS mean differences generally favoured lenvatinib over sorafenib, although the differences were not significant (table 3). For individual symptoms assessed by the QLQ-C30, there was a numerical difference in favour of lenvatinib for diarrhoea, although this difference did not exceed the threshold for clinical significance (LS mean difference –9·67, 95% CI –12·87 to –6·48). The LS mean difference for constipation favoured sorafenib (LS mean difference 2·54, 95% CI 0·23 to 4·85), but this difference was not clinically significant. There were no clinically meaningful or nominally statistically significant differences between the two groups for the QLQ-C30 global health status/QOL and functioning scores, the QLQ-HCC18 symptom scores, the EQ-VAS, or the EQ-5D index (table 3).

The median times to definitive deterioration, using patient-reported thresholds of meaningful change of symptoms, function, and QOL for each PRO scale are displayed in the appendix (p 5); HRs are shown in figure 2. Kaplan-Meier estimates of the distribution of time to definitive deterioration over time for each PRO scale are shown in the appendix (pp 8–32).

Patients treated with lenvatinib experienced nominally statistically significant delays in definitive, meaningful deterioration of fatigue (HR 0·83, 95% CI 0·69–0·99), pain (0·80, 0·66–0·96), and diarrhoea (0·52, 0·42–0·65), as assessed by the QLQ-C30 compared with patients treated with sorafenib. However, there was no difference in time to definitive deterioration for the global health status/QOL score. For symptoms assessed by the QLQ-HCC18, nominally statistically significant delays in favour of lenvatinib were observed for nutrition and for

body image (reflecting patient perceptions of muscle loss and abdomen appearance). It is notable that there was no difference in the time to definitive deterioration for the QLQ-HCC18 pain scale score, in contrast to the QLQ-C30 pain scale score. Similarly, no difference in time to definitive deterioration was observed for the QLQ-HCC18 fatigue score, whereas the QLQ-C30 fatigue score showed a difference in time to definitive deterioration. For the EQ-VAS, the time to definitive deterioration for patients treated for lenvatinib was nominally statistically significantly greater than that for patients treated with sorafenib (figure 2, appendix p 5).

159 patients were classified as clinical responders (lenvatinib group, n=115; sorafenib group, n=44) and the remaining 795 were classified as non-responders (lenvatinib group, n=363; sorafenib group, n=432) based on best overall response assessed during the trial. Differences in overall LS mean estimates for change from baseline in favour of responders were seen for most of the PRO scales (appendix p 6). Among the differences from the EORTC scales, only those for physical functioning, fatigue, dyspnoea, and appetite loss from the QLQ-C30 exceeded 5 points. Neither LS mean difference from the EQ-5D met the threshold for clinical significance. HRs for time to definitive deterioration were in favour of responders across all scales (appendix p 33), and median time to definitive deterioration for responders exceeded those for non-responders by a range of 4·8 to 14·6 months.

In an additional analysis by clinical response, PRO scores for the 159 patients classified as clinical responders (lenvatinib group, n=115; sorafenib group, n=44) were compared with those for the 490 patients with stable disease (lenvatinib group, n=246; sorafenib group, n=244). Overall LS mean differences favoured responders in this analysis, with nominally statistically significant differences between responders and patients with stable disease observed for QLQ-C30 emotional functioning and dyspnoea, QLQ-HCC18 jaundice, and EQ-5D index score (appendix p 7). No LS mean differences exceeded 4 points. All time to definitive deterioration HRs favoured responders over patients with stable disease, and all were nominally statistically significant except for QLQ-C30 social functioning and insomnia and QLQ-HCC18 sex life (appendix p 34).

	Overall LS mean score		Difference (95% CI)
	Lenvatinib (n=468)	Sorafenib (n=463)	
EORTC QLQ-C30 global health status/QOL	60·04	60·78	−0·75 (−3·10 to 1·61)
EORTC QLQ-C30 functional scales			
Physical functioning	74·71	73·55	1·17 (−1·74 to 4·07)
Role functioning	72·35	70·69	1·66 (−1·62 to 4·95)
Emotional functioning	79·69	78·10	1·59 (−1·00 to 4·17)
Cognitive functioning	79·22	80·34	−1·13 (−3·70 to 1·45)
Social functioning	72·51	70·58	1·94 (−1·27 to 5·15)
EORTC QLQ-C30 symptom scales or items			
Fatigue	35·80	36·79	−0·99 (−3·99 to 2·02)
Nausea and vomiting	9·42	10·08	−0·66 (−2·64 to 1·33)
Pain	26·30	27·49	−1·19 (−4·09 to 1·72)
Dyspnoea	19·11	20·14	−1·03 (−3·88 to 1·81)
Insomnia	23·94	21·77	2·17 (−0·61 to 4·95)
Appetite loss	27·54	27·74	−0·20 (−3·67 to 3·27)
Constipation	14·31	11·77	2·54 (0·23 to 4·85)
Diarrhoea	21·10	30·78	−9·67 (−12·87 to −6·48)
Financial difficulties	29·52	30·46	−0·94 (−4·43 to 2·54)
EORTC QLQ-HCC18 scales			
Fatigue	31·42	33·43	−2·01 (−5·04 to 1·02)
Jaundice	11·53	11·47	0·06 (−1·83 to 1·94)
Body image	24·73	26·11	−1·38 (−4·36 to 1·60)
Nutrition	21·99	25·03	−3·04 (−5·64 to −0·44)
Pain	19·63	18·77	0·86 (−1·41 to 3·13)
Fever	9·29	9·39	−0·10 (−1·67 to 1·48)
Sex life	31·34	31·11	0·23 (−3·97 to 4·44)
Abdominal swelling	17·77	18·86	−1·09 (−3·92 to 1·73)
EQ-VAS	65·47	65·11	0·36 (−2·13 to 2·84)
EQ-5D index	0·73	0·72	0·01 (−0·02 to 0·04)

On the EORTC QLQ-C30, higher scores for the summary, global, and functional domains indicate better health-related QOL, whereas higher scores on symptom scales represent higher levels of symptomatology. On the EORTC QLQ-HCC18, higher scores represent more severe symptoms or problems. On the EQ-5D index and EQ-VAS, higher scores represent a better health state. For each scale, the model-adjusted estimates of the mean score (LS means) were computed for each treatment group at the mean postbaseline follow-up time (31·6 weeks [range 31·4–31·7 weeks] after baseline) ; the difference between these estimates is the overall LS mean difference. LS=least-squares. EORTC=European Organisation for the Research and Treatment of Cancer. QLQ-C30=Quality of Life Questionnaire-Core 30. QOL=quality of life. QLQ-HCC18= Quality of Life Questionnaire-Hepatocellular Carcinoma domain. EQ-VAS=EuroQol Visual Analogue Scale. EQ-5D=EuroQol Five-Dimension Scale.

Table 3: **LS mean estimates for patient-reported outcome scales by treatment at mean postbaseline follow-up time**

Discussion

In this study, patients with unresectable hepatocellular carcinoma showed substantial symptom burden and impairments in HRQOL and functioning at the start of treatment in REFLECT, as shown by their baseline PRO scores. Overall LS mean differences in PRO domains generally favoured lenvatinib over sorafenib, although the differences were not clinically meaningful; this could be partly explained by the different safety profiles of sorafenib and lenvatinib, as previously reported.[3] With generally similar scores at baseline, the overall on-treatment mean score was nominally statistically better for diarrhoea with lenvatinib, whereas the LS mean difference in favour of sorafenib for constipation was considerably below the threshold for clinical significance. Although there was no difference between treatment groups in time to definitive deterioration for the global health status/QOL score, treatment with lenvatinib yielded nominally statistically significant delays in time to definitive deterioration of fatigue, pain, and diarrhoea as assessed by the QLQ-C30, in nutrition and body image assessed by the QLQ-HCC18, and EQ-5D VAS, compared with patients treated with sorafenib. Time to definitive deterioration on the QLQ-C30 pain scale, but not the QLQ-HCC18 pain scale, was longer for lenvatinib than for sorafenib. This difference might be

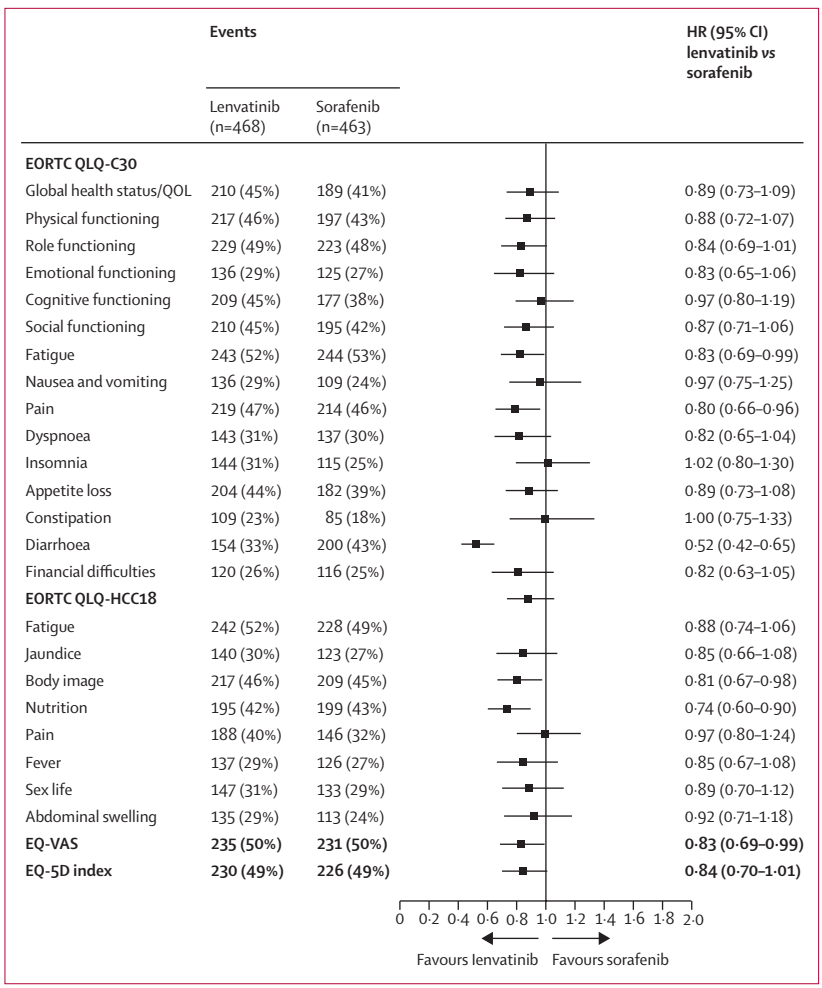

Figure 2: **Time to definitive deterioration**
HR=hazard ratio. EORTC=European Organisation for the Research and Treatment of Cancer. QLQ-C30=Quality of Life Questionnaire-Core 30. QOL=quality of life. QLQ-HCC18=Quality of Life Questionnaire-Hepatocellular Carcinoma domain. EQ-VAS=EuroQol Visual Analogue Scale. EQ-5D=EuroQol Five-Dimension Scale.

because these instruments assess pain differently: the QLQ-HCC18 pain items ask patients to rate their shoulder and abdominal pain, whereas the QLQ-C30 pain items ask patients to rate their pain and if it affected their daily lives. An additional explanation for the favourable lenvatinib QLQ-C30 pain results relative to sorafenib could be related to a greater incidence of treatment-emergent adverse events such as palmar-plantar erythrodysesthesia with sorafenib (26·9% for lenvatinib *vs* 52·4% for sorafenib), as well as a significantly higher objective response rate with lenvatinib relative to sorafenib.[3] Furthermore, time to definitive deterioration on the QLQ-C30 fatigue scale, but not the QLQ-HCC18 fatigue scale, was longer for lenvatinib than for sorafenib, potentially because the QLQ-HCC18 fatigue scale measures comparatively distal aspects of fatigue (eg, the need to sleep, being unable to finish things, and being less active).

Patients on sorafenib were at an increased risk of experiencing a more rapid deterioration in clinically relevant domains including nutrition, body image, and QLQ-C30 pain (specifically interfering with their daily lives). Because body image scores are an indicator of patient-perceived muscle loss and abdomen appearance, more rapid deterioration in this domain in the sorafenib group might reflect weight loss or signs of liver function decompensation (eg, ascites, pruritis, or spider angiomas). By contrast, there was a LS mean difference for the nutrition scale in favour of lenvatinib relative to sorafenib. Lenvatinib's benefit in delaying symptom deteriorations in diarrhoea and nutrition was corroborated in the average treatment effect analyses with a mixed model, attesting to the strengths of the effects in these two symptoms, but not as much in fatigue, pain, body image, and EQ-VAS. It is possible that the numbers of patients remaining on treatment and available for PRO analyses in later cycles were too small to enable robust results for average treatment effect analyses for symptom areas in which the treatment effects were not as strong, whereas time-to-definitive-deterioration analyses were not as limited to the later cycle patient numbers. Despite these limitations, the overall trend of all symptoms and functioning domains favoured lenvatinib, with role functioning, dyspnoea, and EQ-5D index time to definitive deterioration having nominal p values between 0·05 and <0·1. Only LS mean differences in the constipation scale were nominally statistically significantly in favour of sorafenib, which was not clinically significant and not confirmed in the time-to-definitive-deterioration analyses. Further, this difference could have been subject to the confounding effects of diarrhoea, because more patients treated with sorafenib than with lenvatinib experienced treatment-emergent diarrhoea. No other dimension was close to achieving statistical or clinical significance in favour of sorafenib.

Furthermore, the analysis of PROs by clinical response status, pooled across treatments, shows that complete or partial response to treatment is associated with better symptoms, better functioning, and greater overall HRQOL, both in comparison with patients with stable or progressive disease and in comparison with only patients with stable disease. With more than double the objective response rate (24·1% *vs* 9·2%), lenvatinib appears to enable more patients to benefit from lower symptom burdens, less functioning impairment, and better overall HRQOL.

Taken together, these analyses show potentially favourable trends in PRO results for lenvatinib, including delays in definitive, meaningful deterioration in the domains of fatigue, pain, and diarrhoea. Between-group LS mean differences in PRO scores favouring lenvatinib were more pronounced among responders than non-responders.

Patient-reported symptomatic and functional deterioration has been shown to be correlated with early progression in the patient population with unresectable hepatocellular carcinoma.[20] Symptomatic burden,

functional status, and toxicity have been elevated as crucial endpoints of interest for understanding the effect of cancer on patients' lives by all stakeholder groups.[21] As the therapeutic landscape for unresectable hepatocellular carcinoma evolves to include the newly approved combination of atezolizumab plus bevacizumab, as well as investigational therapies for intermediate-stage and advanced hepatocellular carcinoma such as lenvatinib plus pembrolizumab with or without transarterial chemoembolisation,[22] it will be important to understand the effects of the available treatment options on HRQOL. Additional clinical trials (NCT04443049, NCT04227808, NCT04241523) are ongoing to evaluate the HRQOL effect of lenvatinib, in addition to its efficacy and safety, in patients with unresectable hepatocellular carcinoma to further inform treatment decisions and treatment sequencing considerations.

Some limitations of this study must be acknowledged. Owing to the open-label design of REFLECT, the analyses presented here are subject to potential biases. The REFLECT eligibility criteria excluding patients with 50% or greater liver occupation or portal vein invasion at the main portal vein do not correspond to the whole segment of BCLC C stage patients and do not compare fully with the Asia Pacific or SHARP population, and patients at the more advanced end of the BCLC C stage might experience different effects on HRQOL with treatment than the REFLECT study population. In addition, despite the robust sample size in REFLECT, attrition throughout the study was high, particularly in the sorafenib group. An additional limitation is that missing data were assumed to be missing at random. The analyses included no adjustment for multiplicity, resulting in nominal p values. Treatment group was the only stratification factor included in the Cox models, although other stratification factors could have been included. In addition, the PRO measures used lacked a domain or item to capture hand-and-foot skin reaction or anti-VEGF receptor cardiovascular toxicity, key symptoms associated with TKIs. Observed PRO scale scores at the time of permanent treatment cessation were better in both treatment groups for patients having disease progression than for those discontinuing treatment for other reasons, including toxicities. Discontinuation rates due to toxicity and the effect of toxicity on HRQOL will be important focus points for future analyses.

When taken as whole, the symptoms, function, and overall HRQOL effect of lenvatinib and sorafenib treatment on patients were generally favourable to lenvatinib across most functional and symptom areas. Patients treated with lenvatinib experienced delays in definitive deterioration of fatigue, pain, and diarrhoea compared with patients treated with sorafenib, although there were no differences between groups in time to definitive deterioration in global health status/QOL score. Lenvatinib's higher objective response rate might enable more patients to benefit from lower symptom burden, less functional impairment, and better HRQOL.

Contributors

AV, SQ, MK, YS, SH, TY, J-HY, LF, KS, CL, MS, KM, TO, TT, LB, GM, and VB made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Declaration of interests

This analysis was done under a research contract between RTI Health Solutions and Eisai and was funded by Eisai and Merck Sharp & Dohme. LB is an employee of RTI Health Solutions. KM, TO, TT, and GM are employees of Eisai. YS was an employee of Eisai when this research was done and is currently an employee of Bayer. AV reports personal fees from Eisai, Roche, AstraZeneca, Lilly, Bayer, Merck, BMS, MSD, Incyte, PierreFabre, Ipsen, and Sanofi, outside the submitted work. MK reports personal fees from Eisai, Ono, MSD, BMS, Roche, Bayer, and Lilly and grants from Gilead Sciences, Taiho, Sumitomo Dainippon, Takeda, Otsuka, EA Pharma, AbbVie, and Eisai, outside the submitted work. SH reports professional fees from Eisai for PRO analytic planning and execution. TY reports personal fees from Eisai, Bayer, Lilly, Chugai, and Takeda, outside the submitted work. J-HY reports grants from Eisai and Merck and grants from AstraZeneca, Daewoong Pharmaceuticals, and Hanmi Pharmaceuticals, outside the submitted work. CL reports grants and personal fees from BMS, Merck, Eisai, Servier, Sanofi, Roche, Exelixis, Daiichi-Sankyo, Bayer, Amgen, Novartis, Pfizer, Ipsen, and AstraZeneca, outside the submitted work. MS reports personal fees from Eisai, Genentech, Bayer, and Exelixis, outside the submitted work. VB reports personal fees and non-financial support from Roche, BMS, and Ipsen and personal fees from Eisai, MSD, AstraZeneca, and Bayer, outside the submitted work. All other authors declare no competing interests.

Data sharing

The source data for this Article are considered both commercially confidential and Health Insurance Portability and Accountability Act confidential; therefore, the data holders (Eisai) do not intend to publicly post or share the source data. Eisai might, however, consider requests on a case-by-case basis if contacted by an individual researcher (access will be permitted after signature of a data-access agreement). Additional related documents, however, will be made available (ie, the study protocol and statistical analysis plans).

Acknowledgments

This study was sponsored by Eisai, Woodcliff Lake, NJ, USA, and Merck Sharp & Dohme, a subsidiary of Merck & Co, Kenilworth, NJ, USA. Medical writing services were provided by Katherine Lothman of RTI Health Solutions, Research Triangle Park, NC, USA. These services were funded by Eisai and Merck Sharp & Dohme.

References

1 Ferlay J, Colombet M, Soerjomataram I, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer* 2019; **144:** 1941–53.

2 Singal AG, El-Serag HB. Hepatocellular carcinoma from epidemiology to prevention: translating knowledge into practice. *Clin Gastroenterol Hepatol* 2015; **13:** 2140–51.

3 Kudo M, Finn RS, Qin S, et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. *Lancet* 2018; **391:** 1163–73.

4 Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; **68:** 394–424.

5 Balogh J, Victor D, 3rd, Asham EH, et al. Hepatocellular carcinoma: a review. *J Hepatocell Carcinoma* 2016; **3:** 41–53.

6 Finn RS, Qin S, Ikeda M, et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. *N Engl J Med* 2020; **382:** 1894–905.

7 Gandhi S, Khubchandani S, Iyer R. Quality of life and hepatocellular carcinoma. *J Gastrointest Oncol* 2014; **5:** 296.

8 Li L, Yeo W. Value of quality of life analysis in liver cancer: a clinician's perspective. *World J Hepatol* 2017; **9:** 867–83.

9 Xing M, Kokabi N, Camacho JC, Kim HS. Prospective longitudinal quality of life and survival outcomes in patients with advanced infiltrative hepatocellular carcinoma and portal vein thrombosis treated with Yttrium-90 radioembolization. *BMC Cancer* 2018; **18:** 75.

10 Johnson PJ, Qin S, Park JW, 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:** 3517–24.

11 Vilgrain V, Pereira H, Assenat E, 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:** 1624–36.

12 Chow PKH, Gandhi M, Tan SB, et al. SIRveNIB: selective internal radiation therapy versus sorafenib in Asia-Pacific patients with hepatocellular carcinoma. *J Clin Oncol* 2018; **36:** 1913–21.

13 Blazeby JM, Currie E, Zee BC, et al. Development of questionnaire module to supplement the EORTC QLQ-C30 to assess quality of life in patients with hepatocellular carcinoma, the EORTC QLQ HCC18. *Eur J Cancer* 2004; **40:** 2439–44.

14 EuroQol G. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy* 1990; **16:** 199–208.

15 Szende A, Oppe M, Devlin N. EQ-5D value sets: inventory, comparative review and user guide. Dordrecht: Springer, 2007.

16 Liu GG, Wu H, Li M, Gao C, Luo N. Chinese time trade-off values for EQ-5D health states. *Value Health* 2014; **17:** 597–604.

17 Anota A, Hamidou Z, Paget-Bailly S, et al. Time to health-related quality of life score deterioration as a modality of longitudinal analysis for health-related quality of life studies in oncology: do we need RECIST for quality of life to achieve standardization? *Qual Life Res* 2015; **24:** 5–18.

18 Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol* 1998; **16:** 139–44.

19 Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. *Health Qual Life Outcomes* 2007; **5:** 70.

20 Hudgens SC, R; Floden, L; Meier G. Understanding quality of life in hepatocellular carcinoma patients. *J Clin Oncol* 2018; **6:** 4093.

21 Kleutz P, Hudgens S, Batnagar V. Patient reported outcomes in oncology clinical trials: clinical trial design and operational issues toward regulatory grade clinical evidence. 2019 Accelerating Anticancer Agent Development and Validation Workshop; Rockville, MA, USA; May 8–10, 2019.

22 Finn RS, Ikeda M, Zhu AX, et al. A phase Ib study of lenvatinib plus pembrolizumab in patients with unresectable hepatocellular carcinoma. *J Clin Oncol* 2020; **38:** 2960–70.