ORIGINAL RESEARCH

Comparative Efficacy of Cabozantinib and Regorafenib for Advanced Hepatocellular Carcinoma

Robin K. Kelley · Patrick Mollon · Jean-Frédéric Blanc ·

Bruno Daniele · Thomas Yau · Ann-Lii Cheng · Velichka Valcheva ·

Florence Marteau · Ines Guerra · Ghassan K. Abou-Alfa

ABSTRACT

Background: No trials have compared cabozantinib and regorafenib for the second-line treatment of advanced hepatocellular carcinoma (HCC).
Objectives: Conduct a matching-adjusted indirect comparison (MAIC) of the efficacy and safety of second-line cabozantinib and regorafenib in patients with advanced HCC and disease progression after prior sorafenib.

Methods: The CELESTIAL and RESORCE trials were used for indirect comparison of second-line cabozantinib and regorafenib in advanced HCC. Population-level data were available for RESORCE, individual patient data (IPD) for CELESTIAL. To align with RESORCE, the CELESTIAL population was limited to patients who received first-line sorafenib only. To minimize potential effect-modifying population differences, the CELESTIAL IPD were weighted to balance the distribution of clinically relevant baseline characteristics with those of RESORCE. Overall survival (OS) and progression-free survival (PFS) were evaluated for the matching-adjusted second-line CELESTIAL population and compared with those for RESORCE using weighted Kaplan-Meier curves and parametric modeling. Rates of grade 3/4 treatment-

R. K. Kelley (✉)
UCSF, Helen Diller Family Comprehensive Cancer Center, San Francisco, CA, USA
e-mail: katie.kelley@ucsf.edu

P. Mollon · V. Valcheva · F. Marteau
Ipsen Pharma, Boulogne-Billancourt, France

J.-F. Blanc
Hôpital Haut-Lévêque, CHU de Bordeaux, Bordeaux, France

B. Daniele
Azienda Ospedaliera G Rummo, Benevento, Italy

B. Daniele
Ospedale del Mare, Naples, Italy

T. Yau
University of Hong Kong, Pokfulam, Hong Kong

A.-L. Cheng
National Taiwan University Cancer Center, National Taiwan University Hospital, Taipei, Taiwan, Republic of China

I. Guerra
IQVIA Ltd, London, UK

G. K. Abou-Alfa
Memorial Sloan Kettering Cancer Center, New York, NY, USA

G. K. Abou-Alfa
Weill Medical College at Cornell University, New York, NY, USA

emergent adverse events (TEAEs) affecting > 5% of patients in any study arm were compared.

Results: In the matching-adjusted second-line populations (CELESTIAL, effective sample size = 266; RESORCE, n = 573), median (95% confidence interval) OS was similar for cabozantinib and regorafenib (11.4 [8.9–17.0] versus 10.6 [9.1–12.1] months; p = 0.3474, log-rank test). Median PFS was longer for cabozantinib than regorafenib (5.6 [4.9–7.3] versus 3.1 [2.8–4.2] months; p = 0.0005, log-rank test). There was a trend for lower rates of some grade 3/4 TEAEs with regorafenib than with cabozantinib, which may reflect the exclusion of sorafenib-intolerant patients from RESORCE but not from CELESTIAL, a difference that the MAIC methods could not remove. Only diarrhea rates were statistically significantly lower for regorafenib ($p \leq 0.001$).

Conclusions: Cabozantinib may achieve similar OS and prolonged PFS compared with regorafenib in patients with progressive advanced HCC after prior sorafenib.

PLAIN LANGUAGE SUMMARY

Cabozantinib and regorafenib are treatments approved for some patients with advanced hepatocellular carcinoma (HCC), a type of liver cancer, after disease progression despite prior sorafenib treatment. Cabozantinib, regorafenib and sorafenib are tyrosine kinase inhibitors (TKIs), meaning that they slow cancer progression by targeting specific ways that tumors grow. Cabozantinib and regorafenib offer benefits to patients compared with placebo (i.e., no treatment) for those who have progressed despite sorafenib treatment. No clinical studies have compared cabozantinib and regorafenib directly. This study compared the efficacy and safety of cabozantinib and regorafenib using data from trials of each drug versus placebo: CELESTIAL for cabozantinib and RESORCE for regorafenib. These two trials were similar—both involved patients with progressive advanced HCC who had received previous cancer treatment. There were some important differences, but these were minimized using statistical methods (matching and adjustments/"weighting") allowing outcomes to be meaningfully compared. One difference that could not be removed by the statistical methods was that patients who were intolerant to prior sorafenib were excluded from RESORCE but were eligible for the CELESTIAL trial. In the otherwise matched populations, treatment with cabozantinib was associated with similar overall survival and significantly longer progression-free survival than regorafenib. Rates of diarrhea were significantly lower for regorafenib than cabozantinib, suggesting that regorafenib may be better tolerated, but this may reflect the exclusion of sorafenib-intolerant patients from RESORCE. These findings cannot replace a head-to-head study, but may help in guiding decision-making between cabozantinib and regorafenib in patients with progressive advanced HCC after soraftenib treatment.

Key Summary Points

Why carry out this study?

The tyrosine kinase inhibitors cabozantinib and regorafenib are approved for the treatment of patients with advanced hepatocellular carcinoma (HCC) who have progressed despite prior treatment with sorafenib.

No clinical trials have directly compared cabozantinib and regorafenib for the second-line treatment of advanced HCC.

This matching-adjusted indirect comparison (MAIC) used data from the phase 3 CELESTIAL (NCT01908426) and RESORCE (NCT01774344) trials to generate comparative efficacy and safety estimates for cabozantinib versus regorafenib in patients with progressive advanced HCC after prior sorafenib therapy.

What was learned from this study?

In patients with progressive advanced HCC who have received prior sorafenib treatment, cabozantinib may achieve similar overall survival and prolong progression-free survival compared with regorafenib; regorafenib may be associated with lower rates of grade 3 or 4 diarrhea.

A MAIC cannot replace a head-to-head randomized controlled trial, but these findings may help in guiding clinical decision-making between cabozantinib and regorafenib when treating patients with progressive advanced HCC in the absence of direct trial evidence.

INTRODUCTION

The era of targeted therapy for liver cancer began in 2007 with the approval of sorafenib for the first-line management of patients with advanced hepatocellular carcinoma (HCC) [1–4]. More than a decade later, the range of therapeutic options for HCC has broadened to include a number of new targeted therapies with proven survival benefit in phase 3 trials. In the first-line setting, the tyrosine kinase inhibitor (TKI) lenvatinib has now been approved for patients with unresectable disease [5, 6]. Approved second-line treatment options (after prior treatment with sorafenib) now include the TKIs regorafenib [7, 8] and cabozantinib [9, 10] in patients with advanced HCC, the anti-vascular endothelial growth factor receptor 2 (VEGFR2) monoclonal antibody ramucirumab in patients with alpha-fetoprotein levels > 400 ng/ml [11, 12] and the checkpoint inhibitors nivolumab and pembrolizumab (programmed cell death receptor-1 antibodies) as monotherapy or, in the case of nivolumab, in combination with ipilimumab (a cytotoxic T-lymphocyte-associated protein 4 antibody) [13–16].

The arrival of second-line agents for advanced HCC extends the previous therapeutic offering for patients, many of whom present with advanced disease at the time of diagnosis and have had few therapeutic options available to them [17]. Accordingly, 5-year survival rates for localized, regional and distant HCC are typically poor (31%, 11% and 2%, respectively) [18], highlighting the importance of new second-line therapies and the clinical imperative to optimize their use.

The TKIs regorafenib and cabozantinib are both approved as second-line agents for patients with HCC after prior treatment with sorafenib [7–10]. They share a drug class and oral mode of administration, but they differ in their molecular targeting profiles. Regorafenib targets multiple receptor tyrosine kinases, including those involved in tumor angiogenesis (VEGFR-1, -2, -3, TIE2), oncogenesis (KIT, RET, RAF-1, BRAF, BRAF^{V600E}), metastasis (VEGFR3, PDGFR, FGFR) and tumor immunity (CSF1R) [19]. Cabozantinib also has inhibitory activity against tumor angiogenesis and oncogenesis, but it additionally targets the hepatocyte growth factor receptor protein (MET), involved in tumor growth and invasion, and other tyrosine kinases, including those involved in

modulation of tumor immunity (e.g., AXL, MER) [20, 21]. Current HCC management guidelines recommend either agent in advanced disease following progression after sorafenib on the basis of evidence from the pivotal RESORCE (NCT01774344) and CELESTIAL (NCT01908426) phase 3 trials [1, 8, 10, 22, 23].

RESORCE compared once-daily regorafenib 160 mg to placebo for weeks 1–3 of every 4-week cycle (in patients with HCC who were tolerant to sorafenib, but who had progressed during sorafenib treatment; $n = 573$) [8]. CELESTIAL compared once-daily cabozantinib 60 mg to placebo (continuous dosing) in patients with advanced HCC who had progressed after at least one systemic treatment for HCC and may have received up to two previous systemic regimens for advanced HCC ($n = 707$), one of which was required to be sorafenib [10].

In the second-line (sorafenib-tolerant) population included in RESORCE, regorafenib improved median overall survival (OS) and progression-free survival (PFS) compared with placebo (hazard ratio [HR] [95% confidence interval, CI], 0.63 [0.50–0.79] and 0.46 [0.37–0.56], respectively; one-sided $p < 0.0001$ for both) [8]. The most common clinically relevant grade 3 or 4 treatment-emergent adverse events (TEAEs) were hypertension, hand-foot skin reaction, fatigue and diarrhea [8]. In the mixed second- and third-line population included in CELESTIAL, cabozantinib also significantly improved OS and PFS compared with placebo (HR [95% CI], 0.76 [0.63–0.92], two-sided $p = 0.005$ and 0.44 [0.36–0.52], two-sided $p < 0.001$, respectively). Cabozantinib also prolonged median OS and PFS in the subgroup of patients who had only received previous systemic treatment with sorafenib (stratified HR [95% CI], 0.70 [0.55–0.88] and 0.40 [0.32–0.50], respectively). Similar to RESORCE, hand-foot reaction (assessed as palmar-plantar erythrodysesthesia), fatigue and diarrhea were among the most common grade 3 or 4 events; higher rates of hypertension and increased aspartate aminotransferase (AST) level were also recorded with cabozantinib compared with placebo [10].

While the CELESTIAL and RESORCE trials provide robust evidence for the second-line use of either cabozantinib or regorafenib in progressive advanced HCC, at the time of writing, there have been no randomized controlled trials to compare the two agents directly. In the absence of head-to-head trial data, indirect treatment comparisons offer a means of estimating probable treatment outcomes if comparator drugs were used in the same way in similar patients [24, 25]. Standard indirect treatment comparisons estimate the effect sizes of individual treatments relative to a common reference arm (e.g., placebo); these relative estimates are then compared. This approach is used in network meta-analyses, but it relies on the studies involved being sufficiently similar in all respects other than the treatments being compared [26]. When this is not the case and a standard indirect treatment comparison is not valid, population-adjustment methods are used to minimize between-trial population differences. These approaches weight individual patient data (IPD) to reduce differences in the distribution of clinically relevant covariates, thereby reducing the potential for bias in comparative outcome estimates. Propensity score matching is one method of population matching, but it requires IPD to be available for all studies being compared [25]. When IPD are available for only one of two studies being compared, a matching-adjusted indirect comparison (MAIC) method can be used. A MAIC weights the IPD for the available study so that its baseline characteristics match those of a reference comparator study for which only published aggregate-level data are available [25]. MAIC analyses are routinely used for health technology assessment [24, 25], and their potential to inform clinical decision-making in the absence of direct comparative data has been utilized across a range of cancer types (e.g., breast cancer, prostate cancer, basal cell carcinoma) [27–29], including HCC [30–32].

We report here the first MAIC of second-line TKI options for HCC, providing an assessment of the comparative efficacy and safety of cabozantinib and regorafenib for patients with advanced HCC who have received sorafenib as the only prior systemic therapy (see Fig. 1 for a graphical summary of the study).
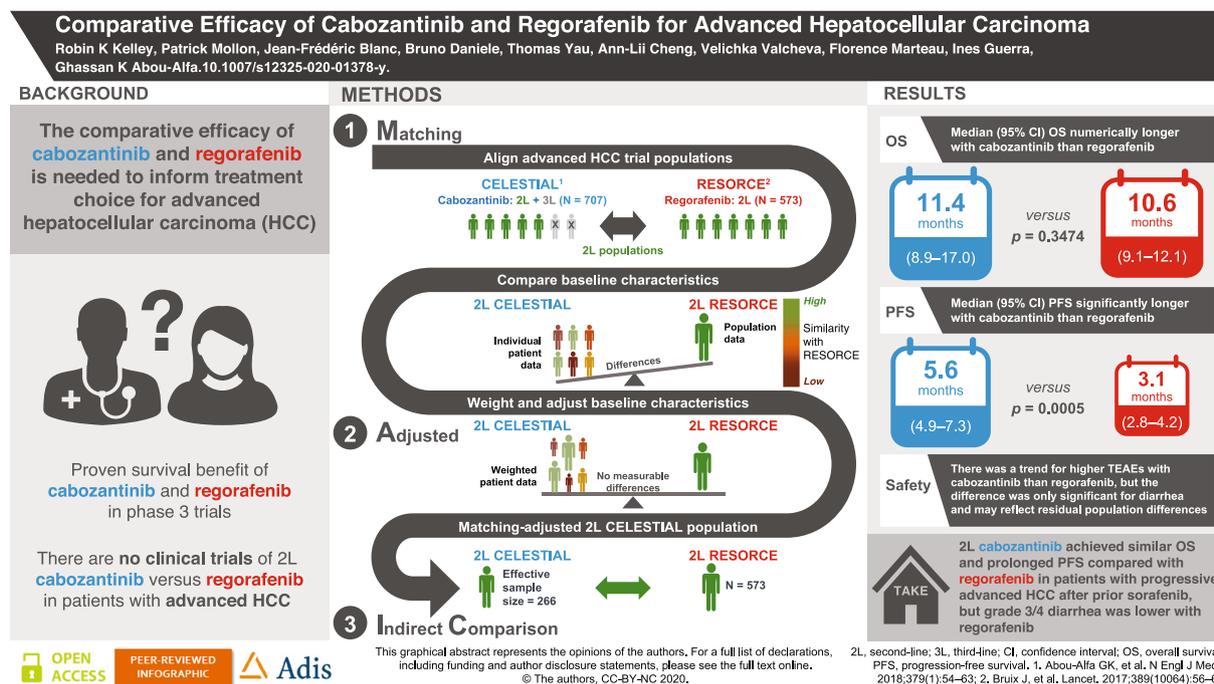
Fig. 1 Overview of the study

METHODS

Data Source Identification And Eligibility

The CELESTIAL trial of cabozantinib (n = 707) [10] and RESORCE trial of regorafenib (n = 573) [8] were identified in the published literature and assessed for their feasibility for a standard indirect treatment comparison (ITC) of TKIs recently approved for the second-line management of advanced progressive HCC following prior sorafenib treatment.

RESORCE included a second-line population of patients with advanced HCC who tolerated sorafenib, but who had progressed following sorafenib treatment only [8, 12]. In contrast, CELESTIAL included a mixed second- and third-line population [10]. While CELESTIAL patients must have received previous treatment with sorafenib (by default), they may also have received a second systemic treatment for advanced HCC prior to recruitment to the trial. Furthermore, CELESTIAL did not exclude patients who were intolerant to sorafenib.

Despite these notable population differences, the availability of IPD for CELESTIAL (obtained from Exelixis with a data cutoff date of 1 June 2017) enabled isolation of a pure second-line CELESTIAL subpopulation for comparison with RESORCE. The similarity of the efficacy and safety outcomes reported for the two trials confirmed their potential for use in an indirect treatment comparison of second-line cabozantinib versus regorafenib in patients with progressive advanced HCC.

To evaluate whether a standard ITC was feasible for RESORCE (regorafenib) and the second-line CELESTIAL (cabozantinib) subpopulation, a panel of expert oncologists convened in June 2018 to review the design similarities and baseline population characteristics of the two trials (Supplemental Tables S1 and S2). The panel identified differences in several effect-modifying covariates with the potential to bias a standard ITC. Therefore, a MAIC analysis approach was selected to compare the safety and efficacy of regorafenib, using the aggregate population data published for RESORCE and

IPD for the second-line subpopulation from CELESTIAL.

Matching-Adjusted Indirect Comparison: Cabozantinib and Regorafenib

The MAIC analysis was conducted in accordance with the established methodology outlined in the guidance from the National Institute for Health and Care Excellence [25, 33] as summarized in Supplemental Fig. S1 (see the electronic supplementary material).

Population Matching

The IPD from CELESTIAL were used to identify a pure second-line population of patients who received cabozantinib after sorafenib only. The IPD also allowed the clinically meaningful differences in potential effect-modifying covariates to be minimized by weighting the baseline characteristics of the second-line CELESTIAL subpopulation so as to match them to those of the published RESORCE population-level statistics (e.g., means, medians, percentages). The baseline differences between the trials that had been identified as being potential effect modifiers by the expert panel were assessed for colinearity, and the following covariates were selected as the matching criteria: age, race, geographical region, Eastern Cooperative Oncology Group Performance Status, Child-Pugh class, duration of prior sorafenib treatment, extrahepatic disease, macrovascular invasion, etiology of HCC (hepatitis B, alcohol use and hepatitis C) and serum alpha-fetoprotein level. Patients in CELESTIAL with missing IPD for any of the selected matching criteria were excluded from the analysis.

Outcome Evaluation: Survival Analysis

Survival outcomes for the matching-adjusted second-line CELESTIAL population were then evaluated and compared with those published for RESORCE. For cabozantinib, median OS and PFS estimates were derived from weighted Kaplan-Meier (KM) curves fitted to the survival data [34]. For cabozantinib, confidence intervals (CIs) for the KM weighted curves were generated from simulations [35] and for median survival using Woodruff's method [36]. For regorafenib, KM curves and median survival estimates were sourced from the RESORCE publication [8].

MAIC analyses can be 'anchored' or 'unanchored,' with methodologic guidelines generally favoring an anchored analysis approach, when feasible [25]. An anchored MAIC generates estimates relative to a common comparator arm, such as placebo. An unanchored analysis involves no common comparator arm and bases estimates on absolute outcomes. An unanchored analysis involves more assumptions than an anchored approach, but is necessary when there is no common comparator arm (e.g., single-arm studies) or when assumptions underpinning an anchored approach are not satisfied. For anchored hazard ratios to be valid, the proportional hazards assumption must be satisfied (i.e., the treatment effect must be proportional over time and the survival curves fitted to each treatment group must have a similar shape) [25, 33].

In the present analysis, the feasibility of conducting an anchored analysis (with placebo as the common comparator arm) was assessed. The proportional hazards assumption was tested for OS and PFS by visual inspection of the log-cumulative hazard plots to ensure that there was no pattern of non-parallelism. The test was conducted for the matching-adjusted and unmatched second-line cabozantinib populations versus placebo from CELESTIAL, for the regorafenib versus placebo populations from RESORCE and for the matching-adjusted and unmatched second-line cabozantinib CELESTIAL population versus the regorafenib RESORCE population. The findings of the visual inspection of the log-cumulative hazard plots were then validated by visual inspection of the scaled Schoenfeld residuals and using the Grambsch-Therneau test (a statistical test based on the scaled Schoenfeld residuals) [33, 37].

Where the proportional hazards assumption was not satisfied and an anchored analysis not supported, an unanchored analysis was conducted by fitting individual parametric survival curves to each treatment arm, in line with best practice guidelines. Parametric model selection was based on an analysis of Akaike's information criterion and Shwarz's Bayesian

information criterion (AIC/BIC), with superior model fit indicated by lower AIC and BIC [33].

Outcome Evaluation: Safety

Incidence of grade 3 or 4 TEAEs affecting > 5% of patients in any of the second-line CELESTIAL or RESORCE treatment arms was compared for cabozantinib versus regorafenib. In an anchored analysis, the estimated relative effects (log odds ratios [ORs]) of cabozantinib versus placebo in the weighted population were generated and compared with log-ORs for regorafenib versus placebo, computed from the published data. Safety outcome estimates of cabozantinib versus regorafenib were constructed in the log-OR scale.

If a TEAE of interest did not occur in either of the placebo arms of the trials (preventing an anchored log-OR analysis), an unanchored analysis of the active treatment arms was conducted. The number of TEAEs occurring in CELESTIAL was used to compute a weighted, unanchored estimate.

Analyses

The analyses were performed using R version 3.5.2 (R Core Team, 2014). The package 'survey' version 3.36 was used to fit weighted survival models with weights computed from the MAIC used as sampling weights.

Compliance with Ethics Guidelines

The results presented in this manuscript are based on previously published studies. All procedures performed in those studies involving human participants were in accordance with the ethical standards of the local Institutional Review Boards for each site and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the CELESTIAL (ClinicalTrials.gov identifier NCT01908426) and RESORCE (ClinicalTrials.gov identifier NCT01774344) trials.

RESULTS

Patient Characteristics

RESORCE involved a total of 573 second-line patients with advanced HCC who were randomized to regorafenib ($n = 379$) or placebo ($n = 194$) following progression on sorafenib. In total, 495 second-line patients were enrolled in CELESTIAL and randomized to cabozantinib ($n = 331$) or placebo ($n = 164$). When limited to the CELESTIAL patients for whom data were available for all effect-modifying characteristics, the pre-matched population size reduced further to 484 patients (cabozantinib, $n = 326$; placebo, $n = 158$) and, following matching, to an effective sample size of 266 patients (non-additive with respect to each treatment arm: cabozantinib, $n = 187$; placebo, $n = 81$) (Table 1).

Application of MAIC weighting to the baseline IPD from CELESTIAL was effective in balancing the effect-modifying baseline characteristics of the second-line CELESTIAL and RESORCE populations, most notably patient ethnicity and geographical region of origin, HCC etiology, proportion of patients with Eastern Cooperative Oncology Group (ECOG) performance status 0 and HCC etiology (Table 2). As a result, there were no discernable clinically relevant differences between the baseline characteristics of the matching-adjusted second-line CELESTIAL population and the RESORCE populations (Table 2).

Survival Outcomes

KM-Derived Estimates

Estimated median (95% CI) OS derived from the weighted KM curves were similar: 11.4 (8.9–17.0) months for the matching-adjusted cabozantinib population and 10.6 (9.1–12.1) months for the regorafenib population ($p = 0.3474$, log-rank test). In comparison, equivalent OS estimates for the placebo arms were 7.2 (6.1–10.8) for the matching-adjusted second-line CELESTIAL population and 7.8 (6.3–8.8) months for RESORCE (Fig. 2 and Table 3).

Table 1 MAIC population sizes

	CELESTIAL population				RESORCE population
	Unmatched			Matching-adjusted	
	n (overall)	*n* (second line)	*n* (second line with non-missing data)	*Effective sample size*	*n* (published)
Active treatment	470	331	326	187	379
Placebo	237	164	158	81	194
Total	707	495	484	266[a]	573

[a] Non-additive with respect to each treatment arm. The ESS is computed separately for each patient group as $\frac{(\sum^{i} \hat{w}_i)^2}{\sum^{i}(\hat{w}_i^2)}$ (the squared sum of weights [numerator] and sum of squared weights [denominator]). The ESS computed for the full population will not equate to the sum of the ESS for each patient group individually, unless all the weights are the same for each treatment arm

n number of patients enrolled and randomized

Estimated median (95% CI) PFS was 5.6 (4.9–7.3) months for the matching-adjusted cabozantinib population compared with 3.1 (2.8–4.2) months for regorafenib, a statistically significant difference ($p = 0.0005$, log-rank test). For the placebo arms, equivalent estimates were 1.9 (1.9–2.1) for the matching-adjusted second-line CELESTIAL population and 1.5 (1.4–1.6) months for RESORCE (Fig. 3 and Table 3).

Parametric Modeling Estimates
Visual inspection of the log of cumulative hazard versus time plots displayed a distinct pattern of non-parallelism for both OS and PFS, indicating that the proportional hazards assumption was not valid, and an anchored analysis was not supported (Supplemental Fig. S2). The plots of scaled Schoenfeld residuals versus time showed a systematic departure from the horizontal for both outcomes, confirming the assessment that the proportional hazards assumption was not supported, as did the non-zero slopes (at the 5% significance level) given by the Grambsch-Therneau test (Supplemental Fig. S3). In line with recommended practice, an unanchored analysis was therefore conducted

by fitting individual parametric survival curves to each treatment arm [33].

Analysis of AIC and BIC for different candidate models identified the log-logistic distribution as the best-fit model for OS and the generalized gamma model as the best-fit model for PFS (Supplemental Tables S3 and S4). The direction of the survival trends generated by the parametric modeling mirrored those of the weighted KM analyses for both the active treatment and the placebo arms (Supplemental Table S5). As for the KM-derived estimates, median (95% CI) OS estimates were similar and the 95% CIs overlapped: 11.40 (10.01–12.96) months for the matching-adjusted second-line cabozantinib population versus 10.29 (9.15–11.56) months for the regorafenib population. For the placebo arms equivalent estimates were 8.27 (7.00–9.76) months for the matching-adjusted second-line CELESTIAL population and 7.30 (6.30–8.47) months for RESORCE (Table S5).

For PFS, the median (95% CI) estimate was longer for the matching-adjusted second-line cabozantinib population (5.49 [4.92–6.13] months) compared with the regorafenib population (3.39 [3.05–3.78] months). There was no

Table 2 Baseline characteristics of the MAIC populations

	CELESTIAL second-line population		RESORCE population
	Unmatched ($n = 495$)[a]	Matching-adjusted ($n = 266$)[b]	Published ($n = 573$)
Age < 65 years, %	53.33	54.97	54.97
Female, %	17.58	18.63	12.04
Asian geographical region, %	22.83	37.70	37.70
White, %	58.18	35.95	35.95
ECOG performance status 0, %	56.97	65.79	65.79
Child–Pugh class A, %	98.79	97.91	97.91
Duration of prior sorafenib treatment, mean (months)	7.65	11.63	11.63
Extrahepatic disease, %	76.16	71.90	71.90
Macrovascular invasion, %	29.41	28.62	28.62
Etiology, %			
Hepatitis B	37.37	37.70	37.70
Alcohol use	21.52	25.31	25.31
Hepatitis C	25.10	20.77	20.77
Alpha-fetoprotein > 400 ng/ml, %	40.81	43.46	43.46

[a] Includes patients with missing data for effect-modifying baseline characteristics
[b] Effective sample size; excludes patients with missing data for effect-modifying baseline characteristics
ECOG Eastern Cooperative Oncology Group

overlap in the CIs for the cabozantinib and regorafenib PFS estimates, suggesting the difference in favor of cabozantinib was statistically significant. For the placebo arms, equivalent estimates were 2.35 (2.11–2.61) months for the matching-adjusted second-line CELESTIAL population and 1.87 (1.68–2.09) months for RESORCE (Table S3).

Safety Outcomes

TEAEs of interest (grade 3 or 4 occurring in more than 5% of patients in any trial arm) were AST increase ("increased AST" in RESORCE), diarrhea, elevated bilirubin, fatigue, hypertension and palmar-plantar erythrodysesthesia syndrome (hand-foot skin reaction in RESORCE).

An anchored log-OR analysis found no significant difference between the matching-adjusted cabozantinib and regorafenib populations in terms of frequency of grade 3 or 4 TEAEs: fatigue ($p = 0.9313$); elevated bilirubin ($p = 0.8558$) or increased AST ($p = 0.2201$). This result was consistent for both the matching-adjusted and unmatched cabozantinib populations (Supplemental Table S6). Although there was a possible trend toward a higher rate of grade 3 or 4 hypertension in the matching-adjusted cabozantinib population compared with the regorafenib population, the confidence intervals for the estimate crossed zero (indicating no difference), and a test of the null hypothesis confirmed that the difference was not statistically significant ($p = 0.0611$) (Fig. 4 and Supplemental Table S6).
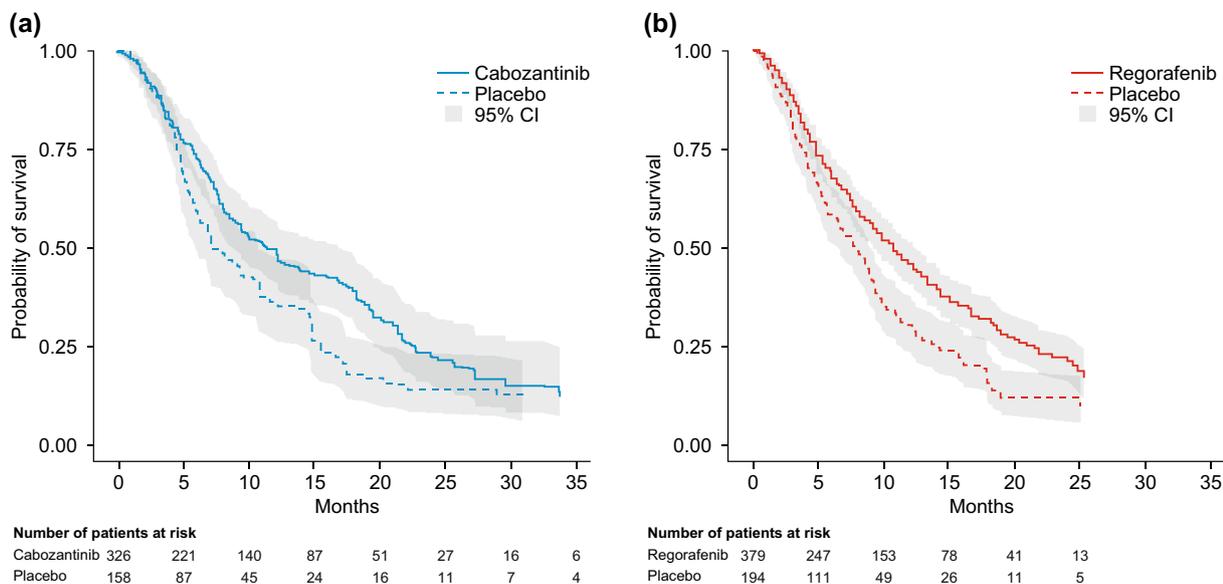
Fig. 2 Kaplan-Meier curves for overall survival in the matching-adjusted second-line CELESTIAL population (**a**) and the RESORCE population (**b**). *CI* confidence interval

Owing to the low frequency of grade 3 or 4 palmar-plantar erythrodysesthesia in the placebo-treated patients (one case in RESORCE; no cases in CELESTIAL) and no cases of grade 3 or 4 diarrhea in the RESORCE placebo arm, a meaningful anchored log-OR estimate could not be constructed for these TEAEs. Unanchored estimates were therefore computed. The unanchored analysis found no significant difference in rates of palmar-plantar erythrodysesthesia between the matching-adjusted cabozantinib and regorafenib populations ($p = 0.848$), but significantly lower rates of diarrhea in the regorafenib population compared with both the unmatched and matching-adjusted cabozantinib populations ($p = 0.001$ and $p < 0.001$ respectively) (Fig. 4 and Supplemental Table S6).

DISCUSSION

Outcome Interpretation

From the present analysis, we report comparative efficacy and safety estimates for cabozantinib and regorafenib for the second-line treatment of patients with advanced HCC after

prior sorafenib, using data from the phase 3 CELESTIAL and RESORCE trials. The treatment landscape for HCC is expanding rapidly, and there is a resultant need for head-to-head clinical trial data to guide second-line HCC treatment decisions. In this setting, indirect treatment comparisons offer standardized methods for generating comparative estimates that are widely accepted for health technology assessment [24, 25] and are increasingly recognized in the clinical sphere for their potential to guide clinicians in their decision-making [27–32].

A *standard* indirect treatment comparison of cabozantinib and regorafenib is not feasible owing to clinically relevant differences in the baseline characteristics of the RESORCE population and the second-line subpopulation from CELESTIAL. Therefore, a MAIC was selected as a more robust method of comparison, with an unanchored approach selected for the survival analysis on the basis of the results of three tests of the proportional hazards (PH) assumption, which indicated that the PH assumption was not satisfied. For OS, this conclusion was further validated by the identification of the log-logistic model as the best-fit parametric model; log-logistic models are accelerated failure time models

Table 3 Median survival estimates for the matching-adjusted second-line CELESTIAL population and the RESORCE population: weighted Kaplan-Meier estimates

		KM-derived estimate, months (median [95% CI])	*p* value
Overall survival			
Active treatment	Cabozantinib (ESS = 187)	11.4 (8.9–17.0)	0.3474[a]
	Regorafenib (*n* = 379)	10.6 (9.1–12.1)	
Placebo	CELESTIAL (ESS = 81)	7.2 (6.1–10.8)	NE
	RESORCE (*n* = 194)	7.8 (6.3–8.8)	
Progression-free survival			
Active treatment	Cabozantinib (ESS = 187)	5.6 (4.9–7.3)	0.0005[a]
	Regorafenib (*n* = 379)	3.1 (2.8–4.2)	
Placebo	CELESTIAL (ESS = 81)	1.9 (1.9–2.1)	NE
	RESORCE (*n* = 194)	1.5 (1.4–1.6)	

CI confidence interval, *ESS* effective sample size, *KM* Kaplan-Meier, *NE* not evaluated

[a] Log-rank test

and do not produce a single hazard ratio, making them incompatible with the PH assumption.

The matching and statistical adjustment steps of a MAIC analysis help to reduce the potential for bias by minimizing population differences that might modify or obscure the ability to discern true treatment effects. In the present analysis, weighting the baseline IPD for second-line CELESTIAL patients to align them with the baseline RESORCE characteristics was effective in reducing discernible population differences. In the unanchored survival analysis, cabozantinib was associated with similar OS and prolonged PFS compared with regorafenib. These findings were consistent across analyses for both the KM-derived and parametric modeling survival estimates. There is, however, a need for caution when interpreting the PFS result owing to differences in the tumor assessment schedules used in the CELESTIAL and RESORCE trials. RESORCE assessed tumor growth every 6 weeks for the first eight cycles and every 12 weeks thereafter during treatment.

In CELESTIAL, tumors were assessed every 8 weeks after randomization and performed until 8 weeks after radiographic progression or treatment/placebo discontinuation, whichever occurred later. Assessment of PFS was, therefore, initially more frequent for patients in RESORCE than in CELESTIAL and, thereafter, more frequent in CELESTIAL than in RESORCE. This difference may have introduced bias into the PFS result in the current analysis. At the individual patient level, the direction of any such bias would depend on the timing of tumor growth. It would, for example, favor cabozantinib if tumor growth occurred at week 10 (assessed at week 12 in RESORCE, but not until week 16 in CELESTIAL), yet favor regorafenib if tumor growth occurred at week 13 (assessed at week 16 in CELESTIAL, but not until week 18 in RESORCE). Thus, the overall direction of bias, if any, remains unclear.

The grade 3 or 4 TEAE profiles for regorafenib and cabozantinib were generally similar for both matching-adjusted and unmatched cabozantinib populations (Fig. 4). The
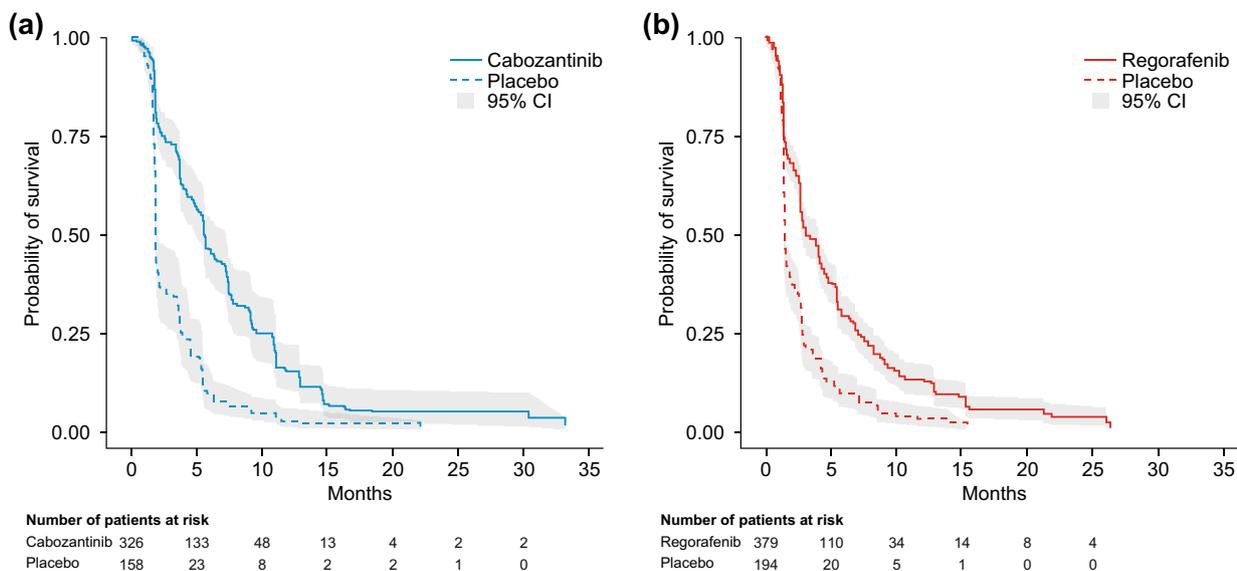
Fig. 3 Kaplan-Meier curves for progression-free survival in the matching-adjusted second-line CELESTIAL population (**a**) and the RESORCE population (**b**). *CI* confidence interval

exception was rates of grade 3 or 4 diarrhea, which were statistically significantly lower in the regorafenib population than in either the unmatched or matching-adjusted cabozantinib populations. There was also a trend toward higher rates of hypertension in the cabozantinib compared with regorafenib populations; however, the difference was not statistically significant. When interpreting the safety results, it is also noteworthy that sorafenib-intolerant patients were excluded from RESORCE, but not from CELESTIAL. The MAIC procedures were not able to adjust for this between-trial difference, resulting in the possibility of a bias toward higher rates of TKI treatment intolerance (and possibly later-stage disease) in the CELESTIAL population. Overall, the wide confidence intervals depicted in the forest plot of log-OR estimates for TEAEs with cabozantinib (versus regorafenib) indicate a high degree of imprecision in the estimates, likely arising from the very low frequency of grade 3 or 4 events in the placebo arms, particularly in CELESTIAL.

Limitations

Specific limitations associated with the PFS and TEAE estimates have been discussed. While MAIC procedures can reduce the impact of potentially effect-modifying baseline characteristics for reported covariates, they were not able to adjust for between-trial differences in assessment schedules or for the presence of sorafenib-intolerant patients in the CELESTIAL population (versus their exclusion from RESORCE). Such differences are unavoidable features of some indirect treatment comparisons [27, 28] and network meta-analyses [38, 39], but are relevant factors to consider when interpreting their results.

There are additional limitations to MAIC analyses that are also worth of consideration. Matching cannot account for all differences between trial populations, and it is possible that the results of this MAIC are affected by some residual between-trial differences, as evidenced by the difference in survival outcomes for the placebo arms despite matching and adjustment (Tables 3 and S5). Additional effect modifiers can exist between comparator trials, despite baseline weighting and matching. For example, treatment adherence following randomization can differ in comparator trials, affecting resultant drug exposure and influencing treatment-related outcomes. There may also be differences in unknown prognostic variables (i.e., covariates that affect outcome but do not alter
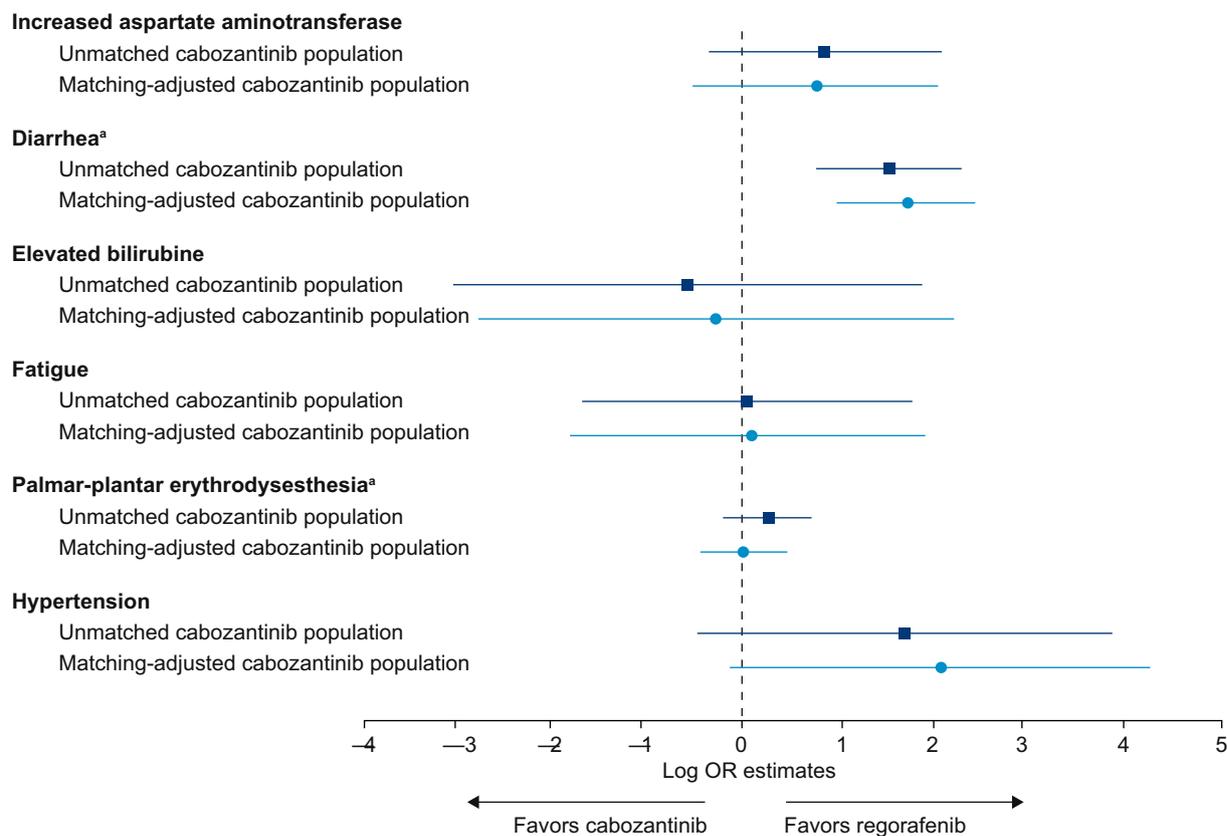
Fig. 4 Forest plot of grade 3 or 4 TEAE log-OR (95% CI) estimates for the unmatched and matching-adjusted second-line CELESTIAL populations compared with the RESORCE population. [a]Unanchored analysis. *CI* confidence interval, *OR* odds ratio, *TEAE* treatment-emergent adverse event

treatment effect). In an anchored analysis, purely prognostic variables do not affect interpretation of the results because they do not affect the relative treatment effects for each drug versus placebo because of within-study randomization. However, in the present unanchored analysis, where the comparison of cabozantinib and regorafenib is based on absolute rather than relative treatment outcomes, the results may be influenced by potential imbalances in unknown prognostic variables at study entry or (particularly relevant to OS) by heterogeneity in post-progression treatments. Heterogeneity in use of downstream therapies is another possible cofounding variable. In CELESTIAL, more than one-quarter of patients (26% cabozantinib; 33% placebo) received subsequent systemic or local liver-directed anticancer therapy [10]. The equivalent proportion

is not reported for RESORCE and is not known for the matching-adjusted second-line CELESTIAL population, but some degree of between-trial heterogeneity in use of post-progression treatments is likely.

Finally, the reduced effective sample size resulting from MAIC matching and adjustments decreases the statistical power of subsequent analyses. A lack of power in this analysis was reflected in the low rates of TEAE, and the resultant imprecision was evident in the high anchored log-ORs and large CIs for some TEAEs. Furthermore, the necessary use of an unanchored survival analysis and related discarding of placebo data can result in artificially narrow interval estimates, with arbitrary implications on statistical significance. For these reasons, the results of a MAIC cannot replace evidence from a randomized controlled trial.

Clinical Interpretation

Despite these limitations, when evidence from head-to-head trials is not available, a MAIC offers insight into the plausible comparative results for different therapies in a common clinical context, which can be valuable for healthcare professionals and health technology decision-makers. Such insights are particularly important when selecting the optimum treatment approach for patients with advanced HCC for whom the prognosis has traditionally been poor, with few therapeutic options available. First-line treatment in advanced HCC frequently fails after a period of time owing to adaptive or intrinsic resistance, disease progression or significant toxicity [17]: thus, there is a growing need for second- and later-line treatment options and for insights that will help to guide clinicians when selecting the optimum treatment sequence for their patients.

From a clinical perspective, the results of this analysis reinforce and further those of a recent network meta-analysis of randomized controlled trials of second-line agents for use in HCC after prior sorafenib [38]. The network meta-analysis included 13 trials and 11 different second-line therapies for advanced HCC and generated anchored efficacy and safety estimates using everolimus as the common comparator with OS as the primary endpoint. Of the 11 treatments included in the analysis, only cabozantinib and regorafenib significantly prolonged OS [38]. Overall, the authors concluded that second-line cabozantinib and regorafenib offered the best combination of efficacy and safety for patients with advanced HCC following prior sorafenib therapy. They further noted that the role of cabozantinib may be of particular importance in patients who were intolerant to sorafenib, because sorafenib-intolerant patients were excluded from the RESORCE trial of regorafenib [38]. In clinical practice, a range of factors must inform decision-making in optimum treatment sequencing. As well as the anticipated effectiveness and tolerability, treatment selection is influenced by patient characteristics (e.g., comorbidities) and their preferences, route of administration, dosing regimen and implications for patient quality of life. Also worthy of consideration are a patient's response to and tolerance of prior treatment as well as the molecular targets of available therapeutic options and whether they complement those of prior treatment and address possible mechanisms of resistance (e.g., AXL, MER) [20, 21].

CONCLUSION

MAIC analyses permit indirect comparisons between clinical trials with heterogeneous populations but with common treatment outcomes. In this MAIC analysis, which used data from the CELESTIAL and RESORCE trials, unanchored survival estimates suggest cabozantinib may be associated with similar OS and prolonged PFS compared with regorafenib in patients with advanced progressive HCC who were receiving second-line treatment following progression after prior sorafenib; regorafenib was associated with lower rates of diarrhea. Although this MAIC provides a useful indication of the comparative efficacy and safety of these two second-line agents, it is not a replacement for a head-to-head comparative trial.

ACKNOWLEDGEMENTS

REFERENCES

1. European Association for the Study of the Liver Clinical Practice Guidelines: management of hepatocellular carcinoma. J Hepatol. 2018;69(1):182–236. https://doi.org/10.1016/j.jhep.2018.03.019.

2. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc J-F, et al. Sorafenib in advanced hepatocellular carcinoma. New Engl J Med. 2008;359(4):378–90. https://doi.org/10.1056/NEJMoa0708857.

3. Bayer Schering Pharma. Nexavar® first FDA-approved drug therapy for liver cancer. 2007. https://www.investor.bayer.de/en/nc/news/archive/investor-news-2007/investor-news-2007/nexavarR-first-fda-approved-drug-therapy-for-liver-cancer/. Accessed 17 April 2020.

4. Kudo M. Targeted therapy for liver cancer: updated review in 2012. Curr Cancer Drug Targets. 2012;12(9):1062–72.

5. Kudo M, Finn RS, Qin S, Han K-H, Ikeda K, Piscaglia F, et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. Lancet. 2018;391(10126):1163–73. https://doi.org/10.1016/S0140-6736(18)30207-1.

6. Merck & Co. Inc. Eisai and Merck announce FDA approval of Lenvima® (lenvatinib) capsules for first-line treatment of unresectable hepatocellular carcinoma (HCC). 2018. https://investors.merck.com/news/press-release-details/2018/Eisai-And-Merck-Announce-FDA-Approval-Of-LENVIMA-lenvatinib-Capsules-For-First-line-Treatment-Of-Unresectable-Hepatocellular-Carcinoma-HCC/default.aspx. Accessed 17 April 2020.

7. Bayer. Bayer receives FDA approval for stivarga® (regorafenib) for the second-line systemic treatment of liver cancer. 2017. https://www.bayer.us/en/newsroom/press-releases/article/?id=123100. Accessed 17 April 2020.

8. 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(10064):56–66. https://doi.org/10.1016/S0140-6736(16)32453-9.

9. Ipsen. Ipsen receives positive CHMP opinion for Cabometyx® (cabozantinib) for the second-line treatment of patients with hepatocellular carcinoma (HCC). 2018. https://www.ipsen.com/press-releases/ipsen-receives-positive-chmp-opinion-for-cabometyx-cabozantinib-for-the-second-line-treatment-of-patients-with-hepatocellular-carcinoma-hcc/. Accessed 17 April 2020.

10. 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. New Engl J Med. 2018;379(1):54–63. https://doi.org/10.1056/NEJMoa1717002.

11. Eli Lilly and Company. Lilly's Cyramza® (ramucirumab) becomes first FDA-approved biomarker-driven therapy in patients with hepatocellular carcinoma. 2019. https://investor.lilly.com/news-releases/news-release-details/lillys-cyramzar-ramucirumab-becomes-first-fda-approved-biomarker. Accessed 17 April 2020.

12. 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;16(7):859–70. https://doi.org/10.1016/S1470-2045(15)00050-9.

13. Bristol-Myers Squibb. Prescribing Informaton: Opdivo (nivolumab) injection, for intravenous use. 2020. https://www.accessdata.fda.gov/drugsatfda_docs/label/2020/125554s078lbl.pdf. Accessed 17 April 2020.

14. Bristol-Myers Squibb. Prescribing Information: Yervoy® (ipilimumab) injection, for intravenous use. 2020 https://www.accessdata.fda.gov/drugsatfda_docs/label/2020/125377s108lbl.pdf. Accessed 17 April 2020.

15. Merck & Co. Inc. FDA Approves Merck's KEYTRUDA® (pembrolizumab) for the Treatment of Patients with Hepatocellular Carcinoma (HCC) Who Have Been Previously Treated with Sorafenib. https://investors.merck.com/news/press-release-details/2018/FDA-Approves-Mercks-KEYTRUDA-pembrolizumab-for-the-Treatment-of-Patients-with-Hepatocellular-Carcinoma-HCC-Who-Have-Been-Previously-Treated-with-Sorafenib/default.aspx. Accessed 17 April 2020.

16. Finn RS, Ryoo BY, Merle P, Kudo M, Bouattour M, Lim HY, et al. Pembrolizumab as second-line therapy in patients with advanced hepatocellular carcinoma in KEYNOTE-240: a randomized, double-

blind. Phase III Trial J Clin Oncol. 2020;38(3): 193–202. https://doi.org/10.1200/jco.19.01307.

17. Ray EM, Sanoff HK. Optimal therapy for patients with hepatocellular carcinoma and resistance or intolerance to sorafenib: challenges and solutions. J Hepatocell Carcinoma. 2017;4:131–8. https://doi.org/10.2147/JHC.S124366.

18. American Cancer Society. Liver cancer survival rates. 2020. https://www.cancer.org/cancer/liver-cancer/detection-diagnosis-staging/survival-rates.html. Accessed 17 April 2020.

19. Bayer Plc. Summary of product characteristics: Stivarga 40 mg film-coated tablets. 2018. https://www.medicines.org.uk/emc/product/1263/smpc. Accessed 17 April 2020.

20. Ipsen Ltd. Summary of Product Characteristics: Cabometyx 20 mg Film-coated Tablets. 2019. https://www.medicines.org.uk/emc/product/4331/smpc. Accessed 17 April 2020.

21. Mukherjee SK, Wilhelm A, Antoniades CG. TAM receptor tyrosine kinase function and the immunopathology of liver disease. Am J Physiol Gastrointest Liver Physiol. 2016;310(11): G899–G905. https://doi.org/10.1152/ajpgi.00382.2015.

22. Vogel A, Cervantes A, Chau I, Daniele B, Llovet J, Meyer T, et al. Hepatocellular carcinoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2018;29(Suppl 4): iv238–iv55. https://doi.org/10.1093/annonc/mdy308.

23. National Comprehensive Cancer Network. NCCN clinical practice guidelines in onclogy (NCCN Guidelines®): hepatobiliary cancers. Version 1. 2018 - February 14, 2018. https://www.nccn.org/professionals/physician_gls/pdf/hepatobiliary.pdf. Accessed 17 April 2020.

24. Thom H, Jugl SM, Palaka E, Jawla S. Matching adjusted indirect comparisons to assess comparative effectiveness of therapies: usage in scientific literature and health technology appraisals. Value Health. 2016;19(3):A100–A101101. https://doi.org/10.1016/j.jval.2016.03.1723.

25. Phillippo DM AA, Dias S, Palmer S, Abrams KR, Welton NJ. NICE DSU technical support document 18: methods for population-adjusted indirect comparisons in submissions to NICE: 2016. https://www.nicedsu.org.uk/wp-content/uploads/2017/05/Population-adjustment-TSD-FINAL.pdf. Accessed 17 April 2020.

26. Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JPT. Evaluating the quality of evidence from a network meta-analysis. PLoS One. 2014;9(7): e99682–e. https://doi.org/10.1371/journal.pone.0099682.

27. Chowdhury S, Oudard S, Uemura H, Joniau S, Pilon D, Lefebvre P, et al. Matching-adjusted indirect comparison of health-related quality of life and adverse events of apalutamide versus enzalutamide in non-metastatic castration-resistant prostate cancer. Adv Ther. 2020;37(1):512–26. https://doi.org/10.1007/s12325-019-01157-4.

28. Tremblay G, Chandiwana D, Dolph M, Hearnden J, Forsythe A, Monaco M. Matching-adjusted indirect treatment comparison of ribociclib and palbociclib in HR+, HER2- advanced breast cancer. Cancer Manag Res. 2018;10:1319–27. https://doi.org/10.2147/CMAR.S163478.

29. Odom D, Mladsi D, Purser M, Kaye JA, Palaka E, Charter A, et al. A matching-adjusted indirect comparison of sonidegib and vismodegib in advanced basal cell carcinoma. J Skin Cancer. 2017. https://doi.org/10.1155/2017/6121760.

30. Venkatachalam M, Stenehjem D, Parikh ND, Singh P, Marett B, Sill B, et al. Cost of treatment-related adverse events (TRAES) in second-line (2l) advanced hepatocellular carcinoma (AHCC): match adjusted indirect comparison (MAIC) of nivolumab and regorafenib. Value Health. 2017;20(9): A502–A503503. https://doi.org/10.1016/j.jval.2017.08.590.

31. Roskell N, Gregory J, Wisniewski T, Thompson GJ, De la Cruz C. Indirect treatment comparision of nivolumab versu regorafenib, cabozantinib and best supportive care after treatent with sorafenib for hepatocellular carcinoma (PCN17). Value Health. 2018;21:S17–S1818. https://doi.org/10.1016/j.jval.2018.09.099.

32. Parikh N, Marshall A, Huff KD, Savidge R, Betts KA, Song J, et al. Comparative efficacy of second-line treatments for advanced hepatocellular carcinoma: a network meta-analysis. J Clin Oncol. 2020;38(4_suppl):545. https://doi.org/10.1200/JCO.2020.38.4_suppl.545.

33. Latimer N. NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data: 2011. https://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf. Accessed 17 April 2020.

34. Cole SR, Hernán MA. Adjusted survival curves with inverse probability weights. Comput Methods Programs Biomed. 2004;75(1):45–9. https://doi.org/10.1016/j.cmpb.2003.10.004.

35. Link CL. Confidence intervals for the survival function using Cox's proportional-hazard model with covariates. Biometrics. 1984:601–9. https://doi.org/10.2307/2530904; https://www.jstor.org/stable/2530904.

36. Woodruff RS. Confidence intervals for medians and other position measures. J Am Stat Assoc. 1952;47(260):635–46. https://doi.org/10.1080/01621459.1952.10483443.

37. Xue Y, Schifano ED. Diagnostics for the Cox model. Commun Stat Appl Methods. 2017;24(6):583–604. https://doi.org/10.29220/CSAM.2017.24.6.583.

38. Bakouny Z, Assi T, El Rassy E, Nasr F. Second-line treatments of advanced hepatocellular carcinoma: systematic review and network meta-analysis of randomized controlled trials. J Clin Gastroenterol. 2019;53(4):251–61. https://doi.org/10.1097/MCG.0000000000001160.

39. Schmidt E, Lister J, Neumann M, Wiecek W, Fu S, Vataire AL, et al. Cabozantinib versus standard-of-care comparators in the treatment of advanced/metastatic renal cell carcinoma in treatment-naive patients: a systematic review and network meta-analysis. Target Oncol. 2018;13(2):205–16. https://doi.org/10.1007/s11523-018-0559-0.