Adv Ther
https://doi.org/10.1007/s12325-021-01700-2

ORIGINAL RESEARCH

Comparative Efficacy of Cabozantinib and Ramucirumab After Sorafenib for Patients with Hepatocellular Carcinoma and Alpha-fetoprotein ≥ 400 ng/mL: A Matching-Adjusted Indirect Comparison

Jörg Trojan · Patrick Mollon · Bruno Daniele · Florence Marteau · Lidia Martín · Yuxin Li · Qing Xu · Fabio Piscaglia · Renata Zaucha · Debashis Sarker · Ho Yeong Lim · Marino Venerito

Received: December 21, 2020 / Accepted: March 6, 2021
© The Author(s) 2021

ABSTRACT

Introduction: Cabozantinib and ramucirumab are approved for the treatment of adults with hepatocellular carcinoma (HCC) following prior sorafenib treatment; ramucirumab is restricted to use in patients with serum alpha-fetoprotein (AFP) ≥ 400 ng/mL. This matching-adjusted indirect comparison evaluated the efficacy and safety of both drugs after sorafenib in patients with HCC and AFP ≥ 400 ng/mL.

Methods: Individual patient data (IPD) from the CELESTIAL trial (cabozantinib) and population-level data from the REACH-2 trial (ramucirumab) were used. To align with REACH-2, the CELESTIAL population was limited to patients who received first-line sorafenib only and had baseline serum AFP ≥ 400 ng/mL. The IPD from CELESTIAL were weighted to balance the distribution of 11 effect-modifying baseline characteristics with those of REACH-2. Overall survival (OS; primary endpoint) and progres-

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12325-021-01700-2.

J. Trojan (✉)
Gastrointestinal Oncology Unit, Medical Clinic 1, Goethe University Hospital Frankfurt, Theodor-Stern-Kai 7, 60590 Frankfurt, Germany
e-mail: trojan@em.uni-frankfurt.de

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

B. Daniele
Oncology Unit, Ospedale del Mare, Naples, Italy

L. Martín
Ipsen Pharma, Barcelona, Spain

Y. Li
IQVIA Ltd, London, UK

Q. Xu
IQVIA Inc, Beijing, China

F. Piscaglia
Department of Digestive Disease and Internal Medicine, University of Bologna, Bologna, Italy

R. Zaucha
Department of Oncology, Medical University of Gdańsk, Gdańsk, Poland

D. Sarker
Comprehensive Cancer Centre, King's College London, London, UK

H. Y. Lim
Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

M. Venerito
Department of Gastroenterology, Otto von Guericke University Hospital, Magdeburg, Germany

Published online: 06 April 2021

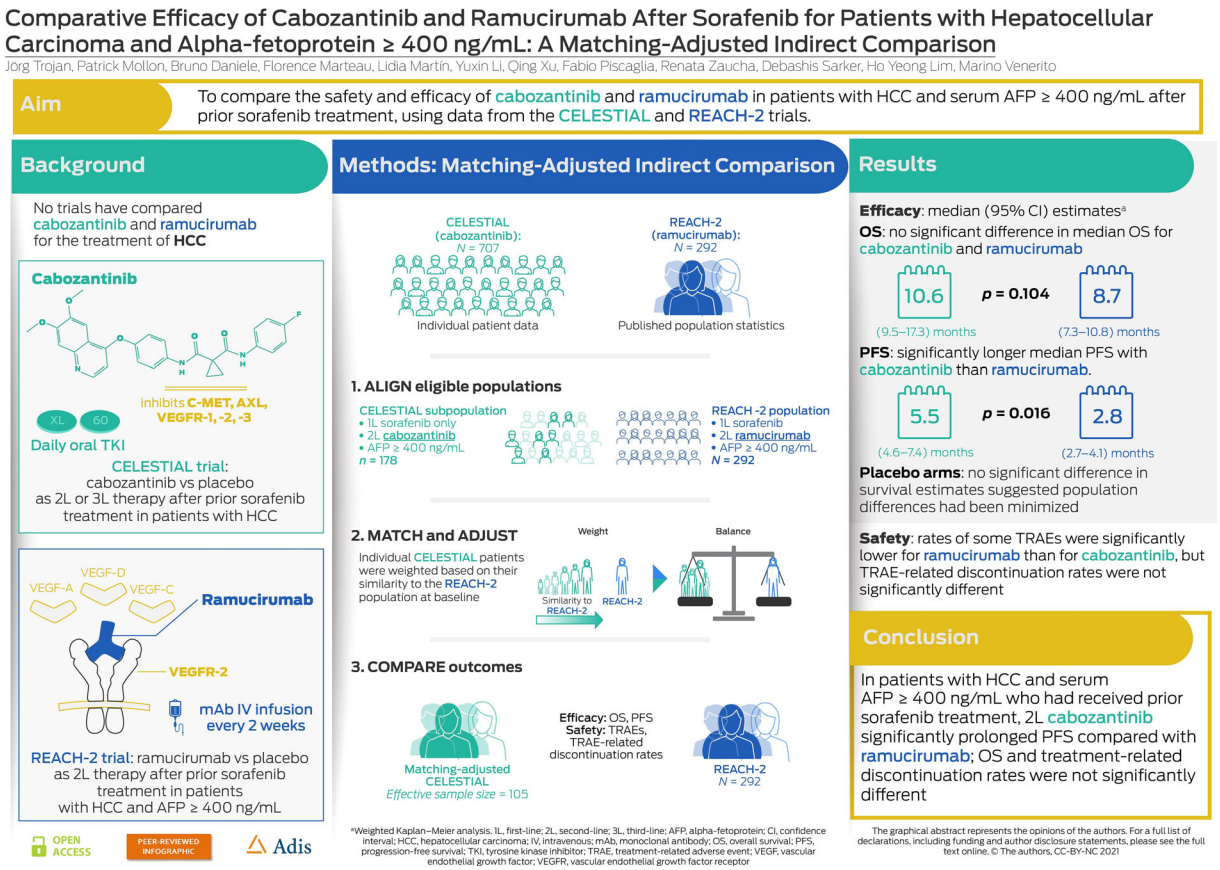
sion-free survival (PFS) were compared for the CELESTIAL (matching-adjusted) and REACH-2 populations using weighted Kaplan-Meier (KM) curves and parametric (OS, Weibull; PFS, log-logistic) modeling. Rates of treatment-related adverse events (TRAEs) and TRAE-related discontinuations were also compared.

Results: After matching and weighting, baseline characteristics were balanced between populations (REACH-2, $N = 292$; CELESTIAL, effective sample size = 105). Weighted KM estimates for OS (median [95% CI]) were not significantly different between cabozantinib and ramucirumab (10.6 [9.5–17.3] months versus 8.7 [7.3–10.8] months; $p = 0.104$), but PFS was significantly longer for cabozantinib than for ramucirumab (5.5 [4.6–7.4] months versus 2.8 [2.7–4.1] months; $p = 0.016$). Parametric modeling results were consistent with the weighted KM analysis. Rates of some grade 3 or 4 TRAEs were lower with ramucirumab than with cabozantinib; however, TRAE-related discontinuation rates were similar ($p = 0.271$).

Conclusion: In this MAIC, cabozantinib significantly prolonged median PFS compared with ramucirumab after prior sorafenib treatment in patients with HCC and AFP $\geq$ 400 ng/mL; rates of some grade 3 or 4 TRAEs were lower with ramucirumab than cabozantinib but related discontinuation rates were not significantly different between treatments.

Trial Registration: Clinical trials.gov identifiers: CELESTIAL trial (NCT01908426) and REACH-2 trial (NCT02435433).

Graphic abstract: These slides can be retrieved under Electronic Supplementary Material.



Keywords: Alpha-fetoprotein (AFP); Cabozantinib; Hepatocellular carcinoma (HCC); Indirect treatment comparison (ITC); Matching-adjusted indirect comparison (MAIC); Monoclonal antibody (mAb); Ramucirumab; second-line treatment / 2L treatment; Tyrosine kinase inhibitor (TKI); Vascular endothelial growth factor (VEGF)

Key Summary Points

Why carry out this study?

Cabozantinib (a once-daily oral tyrosine kinase inhibitor with targets including MET, AXL and vascular endothelial growth factor [VEGF] receptors) and ramucirumab (a monoclonal antibody specific for VEGF receptor-2, administered every 2 weeks by intravenous infusion) are approved for the second-line (2L) treatment of hepatocellular carcinoma (HCC) after sorafenib; however, ramucirumab is restricted to use in patients with serum alpha-fetoprotein (AFP) levels of 400 ng/mL or above.

No clinical trials have compared cabozantinib and ramucirumab directly in patients with HCC and elevated serum AFP. This analysis compared 2L cabozantinib and ramucirumab in patients with HCC and serum AFP of 400 ng/mL or above using a matching-adjusted indirect comparison (MAIC) approach.

What was learned from this study?

In 2L HCC populations matched for prior therapy and clinically relevant baseline characteristics, cabozantinib significantly prolonged median (95% CI) progression-free survival compared with ramucirumab (5.5 [4.6–7.4] months vs. 2.8 [2.7–4.1] months; $p = 0.016$); overall survival (median [95% CI]) was not significantly different for cabozantinib (10.6 [9.5–17.3] months) and ramucirumab (8.7 [7.3–10.8] months) ($p = 0.104$).

Discontinuation rates resulting from treatment-related adverse events were not significantly different for the matching-adjusted cabozantinib population and the ramucirumab population (log odds ratio [95% CI] for cabozantinib vs. ramucirumab, 1.16 [−0.89, 3.20]; $p = 0.271$).

In the absence of direct comparative trial data, this MAIC analysis may help to inform individualized clinical decision-making with respect to 2L treatment for patients with HCC and elevated serum AFP who have received prior sorafenib.

DIGITAL FEATURES

This article is published with digital features, including a summary slide, graphical plain language summary, infographic, author video, and animation video to facilitate understanding of the article. To view digital features for this article, go to https://doi.org/10.6084/m9.figshare.14141435.

INTRODUCTION

As many as one-third of patients treated for advanced hepatocellular carcinoma (aHCC) are eligible for second-line (2L) treatment [1]. The recent approval of a number of 2L vascular endothelial growth factor (VEGF)-targeted therapies with proven benefit over placebo holds the potential to preserve and to prolong the life of a substantial number of patients with HCC [2–7]. Yet, there is limited comparative evidence to inform the selection of VEGF-targeting agents and to guide optimum treatment sequencing for patients with HCC.

The range of VEGF-targeted therapies approved for 2L use after prior sorafenib treatment in patients with HCC includes the tyrosine kinase inhibitors (TKIs) cabozantinib and regorafenib and the monoclonal antibody (mAb) ramucirumab, which is limited to use in

patients with serum alpha-fetoprotein (AFP) of 400 ng/mL or above [2–7]. All three agents target VEGF-related angiogenesis and tumor neovascularization, but there are clinically relevant differences in their mechanisms of action, modes of administration and indications. Cabozantinib and regorafenib are both oral TKIs (administered daily) with activity against multiple receptor kinases [2–5]. Cabozantinib has activity against a broader range of receptor kinases than regorafenib, including MET (hepatocyte growth factor receptor protein), AXL (GAS6 receptor), RET, ROS1, TYRO3, MER, KIT (stem-cell-factor receptor), TRKB, Fms-like tyrosine kinase-3 and TIE-2 as well as VEGF receptors [2–5]. It is noteworthy that overexpression of MET and AXL has a negative impact on HCC prognosis and that elevated MET expression is associated with sorafenib resistance [8]. Based on the results of the phase 3 CELESTIAL trial (NCT01908426), which involved patients with HCC receiving 2L or third-line (3L) treatment, cabozantinib is approved in the USA and Europe for use in patients with HCC who have been previously treated with sorafenib [2, 3, 9].

Ramucirumab, in contrast, is a recombinant monoclonal human immunoglobulin G1 antibody that is specific for VEGF receptor-2 and is administered as an intravenous (IV) infusion once every 2 weeks [6, 7]. The agent is approved as a 2L treatment for patients with HCC after prior sorafenib treatment based on the results of the phase 3 REACH-2 trial (NCT02435433), but only in the subgroup of patients with serum AFP levels of 400 ng/mL or above [6, 7, 10]. Approximately 50% of HCC tumors secrete AFP, and a plasma AFP level > 400 ng/mL is generally considered to support a diagnosis of HCC; elevated levels are associated with poor prognosis [11, 12].

International management guidelines for HCC broadly align with the approved indications for these VEGF-targeting agents, endorsing their use in the 2L setting but with agent-specific caveats reflecting their respective phase 3 trial designs and resultant labels [13–15]. In the USA, guidance from the National Comprehensive Cancer Network also now features checkpoint inhibitor regimens, approved by the US Food and Drug Administration (pembrolizumab or nivolumab with/without ipilimumab) as additional 2L options following prior sorafenib treatment [13]. No HCC randomized controlled trials (RCTs) have compared the available 2L treatment options directly, resulting in limited evidence to inform optimum treatment sequencing in patients with progressive disease.

A recent matching-adjusted indirect comparison (MAIC), using data from the CELESTIAL trial (cabozantinib; NCT01908426) and the RESORCE trial (regorafenib; NCT01774344) provided indirect comparative data for the efficacy and safety of cabozantinib and regorafenib [16]. The MAIC method is a recognized means of addressing comparative evidence gaps that often exist around new treatments and licensed alternatives [17, 18]. Initially used to inform critical reimbursement decisions, the MAIC method of treatment comparison is increasingly being used to guide clinical decision-making [19, 20]. No MAIC data are currently available for cabozantinib and ramucirumab.

An exploratory analysis of CELESTIAL showed that cabozantinib improved overall survival (OS) and progression-free survival (PFS) compared with placebo across a range of serum AFP levels in a mixed 2L and 3L population [21]. Outcomes were stratified by baseline serum AFP level bifurcated at 400 ng/mL. For OS with cabozantinib, the hazard ratio (HR) (95% confidence interval [CI]) was 0.81 (0.62–1.04) for patients with baseline serum AFP < 400 ng/mL and 0.71 (0.54–0.94) for patients with levels of 400 ng/mL or above. HRs for PFS were consistent with those for OS [21].

Building on these data, here we used the MAIC method to compare the efficacy and safety of cabozantinib and ramucirumab using data from the CELESTIAL and REACH-2 pivotal trials in populations matched in terms of their prior therapy and baseline characteristics [9, 10].

METHODS

Data Source and Analysis Approach

The phase 3 CELESTIAL and REACH-2 trial populations were selected for this comparative study based on the similarity of their trial designs and evaluated outcomes (Table S1) [9, 10]. Individual patient data (IPD) were available from the CELESTIAL trial, and the published aggregate population-level statistics were available from REACH-2.

To align with the (post-sorafenib) 2L population of patients with serum AFP levels of 400 ng/mL or above who were eligible for REACH-2, the subgroup of patients from CELESTIAL with baseline serum AFP levels of 400 ng/mL or above and who had received prior sorafenib only was identified. The baseline characteristics of this CELESTIAL subpopulation ($N = 178$) were compared with those of the REACH-2 population ($N = 292$) to assess the feasibility of a meaningful standard indirect treatment comparison (ITC) (Table S2). The presence of residual differences in potentially effect-modifying baseline characteristics (e.g., in race, region, duration of prior sorafenib treatment and HCC etiology) indicated that a standard ITC would be subject to potential confounding. Instead, a MAIC method was selected to reduce the risk of bias that would be associated with a standard ITC [17, 18, 20].

The MAIC method requires IPD to be available for one trial (here CELESTIAL) so that each patient's contribution to the analysis can be weighted based on their similarity to the comparator trial population (here, REACH-2) at baseline. The aim of this patient-level weighting/adjusting was to match the clinically relevant baseline characteristics of the CELESTIAL population to those of the REACH-2 population, thereby minimizing potential effect modifying differences and sources of outcome bias.

Population Matching

Primary Analysis

A panel of clinical experts met (April 17, 2020) to review the baseline characteristics of the aligned CELESTIAL subpopulation and the REACH-2 population. They identified 11 potential effect modifiers of OS between the populations that were deemed clinically relevant for matching: age, sex, duration of prior sorafenib treatment, baseline HCC (extrahepatic spread, macrovascular invasion), etiology of HCC (hepatitis B, hepatitis C, non-viral), AFP level (median of $\log_{10}$(AFP)), Barcelona Clinic Liver Cancer stage B and albumin-bilirubin (ALBI) grade 1.

Sensitivity Analysis

As a sensitivity analysis, effect modifiers of OS were also identified empirically using statistical modeling. This was done by adding or eliminating potential effect-modifying baseline characteristics to/from a regression model using a stepwise approach (forwards, backwards, both directions). In all model directions, the following seven variables were identified as having the lowest Akaike information criterion (AIC) values and therefore as being the most predictive of treatment effect [22, 23] on OS: age, sex, etiology of alcohol use, macrovascular invasion, duration of prior sorafenib treatment, ALBI grade and AFP value.

Validation Analysis

Despite population alignment by subgroup selection, median (interquartile range) baseline AFP level remained substantially higher for the subgroup of 2L CELESTIAL patients with AFP of 400 ng/mL or higher than for the REACH-2 population (8813 [1648–30,751] ng/mL vs. 3394 [1177–16,812] ng/mL, respectively). To aid matching, AFP measurements were therefore log transformed and median $\log_{10}$(AFP) used as

the matching criterion in the primary analysis. Minor residual differences in median $\log_{10}$(AFP) persisted, in part because of the method used to calculate weights (centering moments based on mean rather than median population values) and because baseline AFP values for the eligible CELESTIAL subpopulation did not follow a Normal distribution. To assess the implications of this residual difference, a validation analysis was conducted that repeated the primary analysis but with AFP excluded as a matching criterion. By removing AFP as a matching criterion, it was possible to investigate the extent to which differences in baseline serum AFP level for the comparator populations may have influenced the outcome analysis. If any outcome differences seen in the primary analysis were driven by minor residual differences in baseline AFP level, the impact of much larger baseline AFP differences would be clearly evident in the results of the validation analysis.

Survival Outcomes

Survival outcomes (OS [primary outcome from the trials] and PFS) were evaluated for the matching-adjusted CELESTIAL population and compared with those for the REACH-2 population. Median OS and PFS estimates for the CELESTIAL arms were derived from weighted Kaplan-Meier (KM) curves fitted to the survival data and compared with those fitted to simulated IPD for the REACH-2 population. CIs were generated for the weighted KM analysis using Woodruff's method [24].

Active Treatment Arms

The validity of comparing the survival outcomes for cabozantinib and ramucirumab using an anchored analysis was assessed by testing the proportional hazards (PH) assumption for the OS and PFS outcomes. An anchored analysis derives relative survival estimates for the active treatment arms anchored via a common comparator (here, placebo) and generates a HR for relative effect comparison. For an anchored analysis to be meaningful, however, the PH assumption must hold true (i.e., the ratio of the hazards for any two individuals must be constant over time). In the present analysis, three methods were used to test the PH assumption: (1) visual inspection of the log-cumulative hazard plots for cabozantinib and ramucirumab (a pattern of non-parallelism indicates PH violation); (2) visual inspection of the scaled Schoenfeld residuals over time (systematic departure from the horizontal indicates PH violation); (3) the Grambsch-Therneau statistical test of the scaled Schoenfeld residuals (the test evaluates whether there is a non-zero slope in a generalized linear regression of the scaled residuals as functions of time; non-zero values indicate PH violation) [20, 25, 26].

If the PH assumption does not hold true, best-practice guidance for MAIC analyses advises use of an unanchored analysis and generation of absolute outcomes for each treatment arm (not relative to a common [placebo] comparator) [20]. This approach involves fitting individual parametric survival curves to each treatment arm with best-fit models identified by AIC/Bayesian information criterion (BIC) analysis. Potential models are ranked (based on AIC and BIC) for each treatment arm. The sum of the model rankings for the two treatment arms is used to identify the best-fit model with the superior model for the analysis indicated by the lowest sum rank. Where more than one model is potentially eligible, the sum of the AIC/BIC values is also taken into consideration (the superior model having the lowest sum value) and the choice of model is validated by visual assessment of its fit to weighted KM and log-cumulative hazard curves [20].

Placebo Comparison

To assess the effectiveness of the MAIC weighting method for matching the baseline characteristics of the trial populations, negative outcome controls (i.e., outcomes for the weighted-CELESTIAL and REACH-2 placebo arms) were evaluated. Similarity of outcomes for the matching-adjusted and comparator placebo arms indicates that any differences between the original trial populations at baseline have been reduced, so the potential for bias in the active treatment arm comparison has also been reduced, thus validating the MAIC approach.

For the matching-adjusted CELESTIAL and REACH-2 placebo arms, median (95% CI) estimates were computed for OS and PFS and HR (95% CI) estimates using weighted Cox PH regression models.

Safety Outcomes

The primary publication from the CELESTIAL trial reported all-cause AEs [9], but treatment-related AEs were published for REACH-2 [10]. Using the CELESTIAL IPD, the incidences of treatment-related AEs (TRAEs) of interest (any grade, and grade 3 or 4) were identified for the matching-adjusted cabozantinib population and compared with those for the ramucirumab population. Events of interest were TRAEs reported in both CELESTIAL and REACH-2 that affected at least 5% of patients in any of the trial arms. Rates of treatment discontinuation due to TRAEs were also compared for the matching-adjusted cabozantinib population and the ramucirumab population.

In an anchored analysis, TRAE estimates (log odds ratios [ORs]) were generated for the matching-adjusted CELESTIAL population and compared with log ORs for ramucirumab versus placebo, computed from the published REACH-2 data. Anchored estimates were not feasible for AEs that did not occur in one of the trial arms. In such instances, unanchored estimates of the log ORs were calculated, with the number of TRAEs occurring in CELESTIAL being used to compute a weighted estimate [22]. This method multiplied the weighted mean of the TRAE occurrences by the number of patients in the corresponding trial arm. The delta method was then used to compute variances, and 95% CIs and p values were estimated for the log OR of each TRAE.

Safety outcomes were not evaluated for the matching-adjusted CELESTIAL placebo population and the REACH-2 placebo population because of their low rates of AEs.

Analyses

The analyses were performed using R version 4.0.2 (R Core Team, 2020). The package 'survey' version 4.0 was used to fit weighted survival models with weights computed from the MAIC that were used as sampling weights.

Compliance with Ethics Guidelines

The results presented in this article are based on published studies. All procedures performed in those studies involving human participants were conducted in accordance with the ethical standards of the local Institutional Review Boards for each site and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the CELESTIAL and REACH-2 trials.

RESULTS

Patient Characteristics

In total, 292 REACH-2 patients were included in the analysis, of whom 197 received ramucirumab and 95 received placebo. Of the 707 patients randomized in CELESTIAL, 202 had received prior sorafenib treatment only and had a serum AFP level of 400 ng/mL or above. Of these patients, 178 had complete baseline data and were eligible for matching, of whom 114 received cabozantinib and 64 received placebo. Following baseline weighting and adjustments, the effective sample sizes (ESSs) from CELESTIAL were 63 patients allocated to cabozantinib and 44 to placebo (Table 1).

Application of MAIC weighting to the baseline IPD from CELESTIAL was effective in aligning the potentially effect-modifying baseline characteristics of the 2L CELESTIAL population with serum AFP of 400 ng/mL or above with those of the REACH-2 population (Table 2; Fig. S1).

Survival Outcomes

KM-Derived Estimates

There was no significant difference in the OS estimates (median [95% CI]) derived from the

Table 1 MAIC population sizes

	CELESTIAL population					REACH-2 population
	Unmatched		Matched (primary analysis)	Matched (sensitivity analysis)	Matched (validation analysis)	
	N (overall)	*N* (2L and AFP ≥ 400 ng/mL)	ESS	ESS	ESS	*N* (published)
Active treatment	470	114	63	73	80	197
Placebo	237	64	44	46	49	95
Total	707	178	105[a]	119[a]	128[a]	292

2L second line, *AFP* alpha-fetoprotein, *ESS* effective sample size, *MAIC* matching-adjusted indirect comparison

[a] Total ESS is nonadditive with respect to each treatment arm, calculated as the squared sum of the weights divided by the sum of the squared weights, for each arm

weighted KM analysis for the matching-adjusted cabozantinib population (10.6 [9.5, 17.3] months) and for the ramucirumab population (8.7 [7.3, 10.8] months) ($p = 0.104$, log-rank test) (Table 3; Fig. 1a). Median (95% CI) PFS estimates, however, were significantly longer for the matching-adjusted cabozantinib population than for the ramucirumab population: 5.5 (4.6, 7.4) months versus 2.8 (2.7, 4.1) months ($p = 0.016$, log-rank test) (Table 3; Fig. 1b).

Parametric Modeling Estimates

Tests of the PH assumption indicated that an anchored analysis (including hazard ratio estimates) was not supported and that an unanchored parametric modeling analysis would provide more meaningful survival estimates. This assessment was based on visual inspection of the log of cumulative hazard versus time plots, which displayed patterns of non-parallelism for both OS and (more distinctly) PFS, suggesting a violation of the PH assumption. This assessment was confirmed by plots of scaled Schoenfeld residuals versus time for both outcomes, which showed a systematic departure from the horizontal for both outcomes, as did the non-zero slopes (at the 5% significance level) given by the Grambsch–Therneau test (Fig. S2 and 3). Accordingly, and in agreement with recommended practice [20], an unanchored analysis was conducted by fitting individual parametric survival curves to each treatment arm.

The results of the parametric modeling analysis mirrored those of the weighted KM analysis (Table 3). For OS, the Weibull model had the lowest sum rank and AIC/BIC values and, following validation of model choice by visual inspection, was selected as the best-fit model (Table S3). Estimated median (95% CI) OS was 12.0 (9.6, 14.5) months for the matching-adjusted cabozantinib population versus 9.6 (8.4, 10.8) months for the ramucirumab population (Table 3). For PFS, the log-logistic model was selected as the best-fit model, generating median (95% CI) PFS estimates of 5.2 (4.1, 6.4) months for the matching-adjusted cabozantinib population versus 3.2 (2.8, 3.6) months for the ramucirumab population (Tables 3 and S4).

Negative Control Estimates

There were no significant differences in OS or PFS estimates for the matching-adjusted CELESTIAL and the REACH-2 placebo groups. Estimated median (95%) OS was 5.3 (4.8–8.2) months for the matched-adjusted CELESTIAL placebo population, and 7.4 (5.8–9.4) months for the REACH-2 placebo population. For PFS, median (95% CI) estimates were 1.9 (1.8–2.1) months for the matching-adjusted CELESTIAL placebo population and 1.6 (1.4–2.6) months for the REACH-2 placebo population (Table S5). Similarly, the results of the parametric modeling analysis showed no significant difference in either OS or PFS survival estimates for the

Table 2 Baseline matching characteristics used in the primary MAIC analysis, before and after matching[a]

Baseline characteristics	CELESTIAL 2L population with AFP ≥ 400 ng/mL		REACH-2 population
	Unmatched (N = 202)	Matching-adjusted (N = 105)	Unmatched (N = 292)
Age under 65 years, %	57.87	51.71	51.71
Female, %	21.91	20.21	20.21
Duration of prior sorafenib treatment < 5 months, %[b]	44.94	57.19	57.19
Extrahepatic disease, %	76.40	72.26	72.26
Macrovascular invasion, %	34.83	35.27	35.27
Etiology, %			
Hepatitis B	44.38	36.64	36.64
Hepatitis C	22.47	26.03	26.03
Nonviral[c]	30.34	31.51	31.51
AFP, median $\log_{10}$(AFP)[d,e]	3.94	3.00	3.53
ALBI grade 1, %	34.83	48.97	48.97
BCLC stage B, %	8.99	18.49	18.49

2L second line, *AFP* alpha-fetoprotein, *ALBI*, albumin–bilirubin, *BCLC* Barcelona Clinic Liver Cancer, *HCC* hepatocellular carcinoma, *IQR* interquartile range, *MAIC* matching-adjusted indirect comparison

[a] Matching variables selected as potential effect modifiers by expert clinical panel

[b] Categorization of prior sorafenib treatment use in (and published for) the REACH-2 trial

[c] HCC of nonviral etiology was not recorded directly in the REACH-2 trial. Estimate is derived from the sum of patients with etiology of alcohol use plus nonalcoholic steatohepatitis fatty liver, using the total REACH-2 population size as the denominator. There might be overlap in patients between these two etiology categories

[d] Reported as median of $\log_{10}$(AFP) because of the magnitude of difference in median (IQR) AFP for the pre-matched CELESTIAL and REACH-2 populations: median (IQR) values of 8813 (1648, 30,751) ng/mL for the 2L CELESTIAL population with AFP ≥ 400 ng/mL versus 3394 (1177, 16,812) ng/mL for REACH-2

[e] Median baseline serum AFP level was available from the REACH-2 publication but, owing to the way that weights are calculated (using centering moments based on mean rather than median values) and because baseline AFP values for the eligible CELESTIAL subpopulation did not follow a Normal distribution, median log10(AFP) values for the REACH-2 population and for the 2L CELESTIAL subgroup with AFP ≥ 400 ng/mL could not be exactly matched

matching-adjusted CELESTIAL placebo population and the REACH-2 placebo population (Table S5; Fig. S4).

Safety Outcomes

TRAEs of interest (reported in both trials and occurring in at least 5% of patients in any trial arm) were: increased aspartate aminotransferase (AST; any grade and grade 3 or 4); diarrhea (any grade); fatigue (any grade and grade 3 or 4); decreased appetite (any grade); vomiting (any grade); hypertension (any grade and grade 3 or 4); nausea (any grade) and proteinuria (any grade).

Anchored log OR TRAE estimates found significantly lower rates of any grade diarrhea and

Table 3 Efficacy estimates for the matching-adjusted cabozantinib and the ramucirumab populations

	Weighted KM survival estimate, months, median (95% CI)	*p* value[a]	Parametric modeling survival estimate, months, [b] median (95% CI)
Primary analysis			
OS			
Cabozantinib (ESS = 63)	10.6 (9.5, 17.3)	0.104	12.0 (9.6, 14.5)
Ramucirumab (*N* = 197)	8.7 (7.3, 10.8)		9.6 (8.4, 10.8)
PFS			
Cabozantinib (ESS = 63)	5.5 (4.6, 7.4)	0.016	5.2 (4.1, 6.4)
Ramucirumab (*N* = 197)	2.8 (2.7, 4.1)		3.2 (2.8, 3.6)
Sensitivity analysis			
OS			
Cabozantinib (ESS = 73)	10.6 (9.5, 17.3)	0.133	11.7 (9.2, 14.2)
Ramucirumab (*N* = 197)	8.7 (7.3, 10.8)		9.6 (8.4, 10.8)
PFS			
Cabozantinib (ESS = 73)	5.5 (4.6, 7.4)	0.023	5.2 (4.1, 6.4)
Ramucirumab (*N* = 197)	2.8 (2.7, 4.1)		3.2 (2.8, 3.6)
Validation analysis			
OS			
Cabozantinib (ESS = 80)	9.7 (8.1, 12.4)	0.409	10.6 (8.6, 12.4)
Ramucirumab (*N* = 197)	8.7 (7.3, 10.7)		9.6 (8.4, 10.8)
PFS			
Cabozantinib (ESS = 80)	5.4 (4.0, 6. 9)	0.031	4.9 (4.0, 5.8)
Ramucirumab (*N* = 197)	2.8 (2.7, 4.1)		3.2 (2.8, 3.6)

CI confidence interval, *ESS* effective sample size, *KM* Kaplan-Meier, *OS* overall survival, *PFS* progression-free survival

[a] Log-rank test

[b] Best-fit models: OS, Weibull (for primary, sensitivity and validation analyses); PFS, log-logistic (for primary, sensitivity and validation analyses)

hypertension and grade 3 or 4 hypertension in the ramucirumab population than in the matching-adjusted cabozantinib population. Unanchored log OR estimates also found significantly lower rates of grade 3 or 4 fatigue and increased AST with ramucirumab than with cabozantinib. Frequency of any grade proteinuria (unanchored) was lower in the matching-adjusted cabozantinib population than in the ramucirumab population (Table 4; Fig. S5).

Rates of treatment discontinuation due to TRAEs were not significantly different for the matching-adjusted cabozantinib population and the ramucirumab population (Table 4).

Sensitivity and Validation Analyses

Weighting the baseline IPD of patients in the 2L CELESTIAL population with serum AFP levels of 400 ng/mL or above based on seven empirically-

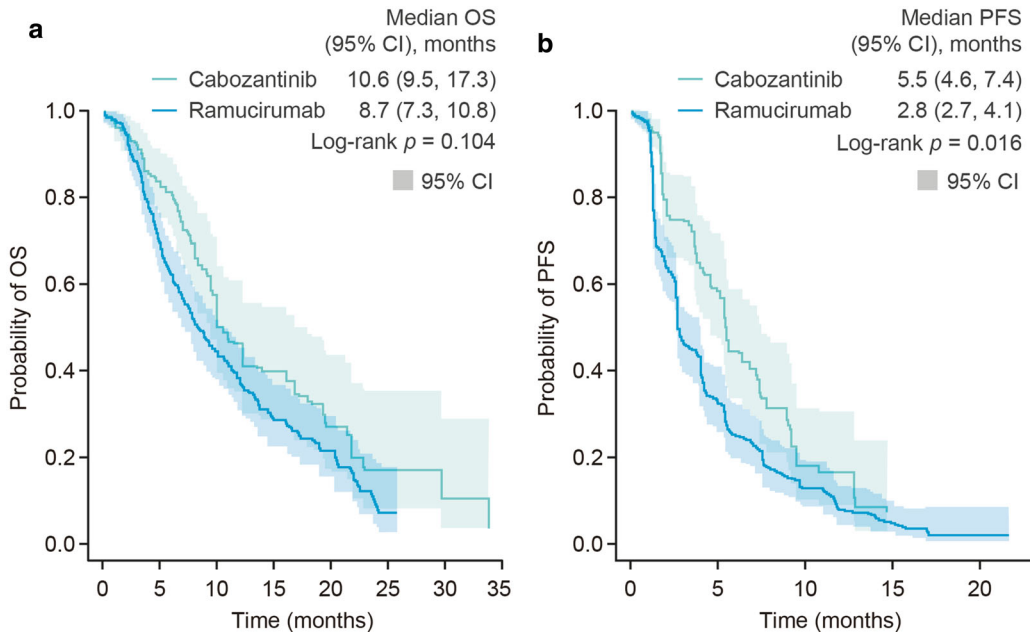


Fig. 1 Kaplan-Meier curves for **a** OS and **b** PFS of the matching-adjusted CELESTIAL population and the REACH-2 populations. **a** OS, **b** PFS. *2L* second line, *CI* confidence interval, *OS* overall survival, *PFS* progression-free survival

identified matching criteria was effective in aligning the potentially effect-modifying baseline characteristics with those of the REACH-2 population (Table S6). Following matching and adjustments for these seven characteristics, the ESSs from CELESTIAL were 73 patients for cabozantinib and 46 for placebo. As was used for the primary analysis, the REACH-2 population included 197 patients randomly allocated to ramucirumab and 95 randomly allocated to placebo (Table 1).

Median (95% CI) survival estimates for the sensitivity analysis mirrored those of the primary analysis, for both the weighted KM analysis (Table 3; Fig. S6) and (following confirmation of PH assumption violation [Tables S3, S4, and S7; Figs. S7 and S8]) for the parametric modeling analysis (Tables S7). Results of the safety sensitivity analysis were also consistent with those of the primary analysis (Table S8; Fig. S5).

The validation analysis, which repeated the primary analysis but with baseline AFP level excluded from the 11 matching criteria, further supported the findings of the primary analysis. Excluding the baseline AFP criterion from the weighting and adjustment steps resulted in slightly larger ESSs from CELESTIAL than were available for the primary analysis (80 patients allocated to cabozantinib and 49 to placebo, Table 1). At baseline, the 2L CELESTIAL population with serum AFP levels of 400 ng/mL and the REACH-2 population were closely matched (Table S9). Compared with the primary analysis, exclusion of the baseline AFP matching criterion had little impact on the comparative PFS and safety outcomes, but did reduce the magnitude of difference in median OS estimates for the matching-adjusted CELESTIAL population and the REACH-2 populations, although the difference was already non-significant (Tables 3, S3, S4, and S10; Fig. S9, S10 and S11).

DISCUSSION

Main Findings

In the absence of head-to-head trials directly comparing cabozantinib and ramucirumab, we conducted a MAIC using data from the pivotal

Table 4 Log OR (95% CI) and *p* values for TRAEs reported in at least 5% of patients in any arm of CELESTIAL or REACH-2 (cabozantinib vs. ramucirumab)

TRAE	Unmatched analysis		Matched-adjusted analysis	
	Log OR (95% CI)	*p* value	Log OR (95% CI)	*p* value
Any grade				
Increased AST	−0.50 (−2.32, 1.32)	0.6019	−0.58 (−2.59, 1.42)	0.5799
Diarrhea	1.61 (0.23, 2.99)	0.0220	1.53 (0.00, 3.05)	0.0499
Fatigue	−0.05 (−1.29, 1.18)	0.9377	0.44 (−0.89, 1.76)	0.5288
Decreased appetite	0.66 (−0.72, 2.05)	0.3519	1.10 (−0.46, 2.66)	0.1691
Vomiting	1.05 (−1.58, 3.69)	0.4430	−0.90 (−3.06, 1.27)	0.4247
Hypertension	1.92 (−0.33, 4.16)	0.0942	2.52 (0.23, 4.81)	0.0305
Nausea	−0.33 (−2.17, 1.52)	0.7413	−0.15 (−2.27, 1.97)	0.8968
Proteinuria[a]	−2.11 (−3.52, −0.70)	0.0034	−1.78 (−2.99, −0.56)	0.0043
Grade 3/4				
Increased AST[a]	2.28 (1.02, 3.55)	0.0004	1.79 (0.47, 3.11)	0.0078
Fatigue[a]	2.24 (0.70, 3.77)	0.0044	2.72 (1.23, 4.22)	0.0004
Hypertension	16.34 (14.73, 17.94)	< 0.0010	16.92 (15.20, 18.65)	< 0.0010
Leading to discontinuation				
Any TRAE	0.40 (−1.57, 2.36)	0.7509	1.16 (−0.89, 3.20)	0.2709

AST aspartate aminotransferase, *CI* confidence interval, *OR* odds ratio, *TRAE* treatment-related adverse event

[a] Unanchored analysis because no AEs occurred in at least one of the placebo arms of the trials

phase 3 CELESTIAL and REACH-2 trials to compare these two VEGF-targeting agents in HCC populations matched in terms of prior therapy and key baseline characteristics. The MAIC method is a recognized means of evaluating comparative outcomes from trials with similar endpoints, but heterogenous populations [19].

In the primary analysis, the baseline data of individual CELESTIAL patients were weighted so that the overall characteristics of the 2L CELESTIAL population with elevated AFP matched those of the REACH-2 population. The assigned weights were intended to minimize residual differences in 11 potentially effect-modifying baseline characteristics selected by a panel of clinical experts. Weighted KM analysis indicated that daily oral administration of cabozantinib was associated with significantly longer PFS than IV administration of ramucirumab every 2 weeks. OS was not significantly different for cabozantinib and ramucirumab. The parametric modeling analysis (undertaken in line with recommendations for MAIC analyses published by the National Institute for Health and Care Excellence's [NICE] Decision Support Unit [20]) mirrored the results of the weighted KM analysis: significantly prolonged PFS, but no significant difference in OS. A sensitivity analysis (using seven effect modifiers selected using stepwise AIC regression) and a validation analysis (examining the impact of removing baseline AFP from the primary analysis' matching criteria) also reinforced these findings, further strengthening confidence in the results.

In terms of TRAEs, the odds of experiencing any grade diarrhea (anchored) or hypertension (anchored), or grade 3 or 4 hypertension (anchored), fatigue (unanchored), or increased AST (unanchored) were significantly lower for ramucirumab than for cabozantinib; however, rates of treatment-related proteinuria (unanchored) were significantly lower for cabozantinib than for ramucirumab, and treatment discontinuation rates due to TRAEs were not significantly different between populations.

When considering the results of the analysis, it should be noted that it was only possible to generate comparative rate estimates for TRAEs that occurred in both trials. The most common adverse event (AE) reported for patients treated with cabozantinib in the CELESTIAL trial, for example, was palmar-plantar erythrodysesthesia (PPE; any grade/grade 3 or 4 in 46%/17% of patients), but this was not reported in REACH-2 and so was not evaluable in the current analysis. Similarly, in REACH-2, treatment infusion-related reactions and bleeding or hemorrhage events occurred in 7 and 10% of patients treated with ramucirumab, respectively, but neither were relevant/reported in CELESTIAL, so they did not feature in the present analysis. The difference in treatment-related TRAEs reported in the CELESTIAL and REACH-2 trials reflects the differing mechanisms of action of the two drugs, with cabozantinib being a TKI with activity against a range of receptor kinases, including VEGF receptors-1, -2 and -3, AXL and MET, and ramucirumab being a mAb that binds specifically to VEGF receptor-2 [2, 3, 6, 7]. Furthermore, when considering the respective efficacy-safety profiles of the two therapies, it is also relevant to note that the occurrence of some TRAEs may be positively correlated with treatment outcomes, as reported for the subgroup of cabozantinib patients from CELESTIAL who experienced any grade PPE or grade 3 or higher hypertension [27].

Strengths of the Approach

Although network meta-analyses (NMAs) or traditional ITCs are perhaps more familiar methods of generating comparative data than MAIC analyses, they are not able to produce meaningful comparisons in all scenarios. The reliability of such methods is compromised, for example, when a common trial comparator is not available or when there is distinct heterogeneity between comparator studies [28]. A review of nearly 181 technology appraisals conducted by NICE found that more than half (54%) of all related assessments did not include a mixed/indirect treatment comparison; trial design heterogeneity was cited as the most common reason for their absence [29]. MAIC analyses can overcome some of the challenges of between-trial population heterogeneity. While NMAs rely on published data, MAIC analyses require patient-level data to be available for at least one of the studies in order to permit adjustments for between-trial population differences and to minimize the potential for outcome bias [28]. Increasing recognition of the MAIC method is reflected in the growing number of related publications apparent within the peer-reviewed medical and cancer-related literature over the past 10 years (Fig. S12).

In the current analysis, weighting of the baseline IPD of the 2L CELESTIAL subgroup with elevated AFP was broadly effective in balancing the distribution of potentially effect-modifying differences between the two trial populations. Although a difference in baseline serum AFP levels persisted between the matching-adjusted 2L CELESTIAL population with serum AFP of 400 ng/mL or higher and the REACH-2 population despite the matching and weighting procedures, the validation analysis provides confidence that this residual difference had minimal impact on the evaluated outcomes. In addition, the similarity of the estimates for the matching-adjusted CELESTIAL placebo arm and REACH-2 placebo arm suggests that the weighting process was successful in reducing clinically relevant baseline differences between the populations and, accordingly, in reducing the potential for associated confounding in the active treatment arm analysis.

It is also relevant and appropriate that the present analysis used published data from the pivotal REACH-2 trial rather than from the earlier REACH trial [10, 30]. In REACH, patients with HCC were randomly allocated to 2L

ramucirumab or placebo after 1L sorafenib treatment; patients with a baseline serum AFP of 400 ng/mL or above constituted a pre-specified subpopulation (n = 250). Overall, the REACH trial was negative: ramucirumab did not significantly improve OS compared with placebo, but the subgroup analyses suggested a possible OS benefit with ramucirumab over placebo in patients with baseline serum AFP of 400 ng/mL or above (vs. < 400 ng/mL) [30]. This signal from the REACH trial was explicitly investigated and validated by the subsequent pivotal REACH-2 trial, which led to the regulatory approval of ramucirumab for patients with HCC and serum AFP of 400 ng/mL following prior sorafenib treatment [4, 5, 10].

Limitations of the Approach

Despite their strengths, MAIC analyses cannot offer the quality of evidence generated by a head-to-head RCT and cannot adjust for all potential differences in trial populations and designs. The tumor assessment schedules, for instance, were different for the CELESTIAL trial (every 8 weeks) and the REACH-2 trial (every 6 weeks during the first 6 months; every 9 weeks thereafter). The PFS estimates in the current analysis may, therefore, be subject to resulting bias, although the direction of any such bias is primarily discernable at the individual patient level and depends on the relative timing of tumor growth and scheduled assessment. For instance, any tumor growth that occurred at week 5 would have been detected at week 6 in REACH-2 but not until week 8 in CELESTIAL, favoring cabozantinib. Yet, tumor growth occurring just after the first assessment in REACH-2 (e.g., during week 7) would have been assessed at week 8 in CELESTIAL but not until week 12 in REACH-2, favoring ramucirumab. The overall impact of the difference in assessment schedules remains unclear, but may favor cabozantinib because of the lower frequency of tumor assessments conducted in CELESTIAL (versus REACH-2) in the initial 6 months of follow-up.

For OS, the median estimates may have been influenced by subsequent treatment use, something that is particularly relevant when comparing therapies within a rapidly evolving treatment landscape, like that of HCC. It is plausible that temporal changes in the availability of therapies may have favored the ramucirumab population over the cabozantinib population because of the later start date of the REACH-2 trial (vs. the CELESTIAL trial).

Although it is possible that residual differences persisted between the trial populations despite the matching and weighting procedures, the consistent results of the primary, sensitivity and validation analyses provides reassurance that factors predictive of treatment effect were generally well balanced. In terms of prognostic factors, in an anchored analysis, any variables that are prognostic of disease course do not inhibit interpretation of the results because randomization should ensure equal distribution of any such factors between the active and placebo arms of each trial; between-trial comparison of results relative to the common placebo arm are therefore unaffected. Unanchored analyses that use parametric modeling (as in the present analysis) are open to greater potential bias because they compare absolute rather than relative estimates. Nevertheless, the similarity of the placebo arm analysis in the present study provides reassurance that the matching and adjustment steps successfully minimized clinically relevant differences between the comparator populations.

It is also relevant to note that MAIC analyses can only compare similar outcomes reported in both trials, prohibiting the evaluation of some TRAEs of potential relevance or the comparison of quality-of-life outcomes, which were evaluated differently for the two trials and so was beyond the scope of the current analysis [31, 32]. Finally, selection of the 2L CELESTIAL subpopulation with serum AFP of 400 ng/mL or above, and the subsequent weighting and adjustment procedures, unavoidably reduced the number of patients eligible for inclusion in the analysis. These steps were necessary to compare 'like with like,' but reduced the size of the original CELESTIAL trial population (powered to demonstrate treatment effect versus placebo) from 707 individuals to an ESS of 199

individuals (73 in the cabozantinib arm), limiting the overall statistical power of the analysis.

Interpretation

This MAIC analysis provides further insights into the therapeutic options available for patients with HCC who have progressed despite prior sorafenib therapy, particularly those with elevated serum AFP and potentially poor prognosis [12].

The results reinforce and build on those of a recent ITC that reported clinical equivalence of cabozantinib, ramucirumab and regorafenib with respect to OS in patients with serum AFP levels of 400 ng/mL or above [33] and of an NMA of available phase 3 trials of 2L agents for advanced HCC [34]. The NMA subgroup analysis of patients with AFP > 400 ng/mL suggested that cabozantinib may significantly prolong PFS compared with ramucirumab (HR [95% CI], 0.59 [0.40, 0.88]), but it found no significant differences in OS estimates for cabozantinib, ramucirumab or regorafenib for patients with AFP > 400 ng/mL. AEs were not explicitly reported in the NMA for AFP subgroups, but any grade diarrhea, fatigue, nausea and decreased appetite were reported in at least 10% of patients receiving cabozantinib or ramucirumab in the included trials. Diarrhea was the most common AE reported for cabozantinib (54% of patients compared with 18% for ramucirumab), and peripheral edema was the most common for ramucirumab (36% of patients compared with 0% for cabozantinib). Overall, the authors considered the AEs in patients treated with ramucirumab to be relatively mild, with 36% of patients experiencing grade 3 or 4 AE rates compared with 68% of those treated with cabozantinib [34].

A recent MAIC analysis of 2L cabozantinib and regorafenib in patients with HCC who had received sorafenib therapy has provided some indirect comparative data for the two VEGF-targeting TKIs [16]. Using data from the phase 3 CELESTIAL and RESORCE trials, the authors found no significant difference in OS estimates (median [95% CI]) for cabozantinib and regorafenib (11.4 [8.9–17.0] vs. 10.6 [9.1–12.1] months; $p = 0.3474$), but significantly prolonged PFS (median [95% CI]) with cabozantinib compared with regorafenib (5.6 [4.9–7.3] months vs. 3.1 [2.8–4.2] months; $p = 0.0005$) [16]. Perhaps unsurprisingly, given its focus on patients with elevated serum AFP levels (known to be associated with poor prognosis [12]), median survival estimates were numerically shorter for the current analysis than for the MAIC of cabozantinib and regorafenib, but absolute differences were similar: a median PFS benefit of 2.7 months with cabozantinib (vs. ramucirumab) in the present analysis compared with a 2.5-month benefit for cabozantinib versus regorafenib [16]. In terms of AEs, the MAIC of the two TKIs reported a trend for lower rates of treatment-emergent grade 3 or 4 hypertension with regorafenib than with cabozantinib, but only differences in rates of grade 3 or 4 diarrhea reached statistical significance. The authors noted that the observed differences in treatment-emergent AEs may, at least in part, reflect that the MAIC methods were not able to adjust for the fact that sorafenib-intolerant patients were excluded from the RESORCE trial, but not from CELESTIAL [16].

The present analysis builds on these data, addressing the evidence gap with respect to the scant availability of comparative data for cabozantinib and ramucirumab and utilizing the MAIC analysis method to minimize potential sources of confounding by matching and weighting the baseline data for individual CELESTIAL patients to those of the REACH-2 population. It offers insights into the efficacy and tolerability of cabozantinib in patients with serum AFP levels of 400 ng/mL or above and focuses the analysis on the pure 2L CELESTIAL population rather than on the mixed 2L and 3L population used in previous indirect comparisons.

The results suggest that cabozantinib may offer an efficacious and tolerable alternative to 2L ramucirumab in patients with serum AFP levels of 400 ng/mL or above. In addition to informing efficacy-safety considerations when treating patients with elevated serum AFP, the findings may also provide insights for clinicians seeking to optimize treatment options within the context of local healthcare resourcing, given

the potential implications of administering daily oral therapy (cabozantinib) versus fortnightly IV therapy (ramucirumab). Similarly, the findings may be of interest to clinicians seeking to individualize care decisions for patients with clear attitudes and preferences for particular treatment regimens and methods of administration (e.g., high frequency [daily] oral versus lower-frequency [fortnightly] IV therapy).

CONCLUSION

In this MAIC analysis of 2L cabozantinib and ramucirumab after prior sorafenib therapy in patients with HCC and elevated serum AFP, the alignment and weighting processes effectively balanced the distribution of effect-modifying baseline characteristics between the matching-adjusted CELESTIAL population and REACH-2 population. The similarity of the placebo arm survival analysis reinforces the assertion that clinically relevant baseline characteristics were largely balanced for the comparator populations. OS estimates were not significantly different for the cabozantinib and ramucirumab treatment arms, but PFS was significantly longer in the cabozantinib group (by an additional 2.7 months), almost double that of the ramucirumab group. These findings were consistent for the primary analysis and the sensitivity analysis. Rates of some grade 3 or 4 TRAEs were lower with ramucirumab than with cabozantinib, likely reflecting the different mechanisms of action of the two drugs, but there was no significant difference in discontinuation rates resulting from TRAEs. These MAIC results do not replace those of a head-to-head RCT, but, in the absence of RCT evidence, they contribute indirect comparative efficacy-tolerability data to inform clinical decision-making around optimum 2L treatment for patients with elevated AFP who have received prior sorafenib.

ACKNOWLEDGEMENTS

Funding. Open Access funding enabled and organized by Projekt DEAL. This study and the journal's Rapid Service Fee was funded by Ipsen. The CELESTIAL study was sponsored by Exelixis. Ipsen was involved in the design of the study and the analysis and interpretation of the data as well as the review of the manuscript.

Medical Writing, Editorial, and Other Assistance. The authors thank Alison Chisholm, MPH, of Oxford PharmaGenesis, Oxford, UK, who provided medical writing and editorial support, which was sponsored by Ipsen in accordance with Good Publication Practice guidelines. The authors thank all patients involved in these studies, as well as their caregivers, care team, investigators, and research staff in participating institutions.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

Prior Presentation. The preliminary results of the study were presented at the American Association for the Study of Liver Diseases (AASLD) Liver Meeting, Digital Experience, November 13–16, 2020.

Disclosures. Jörg Trojan has received research grants from Ipsen and Roche, and speaker and consulting honoraria from Amgen, AstraZeneca, Bayer Healthcare, Bristol Myers Squibb, Eisai, Eli Lilly, ImClone Systems, Ipsen, Merck Serono, Merck Sharp & Dohme and Roche. Patrick Mollon, Florence Marteau and Lidia Martín are employees of Ipsen. Bruno Daniele has received personal fees and nonfinancial support from Bayer, Ipsen and Sanofi, and personal fees from AstraZeneca, Eli Lilly, Eisai, Incyte, Merck Sharp & Dohme and Roche. Yuxin Li and Qing Xu are employees of IQVIA, which was contracted by Ipsen to conduct the analysis. Fabio Piscaglia has received personal fees for speakers' bureaus and advisory boards and for being a consultant from Alkermes, AstraZeneca, Bayer, Bracco, Bristol Myers Squibb, Eisai, GE Healthcare, Ipsen, La Force Guerbet, Roche and Tiziana Life Sciences.

Renata Zaucha has received speaker honoraria from Amgen, Bristol Myers Squibb and Roche; has been an advisory board member for Bristol Myers Squibb, Celgene and Pierre Fabre; and has received travel support from Amgen, Bristol Myers Squibb and Roche. Debashis Sarker has received speakers' bureau honoraria from AstraZeneca, Bayer, Eisai and Merck Sharp & Dohme; has been an advisory board member for Eisai, Ipsen, Novartis and Surface Oncology; and has received travel Grants from Eisai, Ipsen and MiNA Therapeutics. Ho Yeong Lim has received speaker honoraria from Bayer, and has been an advisory board member (or had a consulting role) for Bayer, Eisai, Ipsen, Bristol Myers Squibb, Merck Serono and Roche. Marino Venerito has received honoraria from Bayer Vital, Eli Lilly, Merck Serono, Nordic Pharma and Sirtex, and has been an advisory board member for Amgen, Bristol Myers Squibb, Eisai, Eli Lilly, Ipsen, Merck Sharp & Dohme and Nordic Pharma.

Compliance with Ethics Guidelines. The results presented in this manuscript are based on published studies. All procedures performed in those studies involving human participants were conducted in accordance with the ethical standards of the local Institutional Review Boards for each site and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the CELESTIAL and REACH-2 trials.

Data Availability. Where patient data can be anonymized, Ipsen will share all individual participant data that underlie the results reported in this article with qualified researchers who provide a valid research question. Study documents, such as the study protocol and clinical study report, are not always available. Proposals should be submitted to DataSharing@Ipsen.com and will be assessed by a scientific review board. Data are available beginning 6 months and ending 5 years after publication; after this time, only raw data may be available.

Open Access. This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES

1. Rimassa L, Wörns MA. Navigating the new landscape of second-line treatment in advanced hepatocellular carcinoma. Liver Int. 2020;40:1800–11. https://doi.org/10.1111/liv.14533.

2. Exelixis I. Prescribing information: CABOMETYX® (cabozantinib) tablets, for oral use. US Food and Drug Administration, 2020. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2019/208692s003lbl.pdf. Accessed 2 Dec 2020.

3. Ipsen Pharma. Summary of product characteristics, CABOMETYX. European Medicines Agency, 2020. https://www.ema.europa.eu/en/documents/product-information/cabometyx-epar-product-information_en.pdf. Accessed 2 Dec 2020.

4. Bayer AG. Summary of product characteristics, Stivarga 40 mg film-coated tablets. European Medicines Agency, 2018. Available from: https://www.ema.europa.eu/en/documents/product-information/stivarga-epar-product-information_en.pdf. Accessed 2 Dec 2020.

5. Bayer HealthCare Pharmaceuticals Inc. Prescribing Information: STIVARGA® (regorafenib) tablets, for oral use. US Food and Drug Administration, 2017. Available from: https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/203085s007lbl.pdf. Accessed 2 Dec 2020.

6. Eli Lilly and Company. Prescribing information, CYRAMZA (ramucirumab) injection, for intravenous use. US Food and Drug Adminstration, 2019. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2019/125477s029lbl.pdf. Accessed 2 Dec 2020.

7. Eli Lilly Nederland B.V. Summary of product characteristics, Cyramza 10 mg/mL concentrate for solution for infusion. European Medicines Agency, 2019. Available from: https://www.ema.europa.eu/en/documents/product-information/cyramza-epar-product-information_en.pdf. Accessed 2 Dec 2020.

8. Marino D, Zichi C, Audisio M, Sperti E, Di Maio M. Second-line treatment options in hepatocellular carcinoma. Drugs Context. 2019;8:212577. https://doi.org/10.7573/dic.212577.

9. Abou-Alfa GK, Meyer T, Cheng AL, et al. Cabozantinib in patients with advanced and progressing hepatocellular carcinoma. N Engl J Med. 2018;379(1):54–63. https://doi.org/10.1056/NEJMoa1717002.

10. Zhu AX, Kang YK, Yen CJ, et al. Ramucirumab after sorafenib in patients with advanced hepatocellular carcinoma and increased α-fetoprotein concentrations (REACH-2): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol. 2019;20(2):282–96. https://doi.org/10.1016/s1470-2045(18)30937-9.

11. European Association for the Study of the Liver. EASL clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol. 2018;69(1):182–236. https://doi.org/10.1016/j.jhep.2018.03.019.

12. Bai D-S, Zhang C, Chen P, Jin S-J, Jiang G-Q. The prognostic correlation of AFP level at diagnosis with pathological grade, progression, and survival of patients with hepatocellular carcinoma. Sci Rep. 2017;7(1):12870. https://doi.org/10.1038/s41598-017-12834-1.

13. Benson AB, D'Angelica MI, Abbott DE, et al. National Comprehensive Cancer Network (NCCN) Clinical practice guidelines in oncology (NCCN Guidelines®) hepatobiliary cancers. Version 5.2020. August 4, 2020. Available at: http://www.nccn.org. Accessed 2 Dec 2020.

14. Benson III A DAM, Abbott D, Abrams TA, Alberts SR, Anaya DA,. Hepatobiliary cancers ver. 1. 2018. In: NCCN clinical practice guidelines in oncology (NCCN Guidelines®) Fort Washington: National Comprehensive Cancer Network; 2018.

15. Vogel A, Cervantes A, Chau I, et al. Hepatocellular carcinoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2018;29(Suppl 4):iv238–55. https://doi.org/10.1093/annonc/mdy308.

16. Kelley RK, Mollon P, Blanc JF, et al. Comparative efficacy of cabozantinib and regorafenib for advanced hepatocellular carcinoma. Adv Ther. 2020;37(6):2678–95. https://doi.org/10.1007/s12325-020-01378-y.

17. Signorovitch JE, Sikirica V, Erder MH, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. Value Health. 2012;15(6):940–7. https://doi.org/10.1016/j.jval.2012.05.004.

18. Signorovitch JE, Wu EQ, Yu AP, et al. Comparative effectiveness without head-to-head trials: a method for matching-adjusted indirect comparisons applied to psoriasis treatment with adalimumab or etanercept. Pharmacoeconomics. 2010;28(10):935–45. https://doi.org/10.2165/11538370-000000000-00000.

19. Thom H, Jugl SM, Palaka E, Jawa S. Matching adjusted indirect comparisons to assess comparative effectiveness of therapies: usage in scientific literature and health technology appraisals. Value Health. 2016;19(3):A100–1.

20. Latimer N. NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data: 2011. Available at: http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf. Accessed 2 Dec 2020.

21. Kelley RK, Meyer T, Rimassa L, et al. Serum alpha-fetoprotein levels and clinical outcomes in the phase III CELESTIAL study of cabozantinib versus placebo in patients with advanced hepatocellular carcinoma. Clin Cancer Res. 2020;26(18):4795–804. https://doi.org/10.1158/1078-0432.Ccr-19-3884.

22. Phillippo D AT, Dias S, Palmer S, Abrams KR, Welton N. NICE DSU Technical Support Document 18: methods for population-adjusted indirect comparisons in submissions to NICE. 2016. Available at: https://research-information.bris.ac.uk/files/94868463/Population_adjustment_TSD_FINAL.pdf. Accessed 2 Dec 2020.

23. Mohammed A, Naugler C, Far BH. Emerging business intelligence framework for a clinical laboratory through big data analytics. In: Emerging trends in computational biology, bioinformatics, and systems biology: algorithms and software tools. New York: Elsevier/Morgan Kaufmann; 2015. p. 577–602.

24. Woodruff RS. Confidence intervals for medians and other position measures. J Am Stat Assoc. 1952;47(260):635–46.

25. 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.

26. Sestelo M. A short course on survival analysis applied to the Financial Industry. BBVA Data and Analytics, Madrid. 27–28/11/2017, v1.2. https://bookdown.org/sestelo/sa_financial/. Accessed 1 Dec 2020.

27. Abou-Alfa GK, Meyer T, Cheng A-L, et al. Association of adverse events (AEs) with efficacy outcomes for cabozantinib (C) in patients (pts) with advanced hepatocellular carcinoma (aHCC) in the phase III CELESTIAL trial [abstract 4088]. J Clin Oncol. 2019;37(Suppl 15):4088. https://doi.org/10.1200/JCO.2019.37.15_suppl.4088.

28. Ishak K, Proskorovsky I, Benedict A, Chen C. Novel methods for indirect comparison of treatments: when are they needed and how do they work? Value Outcomes Spotlight. 2016;2(1):20–3.

29. Martin A, Rizzo M, Iheanacho I. Faulty connections: can criticisms of network meta-analysis in NICE submissions be avoided? [abstract PRM188]. Value Health. 2013;16(7):A323–636.

30. Zhu AX, Park JO, Ryoo BY, 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.

31. Abou-Alfa GK, Mollon P, Meyer T, et al. Quality-adjusted life years assessment using cabozantinib for patients with advanced hepatocellular carcinoma (aHCC) in the CELESTIAL trial. J Clin Oncol. 2019;37(4_suppl):207–307. https://doi.org/10.1200/JCO.2019.37.4_suppl.207.

32. Zhu AX, Nipp RD, Finn RS, et al. Ramucirumab in the second-line for patients with hepatocellular carcinoma and elevated alpha-fetoprotein: patient-reported outcomes across two randomised clinical trials. ESMO Open. 2020;5(4):e000797. https://doi.org/10.1136/esmoopen-2020-000797.

33. Camean-Castillo M, Fenix-Caballero S, Salmeron-Navas F, et al. 2SPD-006 Assessment of regorafenib, ramucirumab and cabozantinib as secondline therapy in hepatocarcinoma and alfa-fetoprotein value ≥400 ng/mL [abstract]. Eur J Hosp Pharm. 2020;27(Suppl 1):A12. https://doi.org/10.1136/ejhpharm-2020-eahpconf.25.

34. Wang D, Yang X, Lin J, et al. Comparing the efficacy and safety of second-line therapies for advanced hepatocellular carcinoma: a network meta-analysis of phase III trials. Therap Adv Gastroenterol. 2020;13:1756284820932483. https://doi.org/10.1177/1756284820932483.