Recurrence-free survival as a surrogate endpoint for overall survival in adults with early-stage hepatocellular carcinoma: A correlation meta-analysis of randomized controlled trials

CO157

Leung L,[1] Singh P,[2] Ademisoye E,[3] Hofer K,[1] Qian A[1], Kim I[2]

[1]Evidinno Outcomes Research Inc., Vancouver, BC, Canada; [2]Bristol Myers Squibb, Princeton, NJ, USA; [3]Bristol Myers Squibb, Uxbridge, UK

Introduction

- Liver cancer is currently the sixth leading cancer globally and the third leading fatal cancer.[1,2] The most common form of liver cancer is hepatocellular carcinoma (HCC), representing approximately 90% of cases[3]
- Overall survival (OS) is generally the standard endpoint for oncology trials. However, observing a benefit on OS may require considerable follow-up time. Thus, research on earlier endpoints such as recurrence-free survival (RFS) as surrogates of OS could potentially benefit patients by accelerating the approval and access to novel therapies
- Past work on surrogacy in HCC includes Huan et al 2017,[4] who investigated the surrogacy relationship between disease-free survival (DFS) and OS
  - The authors concluded that DFS and OS were strongly correlated at both individual and trial levels without specifying what qualified as "strong"
- Currently, there are a number of trials underway testing drug treatments in the adjuvant setting for early-stage HCC following ablation or resection, including CheckMate 9DX (NCT03383458)[5]

Objective

- To evaluate the appropriateness of RFS as a surrogate for OS in adults with early-stage HCC who underwent curative ablation or resection

Methods

Systematic literature review

- A systematic review was conducted using standard methodologies[6]
- MEDLINE®, Embase, and Cochrane Central Register of Controlled Trials (CENTRAL) were searched from database inception to July 20, 2022. Searches were limited to the English language
- Included articles were randomized controlled trials (RCTs) on adults with early-stage HCC who underwent curative ablation or resection
- Outcomes of interest were OS and RFS (or its analogs such as DFS or progression-free survival). To be included in the evidence base, RCTs must have reported relative treatment effects for both OS and RFS either in the form of hazard ratios (HR_{OS} and HR_{RFS}) or Kaplan-Meier curves

Data analysis

Trial-level surrogacy models and analysis sets

- The surrogacy of RFS for OS at the trial level was assessed using two meta-analysis models. HRs were log-transformed to be consistent with the linearity assumption for the relationship between the treatment effects
- The first model was based on an alternative bivariate random-effects meta-analysis (BRMA) model proposed by Riley et al 2008,[7] which provides an overall correlation measure between the log-transformed HRs of RFS and OS (log HR_{RFS} and log HR_{OS})
- The second model was a weighted linear regression (WLR) model weighted by the sample size of comparison. The association between log HR_{RFS} and log HR_{OS} was measured by the Pearson correlation coefficient
- Six sensitivity analyses were conducted by:
  1. Omitting trials that failed the proportionality test
  2. Restricting to trials that explicitly included death in the definition of RFS
  3. Omitting trials that permitted treatment crossover
  4. Restricting to trials with 50% or more patients undergoing curative resection
  5. Restricting to trials in the adjuvant setting
  6. Restricting to trials in the adjuvant setting that explicitly included death in the definition of RFS

Assessing the surrogacy equation and the correlation estimates

- The validity of the model was assessed by using a leave-one-out cross-validation (LOOCV) approach based on WLR models. Based on the WLR models, prediction accuracies on the statistical significance of HR_{OS} as a binary outcome were also reported for each analysis set
- The National Institute for Health and Care Excellence (NICE) Decision Support Unit Technical Support Document 20 was used as a guide to assess model validity[8]

Results

Study selection

- Of 13,137 records identified, 48 publications pertaining to 47 unique RCTs were included in the literature review and subsequent correlation meta-analysis (**Figure 1**)
- Sample size of the included trials ranged from 23 to 1114 patients. Most (43 out of 47) trials were conducted in East Asia (China, Japan, South Korea, and Taiwan). More details are presented in **Table 1**

Figure 1. PRISMA flow diagram

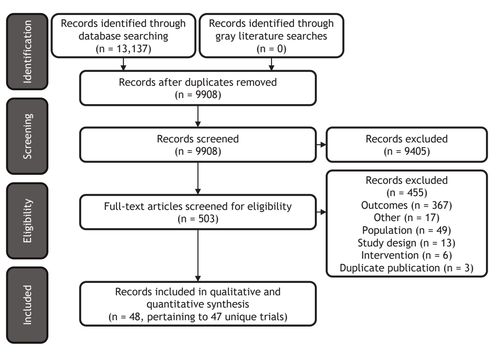


Table 1. Trial characteristics

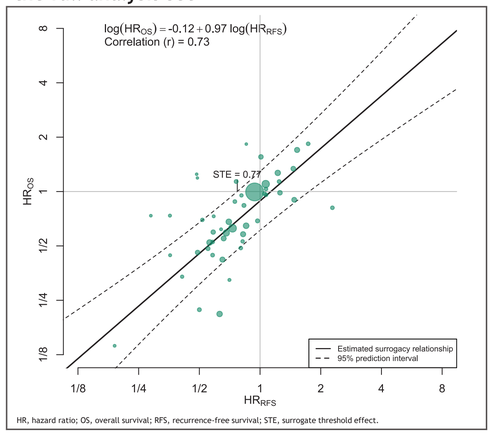
Trial	Sample size, N	Region	Included in which sensitivity analyses
Akamatsu et al 2004	42	Japan	1, 3
Chen et al 2006	180	China	1, 3
Chung et al 2013	103	Singapore	All
Dumortier et al 2014	75	France	1, 3, 4, 5
Fang et al 2014	120	China	1, 3
Feng et al 2012	168	China	1, 3
Feng et al 2017	105	China	1, 3
Ge et al 2019	23	China	1, 3, 5
Hao et al 2016	218	China	1, 3
Hasegawa et al 2006	160	Japan	1, 3, 4, 5
Hirokawa et al 2020	114	Japan	1, 3, 4, 5
Huang et al 2015	200	China	4, 5
Hui et al 2009	127	China	3, 4, 5
Ishizuka et al 2016	117	Japan	2, 3, 4, 5, 6
Izumi et al 1994	50	Japan	1, 3, 5
Kaibori et al 2012	124	Japan	1, 3
Kawata et al 1995	24	Japan	1, 3, 4, 5
Kitahara et al 2020	30	Japan	1, 3, 5
Kuang et al 2004	41	Japan	1, 4, 5
Lee et al 2015	230	South Korea	1, 3, 5
Li et al 2020 A	156	China	2, 3, 4, 5, 6
Li et al 2020 B	127	China	All
Liao et al 2017	96	China	1, 3
Liao et al 2022	385	China	3
Lin et al 2004	157	Taiwan	1, 2, 3
Lin et al 2005	187	Taiwan	1, 3
Liu et al 2016	200	China	1, 2, 3, 5, 6
Lo et al 2007	86	China	All
Luo et al 2022	223	China	1
McRFA 2020	96	China	1, 2, 3
Peng et al 2013	189	China	1, 2, 3
Shibata et al 2006	74	Japan	1, 2, 3
Shibata et al 2009	89	Japan	1, 2, 3
Shiina et al 2005	232	Japan	1, 3
STORM 2015	1114	Multinational	2, 3, 4, 5, 6
Sun et al 2006	236	China	2, 3, 4, 5, 6
Sun et al 2019	52	China	3, 4, 5
Takayama et al 2000	155	Japan	1, 2, 4, 5, 6
Wang et al 2015	360	China	1, 3
Wang et al 2018	280	China	1, 2, 3, 5, 6
Wei et al 2018	250	China	2, 3, 4, 5, 6
Xu et al 2016	200	China	2, 3, 5, 6
Yamamoto et al 1996	76	Japan	1 (stage I subpopulation only), 4[a], 5[a]
Yamasaki et al 1996	97	Japan	1, 3
Yi et al 2014	94	China	1, 3
Zaitoun et al 2021	188	Saudi Arabia	3
Zhang et al 2018	94	China	3

[a]Yamamoto et al 1996 had two separate subpopulations (stage I and stage II) and "*" indicates that both subpopulations were included in the analysis.

Full analysis set

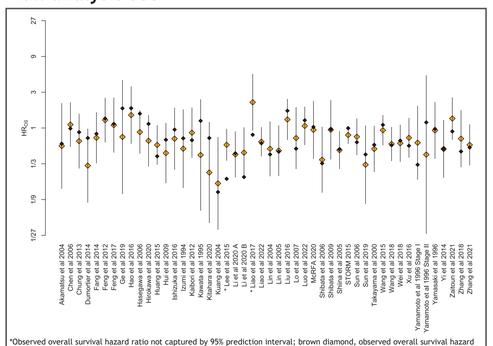
- The surrogacy equation was $\log(HR_{OS}) = -0.12 + 0.97 \times \log(HR_{RFS})$
- Using LOOCV, 46 out of 48 (95.8%) comparisons lay inside the 95% prediction interval, indicating the substantial validity of the surrogacy equation. LOOCV was also done for sensitivity analyses which ranged from 91.7% (11/12) to 96.2% (25/26)
- The WLR model and LOOCV on the full analysis set are presented in **Figure 2** and **Figure 3**, respectively. A summary of the results of all the analyses is presented in **Table 1**
- A surrogate threshold effect (STE) was calculated using WLR, defined as the minimum value of HR_{RFS} to predict a statistically significant and positive treatment effect on HR_{OS}

Figure 2. Weighted linear regression model for the full analysis set



HR, hazard ratio; OS, overall survival; RFS, recurrence-free survival; STE, surrogate threshold effect.

Figure 3. Leave-one-out cross-validation on the full analysis set



*Observed overall survival hazard ratio not captured by 95% prediction interval; brown diamond, observed overall survival hazard ratio; yellow diamond, predicted overall survival hazard ratio. The 95% prediction intervals are presented as whiskers.
OS, overall survival.

Table 2. Summary of meta-analysis results

Analysis	N	R_{BRMA} (95% CI)	R_{WLR} (95% CI)	LOOCV (% validated)	LOOCV on the accuracy on HR_{OS} significance[a]
Primary analysis					
Full analysis set	47	0.67 (0.52-0.79)	0.73 (0.61-0.85)	46/48 (95.8%)	33/48 (68.8%)
Sensitivity analyses					
1) Omitting trials that failed the proportionality test	35	0.68 (0.50-0.81)	0.75 (0.60-0.88)	33/35 (94.3%)	26/35 (74.3%)
2) Trials that explicitly included death in the definition of RFS	17	0.91 (0.79-0.96)	0.90 (0.74-0.96)	16/17 (94.1%)	13/17 (76.5%)
3) Omitting trials that permitted crossover	42	0.67 (0.49-0.79)	0.70 (0.57-0.83)	40/42 (95.2%)	30/42 (71.4%)
4) Trials with 50% or more patients undergoing curative resection	18	0.64 (0.35-0.82)	0.74 (0.36-0.87)	18/19 (94.7%)	12/19 (63.2%)
5) Trials in the adjuvant setting	25	0.61 (0.37-0.77)	0.69 (0.36-0.84)	25/26 (96.2%)	17/26 (65.4%)
6)Trials in the adjuvant setting that explicitly included death in the definition of RFS	12	0.88 (0.70-0.96)	0.88 (0.58-0.96)	11/12 (91.7%)	7/12 (58.3%)

[a]Accuracy is defined as the proportion of the HR_{OS} significance correctly predicted by the model, out of all predicted HR_{OS}.
BRMA, bivariate random-effects meta-analysis; CI, confidence interval; HR, hazard ratio; LOOCV, leave-one-out cross-validation; OS, overall survival; R_{BRMA}, Pearson correlation estimate based on BRMA; RFS, radiographic progression-free survival; R_{WLR}, Pearson correlation estimate based on WLR; WLR, weighted linear regression.

Conclusions

- Meaningful and consistent correlations between the treatment effects on RFS and OS in early-stage HCC were observed in the primary analysis (BRMA and WLR) and sensitivity analyses
- The highly accurate surrogacy equation between the treatment effects may enable earlier assessments of OS benefit from the RFS benefit for early-stage HCC, which is further supported by a high STE value
- Restricting the evidence base to trials explicitly reporting death in the definition of RFS led to stronger correlation estimates

References

1. Akinyemiju T, et al. *JAMA Oncol*. 2017;3(12):1683-1691.
2. Villanueva A, et al. *N Engl J Med*. 2019;380(15):1450-1462.
3. Llovet JM, et al. *Nat Rev Dis Primers*. 2021;7(1):6.
4. Huan HB, et al. *Oncotarget*. 2017;8(52):90291-90300.
5. U.S. National Library of Medicine. ClinicalTrials.gov website. https://clinicaltrials.gov/ct2/show/NCT03383458. Published December 26, 2017. Updated January 23, 2023. Accessed April 11, 2023.
6. Cochrane. Cochrane Training website. https://training.cochrane.org/handbook. Published September 22, 2008. Updated August 4, 2022. Accessed April 11, 2023.
7. Riley RD, et al. *Biostatistics*. 2008;9(1):172-86.
8. NICE Decision Support Unit. Full list of technical support documents (TSDs) website; http://nicedsu.org.uk/wp-content/uploads/2020/10/TSD-20-mvmeta-final.pdf. Published October 22, 2019. Updated October 22, 2019. Accessed April 11, 2023.

Acknowledgments

- This study was funded by Bristol Myers Squibb and conducted by Evidinno Outcomes Research Inc. L.L., K.H., and A.Q. report employment with Evidinno Outcomes Research Inc. P.S., E.A., and I.K. report employment with Bristol Myers Squibb. Authors report no other conflicts of interest.

Presented at ISPOR 2023; May 7-10, 2023; Boston, MA

Email: lleung@Evidinno.com