Assessing surrogacy using restricted mean survival time ratio for overall survival in liver cancer: a narrative review

Tiffany H. Leung[1], James C. Ho[1], Xiaofei Wang[2], Herbert Pang[3]

[1]Department of Medicine, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China; [2]Department of Biostatistics and Bioinformatics, School of Medicine, Duke University, Durham, NC, USA; [3]School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China

Contributions: (I) Conception and design: JC Ho, X Wang, H Pang; (II) Administrative support: H Pang, JC Ho; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: TH Leung; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Dr. Herbert Pang, PhD. School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, G/F, Patrick Manson Building, (North Wing), 7 Sassoon Road, Pokfulam, Hong Kong, China. Email: herbpang@hku.hk.

Background and Objective: The application of immunotherapy in cancers, including liver cancer, has been increasing. However, non-proportional hazard (NPH) is often observed in cancer immunotherapy trials. In presence of violation of proportional hazard (PH) assumption, restricted mean survival time (RMST) ratio was proposed as an alternative to hazard ratio (HR) for evaluating the treatment effects of such trials. To shorten the total study duration, an intermediate endpoint with shorter follow-up such as progression-free survival (PFS) is used as the primary endpoint. Our aim is to evaluate the applicability of RMST ratio in addition to the HR in assessing the level of PFS serving as a surrogacy of overall survival (OS).

Methods: Phase II or phase III hepatocellular carcinoma (HCC) immunotherapy studies that were published between January 2013 and August 2022 were identified via the search in PubMed. Weighted least-square regression (WLSR) was applied to analyze the trial level data with the sample size of study being set as the weight. The evaluation was conducted twice with RMST ratio and HR being applied in respective evaluation to examine the level of PFS as a surrogacy for OS.

Key Content and Findings: Based on the results of eight included trials, the R-square values of WLSR with either HR or RMST ratio being applied were 0.31 and 0.16 separately, indicating a moderate and low correlation between PFS and OS respectively.

Conclusions: In this study, our results demonstrated the potential of RMST ratio in addition to HR for evaluating the level of surrogacy in immunotherapy trials. Furthermore, including more large scale and homogeneous studies into the research may help better understand the level of surrogacy in liver cancer.

Keywords: Immunotherapy; liver cancer; progression-free survival (PFS); restricted mean survival time (RMST); surrogate endpoint

Submitted May 03, 2023. Accepted for publication Sep 05, 2023. Published online Sep 22, 2023.
doi: 10.21037/cco-23-48
View this article at: https://dx.doi.org/10.21037/cco-23-48

Introduction

Liver cancer, which ranked the 6th in terms of global cancer incidence, is one of the major types of cancer. With its incidence being estimated to exceed 1.4 million globally by 2040 (1), it is crucial that more advanced treatments are being developed. Among various subtypes of liver cancer, hepatocellular carcinoma (HCC) is the major type of primary liver cancer, and often occurs among patients with hepatitis B/C virus infection or alcohol abuse (2).

To assess the treatment effects of the designed treatment plan, overall survival (OS) and progression-free survival (PFS) often serve as the main outcomes of clinical trials. In

clinical trials, OS refers to the period from randomization to death. PFS is the period from randomization to the time point when either disease progression or death is observed. Meanwhile, censored cases refer to those either alive at the data cutoff date or lost to follow-up after their last date of follow-up. Currently, OS often serves as the gold standard endpoint for the assessments of new cancer therapies. However, using OS as the primary endpoint is frequently accompanied by some shortcomings, such as longer follow-up time and a larger required sample size to meet the study design power (3). Another issue in using OS as the endpoint arises when delayed treatment effect due to immunotherapies has been observed (4). Furthermore, given that patients may receive subsequent lines of treatment, the OS may be confounded. Therefore, PFS can better reflect the treatment effect without being biased. For example, PFS is well accepted in global registration of various estimated glomerular filtration rate-targeted (eGFR-targeted) lung cancer treatments (5,6). Despite that there is yet to be effective systematic treatment after immunotherapy in HCC, we believe that it is still important to investigate the potential of PFS serving as a surrogacy for OS because of the differences in survival benefits between immunotherapy and other treatments.

Hazard ratio (HR) is often applied in clinical trials to compare the occurrence of events between different arms in the trial. Although it has been widely employed, it is less suitable to be used for non-proportional hazard (NPH) studies, which are the ones where the hazards are expected to be inconsistent over time. As described earlier, immunotherapy-based trials often experience delayed treatment effect and therefore violate the proportional hazard (PH) assumption, making HR less appropriate to be applied to these trials. Under such circumstances, restricted mean survival time (RMST) ratio can be an alternative. While RMST refers to the area under the Kaplan-Meier curve up to the timepoint of interest, ratio refers to the comparison of the two areas of treatment arm and control arm. One of the greatest benefits of RMST ratio is that it is less affected by the issue of NPH. Application of RMST ratio can better reflect the benefits in long-term survival brought by immunotherapy than HR does (4).

The main goal of this study is to examine the use of RMST ratio in addition to the typical HR in evaluating PFS as a surrogate for OS. We present this article in accordance with the Narrative Review reporting checklist (available at https://cco.amegroups.com/article/view/10.21037/cco-23-48/rc).

Methods

We conducted a literature search using PubMed to identify phase II or phase III HCC studies with the publication date between January 2000 and August 2022 (Figure S1). The drugs given to the treatment arms should be either cytokine-induced killer (CIK) cell agent or checkpoint inhibitors. During the data collection process, studies that were excluded by Leung TH due to reasons such as missing essential information for data analysis will be double-checked by Pang H in a duplicate manner to ensure the reproducibility of our results. After identifying the eligible studies, required variables were extracted from the articles. These variables include the sample size of the study, disease stage of the patients at diagnosis, the histology of cancer, whether patients received pre-treatment or not, and the types of primary endpoints.

In addition to the mentioned variables that are relevant to the study characteristics, HRs and RMST ratios are also essential to our study. While HRs were all reported in the eligible studies, RMST ratios were calculated. The first step to calculate RMST ratio using the published articles was to transfer the provided Kaplan-Meier survival curves of both OS and PFS into time and respective survival probability using the software "WebPlotDigitizer 4.4" (7). The next step is to reconstruct the Kaplan-Meier curves. This was achieved using the method proposed by Guyot *et al.* [2012] based on the provided number at risk and the previously extracted data (8). With the reconstructed Kaplan-Meier curve, RMST ratios and the corresponding confidence intervals (CIs) were calculated using R 4.2.2 and the two R packages, "survRM2" and "survival".

To evaluate the association between PFS and OS, weighted least-square regression (WLSR) was conducted with the study sample size being used as the weight. The WLSR analysis was conducted twice, with one for the comparison between log HR of OS and PFS, and another one for the comparison between log RMST ratio of OS and PFS.

Results

A literature search was conducted in PubMed to identify eligible studies for analysis. The following search terms: ("Pembrolizumab" [Title/Abstract] OR "Ipilimumab" [Title/Abstract] OR "Atezolizumab" [Title/Abstract] OR "Nivolumab" [Title/Abstract] OR "Immunotherapy" [Title/Abstract] OR "Camrelizumab" [Title/Abstract])

Table 1 Summary of HR and RMST ratio for PFS and OS

Immunotherapy trial	Sample size	PFS		OS	
		HR (95% CI)	RMST ratio (95% CI)	HR (95% CI)	RMST ratio (95% CI)
CIK cell					
Lee *et al.*, [2015] (9)	226	0.63 (0.43–0.94)	0.81 (0.69–0.95)	0.21 (0.06–0.75)	0.94 (0.90–0.99)
Takayama *et al*., [2000] (10)	200	0.57 (0.37–0.87)	0.71 (0.55–0.91)	0.32 (0.19–0.56)	0.91 (0.80–1.02)
Xu *et al.*, [2016] (11)	150	0.83 (0.54–1.27)*	0.92 (0.78–1.10)	0.70 (0.40–1.23)**	0.90 (0.81–1.00)
Checkpoint inhibitor					
Finn *et al*., [2020a] (12)	501	0.59 (0.42–0.79)	0.91 (0.79–1.05)	0.58 (0.47–0.76)	0.85 (0.76–0.94)
Finn *et al*., [2020b] (13)	413	0.72 (0.57–0.90)	0.32 (0.26–0.40)	0.78 (0.61–1.00)	0.85 (0.73–0.99)
Kelley *et al*., [2022] (14)	649[†]	0.63 (0.44–0.91)*	0.78 (0.65–0.94)	0.90 (0.68–1.18)*	0.97 (0.89–1.06)
Qin *et al*., [2020][‡] (15)	217	0.87 (0.63–1.18)	1.07 (0.79–1.45)	1.17 (0.81–1.70)	0.94 (0.81–1.10)
Yau *et al*., [2022] (16)	743	0.93 (0.79–1.10)*	0.82 (0.68–0.99)	0.85 (0.72–1.00)	0.89 (0.80–1.00)

[†], the sample size for obtaining OS and PFS were different. The sample size reported in this table is the one for OS. [‡], the HRs (95% CI) of this study was recalculated based on the available information provided in the paper. *, PH violation with NPH test P value <0.05; **, strong PH violation with NPH test P value <0.01. HR, hazard ratio; RMST, restricted mean survival time; PFS, progression-free survival; OS, overall survival; CI, confidence interval; CIK, cytokine-induced killer; PH, proportional hazard; NPH, non-proportional hazard.

AND ((“phase III”[Title/Abstract] OR “phase II”[Title/Abstract] OR “phase 3”[Title/Abstract] OR “phase 2”[Title/Abstract]) AND “hepatocellular carcinoma”[Title/Abstract]) AND 2000/01/01:2022/08/31[Date - Publication] was applied and resulted in 40 potentially relevant articles. After excluding 25 articles that are either non-clinical trials, phase I trials, single arm trials, updates or non-survival-related studies, 15 relevant articles remained. Among these 15 studies, there were 9 liver cancer cytokine-induced killer cell immunotherapy studies with 1,188 liver cancer patients and 6 checkpoint inhibitors immunotherapy studies with 3,043 patients. Among all the studies, 2 CIK cell agent studies provided neither the Kaplan-Meier plots nor the numbers at risk and another 4 failed to specify the number at risk. In addition, 1 checkpoint inhibitor study was removed from the list due to the absence of PFS results. In the end, 3 CIK cell agent studies and 5 checkpoint inhibitor studies with 3,099 patients in total were included for our analysis. The HR and the RMST ratios for PFS and OS of the 8 studies were summarized in *Table 1*. In addition, a more comprehensive description of these 8 studies is provided in the Table S1. It is worth noticing that 2 out of 8 studies were supported by governmental organizations while others were funded by industry.

Figure 1 illustrates the WLSR line and provided the R-squared values between OS and PFS for HR and RMST ratios. While different colors represent different trials, the size of the dots indicates the number of patients in each trial. The R-squared values for HR and RMST ratio were 0.31 and 0.16 respectively, which indicate a moderate correlation and low correlation respectively. Notably, given that the OS and PFS were assessed using data obtained from two different groups of patients in Kelley *et al*. [2022], the weight of this study that was applied for WLSR was defined as the sample size of PFS analysis (14). Furthermore, the result of WLSR using the average sample size of OS and PFS is provided as Figure S2 as a sensitivity analysis. The results of several chi-square tests that were designed to compare the characteristics of PFS population and the additional population in OS group are summarized in *Table 2*. Another sensitivity analysis we conducted was to remove the study of Kelley *et al*. The plot is presented as Figure S3. After excluding this study, the R-squared values of WLSR using HR and RMST ratio were 0.48 and 0.24 respectively. The values indicated moderate correlation and low correlation when assessed using HR and RMST respectively.

In addition, the possible impact of tau value on RMST was investigated in this study. Through examining the changes in RMST after applying several different tau values, it is observed that although changing the tau value could affect the RMST values, the magnitude of impact was small. Based on this reason, all the default tau values were applied

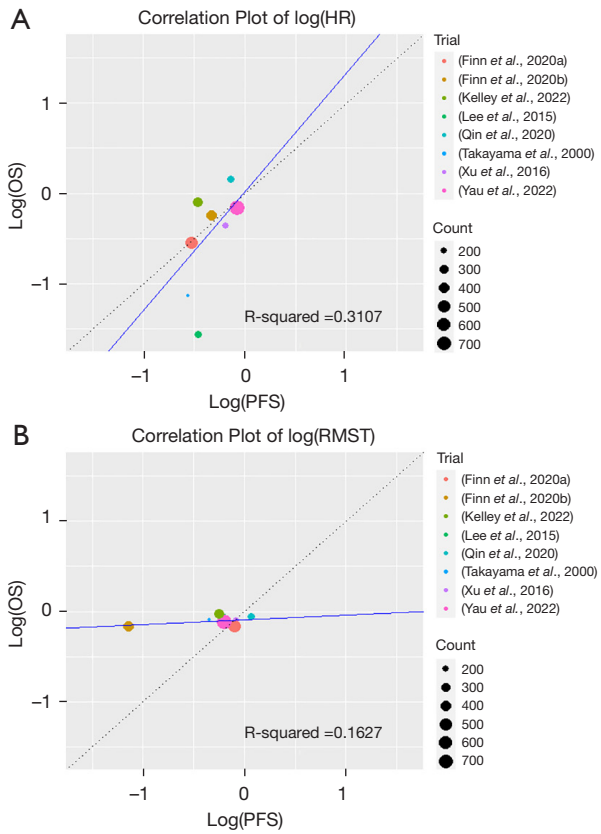

Figure 1 Correlations between PFS and OS based on HCC trial data. (A) HR; (B) RMST ratio. HR, hazard ratio; OS, overall survival; PFS, progression-free survival; RMST, restricted mean survival time; HCC, hepatocellular carcinoma.

for calculating RMST ratios of respective trials instead of setting a desired one by ourselves.

Discussion

In this study, the value of PFS serving as a surrogate for OS was examined. As our results showed, the correlations between PFS and OS were moderate and low when the results were examined using HR and RMST ratio respectively. Given that different levels of correlation were observed, more in-depth consideration is needed, and further evaluation should be conducted. In addition to PFS, overall response rate (ORR) can also be considered as a surrogate endpoint for OS. However, given that it is not relevant to survival outcome it does not have hazards estimation for survival curves and thus no need to make any PHs assumption. Among various kinds of immunotherapies, our study included CIK cell immunotherapy studies and checkpoint inhibitor immunotherapy studies. CIK cell studies mainly targeted stage A patients with some stage B and C patients being involved. As for checkpoint inhibitor studies, they mainly focused on stage B and C patients. Such design is believed to help us ensure that our study covered different stages of HCC patients.

Table 2 Comparison of two groups of patients in (14)

Characteristics	PFS population	Additional population	P value
Stage			0.6136
Stage B	125	87	
Stage C	247	190	
Region			<0.001
Asia	96	87	
Europe	153	65	
Others	123	125	
Race			<0.001
Asian	103	96	
Other	203	170	
Not reported	66	11	
ECOG score			0.5562
0	237	184	
1	134	93	
Albumin-Bilirubin score			0.6749
1	216	156	
2	152	119	
Extra hepatic spread or macrovascular invasion, or both			0.5091
Yes	260	186	
No	112	91	
Alpha-fetoprotein (ng/mL)			0.0904
≥400	252	169	
<400	120	108	
Numbers of sites			0.7854
1	101	76	
2	185	132	
3+	83	68	

PFS, progression-free survival; ECOG, Eastern Cooperative Oncology Group.

In addition, the possible impact of tau value on RMST was also investigated during the RMST value calculation. The impact of changing tau value on RMST ratio was small enough to be ignored, which therefore supports our decision to use the default tau value in our calculations. This is one of the strengths of this study because with the default tau value, which is defined as the maximum follow up time in R, being applied, we were able to make full use of the extracted data.

The small number of included studies is one of the limitations of this study. During the screening process, there were 6 studies that failed to provide the Kaplan-Meier plot and/or the number at risk, making it impossible to reproduce the curves for further procedures. This might be a potential factor that affects the correlation between OS and PFS. In fact, insufficient reporting in essential parameters has also brought difficulties in assessing surrogacy comprehensively (17). Similar issues in small number of included studies were also reported in another paper investigating the level of surrogacy in lung cancer (18). Furthermore, another limitation of this study is that we did not assess the level of surrogacy using individual-level data. As previous research has pointed out, individual-level surrogacy and trial-level surrogacy are distinct concepts and both approaches are recommended to be conducted for investigating the level of surrogacy (19).

Conclusions

In conclusion, the strength of PFS surrogacy for OS for RMST ratio is numerically lower than HR. RMST ratio analysis should be considered in addition to HR when assessing the level of surrogacy. In addition, analyses based on individual patient data are recommended to better assess the surrogacy in liver cancer immunotherapy trials.

Acknowledgments

An electronic abstract of this study has been published in the American Society of Clinical Oncology (ASCO) Annual Meeting 2022.

Funding: The research work was partially supported by HMRF grant of Hong Kong 16172901 (to H Pang, JC Ho) and Postgraduate scholarship of the University of Hong Kong (to TH Leung).

Footnote

Reporting Checklist: The authors have completed the Narrative Review reporting checklist. Available at https://cco.amegroups.com/article/view/10.21037/cco-23-48/rc

Peer Review File: Available at https://cco.amegroups.com/article/view/10.21037/cco-23-48/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://cco.amegroups.com/article/view/10.21037/cco-23-48/coif). XW serves as an unpaid editorial board member of *Chinese Clinical Oncology* from January 2023 to December 2024. THL reports support for attending meeting from University Postgraduate Fellowship of the University of Hong Kong. HP reports HMRF grant of Hong Kong 16172901, an NIHU01 grant from FDA, stock options from Roche, and personal fees from Genentech, outside the submitted work. JCH reports HMRF grant of Hong Kong 16172901. The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Rumgay H, Arnold M, Ferlay J, et al. Global burden of primary liver cancer in 2020 and predictions to 2040. J Hepatol 2022;77:1598-606.
2. Villanueva A. Hepatocellular Carcinoma. N Engl J Med 2019;380:1450-62.
3. Saad ED, Buyse M. Statistical controversies in clinical research: end points other than overall survival are vital for regulatory approval of anticancer agents. Ann Oncol 2016;27:373-8.
4. Bai R, Li W, Du N, et al. Challenges of evaluating immunotherapy efficacy in solid tumors. Chin J Cancer Res 2019;31:853-61.

5. Paz-Ares L, Tan EH, O'Byrne K, et al. Afatinib versus gefitinib in patients with EGFR mutation-positive advanced non-small-cell lung cancer: overall survival data from the phase IIb LUX-Lung 7 trial. Ann Oncol 2017;28:270-7.
6. Zheng L, Wang Y, Xu Z, et al. Concurrent EGFR-TKI and Thoracic Radiotherapy as First-Line Treatment for Stage IV Non-Small Cell Lung Cancer Harboring EGFR Active Mutations. Oncologist 2019;24:1031-e612.
7. Rohatgi A. Webplotdigitizer: Version 4.4. Available online: https://automeris.io/WebPlotDigitizer. 2020;411.
8. Guyot P, Ades AE, Ouwens MJ, et al. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol 2012;12:9.
9. Lee JH, Lee JH, Lim YS, et al. Adjuvant immunotherapy with autologous cytokine-induced killer cells for hepatocellular carcinoma. Gastroenterology 2015;148:1383-91.e6.
10. Takayama T, Sekine T, Makuuchi M, et al. Adoptive immunotherapy to lower postsurgical recurrence rates of hepatocellular carcinoma: a randomised trial. Lancet 2000;356:802-7.
11. Xu L, Wang J, Kim Y, et al. A randomized controlled trial on patients with or without adjuvant autologous cytokine-induced killer cells after curative resection for hepatocellular carcinoma. Oncoimmunology 2015;5:e1083671.
12. Finn RS, Qin S, Ikeda M, et al. Atezolizumab plus Bevacizumab in Unresectable Hepatocellular Carcinoma. N Engl J Med 2020;382:1894-905.
13. Finn RS, Ryoo BY, Merle P, 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;38:193-202.
14. Kelley RK, Rimassa L, Cheng AL, et al. Cabozantinib plus atezolizumab versus sorafenib for advanced hepatocellular carcinoma (COSMIC-312): a multicentre, open-label, randomised, phase 3 trial. Lancet Oncol 2022;23:995-1008.
15. Qin S, Ren Z, Meng Z, et al. Camrelizumab in patients with previously treated advanced hepatocellular carcinoma: a multicentre, open-label, parallel-group, randomised, phase 2 trial. Lancet Oncol 2020;21:571-80.
16. Yau T, Park JW, Finn RS, et al. Nivolumab versus sorafenib in advanced hepatocellular carcinoma (CheckMate 459): a randomised, multicentre, open-label, phase 3 trial. Lancet Oncol 2022;23:77-90.
17. Belin L, Tan A, De Rycke Y, et al. Progression-free survival as a surrogate for overall survival in oncology trials: a methodological systematic review. Br J Cancer 2020;122:1707-14.
18. Pang H, Yang G, Ho JC, et al. Assessing surrogacy using restricted mean survival time ratio for overall survival in non-small cell lung cancer immunotherapy studies. Chin Clin Oncol 2022;11:7.
19. Buyse M, Saad ED, Burzykowski T, et al. Surrogacy Beyond Prognosis: The Importance of "Trial-Level" Surrogacy. Oncologist 2022;27:266-71.

Cite this article as: Leung TH, Ho JC, Wang X, Pang H. Assessing surrogacy using restricted mean survival time ratio for overall survival in liver cancer: a narrative review. Chin Clin Oncol 2023;12(5):53. doi: 10.21037/cco-23-48

Supplementary

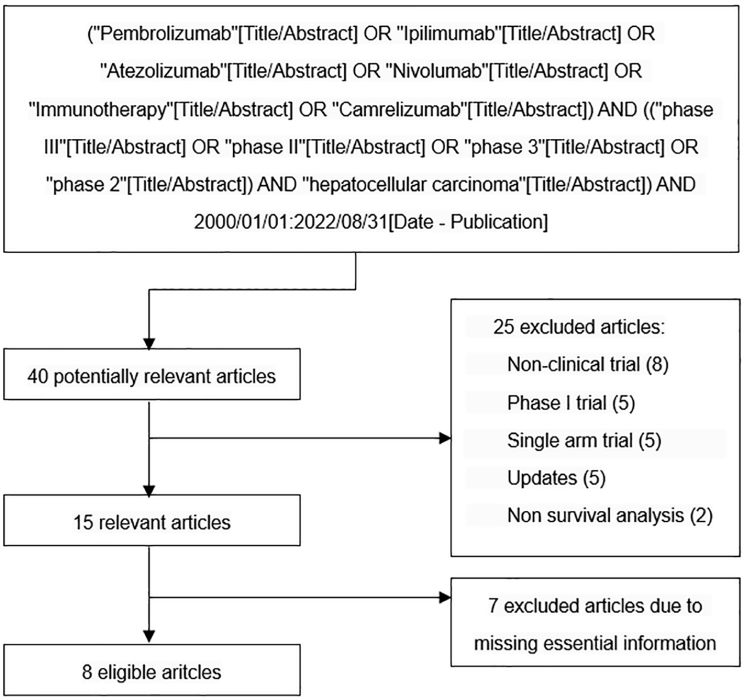

Figure S1 Literature search process.

Table S1 Characteristics of the included studies

Immunotherapy trial	Study type	Funding source	Analysis population	HR estimate	Stratified HR
CIK cell					
Lee *et al*., [2015] (9)	Interventional	Industry	ITT	Cox PH model	Reported
Takayama *et al*., [2000] (10)	Interventional	Government	ITT	Cox PH model	Not reported
Xu *et al*., [2016] (11)	Interventional	Government	ITT	Aalen's linear hazard models	Not reported
Checkpoint inhibitor					
Finn *et al*., [2020a] (12)	Interventional	Industry	ITT	Cox PH model	Not reported
Finn *et al*., [2020b] (13)	Interventional	Industry	ITT	Cox PH model	Reported
Kelley *et al*., [2022] (14)	Interventional	Industry	ITT	Cox PH model	Not reported
Qin *et al*., [2020] (15)	Interventional	Industry	ITT	Cox PH model	Reported
Yau *et al*., [2022] (16)	Interventional	Industry	ITT	Cox PH model	Not reported

HR, hazard ratio; CIK, cytokine-induced killer; ITT, intention-to-treat; PH, proportional hazards.

https://dx.doi.org/10.21037/cco-23-48

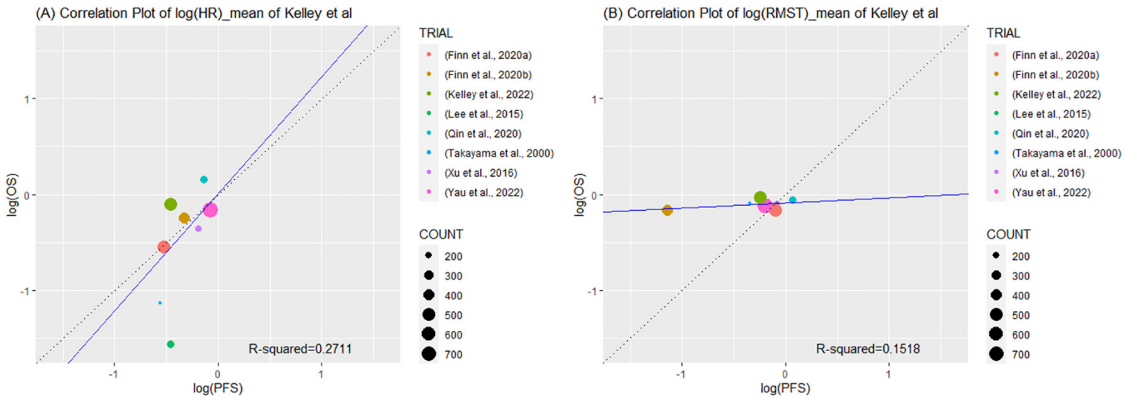

Figure S2 Correlations between PFS and OS with the average sample size being the weight of the study conducted by (14). (A) HR; (B) RMST ratio. PFS, progression-free survival; OS, overall survival; HR, hazard ratio; RMST, restricted mean survival time.

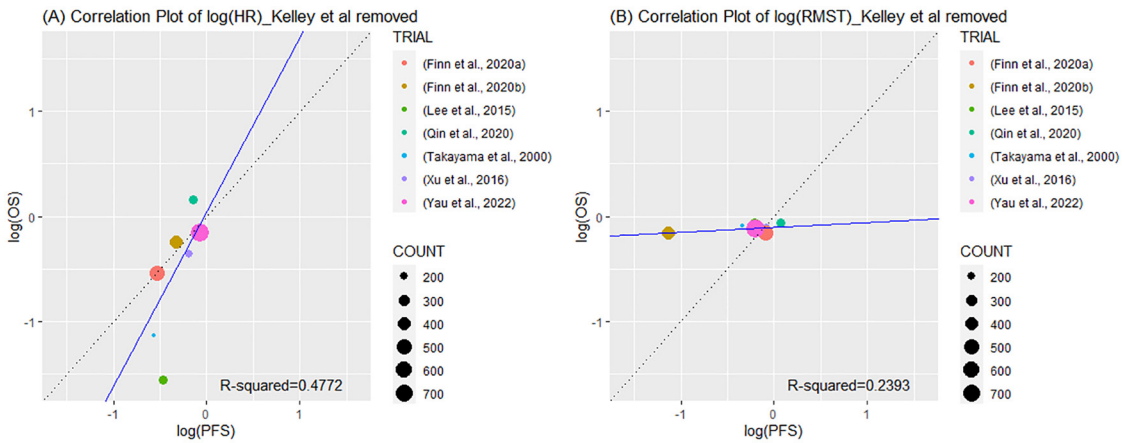

Figure S3 Correlations between PFS and OS after excluding the study conducted by (14). (A) HR; (B) RMST ratio. PFS, progression-free survival; OS, overall survival; HCC, hepatocellular carcinoma; HR, hazard ratio; RMST, restricted mean survival time.

https://dx.doi.org/10.21037/cco-23-48