CO153

Predicting Overall Survival (OS) Benefit in Previously Untreated Metastatic Melanoma from Improvements in Response Outcomes: A Correlation Meta-Analysis of Randomized Controlled Trials (RCTs)

Lisa Leung[1], Paul Serafini[1], Swetha Srinivasan[2], Flavia Ejzykowicz[2], Mir-Masoud Pourrahmat[1], Andriy Moshyk[2], Murat Kurt[2]

[1]Evidinno Research Outcomes Inc., Vancouver, BC, Canada; [2]Bristol Myers Squibb, Lawrenceville, NJ, USA

Background

- Skin cancers are one of the most diagnosed cancers worldwide with melanoma accounting for the majority of mortality.[1]
- Overall survival (OS) is universally recognized as the most unambiguous endpoint with paramount clinical relevance in randomized controlled trials (RCTs) of oncology. However, observing a statistically mature OS benefit may require considerable follow-up time. Therefore, establishing intermediate endpoints that may reach statistical maturity sooner than OS as valid surrogates could expedite drug development and improve patient access to treatments.
- Although progression-free survival (PFS) is a commonly studied surrogate endpoint in metastatic settings of oncology, objective response (OR) and complete response (CR) could enable earlier assessment of emerging treatments as time to achieve partial response or CR is typically shorter than the time to observe statistically mature PFS benefit.
- OR-OS surrogacy in advanced melanoma has been previously studied in the literature.
  - One study in advanced melanoma investigating OR as a surrogate endpoint for OS did not find a strong correlation between OR and OS. However, the analysis was restricted to studies of anti-PD-1/PD-L1s and therefore was based on limited evidence, and it included both studies of previously treated patients and studies of previously untreated patients.[2]
  - Another study investigating OR-OS surrogacy in the treatment of previously untreated advanced melanoma with immune checkpoint inhibitors (ICI) found a strong association at the patient level, but only a moderate association at the trial level.[3]
- To our knowledge, CR-OS surrogacy has not been studied yet for previously untreated advanced melanoma settings.

Objectives

- To evaluate OR and CR as surrogate endpoints for OS by modelling the association between the treatment effects of each surrogate endpoint and OS with aggregate-level data from RCTs investigating first-line (1L) therapies for advanced melanoma.
- To investigate the predictive accuracy of the surrogacy equations for the utility and validity of the models.

Methods

Systematic Literature Review

- A systematic literature review was conducted to search MEDLINE®, Embase, and Cochrane Central Register of Controlled Trials (CENTRAL) from database inception to October 2020. Searches were limited to references in the English language.
- Included articles were RCTs of patients ≥18 years old with advanced, unresectable melanoma undergoing 1L therapy.
- Outcomes of interest were OR, CR, and OS. To be included in the analysis, the trials must have reported the hazard ratio of OS (HR_{OS}) or Kaplan-Meier curves for OS, and the OR rate or CR rate by arm or the odds ratio of OR (OR_{OR}) or odds ratio of CR (OR_{CR}) between the arms.

Trial-level surrogacy models and analysis sets

- The OR-OS and CR-OS surrogacy at the trial level was assessed using two meta-analysis models. HR_{OS}, OR_{OR}, and OR_{CR} 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,[4] which provides an overall correlation measure between $\ln(OR_{OR})$ and $\ln(HR_{OS})$ and between $\ln(OR_{CR})$ and $\ln(HR_{OS})$.
- The second model was a weighted linear regression (WLR) model where each study was weighted by its corresponding sample size in a regression model estimating $\ln(HR_{OS})$ from $\ln(OR_{OR})$ and $\ln(OR_{CR})$. The association between $\ln(OR_{OR})$ and $\ln(HR_{OS})$ and between $\ln(OR_{CR})$ and $\ln(HR_{OS})$ was measured by the Pearson correlation coefficient.
- In addition to the primary analysis, sensitivity analyses were conducted by restricting the evidence base to: (1) Studies in which the experimental arm investigated an ICI, (2) Studies in which the experimental arm investigated an ICI or BRAF/mitogen-activated protein kinase inhibitor (MEKi), (3) Phase III studies, (4) Studies that adjusted for or did not allow treatment crossover, and (5) Studies that assessed response according to RECIST v1.1 criteria.
- A meta-regression analysis was also conducted by including the proportion of BRAF-mutant patients in each study as a continuous covariate to explore the impact of BRAF-mutation status on the strength of correlation.

Assessing the validity of surrogacy equation and the correlation estimates

- The validity of the surrogacy equation derived from WLR was assessed by using a leave-one-out cross-validation (LOOCV).
- The National Institute for Health and Care Excellence Decision Support Unit Technical Support Document 20 was used as a guide to assess model validity.[5] A surrogacy model can be considered to be valid if the observed HR_{OS}'s in the trials were covered by the 95% prediction intervals (PIs) obtained from the model for at least 95% of the contrasts.
- The utility of the WLR model was assessed by estimating the surrogate threshold effect (STE), which is defined as the minimum treatment effect on the surrogate endpoint that would predict a positive treatment effect on OS with 95% probability.
  - In statistical terms,[6] STE corresponds to the OR_{OR} or OR_{CR} at which the upper bound of 95% PI of the HR_{OS} crosses 1.
  - As the STE depends on the sample size of a prospective RCT, it was estimated for two hypothetical trials with sample sizes of 400 and 600 patients. Because STE is monotonic with respect to sample size, the range of STEs obtained from these two hypothetical settings also provide insights on the STEs for trials with sample sizes between 400 and 600 patients.
- The German Institute of Quality and Efficiency in Health Care (IQWiG) guidelines were used to assess the strength of the correlation estimate.[7]
  - According to the IQWiG criteria, a correlation is considered "high" if the lower limit of the 95% confidence interval (CI) of the estimated correlation coefficient ≥ 0.85, "low" if the upper limit of the 95% CI of the estimated correlation coefficient ≤ 0.7, and medium otherwise.

Results

Systematic Literature Review

- Among the identified 63 records which represent 28 unique RCTs included in the literature review, 26 RCTs published between 2000 to 2020 were included in the correlation meta-analysis.
- The RCTS included 69 to 945 patients (median: 381).
- Treatments studied in the RCTs involved mostly immunotherapies +/- chemotherapy (n = 10) or kinase inhibitors +/- chemotherapy (n = 9). In four studies the experimental arm was chemotherapy monotherapy, and two administered a combination of therapies including both an immunotherapy and a kinase inhibitor. The RCTs included in the evidence base and the therapies they investigated are listed in **Table 1.**

Table 1. List of RCTs included in the evidence base and the therapies they investigated.

Trial (or Publication)	N	Phase	Experimental Arm(s)	Control Arm(s)	BRAF-MT (%)
Algazi (2020)	206	II	DAB + TRM (Intermittent)	DAB + TRM (Continuous)	100
Ascierto (2017)	727	III	IPI 10mg	IPI 3mg	21.9
Avril (2004)	229	III	FTMS	DTIC	Not reported
BREAK-3	250	III	DAB	DTIC	100
BRIM-3	675	III	VEM	DTIC	100
CheckMate 064	140	II	NIVO followed by IPI	IPI followed by NIVO	28.2
CheckMate 066	418	III	NIVO	DTIC	0
CheckMate 067	945	III	NIVO + IPI	IPI	31.5
CheckMate 069	142	II	NIVO + IPI	IPI	23.2
CheckMate 511	360	IIIb/IV	NIVO 3mg + IPI 1mg	NIVO 1mg + IPI 3mg	41.9
coBRIM	495	III	VEM + CBM	VEM	100
COLUMBUS	577	III	ENC + BIN	ENC, VEM	100
COMBI-d	423	III	DAB + TRM	DAB	100
COMBI-v	704	III	DAB + TRM	VEM	100
Hersh (2011)	76	II	DTIC + IPI	IPI	Not reported
IMspire150	514	III	ATZ + VEM + CBM	VEM + CBM	100
KEYNOTE-006	834	III	PEM Q3W	IPI	36.2
KEYNOTE-022	120	II	PEM + DAB + TRM	DAB + TRM	100
KEYNOTE-029	102	I/II	PEM + IPI 50mg	PEM + IPI 100mg	34.3
Lebbe (2020)	194	II	PIM	DTIC	41.9
Middleton (2000)	305	III	DTIC	TMZ	Not reported
NEMO	402	III	BIN	DTIC	Not reported
PACMEL	111	II	PAC + PAZ	PAC	0
Patel (2011)	859	III	TMZ	DTIC	Not reported
Robert (2011)	502	III	IPI + DTIC	DTIC	Not reported
Weide (2019)	69	IIa	L19IL2 + DTIC	DTIC	Not reported

ATZ: Atezolizumab; BIN: Binimetinib; CBM: Cobimetinib; DAB: Dabrafenib; DTIC: Dacarbazine; ENC: Encorafenib; FTMS: Fotemustine; IPI: Ipilimumab; NIVO: Nivolumab; PAC: Paclitaxel; PAZ: Pazopanib; PEM: Pembrolizumab; PIM: Pimasertib; TMZ: Temozolomide; TRM: Trametinib; VEM: Vemurafenib.

Correlation Meta-Analysis

Primary Analysis: OR

- WLR (**Figure 1A**) estimated a correlation of -0.61 (95% CI: -0.82 – -0.25) and a surrogacy equation of $\ln(HR_{OS}) = -0.13 - 0.17 \times \ln(OR_{OR})$.
- The estimated STEs were 2.97 and 4.47 for RCTs including 600 and 400 patients, respectively (the larger the sample size, the lower the STE).
- In LOOCV, in 25 out of 26 (96%) contrasts the observed HR_{OS}'s were covered by the 95% PIs generated by the WLR for the predicted HR_{OS}'s.
- When the proportion of BRAF-mutant patients was included as a continuous covariate in the WLR, the correlation was 0.74 (95% CI: 0.38 – 0.91) with a coverage rate of 94% where the estimated surrogacy equation was $\ln(HR_{OS}) = -0.03 - 0.34 \times \ln(OR_{OR}) - 0.22 \times BRAF + 0.27 \times \ln OR_{OR} \times BRAF$, where "BRAF" represents the proportion of BRAF-mutant patients in the study.

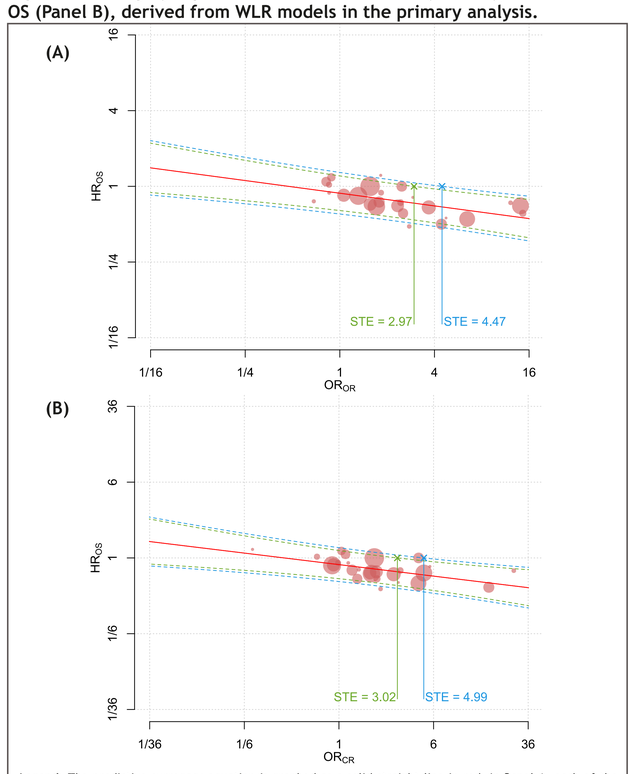
Primary Analysis: CR

- WLR (**Figure 1B**) estimated a correlation of -0.55 (95% CI: -0.79 – -0.16) and a surrogacy equation of $\ln(HR_{OS}) = -0.16 - 0.15 \times \ln(OR_{CR})$.
- The estimated STEs were 3.02 and 4.99 for RCTs including 600 and 400 patients, respectively.
- In LOOCV, in 24 out of 26 (92%) contrasts the observed HR_{OS}'s were covered by the 95% PIs generated by the WLR for the predicted HR_{OS}'s.
- When the proportion of BRAF-mutant patients was included as a continuous covariate in the WLR, the correlation was 0.71 (95% CI: 0.31 – 0.89) with a coverage rate of 94% where the estimated surrogacy equation was $\ln(HR_{OS}) = -0.17 - 0.18 \times \ln(OR_{CR}) - 0.07 \times BRAF + 0.04 \times \ln OR_{CR} \times BRAF$.

Sensitivity Analyses

- The results of the sensitivity analyses are summarized in **Table 2.**
- Sensitivity analyses for OR-OS and CR-OS surrogacy produced slightly worse or better BRMA correlations, WLR correlations, and LOOCV coverage than the primary analysis.
- STEs from the sensitivity analyses were not consistently greater or lesser than STEs from the primary analysis for either OR-OS and CR-OS.

Figure 1. Surrogacy equation between OR and OS (Panel A), and CR and OS (Panel B), derived from WLR models in the primary analysis.



Legend: The predictive surrogacy equation is graphed as a solid straight line in red. In Panel A, each of the plotted red circles represent the (OR_{OR}, HR_{OS}) pair from a trial, and Panel B each of the plotted red circles represent the (OR_{CR}, HR_{OS}) pair from a trial. Sizes of the circles are proportional to the total number of patients within each trial. The dotted curves refer to the 95% PIs for the HR_{OS} for a range of OR_{OR} or OR_{CR} for hypothetical trials with sample sizes of 400 and 600. Solid lines connecting the green- and blue-colored crosses to the x-axis indicate the STEs calculated for two hypothetical trials with sample sizes 400 (blue) and 600 patients (green). Both axes are on the logarithmic scale. HR: hazard ratio; OR_{OR}: odds ratio of objective response; OR_{CR}: odds ratio of complete response; OS: overall survival; STE: surrogate threshold effect; WLR: weighted linear regression.

Table 2. Summary of results from BRMA and WLR models.

Analysis/ Data Set	BRMA Correlation Coefficient		WLR Correlation Coefficient		STE (N = 400 / N = 600)		Prediction Accuracy in LOOCV	
	Surrogate Endpoint		Surrogate Endpoint		Surrogate Endpoint		Surrogate Endpoint	
	OR	CR	OR	CR	OR	CR	OR	CR
Primary Analysis	-0.59	-0.65	-0.61	-0.55	4.47/2.97	4.99/3.02	96%	92%
BRAF-adjusted* Analysis	0.64	0.75	0.74	0.71	2.71/1.97	2.55/1.69	94%	94%
ICI Studies	-0.64	-0.66	-0.80	-0.71	3.96/2.94	9.08/4.74	100%	100.0%
ICI or BRAF/MEKi Studies	-0.53	-0.65	-0.57	-0.57	3.77/2.36	3.21/1.99	100%	96%
Phase III Studies	-0.63	-0.65	-0.64	-0.56	6.47/3.96	7.85/4.31	94%	94%
Crossover-adjusted Studies	-0.78	-0.92	-0.61	-0.79	N/E	2.20/1.51	100%	83%
RECIST v1.1	-0.53	-0.65	-0.56	-0.55	4.34/2.66	3.66/2.22	100%	95%

Note: STE corresponds to the odds ratio of OR or CR at which the upper bound of 95% PI of the HR_{OS} crosses 1.
**: The STE for the BRAF-adjusted analysis is for a hypothetical study with 62% BRAF-mutant patients (62% was the mean from the evidence base). BRAF/MEKi: BRAF and mitogen-activated protein kinase inhibitor, BRMA: bivariate random-effects meta-analysis; CR: complete response; ICI: immune checkpoint inhibitor; LOOCV: leave one out cross validation; N: sample size; N/E: not estimable; OR: objective response; STE: Surrogate Threshold Effect; WLR: weighted linear regression.*

External Validation

- Predictions from the WLR were validated against published HRs of OS from the IMspire170,[8] PIVOT IO 001,[9] and RELATIVITY-047[10] trials (**Table 3**).
- All model predictions approximated the published HRs, but predictions from OR were generally closer to the reported HRs than the predictions from CR.
- Predictions from the BRAF-adjusted OR model were highly accurate (up to two decimal places) for IMspire170 and RELATIVITY-047, and nearly-so for PIVOT IO 001 (0.94 predicted vs. 0.92 observed).

Table 3. External validation against trials that were published recently after the finalization of the evidence base.

Trial	N	BRAF-MT (%)	Observed HR_{OS} (95% CI)	Predicted HR_{OS} (95% PI)	
				BRAF-Unadjusted model	BRAF-Adjusted model
IMspire170[8]	446	0.0	**1.06** (0.69 - 1.61)	*OR as surrogate* **0.92** (0.64 - 1.33)	*OR as surrogate* **1.06** (0.68 - 1.66)
				CR as surrogate **0.87** (0.59 - 1.28)	*CR as surrogate* **0.86** (0.58 - 1.29)
PIVOT IO 001[9]	783	41.0	**0.94** (0.71 - 1.24)	*OR as surrogate* **0.94** (0.70 - 1.26)	*OR as surrogate* **0.97** (0.71 - 1.32)
				CR as surrogate **0.92** (0.64 - 1.32)	*CR as surrogate* **0.88** (0.65, 1.19)
RELATIVITY-047[10]	714	38.3	**0.80** (0.60 - 1.00)	*OR as surrogate* **0.82** (0.61 - 1.09)	*OR as surrogate* **0.80** (0.61 - 1.06)
				CR as surrogate **0.83** (0.61 - 1.13)	*CR as surrogate* **0.80** (0.60 - 1.06)

BRAF-MT: proportion of patients with BRAF-mutant status; CI: confidence interval; N: sample size; OR: objective response; OS: overall survival; PI: prediction interval.

Conclusions

- Statistically meaningful correlations were found between the treatment effects on OR and OS, and between the treatment effects on CR and OS in patients with advanced melanoma receiving 1L therapy. Analyses adjusting for BRAF-mutation status were confirmatory to the main findings from the primary analyses.
- The surrogacy equation between the treatment effects for response outcomes and OS may enable earlier assessments of OS benefit from the OR/CR benefit in previously untreated melanoma, even relative to other surrogate endpoints such as PFS: in our evidence base, the median time to OR was only 2.8 months, compared with 5.3 months for PFS. This surrogacy equation can also be used to validate other surrogates for OS such as PFS, duration of response, and time to next treatment.
- Sensitivity analyses generated similar correlation estimates to those in the primary analysis from both models, indicating their robustness.
- To our knowledge, this study presents the most-recent and comprehensive analysis of OR-OS surrogacy using the broadest evidence base in previously untreated advanced melanoma. Furthermore, it presents the first analysis on CR-OS surrogacy study in advanced melanoma.
- A limitation of this study is that it was limited to the prediction of treatment effects on OS from OR/CR; the validation of OR/CR as surrogates for OS would require additional investigation, including individual-level association and biological plausibility of a causal relationship. Additionally, the results may not generalize beyond metastatic melanoma.
- With more trials published in the future, updates to models for both surrogate endpoints are warranted to explore the impact of the changing treatment landscape in metastatic melanoma on the predictive ability of response outcomes.

References

1. Arnold M, et al. *JAMA Dermatol* 2022;158(5):495-503.
2. Nie RC, et al. *Ther Adv Med Oncol* 2020;12:1758835920929583.
3. Larkin J, et al. *Ann Oncol* 2022;33(Supplement 7):S919-S920.
4. Riley RD, et al. *Biostatistics* 2008;9(1):172-186.
5. Bujkiewicz S, et al. NICE DSU technical support document 20. 2019.
6. Burzykowski T, et al. *Pharm Stat* 2006;5(3):173-186
7. IQWiG. Validity of surrogate endpoints in oncology: Executive summary of rapid report A10-05, Version 1.1. In: Institute for Quality and Efficiency in Health Care: Executive Summaries. Cologne, Germany: Institute for Quality and Efficiency in Health Care (IQWiG); 2005.
8. Gogas H, et al. *Ann Oncol* 2021;32(3):384-394.
9. Diab A, et al. *J Clin Oncol* 2023;41(30):4756-4767.
10. Long GV, et al. *J Clin Oncol* 2022;40(Supplement 36):360385.

Acknowledgments

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

Presented at ISPOR 2024; May 5–8, 2024; Atlanta, GA

Email: mpourrahmat@evidinno.com