SUPPLEMENTARY MATERIAL

Two-Stage Estimation of Overall Survival in the Phase 3 CheckMate 9ER Trial, Adjusting for the Impact of Subsequent Therapy

Marc-Oliver Grimm,[1]* Emma Karim,[2] Venediktos Kapetanakis,[3] Mickael Lothgren,[4] Alessia Ogareva,[5] Pragya Shukla,[6] James Truscott,[3] Camillo Porta[7,8]

[1]Department of Urology, Jena University Hospital, Friedrich-Schiller University Jena, Jena, Germany

[2]Ipsen, London, UK

[3]Evidera, London, UK

[4]Ipsen, Zug, Switzerland

[5]Ipsen, Paris, France

[6]Evidera, Montreal, QC, Canada

[7]Interdisciplinary Department of Medicine, University of Bari "A. Moro", Bari, Italy

[8]Division of Medical Oncology, A.O.U. Consorziale Policlinico di Bari, Bari, Italy

***Corresponding author**: Professor Marc-Oliver Grimm

Address: Department of Urology, Jena University Hospital, Friedrich-Schiller University Jena, Jena, Germany

Email: Marc-Oliver.Grimm@med.uni-jena.de

Telephone number: +49 3641 9-329901

ORCID iDs: Marc-Oliver Grimm (0000-0003-3084-8654), Camillo Porta (0003-2412-1563)

Table S1 Full list of covariates considered for inclusion in the AFT modeling

Covariate	Categorization
Age at discontinuation	< 65 years, ≥ 65 years
Sex	Male, female
Race	White, non-white[a]
Region	USA, Canada, Western Europe, Northern Europe, rest of world
KPS score[b]	70–80, 90–100
Tumor PD-L1 expression[b]	≥ 1%, < 1%, or indeterminate
IMDC prognostic score[b]	0, 1–2, 3–6
Sarcomatoid features[b]	Yes, no
Previous radiotherapy	Yes, no
Previous nephrectomy	Yes, no
Disease duration	NA[c]
Metastatic sites[d,e]	Lung, lymph node, liver, bone, adrenal gland
Number of metastatic organs/sites[d,e]	NA[c]
Sum of diameters of target lesions[d,e]	NA[c]
EQ-5D utility values[d,f]	NA[c]
Time to discontinuation	NA[c]

Progression status at discontinuation	PFS censored before/after discontinuation, progression event before/after discontinuation, did not discontinue
Reason for discontinuation	Progression, toxicity, other

[a]Non-white races include American Indian or Alaska native, Asian, Black or African American native, Hawaiian or other Pacific Islander, or other
[b]Baseline values were used to represent values at secondary baseline
[c]Continuous variable
[d]Imputed from the closest time point to secondary baseline
[e]No cut-off window was used for the assumption that they were less unlikely to change rapidly over time
[f]A cut-off window of 30 days before or after the secondary baseline was used to identify measurements suitable for imputation
AFT, accelerated failure time; EQ-5D, 5-dimension EuroQol questionnaire; IMDC, International Metastatic Renal Cell Carcinoma Database Consortium; KPS, Karnofsky performance status; NA, not applicable; PD-L1, programmed death-ligand 1; PFS, progression-free survival

Table S2 Characteristics of the ITT population and patients included in TSE at the secondary baseline

Patient characteristic	ITT			Patients included in TSE		
	Overall (*n* = 651)	**CaboNivo (*n* = 323)**	**SUN (*n* = 328)**	**Overall (*n* = 540)**	**CaboNivo (*n* = 264)**	**SUN (*n* = 276)**
Age, ≥ 65 years at discontinuation, *n* (%)	274 (42.1)	148 (45.8)	126 (38.4)	238 (44.1)	125 (47.3)	113 (40.9)
White / non-white[a], *n* (%)	533 (81.9) / 118 (18.1)	267 (82.7) / 56 (17.3)	266 (81.1) / 62 (18.9)	440 (81.5) / 100 (18.5)	216 (81.8) / 48 (18.2)	224 (81.2) / 52 (18.8)
Male / female, *n* (%)	480 (73.7) / 171 (26.3)	248 (76.8) / 75 (23.2)	232 (70.7) / 96 (29.3)	399 (73.9) / 141 (26.1)	203 (76.9) / 61 (23.1)	196 (71.0) / 80 (29.0)
Region, *n* (%)						
ROW	332 (51.0)	165 (51.1)	167 (50.9)	273 (50.6)	134 (50.8)	139 (50.4)
USA/Canada/ W. Europe/N. Europe	319 (49.0)	158 (48.9)	161 (49.1)	267 (49.4)	130 (49.2)	137 (49.6)
KPS at baseline, *n* (%)						
70	32 (4.9)	14 (4.3)	18 (5.5)	22 (4.1)	11 (4.2)	11 (4.0)
80	118 (18.1)	52 (16.1)	66 (20.1)	97 (18.0)	44 (16.7)	53 (19.2)
90	224 (34.4)	111 (34.4)	113 (34.5)	184 (34.1)	88 (33.3)	96 (34.8)
100	275 (42.2)	146 (45.2)	129 (39.3)	237 (43.9)	121 (45.8)	116 (42.0)
Not reported	2 (0.3)	0 (0.0)	2 (0.6)	0 (0.0)	0 (0.0)	0 (0.0)

IMDC at baseline, *n* (%)						
0	146 (22.4)	74 (22.9)	72 (22.0)	125 (23.1)	61 (23.1)	64 (23.2)
1–2	376 (57.8)	188 (58.2)	188 (57.3)	313 (58.0)	153 (58.0)	160 (58.0)
3–6	129 (19.8)	61 (18.9)	68 (20.7)	102 (18.9)	50 (18.9)	52 (18.8)
PD-L1 status at baseline, *n* (%)						
< 1% or indeterminate	485 (74.5)	240 (74.3)	245 (74.7)	401 (74.3)	196 (74.2)	205 (74.3)
≥ 1%	166 (25.5)	83 (25.7)	83 (25.3)	139 (25.7)	68 (25.8)	71 (25.7)
Sarcomatoid features at baseline, *n* (%)						
0	557 (85.6)	279 (86.4)	278 (84.8)	465 (86.1)	232 (87.9)	233 (84.4)
1	75 (11.5)	34 (10.5)	41 (12.5)	64 (11.9)	26 (9.8)	38 (13.8)
Missing	19 (2.9)	10 (3.1)	9 (2.7)	1 (2.0)	6 (2.3)	5 (1.8)
Prior radiotherapy, *n* (%)	91 (14.0)	46 (14.2)	45 (13.7)	77 (14.3)	39 (14.8)	38 (13.8)
Prior nephrectomy, *n* (%)	483 (74.2)	238 (73.7)	245 (74.7)	403 (74.6)	194 (73.5)	209 (75.7)
Disease duration at discontinuation, years, mean (SD)	3.81 (4.47)	4.30 (4.30)	3.32 (4.58)	3.72 (4.12)	4.22 (3.96)	3.25 (4.21)
Metastasis at discontinuation, *n* (%)						
Lung	491 (75.4)	240 (74.3)	251 (76.5)	412 (76.3)	200 (75.8)	212 (76.8)

Lymph node	274 (42.1)	135 (41.8)	139 (42.4)	228 (42.2)	111 (42.0)	117 (42.4)
Bone	152 (23.3)	78 (24.1)	74 (22.6)	123 (22.8)	61 (23.1)	62 (22.5)
Liver	126 (19.4)	72 (22.3)	54 (16.5)	102 (18.9)	55 (20.8)	47 (17.0)
Adrenal gland	73 (11.2)	35 (10.8)	38 (11.6)	59 (10.9)	26 (9.8)	33 (12.0)
No organs with metastasis at discontinuation, mean (SD)	2.36 (1.20)	2.35 (1.18)	2.37 (1.22)	2.35 (1.20)	2.32 (1.13)	2.38 (1.26)
Sum of diameter of target lesion at discontinuation, mean (SD)	73.73 (70.14)	65.11 (65.83)	82.25 (73.27)	71.05 (70.27)	61.71 (64.60)	79.99 (74.32)
TRAE leading to discontinuation, *n* (%)	125 (19.2)	90 (27.9)	35 (10.7)	113 (20.9)	79 (29.9)	34 (12.3)
EQ-5D utility (UK) at discontinuation,[b] mean (SD)	0.68 (0.32)	0.71 (0.31)	0.65 (0.34)	0.69 (0.33)	0.72 (0.31)	0.65 (0.34)
Progression[c] status at discontinuation						
PFS censored after discontinuation	38 (5.8)	28 (8.7)	10 (3.0)	38 (7.0)	28 (10.6)	10 (3.6)
PFS censored before discontinuation	90 (13.8)	32 (9.9)	58 (17.7)	81 (15.0)	27 (10.2)	54 (19.6)
Progression event after discontinuation	73 (11.2)	32 (9.9)	41 (12.5)	73 (13.5)	32 (12.1)	41 (14.9)
Progression event before discontinuation	403 (61.9)	207 (64.1)	196 (59.8)	348 (64.4)	177 (67.0)	171 (62.0)

Did not discontinue	47 (7.2)	24 (7.4)	23 (7.0)	-	-	-
Reason for discontinuation, *n* (%)						
Progression	388 (59.6)	176 (54.5)	212 (64.6)	371 (68.7)	164 (62.1)	207 (75.0)
Drug toxicity	77 (11.8)	38 (11.8)	39 (11.9)	73 (13.5)	35 (13.3)	38 (13.8)
Others[d]	139 (21.4)	85 (26.3)	54 (16.5)	96 (17.8)	65 (24.6)	31 (11.2)
Did not discontinue	47 (7.2)	24 (7.4)	23 (7.0)	–	–	–

[a]Non-white races include American Indian or Alaska native, Asian, Black or African American native, Hawaiian or other Pacific Islander or other
[b]Imputed using cut-off of ± 30 days
[c]Progression by IRC; primary definition
[d]Other reasons include adverse events unrelated to study drug, completed treatment, poor/non-compliance, patient requested to discontinue, patient withdrew consent, and others
CaboNivo, cabozantinib plus nivolumab; EQ-5D, 5-dimension EuroQol questionnaire; IMDC, International Metastatic Renal Cell Carcinoma Database Consortium; IRC, independent review committee; ITT, intention-to-treat; KPS, Karnofsky performance status; N. Europe, Northern Europe; PD-L1, programmed death-ligand 1; PFS, progression-free survival; ROW, rest of the world; SD, standard deviation; SUN, sunitinib; TRAE, treatment related adverse event; TSE, two-stage estimation; W. Europe, Western Europe

Table S3 Goodness-of-fit for the parametric AFT models

	CaboNivo		SUN	
Distribution	**AIC**	**BIC**	**AIC**	**BIC**
Exponential	1513.021	1516.990	1717.960	1722.047
Weibull	1501.689	1509.627	1692.641	1700.815
Gamma	1502.890	1510.828	1696.877	1705.050
Log-logistic	1503.713	1511.651	1686.260	1694.434
Log-normal	1503.313	1511.250	1683.287	1691.460
Generalized gamma	1501.824	1513.730	1685.139	1697.399

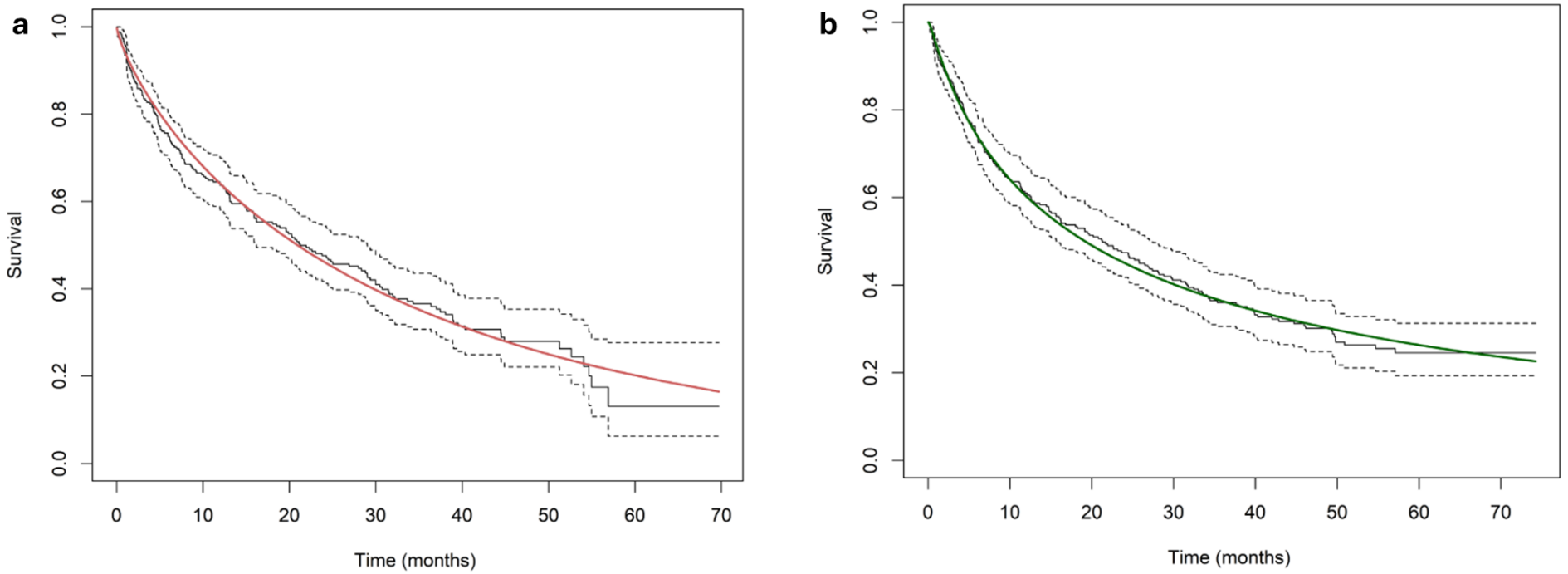
AFT, accelerated failure time; AIC: Akaike information criterion; BIC: Bayesian information criterion; CaboNivo, cabozantinib plus nivolumab; SUN, sunitinib

Re-censoring process and analysis

Re-censoring is a corrective process for the informative censoring bias that is associated with the two-stage estimation (TSE) method. Some patients who receive subsequent therapy may have censored survival time, so TSE includes adjustment of the censoring times for patients who switch treatment. However, patients who do not switch treatment retain an unadjusted survival time, resulting in informative censoring bias if treatment switching is associated with prognosis. In re-censoring, counterfactual survival times associated with a given acceleration factor value are re-censored for all patients at the minimum of the administrative censoring time, C_i. In this case, administrative censoring is the right-censoring that occurs at the end of the study follow-up period for those who have remained event-free. The earliest possible censoring time over all treatment trajectories is then represented by $C_i \times \frac{1}{\alpha}$. In contrast, re-censoring reduces informative censoring bias but may create bias due to lost long-term information. Thus, analyses with re-censoring may underestimate long-term survival of patients who received subsequent therapy, whereas analyses without re-censoring may overestimate their long-term survival [1].

When re-censoring, the larger acceleration factor calculated for the sunitinib (SUN) arm (which reflected a stronger effect of subsequent therapy) resulted in more lost events than in the cabozantinib plus nivolumab (CaboNivo) arm (98 of 217 [45.2%] versus 18 of 183 [9.8%], respectively). There was no difference in the adjusted median overall survival (OS) calculated with and without re-censoring for the CaboNivo arm, but when re-censored, median OS was not reached for SUN. With re-censoring, the adjusted OS hazard ratio (95% CI) for CaboNivo versus SUN was 0.40 (0.29–0.55; $p < 0.001$).

Fig. S1 Observed and predicted survival plots for the CaboNivo (Weibull) (a) and SUN (Log-normal) (b) models



Black lines: solid, observed survival; dashed, 95% confidence intervals
Red line, predicted survival for CaboNivo
Green line, predicted survival for SUN

CaboNivo, cabozantinib plus nivolumab; SUN, sunitinib

References

1. Latimer NR, White IR, Abrams KR, Siebert U. Causal inference for long-term survival in randomised trials with treatment switching: should re-censoring be applied when estimating counterfactual survival times? Stat Methods Med Res. 2019;28(8):2475–93. doi: 10.1177/0962280218780856.