Mixture-cure modelling of overall survival of patients with metastatic non-small cell lung cancer receiving nivolumab + ipilimumab in CheckMate 227 – updated with 5 years of follow-up

MSR16

Young R,[1] Pritchard C,[2] Yuan Y,[3] Chaudhary MA,[3] Lee A,[4] Gordon J,[1] McEwan P[1]

[1]Health Economics and Outcomes Research Ltd, Cardiff, UK; [2]ICON plc, Reading, UK;
[3]Bristol Myers Squibb, Princeton, NJ, USA; [4]Bristol Myers Squibb, Uxbridge, London, UK

Background

- Overall survival (OS) benefit of the immune checkpoint inhibitor (ICI) combination nivolumab + ipilimumab (NIVO + IPI) has been measured versus platinum doublet chemotherapy (Chemo) in patients with previously untreated metastatic non-small cell lung cancer (NSCLC) in the CheckMate 227 study, an open-label, randomised Phase 3 trial of NIVO + IPI versus Chemo[1]
- Mixture-cure modelling (MCM) of OS has previously been undertaken based on 3-year follow-up data from CheckMate 227.[2] Additional data to a minimum of 5 years of follow-up are now available to update estimated survival outcomes for for all first-line patients with NSCLC receiving NIVO + IPI or Chemo from Part 1 of this trial
- Extrapolation of OS outcomes to a lifetime time horizon is required to estimate total benefits of NIVO + IPI for cost-effectiveness analysis
- ICI survival outcomes have complex hazard profiles due to phenomena such as delayed, heterogenous, and durable response, as well as pseudo-progression[3]
- Durable response and long-term survival have been associated with ICI treatment in renal cell carcinoma,[4] melanoma,[5] and lung cancer[6]
- MCM can help estimate the long-term OS benefit of NIVO + IPI versus Chemo by explaining the complex OS hazard profile as a mixture between patients with additional hazard of mortality due to disease, and a fraction who benefit from treatment and are no longer at additional risk of mortality versus the general population

Objectives

- To improve precision of OS models fitted to 3-year data and investigate sensitivity to data maturity under an MCM framework for the NIVO + IPI and Chemo arms of CheckMate 227 with a minimum 5 years of follow-up
- To investigate the impact of MCMs with fewer degrees of freedom using alternative assumptions for the relationship of the survival time distribution of the uncured between treatment arms

Methods

Data

- Patient data were obtained from the February 2022 database lock of CheckMate 227. The full population randomised to receive NIVO + IPI or Chemo was analysed
- Baseline hazard was informed by nationality, age, and gender-specific WHO life tables matched to each patient

Mixture-cure model

- An additional hazard predicted by a MCM was assumed additive to the baseline hazard. Model consisted of:
  - A 'cured' fraction of the population assumed at no additional hazard of death
  - A complementary 'uncured' fraction experiencing additional hazard predicted by a parametric survival distribution – exponential, Weibull, Gamma, Gompertz, log-logistic, lognormal or generalised Gamma
- The cure fraction and parameters of the additional hazard model were jointly varied to find a maximum likelihood parameter estimates
- Models using distributions listed above were fitted by maximum likelihood to describe the hazard of the uncured, and goodness of fit was assessed by considering the Akaike information criteria (AIC), Bayesian information criteria (BIC), and through visual inspection of deviation from the Kaplan-Meier survival estimator and hazard profiles. Models were classified according to differences in information criteria relative to the lowest AIC model[7]

Validation

- Previous models fitted to data with a minimum of 3 years of follow-up and plotted over Kaplan-Meier estimator of data with a minimum of 5 years of follow-up to validate extrapolations
- Hazards predicted from the models were compared with hazard of death (per month) estimated from conditional survival rates from external sources, including the longer follow-up of the CheckMate 017 and 057 studies investigating nivolumab in pre-treated NSCLC, and national registries

Alternative models

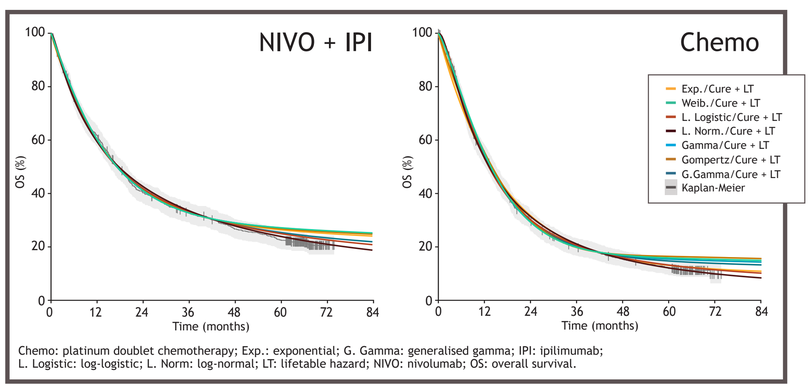
- In the base case, independent models of OS for NIVO + IPI and Chemo were produced. Alternative model structures with fewer degrees of freedom were fitted to investigate sensitivity to scaling rules among the uncured patients:
  - Independent cure fraction and location parameter of uncured distribution – Log-odds of cure dependent upon treatment; uncured distribution in accelerated failure time relationship between treatments
  - Independent cure fraction; all uncured distribution parameters shared – Log-odds of cure dependent upon treatment, survival time distribution of uncured identical between treatments

Results

Independent models

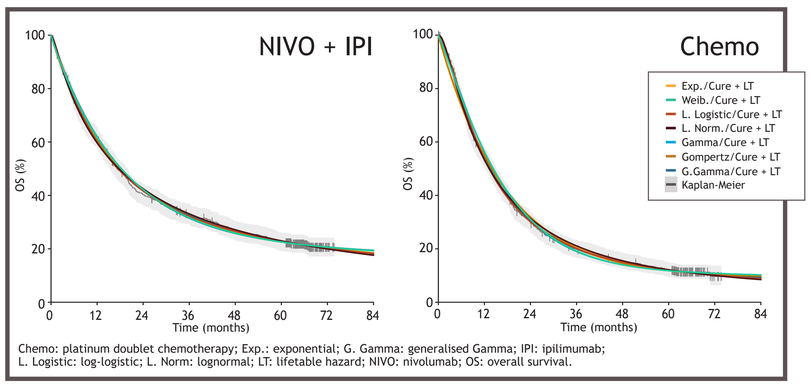
- Amongst models fitted to 3-year data, models with higher cure fractions tended to over-predict OS for both NIVO + IPI and Chemo (**Figure 1**)
- Models fitted to the updated database lock had good visual fit versus the Kaplan-Meier estimator (**Figure 2**)

Figure 1. Comparison of models fitted to 3-year DBL to Kaplan-Meier estimator from 5-year DBL



Chemo: platinum doublet chemotherapy; Exp.: exponential; G. Gamma: generalised gamma; IPI: ipilimumab; L. Logistic: log-logistic; L. Norm: log-normal; LT: lifetable hazard; NIVO: nivolumab; OS: overall survival.

Figure 2. Comparison of models fitted to 5-year DBL versus Kaplan-Meier estimator from same DBL



Chemo: platinum doublet chemotherapy; Exp.: exponential; G. Gamma: generalised Gamma; IPI: ipilimumab; L. Logistic: log-logistic; L. Norm: lognormal; LT: lifetable hazard; NIVO: nivolumab; OS: overall survival.

- Increased follow-up has:
  - **Increased the spread of information criteria** over differing models of the 'uncured' fraction (**Table 1**)
    - At 3 years: Exponential and Gompertz models >10 units of AIC and BIC from lowest model
    - At 5 years: Exponential, Weibull, Gamma, and Gompertz models >10 units AIC and BIC from lowest (log-logistic), Generalised Gamma >10 units BIC from log-logistic
    - Models deviating by >10 units would be considered likely to be a worse model than the best-fitting model and are not recommended to inform lifetime predictions
  - **Decreased the spread of predicted cure fractions** over differing models of the 'uncured' fraction (**Table 1**)
    - At 3 years: NIVO + IPI cure fraction: 29.1% (95% CI 24.0%, 34.0%) (highest - Weibull model) to 11.1% (95% CI 0.1%, 21.8%) (lowest - Lognormal model)
    - At 5 years: NIVO + IPI cure fraction: 21.6% (95% CI 17.4%, 26.0%) (highest – Gamma model) to 8.2% (95% CI 0.1%, 16.3%) (lowest – Lognormal model)

Validation

- Models rejected at 5 years that were acceptable at 3 years (Weibull, Gamma) predict hazards near general population mortality within 8 years (**Figure 3**)
- The log-logistic and lognormal models are consistent with long-term historical observations of mortality hazard for Chemo. To 15 years, the hazard progression of these models is plausible
- All models for NIVO + IPI predict lower long-term hazard than historical observations, but the log-logistic, lognormal and generalised Gamma models are only slightly lower than long-term observations from SEER. Hazard decreases more slowly than for historical sources
- Initial hazard for NIVO + IPI was lower than CheckMate 017 and 057. By 5 years, hazard from all models was consistent with these data

Alternative models

- Independent modelling of treatment arms results in the lowest AIC and BIC
  - However, some models with fewer degrees of freedom which assume that survival times of the uncured are identical between treatment arms do not greatly differ from the best-fitting fully independent model with respect to these information criteria (**Table 2**)
- Assumption of identical distribution of survival times in the uncured fraction, or an accelerated failure time relationship of survival times in the uncured fraction resulted in higher cure fractions for NIVO + IPI and lower cure fractions for Chemo (**Table 2**).
- Deviation of the best-fitting reduced degrees of freedom models from the Kaplan-Meier estimator was low and within the 95% confidence interval of this estimator (**Figure 4**). However, within follow-up, their deviation was generally below the Kaplan-Meier estimator for NIVO + IPI and above for Chemo

Table 1. Comparison of information criteria and cure fractions between 3-year and 5-year follow-up

Distribution of uncured	3-year follow-up				5-year follow-up			
			Cure fraction, % (95% CI)				Cure fraction, % (95% CI)	
	AIC	BIC	NIVO + IPI	Chemo	AIC	BIC	NIVO + IPI	Chemo
Exponential	7277.28	7294.76	**27.6** (22.7, 32.5)	**12.0** (7.5, 16.7)	8144.64	8162.11	**21.6** (17.7, 25.8)	**10.2** (7.0, 13.4)
Weibull	7254.23	7280.44	**29.1** (24.0, 34.0)	**17.3** (13.6, 21.1)	8136.88	8163.09	**21.2** (16.8, 25.7)	**11.7** (8.5, 14.9)
Gamma	7250.83	7277.04	**28.9** (24.0, 33.6)	**16.5** (12.8, 20.4)	8131.83	8158.04	**21.6** (17.4, 26.0)	**11.6** (8.5, 14.8)
Gompertz	7270.35	7296.56	**28.6** (12.6, 34.2)	**18.1** (14.3, 21.8)	8144.14	8170.35	**11.2** (2.5, 22.8)	**10.9** (6.4, 14.3)
Log-logistic	**7250.52**	**7276.73**	**16.6** (7.7, 24.3)	**7.5** (1.7, 13.3)	**8112.10**	**8138.31**	**11.0** (4.2, 17.5)	**5.3** (1.0, 9.4)
Lognormal	7257.79	7283.99	**11.1** (0.1, 21.8)	**5.0** (0.0, 12.6)	8117.22	8143.43	**8.2** (0.1, 16.3)	**5.1** (0.1, 9.7)
Gen. Gamma	7251.37	7286.31	**23.2** (0.2, 32.1)	**15.1** (8.4, 19.9)	8115.30	8150.25	**13.0** (0.1, 21.9)	**9.3** (4.0, 13.2)

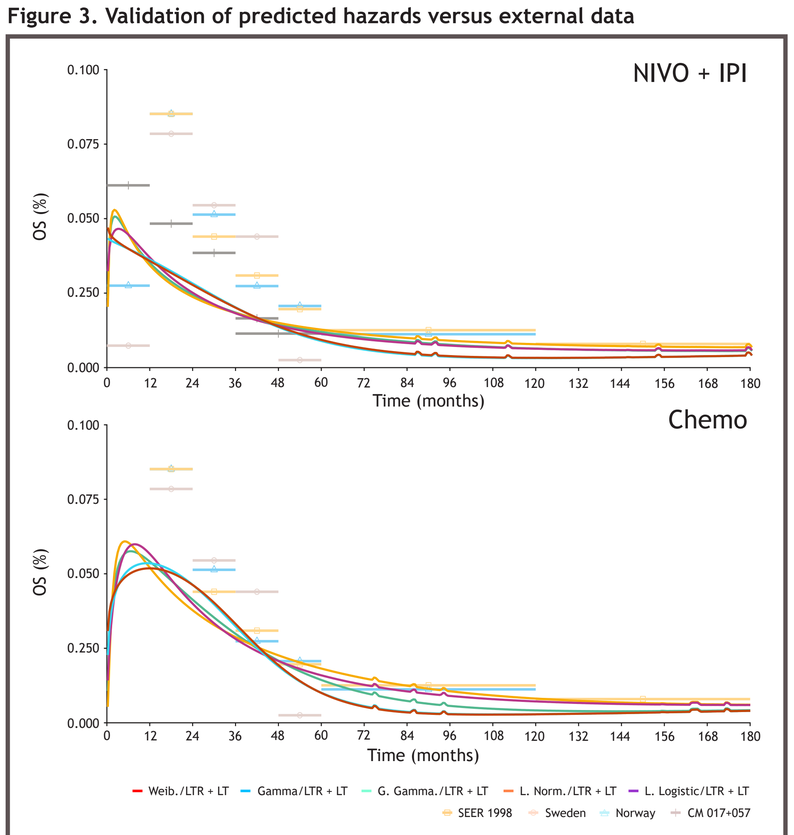
Minimum AIC/BIC; >10 units from minimum AIC/BIC model.
AIC: Akaike information criterion; BIC: Bayesian information criterion; Chemo: platinum doublet chemotherapy; Gen. Gamma: Generalised Gamma; IPI: ipilimumab; NIVO: nivolumab.

Table 2. Comparison of information criteria and cure fractions between MCM types with 5-year follow-up

Distribution of uncured	Independent models				Independent cure and location				Independent cure only			
			Cure fraction, %				Cure fraction, %				Cure fraction, %	
	AIC	BIC	NIVO + IPI	Chemo	AIC	BIC	NIVO + IPI	Chemo	AIC	BIC	NIVO + IPI	Chemo
Exponential	8144.64	8162.11	21.6	10.2	8144.64	8164.88	21.6	10.2	8143.34	8158.52	22.1	9.9
Weibull	8136.88	8163.09	21.2	11.7	8142.41	8167.71	22.5	11.0	8141.26	8161.50	22.9	10.7
Gamma	8131.83	8158.04	21.6	11.6	8138.48	8163.78	22.6	11.0	8137.19	8157.44	23.0	10.7
Gompertz	8144.14	8170.35	11.2	10.9	8145.90	8171.20	20.3	9.3	8144.44	8164.68	20.9	8.5
Log-logistic	**8112.10**	**8138.31**	11.0	5.3	**8120.82**	**8146.12**	15.5	1.9	**8118.86**	**8139.11**	15.4	2.0
Lognormal	8117.22	8143.43	8.2	5.1	8127.29	8152.59	14.9	0.1	8126.43	8146.67	14.6	1.9
Gen. Gamma	8115.30	8150.25	13.0	9.3	8123.15	8153.52	19.6	6.7	8121.17	8146.48	19.5	6.9

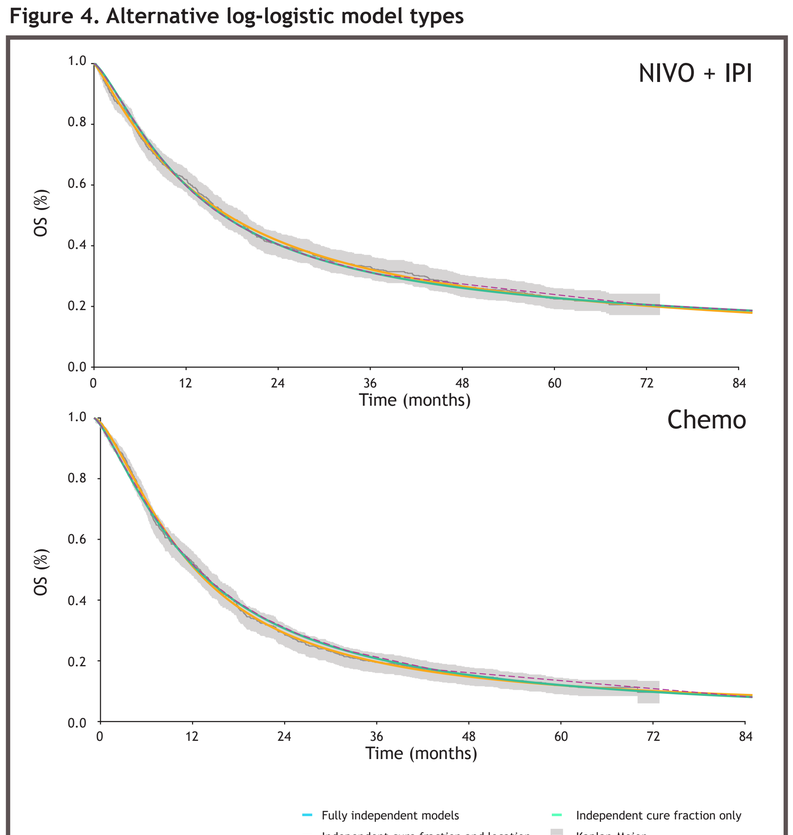
Minimum AIC/BIC; >10 units from minimum AIC/BIC model.
AIC: Akaike information criterion; BIC: Bayesian information criterion; Chemo: platinum doublet chemotherapy; Gen. Gamma: Generalised Gamma; IPI: ipilimumab; NIVO: nivolumab.

Figure 3. Validation of predicted hazards versus external data



CM 017+057: Pooled CheckMate 017 and 057 nivolumab arms; Sweden: Swedish registry (NSCLC Stage IV); Norway: Norwegian registry (LC Stage IV); SEER 1998: Surveillance, Epidemiology, and End Results Program (NSCLC Satage IIIB-IV)
Chemo: platinum doublet chemotherapy; IPI: ipilimumab; L. Logistic: log-logistic; L. Norm: log-normal; LC, lung cancer; LT: lifetable hazard; LTR, long-term response; NIVO: nivolumab; NSCLC, non-small cell lung cancer; OS: overall survival, Weib.: Weibull.

Figure 4. Alternative log-logistic model types



Chemo: platinum doublet chemotherapy; IPI: ipilimumab; NIVO: nivolumab; OS: overall survival.

Discussion

- MCM of 5 year OS data is supportive of modelling a larger cure fraction for patients treated with NIVO + IPI than Chemo
- Additional follow-up improved precision of model estimates, in part through more discrimination between model goodness-of-fit allowing for exclusion of a larger number of models
  - Limited difference in AIC/BIC at 3 years to aid model selection
  - Greater differences at 5-years resulted in more models excluded
  - Improved methods to select mixture-cure models with immature data could aid in early benefits estimation; research into this is ongoing
  - Selected models plausibly bounded by historical hazards and general population hazards
- Goodness of fit is not sensitive to assuming scaled/equal distribution of survival of the uncured between NIVO + IPI and Chemo
  - Joint modelling resulted in increased cure fraction for NIVO + IPI; decreased cure fraction for Chemo, when compared to fully independent models
  - As goodness of fit of lower degrees-of-freedom joint models was not markedly worse than independent models, these models may be used for treatment effect estimation in indirect treatment comparison

Conclusions

- 5-year data have confirmed the OS advantage of NIVO + IPI versus Chemo
- MCMs incorporating this additional follow-up are confirmatory of 3-year analysis and provide improved precision and higher certainty of the expected long-term survival with NIVO + IPI driven by a higher 'cure' fraction
- Research into optimal MCM selection from less mature data would be beneficial in early estimation of long-term benefits from ICIs

References

1. Hellmann, MD, et al. NEJM 2019; 381 #:2020-2031
2. Yong, Y, et al. ViH 2022; 25(7) #:S531
3. Chen, TT. J. immunotherapy cancer 2013; 1 #:18
4. Motzer, RJ, et al. Cancer 2022; 128(11); 2085-2097
5. Wolchok, J.D, et al. JCO 2021; 39(15); abstr 9506
6. Gettinger, SN, et al. JCO 2015; 33(18):2004-2012
7. Burnham KP, Anderson DR, SMR 2004;33:261-304

Disclosures

- This research was supported by Bristol Myers Squibb.

Presented at ISPOR Europe 2022; November 6-9, 2022; Vienna, Austria, and virtual

Email: rob.young@heor.co.uk