Pembrolizumab for adjuvant treatment of resected non-small-cell lung cancer [ID3907]

For presentation/zoom – confidential information redacted

Technology appraisal committee D – Second committee meeting [3 October 2024]

Chair: Dr Raju Reddy

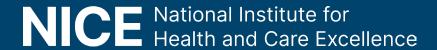
External assessment group: BMJ

Technical team: Lauren Elston, Samuel Slayen, Ian Watson

Company: Merck Sharp & Dohme

Pembrolizumab for treating adjuvant treatment of resected non-small-cell lung cancer

- ✓ Preliminary recommendations and conclusions (DG recap)
- Consultation responses
- Company response and EAG critique
- Other considerations
- Summary



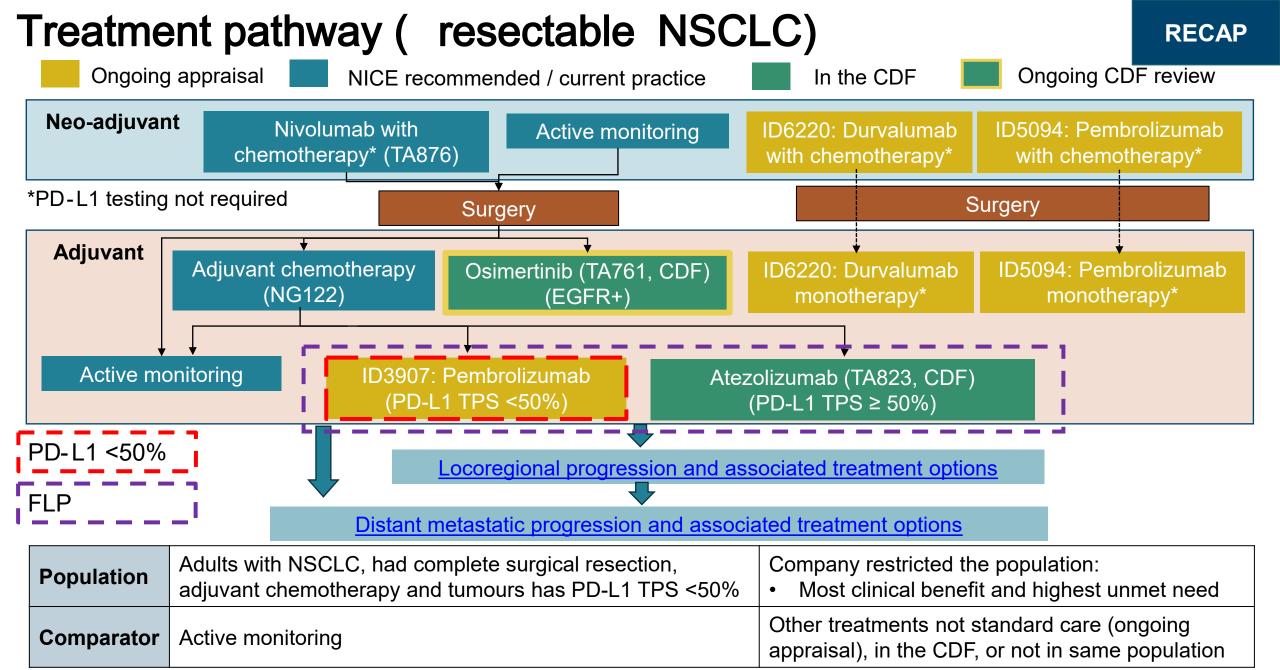
DG: preliminary recommendations

Pembrolizumab is not recommended, within its marketing authorisation, for adjuvant treatment of non-small-cell lung cancer (NSCLC) that is at high risk of recurrence after complete resection and platinum-based chemotherapy in adults.

Committee made this decision as it could not establish a preferred cost — effectiveness estimate and threshold as it had not seen an analysis in the full licensed population. It considered that there was uncertainty around the justification for positioning adjuvant pembrolizumab in a different subgroup population to the marketing authorisation (adults whose tumours express the PD — -L1 biomarker on less than 50% of tumour cells) and around the clinical benefit of pembrolizumab in that population.

Consultation responses have been received from:

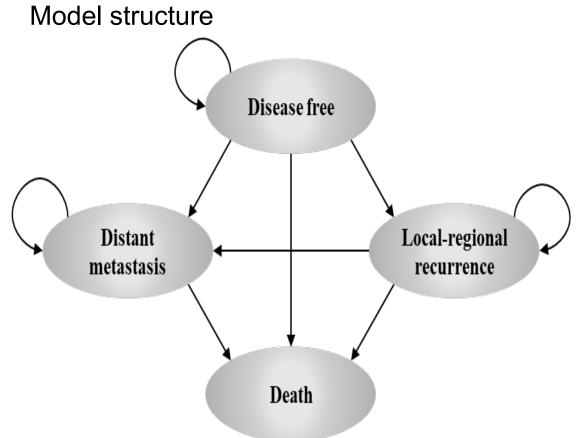
- Roy Castle Lung Cancer Foundation
- British Thoracic Oncology Group
- Merck Sharp & Dohme



NICE

Abbreviations: NSCLC, non-small-cell lung cancer; CDF, cancer drugs fund; PD-L1, programmed death-ligand 1; TPS, tumour proportion score; EGFR, epidermal growth factor receptor; FLP, full licensed population

Company's model overview



- Evidence informing the model
- Impacts of pembrolizumab on costs, QALYs, ICERs

Markov structure implies modelled survival is primarily a function of DFS

- Improved DFS results in improved OS
- OS is the primary driver of modelled benefits

Cure assumption

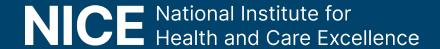
- A cure assumption was applied among people who achieve long-term DFS
- Base case assumes a cure point of 5 7 years.
- After 5 years, risk of recurrence reduces linearly to maximum of 95%
- Modelled patients therefore have a long-term residual risk of recurrence
- Proportion of people reaching the cure point, as determined by DFS curve, is the primary factor influencing the magnitude of incremental QALY benefits

Key issues from ACM1

Issue	Resolved?	ICER impact
Relevant population (PD-L1<50% or full licensed population)	No - for discussion	Large
DFS model selection and treatment effect waning	No - for discussion	Large
Mortality in the cured population and calibration of cure	No - for discussion	Moderate
Validation and calibration of modelled OS to trial OS	No - for discussion	Unknown
Model baseline age	Yes (67 years)	Moderate

Pembrolizumab for treating adjuvant treatment of resected non-small-cell lung cancer

- Preliminary recommendations and conclusions (DG recap)
- ✓ Consultation responses
- Company response and EAG critique
- Other considerations
- □ Summary



Consultation response summary [1/2]

Comments received from BTOG and RCLCF, comments and additional evidence received from MSD

Issue (DG section)	Consultee response summary	EAG comments summary
Population (PD-L1 <50% vs FLP)	RCLCF:- It would be an advantage to patients in the PD -L1 ≥50% subgroup to have pembrolizumab 6 weekly over atezolizumab 3/4 weekly. BTOG:- PD-L1<50% subgroup is larger, misleading to suggest it is small. - Assumptions on PD-L1 and subgroup designation are from advanced/metastatic disease and applied to operable disease. MSD:- PD-L1<50% group was as expected but the PD-L1≥50% group performed worse than expected (possible sampling bias). - Company positioned in the PD-L1<50% subgroup to increase certainty and applicability to UK clinical practice - Sample size of subgroup relatively large, reduces risk of chance findings.	- PD-L1 < 50% focus may be driven by data over biological/ clinical plausibility No evidence of differences in DFS HRs across FLP and < 50% subgroups Sampling bias cannot explain differences.
DFS curve selection and treatment effect waning	 MSD: - The model does capture treatment effect waning. Equalisation of hazards occurs due to the cure assumption. - EAG ACM1 curve selection (PD-L1<50%) has clinically implausible hazard ratio (models benefit for placebo for most of time horizon). - Have provided an additional treatment effect waning mechanism in model which can be applied to force HR to 1. 	 EAG issue was more on fit. Adopted company alternative base case curves, plus additional scenario.

NICE

Abbreviations: BTOG, British Thoracic Oncology Group; DFS, disease free survival; EAG, external assessment group; FLP, full licensed population; ICER, incremental cost-effectiveness ratio; MSD, Merck Sharpe & Dohme; OS, overall survival; PD-L1, programmed death ligand 1; RCLCF, Roy Castle Lung Cancer Foundation.

Consultation response summary [1/2]

Issue (DG section)	Consultee response summary	EAG comments summary
Cure	 MSD - SMR has been added as requested although considered to be double counting Scenarios exploring calibration of ultra - late recurrence to external literature explored Mixture cure modelling not compatible with competing risks structure of model. Additional experimental survival analysis on DFS not necessary. 	 Agrees on "SMR double counting". Recommends adjusted cure rate to ensure 10-20 year recurrences match literature.
Validation and calibration of modelled OS to KEYNOTE- 091 OS	 MSD – Provided visual overlays of post -calibration OS compared to KEYNOTE-091 OS Model not programmed to include time -varying transitions from LR and DM (too complex) no obvious advantages as LR is similar between arms so would not drive model differences Sensitivity analysis exploring cure in LR health state indicate that time dependent probabilities in LR would favour pembrolizumab (especially when OS calibration applied) For DM patients there is no strong evidence that exponential distributions are inaccurate 	- LR: company scenarios limit the uncertainty - DM: less certain, small differences in selected models does not mean small impact. Extrapolation in squamous trial has large effect.

Abbreviations: DFS, disease free survival; DM, distant metastases; EAG, external assessment group; LR, locoregional recurrence; MSD, Merck Sharpe & Dohme; OS, overall survival.

Link to full consultation responses

Summary of new evidence

Additional and updated evidence includes:

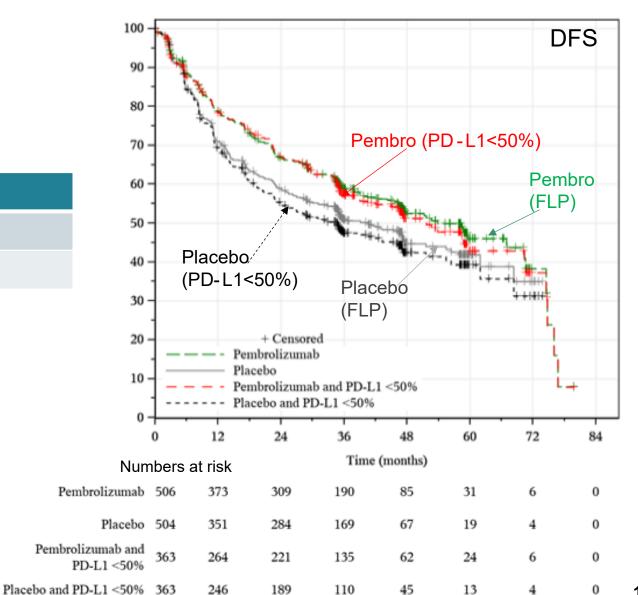
- Additional model supplied in which the survival curve parameters, time on treatment and subsequent treatments have been adjusted to reflect the full licensed population
- Treatment waning functionality added
- Sensitivity analysis examining possibility for cure in LR state added
- Baseline age changed to 67
- SMR of 1.453 added to cured patients (derived from <u>Janssen-Heijen et al., 2012</u>)
- Sensitivity analyses relating to cure proportions and long -term recurrence conducted



Updated clinical evidence (DFS)

Pembrolizumab improves DFS compared to placebo in both populations (FLP and PD-L1 <50%)

Population	DFS HR (95% CI)
Full licensed population (FLP)	0.76 (0.64 to 0.91)
PD-L1 <50%	0.72 (0.58 to 0.89)

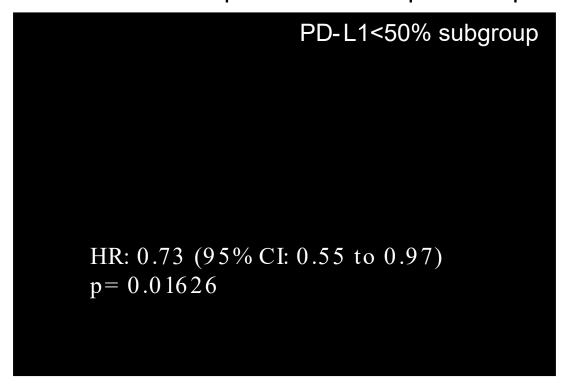


Abbreviations: DFS, disease free survival; FLP, full licensed population; HR, hazard ratio; PD -L1, programmed death ligand 1.

NICE

Updated clinical evidence (OS)

Pembrolizumab improves OS compared to placebo in both populations (FLP and PD-L1 <50%)



90 -				·-\+-	-+-/		The state of the s	banna.	Į	Pem	hrol	izur	nah		
80 —						H-1-46	**************************************			No.			пио		
70					Pl	ace	bo			······································		^{-†} ###### 	 	10 	
-												, 18-18-18-18-18-18-18-18-18-18-18-18-18-1	···· · · · · · · · · · · · · · · · · ·	+++++	
60 —															
-															
50 —															
40 —															
-															
30 —	HI	R:	0.7	9 (95	%	CI	:0	.62	to	1.0	1)			
20					`				_			-)			
20	p=	=U	.03	22	4										
10 —															
-															
0															

OS rate at:	Pembro (95% CI)	Placebo (95% CI)
24 months		
36 months		
60 months		

OS rate at:	Pembro (95% CI)	Placebo (95% CI)
24 months		
36 months		
60 months		

Key Issue – Relevant population [1/3]

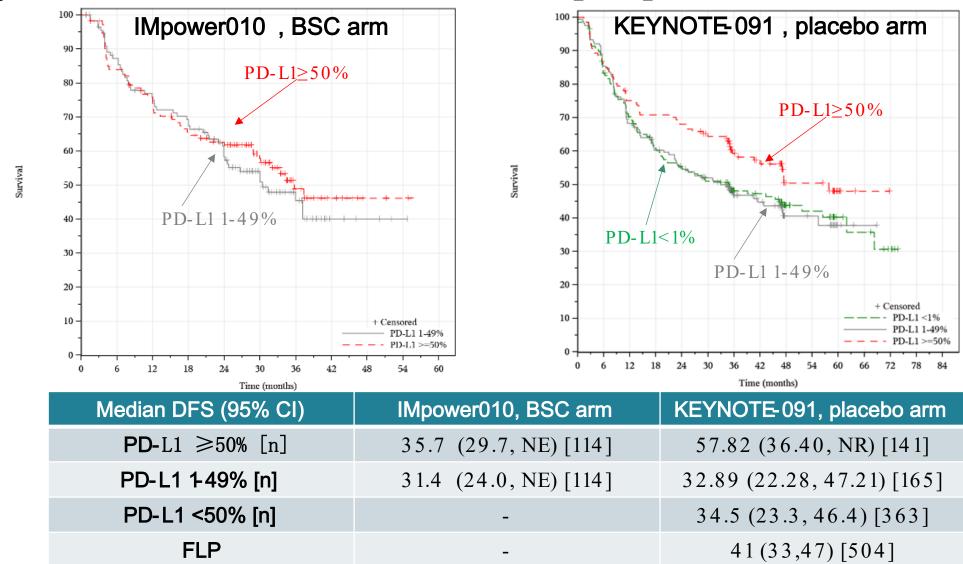
Background (ACM1 committee conclusions)

- Unmet need in entire population, subgroups should be based on known biologically plausible mechanisms or other clear factors (NICE manual) and decision to restrict population was not.
- PD-L1 <50% subgroup results not convincing due to post-hoc nature (results could be due to chance).
- Committee requested analyses with the full licensed population.

Company (DG response)

- Provided additional model using full licensed population (FLP).
- Not seeking reimbursement for the ≥50% subgroup excluded from original submission to increase the certainty and applicability of the results to UK clinical practice.
- New KEYNOTE-091 data suggests overperformance of the placebo in the PD-L1 ≥50% subgroup, (not observed in the IMpower010 study PD-L1 ≥50% subgroup)
- Overperformance in the PD-L1 ≥50% subgroup (median DFS is two years longer than <50% subgroup) could be due to sampling error (i.e. more people "cured at resection" than the 30-40% expected).
- DG implication that PD-L1 <50% results unexpected inaccurate, may lead to inaccurate interpretation about effectiveness of pembrolizumab in this group. PD-L1 ≥50% subgroup did have unexpected results (i.e. pembrolizumab was comparatively less effective than expected in this arm).
- Meta-analysis suggests that PD-L1 expression has negative prognostic value in stage 1 to 3 disease. The results of the PD-L1≥50% subgroup should not be considered biologically plausible
- PD-L1<50% subgroup (n=726), larger sample than trials from NICE NSCLC appraisals (TA823, TA876)

Key Issue – Relevant population [2/3]





Key Issue – Relevant population [3/3]

EAG comments

- While PD-L1 <50% subgroup 'relatively large', KN-091 not designed to be powered for PD-L1 <50% TPS; inherently more subject to bias and choice to focus on this 'unexpected and unexplained' result may be data-driven rather than driven by biological or clinical plausibility (regardless of result validity).
- DFS consistent across FLP / PD-L1 subgroups (<50%/1-49%/<1%); no robust evidence to show differences.
- Overperformance of placebo in PD-L1 ≥50% subgroup not conclusively shown; could be due to other factors.
- Sampling bias cannot explain differences in placebo results; randomisation should ensure balanced populations. Sampling bias causing placebo overperformance in PD-L1 ≥50% subgroup would cause placebo "underperformance" in PD-L1 <50% subgroup. In this scenario, ITT may be the only robust DFS estimate.
- Uncertainty around PD-L1 prognostic meta-analysis, that not all studies in stage subgroups focus on NSCLC
- No routinely commissioned treatments available for either PD-L1 <50% or PD-L1 ≥50%.

Other considerations (DG consultee responses)

- Unmet need for PD-L1 < 50%, as no other adjuvant immunotherapies for this population.
- **BTOG:** PD-L1 <50% subgroup was 726/1,117 (65%) participants in KN -091; use of 'narrower' misleading. Evidence was not reviewed in context of how and why MSD focused on the PD -L1 <50% subgroup.
- RCLCF: Pembrolizumab would have an advantage over atezolizumab in the ≥50% PD-L1 group due to longer dosing window (Q6W vs Q4W); would expect to see a greater effect in patients with higher PD-L1 expression



Which population is most appropriate for decision-making?

NICE

Key Issue – DFS and treatment effect waning

Background (ACM1)

- **Company:** applied sustained treatment effect for lifetime of model; same parametric model applied for DFS transitions to LR and DM (log-normal). Company stated that applying different curves was inappropriate.
- **EAG:** evidence of treatment effect waning in trial, best fitting parametric curves should be used per arm
- **Committee:** EAG-preferred curves fitted better, but not enough evidence to deviate from <u>TSD14</u> (using the same parametric model in both arms); also not enough evidence to select company-proposed models.
- Committee requested DFS modelled using FLP, also noted that selected curves should reflect cure calibration

Company (DG response)

- Treatment effect waning already implicit in model (hazards converge to 1 soon after follow-up); DG is inaccurate saying treatment effect waning not accounted for.
- Updated model to include explicit treatment effect waning mechanism; minimal effect on company base case, but EAG ACM1 approach not clinically plausible
- Base case and alternative for PD-L1<50% at ACM1. Curve choice for FLP uncertain with 3 possible choices

EAG comments

- Issue was more on fit: company base case appears to have good match for pembrolizumab, but not placebo.
- Following committee conclusion from ACM1 (and TSD14), adopted alternative company base case curves; partially resolves the issue but still a poor fit for the final years in placebo arm.
- Includes scenario with best fitting curves, by MSE, for placebo; results in similar issues but for pembrolizumab

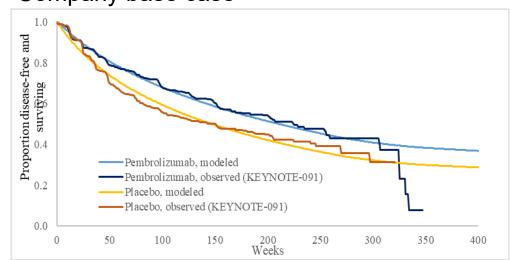
Key Issue – DFS and treatment effect waning - PD-L1<50%

Summary of preferred parametric curves for PD -L1 <50%:

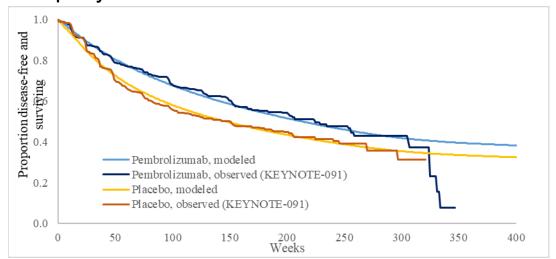
Arm	DF to LR		DF to DM		
	Pembrolizumab	Active monitoring	Pembrolizumab	Active monitoring	
Company base case	Log normal	Log normal	Log normal	Log normal	
Company alternative (ACM1) / EAG base case (ACM2)	Gen. gamma	Gen. gamma	Log normal	Log normal	

See Tables 6 and 7 of DG response appendix for curve fits and predicted DFS and OS

Company base case



Company alternative / EAG base case



DFS graphs lifetime horizon

OS graphs

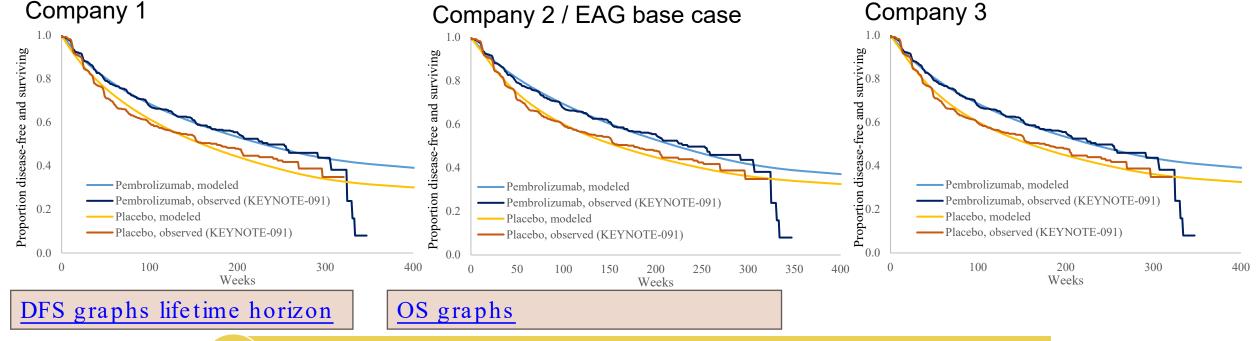


Key Issue – DFS and treatment effect waning – FLP [1/2]

Summary of preferred parametric curves:

Population		DF to LR		DF to DM	
	Arm	Pembrolizumab	Active monitoring	Pembrolizumab	Active monitoring
Full licence	Company 1 (pembro fit)	Log normal	Log normal	Log normal	Log normal
•	Company 2 EAG BC(placebo fit)	Gen. gamma	Gen. gamma	Log normal	Log normal
(FLP)	Company 3 (separate fits)	Log normal	Gen gamma	Log normal	Log normal
	EAG scenario	Gen. gamma	Gen. gamma	Gompertz	Gompertz

See Tables 6 and 7 of company DG response appendix for curve fits and predicted DFS and OS Please note all graphs below show curve choice only and use the 95% cure proportion



What distributions does the committee prefer to use to extrapolate DFS?

NICE

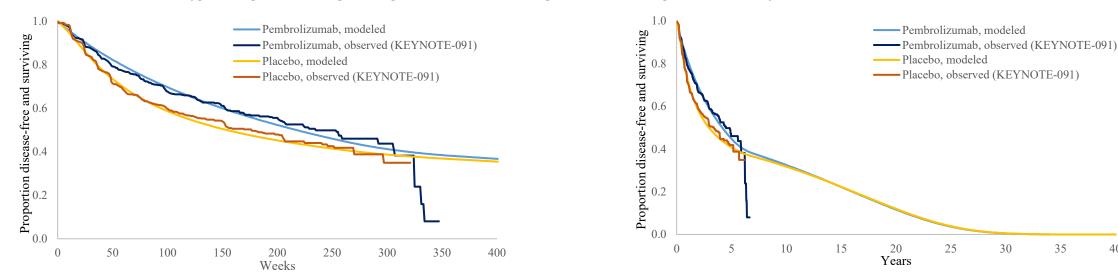
Abbreviations: ACM, appraisal committee meeting; DF, disease free; DM, distant metastases; EAG, external assessment group; gen. gamma, generalised gamma; LR, locoregional recurrence PD - L1, programmed cell death ligand 1.

Key Issue – DFS and treatment effect waning





EAG scenario (gen. gamma/gen. gamma and gompertz /gompertz)



Note: EAG and company disagreed on best choice for placebo DF to DM. Company excluded Gompertz as it allows 0 DM recurrences in the placebo arm and due to crossing DFS curves. EAG considered curves did not cross and that pembrolizumab had higher DFS than placebo across the model time horizon. However, they applied a rule in the model where DFS for pembrolizumab cannot fall below placebo to avoid potentially implausible results.

What distributions does the committee prefer to use to extrapolate DFS?

Key Issue – Modelling of cure

Background (ACM1 committee conclusions)

- 95% cure proportion at 7 years (increasing from 0% at 5 years) appropriate but wanted to see exploration of SMR to ensure mortality in cured portion aligned with clinical opinion
- Requested calibration of model so ultra late recurrence aligned with NSCLC literature (0.8% Sonada et al*)
- Would have liked to see mixture cure models explored

Company (DG response)

- PD-L1<50% base case broadly aligned. Scenarios explored calibrating cure in both populations to 0.8%.
- Sonada mean age is 64 and resections done 20 years ago, 0.8% estimate could be too high and is uncertain
- EAG ACM1 scenario with 75% cure proportion not clinically plausible due to crossing DFS curves
- Applied SMR but used 1.453 in line with literature-based estimate. Considers there is double counting as excess mortality of recurred population already modelled
- Validatory MCMs provided at (clarification) but not compatible with competing risks structure of the model.

EAG comments

- Company ACM1 base case was not calibrated using SMR 1.453 and baseline age of 67.
- Recommends using adjusted cure rate: 92.5% for PD-L1 <50% and 93% for FLP for company 1 (pembro fit) to ensure 10-20 year recurrences match literature.
- Agrees SMR is double counting and should be upper estimate. However, included in base case as no other appropriate value available.

Abbreviations: DFS, disease free survival; DG, draft guidance; EAG, external assessment group; FLP, full licensed population; MCM, mixture cure model; PD -L1, programmed death ligand 1; SMR, standardised mortality ratio.

NICE *Sonoda D, Matsuura Y, Ichinose J et al. Ultra-late recurrence of non -small cell lung cancer over 10 years after curative resection. 20

Cancer Manag Res. 2019 Jul

Key Issue – Modelling of cure

Population	Model Selection	10-20y incidence pembro	10-20y incidence placebo
PD-L1 <50%	Company original base case	0.77%	0.73%
	Age 67 and SM	R 1.453 applied	
PD-L1 <50%	Company base case	0.59%	0.55%
PD-L1 <50%	Company alternative	0.53%	0.37%
PD-L1 <50%	EAG base case	0.95%	0.13%
	Company base case and		
PD-L1 <50%	Cure=92.5%	0.86%	0.80%
	Company alternative and		
PD-L1 <50%	Cure=88.8%	1.20%	0.80%
FLP	Company 1 (best fitting pembro)	0.52%	0.57%
FLP	Company 2 (best fitting placebo)	0.84%	0.38%
FLP	Company 3 (best fitting both)	0.52%	0.38%
FLP	Best fitting pembro and Cure = 93%	0.73%	0.80%
	Best fitting placebo and Cure		
FLP	=89.3%	1.70%	0.80%
FLP	Best fitting both and Cure=89.3%	1.10%	0.80%

EAG: Incidence of ultra -late recurrence should match an external literature estimate for NSCLC. et al. (2019) study reported a 0.8% ultra -late recurrence incidence.

Sonoda

Which cure proportion is appropriate to use? Is the modelling of mortality in the cured population appropriate?

F

NICE Abbreviations: EAG, external assessment group; FLP, full licensed population; NSCLC, non programmed death ligand 1; SMR, standardised -mortality ratio.

Key Issue – Validation and calibration of modelled OS [1/3]

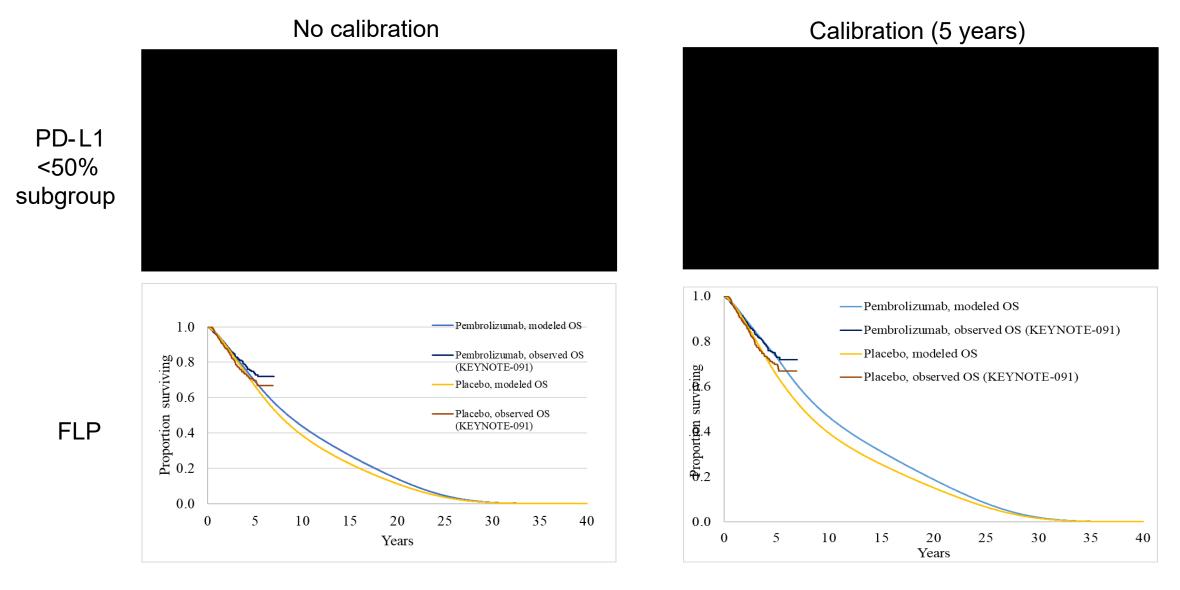
Background (ACM1 committee conclusions)

- Modelled OS data lower than OS data from KN-091. Company consider a post-progression benefit.
- Committee not satisfied with company calibration of OS (applying single modifier to transitions out of LR and DM to make modelled OS fit KN-091 OS better
- Requested additional analyses to validate OS including: visual fits of post-calibration extrapolations, making
 post-recurrence cure and transitions time-dependent and applying modifiable risk ratios to transitions from LR
 and DM to death

Company (DG response)

- Accepted EAG advice to end calibration at 5 years to reflect end of KN-091 trial and have included graphs
 overlaying non-calibrated and calibrated OS with KN-091 results (see next slide).
- Modifiable risk ratios already present in the model.
- Time-dependent transitions out of LR and DM computationally complex, insufficient time in consultation.
- Do not consider that modelling them would substantially change the cost-effectiveness estimates because incidence of LR similar between arms so changing modelling would have limited effect. Also, sensitivity analyses demonstrate that modelling cure in LR state would benefit pembrolizumab (post progression benefit).
- No evidence to suggest the exponential distributions used to extrapolate survival in DM are inaccurate. In scenarios where DM is not adjusted, ICER for pembrolizumab falls.

Key Issue – Validation and calibration of modelled OS [2/3]





Key Issue – Validation and calibration of modelled OS [3/3]

EAG comments

- LR Company scenarios exploring cure in LR limit uncertainty for assumptions made about LR state (e.g. omission of time-varying models).
- DM Less clear: company states small difference in models selected, but this is also true for DFS models that have large impact on cost-effectiveness results.
- In KEYNOTE-407 (which represents ~25% of DM patients in the model) curve choice has a significant effect on long term outcomes

OS of pembrolizumab + chemo in PD-L1 <50%, KN-189

OS of pembrolizumab + chemo in PD-L1 <50%, KN-407



Company and EAG assumptions and results summary

PD-L1 <50% subgroup

Assumption	Company base case	EAG base case	
DFS curves (DF-LR) / (DF-DM) Both arms: log-normal / log-normal		Both arms: gen. gamma / log normal	
SMR	Applied (considered an upper estimate)	Applied (agree as an upper estimate)	
Cure (placebo ULR)	95% (0.55%)	91% (0.75%)	
OS calibration / extrapolation	LR/DM capping 5 years (temporary) / Exponential	LR/DM capping 5 years (temporary) / Exponential	

Full licensed population

Assumption	Company	EAG base case		
DFS curves (DF-LR) / (DF-DM)	 Both arms: log-normal / log-normal Both arms: gen.gamma / log-normal Pembrolizumab: log-normal / log-normal Placebo: generalised gamma / log normal 	Both arms: gen. gamma / log normal		
SMR	Applied (considered an upper estimate)	Applied (agree as an upper estimate)		
Cure (placebo ULR)	95% (1. 0.57%; 2. 0.38%; 3. 0.38%)	89% (0.8% for company scenario 1)		
OS calibration / extrapolation	LR/DM capping 5 years (temporary) / Exponential	LR/DM capping 5 years (temporary) / Exponential		

Abbreviations: DF, disease free; DFS, disease free survival; DM, distant metastases; EAG, external assessment group; gen. Gamma,

NICE generalised gamma; LR, locoregional recurrence; PD-L1, programmed death ligand 1; SMR, standardised mortality rate; ULR, ultra-late 25

recurrence.

Cost-effectiveness results

Cost effectiveness results cannot be reported here due to presence of confidential discounts for included technologies

PD-L1 <50% subgroup

Company base case ICER is between £20,000 and £30,000 EAG base case ICER is above £30,000

Full licensed population

Company ICERs are above £20,000 for all 3 curve options EAG base case ICER is above £30,000

All results are presented in Part 2 slides for committee consideration

Pembrolizumab for treating adjuvant treatment of resected non-small-cell lung cancer

- Preliminary recommendations and conclusions (DG recap)
- Consultation responses
- Company response and EAG critique
- Other considerations
 - Severity / equality / innovation / uncaptured benefits not raised
- Summary

NICE National Institute for Health and Care Excellence

Pembrolizumab for treating adjuvant treatment of resected non-small-cell lung cancer

- Preliminary recommendations and conclusions (DG recap)
- Consultation responses
- Company response and EAG critique
- Other considerations
- ✓ Summary



Key issues for discussion

Issue	For committee	ICER impact
Relevant population (PD-L1<50% or full licensed population)	Which population is most appropriate for decision-making?	Large
DFS model selection and treatment effect waning	What distributions does the committee prefer to use to extrapolate DFS?	Large
Mortality in the cured population and calibration of cure	Which cure proportion is appropriate to use? Is the modelling of mortality in the cured population appropriate?	Moderate
Validation and calibration of modelled OS to trial OS	Should OS in the model be calibrated to KEYNOTE-091 and if so, how?	Unknown

