

**NATIONAL INSTITUTE FOR HEALTH AND CARE
EXCELLENCE**

Final draft guidance consultation

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

1 Recommendation

- 1.1 Pembrolizumab is recommended, within its marketing authorisation, as an option for the adjuvant treatment of non-small-cell lung cancer (NSCLC) with a high risk of recurrence after complete resection and platinum-based chemotherapy in adults. Pembrolizumab is only recommended if the company provides it according to the commercial arrangement (see [section 2](#)).

Why the committee made this recommendation

Usual treatment for completely resected NSCLC after platinum-based chemotherapy in adults is active monitoring. A completely resected tumour is one that has been fully removed surgically.

Clinical trial evidence shows that, compared with placebo, people with NSCLC who have had complete resection of their tumour and adjuvant (after surgery) pembrolizumab have longer before their cancer gets worse and may live longer.

The most likely cost-effectiveness estimate for pembrolizumab is within the range NICE considers an acceptable use of NHS resources. So, pembrolizumab is recommended.

2 Information about pembrolizumab

Marketing authorisation indication

- 2.1 Pembrolizumab (Keytruda, Merck Sharp & Dohme) is indicated for ‘the adjuvant treatment of adults with non-small cell lung carcinoma who are at high risk of recurrence following complete resection and platinum-based chemotherapy’.

Dosage in the marketing authorisation

- 2.2 The dosage schedule is available in the [summary of product characteristics for pembrolizumab](#).

Price

- 2.3 The list price is £2,630 per 100-mg vial (excluding VAT; BNF online accessed October 2024).
- 2.4 The company has a commercial arrangement. This makes pembrolizumab available to the NHS with a discount. The size of the discount is commercial in confidence.

3 Committee discussion

The [evaluation committee](#) considered evidence submitted by Merck Sharp & Dohme, a review of this submission by the external assessment group (EAG), and responses from stakeholders. See the [committee papers](#) for full details of the evidence.

The condition

Treatment options

- 3.1 Non-small-cell lung cancer (NSCLC) is staged from 1A to 4B according to the size and extent of the tumour, location of involved lymph nodes and the presence of distant metastases. This is based on the American Joint Committee on Cancer eighth edition staging system. Resectable NSCLC is usually considered to be early to locally advanced cancer (stage 1A to

3B). Standard care for people with resectable NSCLC is complete surgical resection. Surgery can cure the cancer, but recurrence is common and can either be local-regional (within the lungs and nearby lymph nodes) or distant metastatic (other part of the body). Before surgery, people have the option of neoadjuvant therapy or active monitoring. After complete surgical resection, people have the following options:

- active monitoring
- osimertinib, which is available through the Cancer Drugs Fund (CDF) for people with epidermal growth factor receptor (EGFR) mutation-positive NSCLC (see [NICE's technology appraisal guidance on osimertinib](#))
- adjuvant chemotherapy
- adjuvant chemotherapy followed by maintenance treatment with atezolizumab, which is available through the CDF for people with NSCLC whose tumours express the biomarker PD-L1 on 50% or more of their tumour cells (from now on referred to as PD-L1 tumour proportion score [TPS] 50% or more; see [NICE's technology appraisal guidance on atezolizumab](#)).

Clinical expert submissions stated that the aim of adjuvant treatment is to reduce the risk of recurrence after surgery for people with potentially curable NSCLC. The committee considered data presented that showed that 41% of people with stage 1 to 3 lung cancer with complete resection develop recurrence within 23 months. The patient organisation submission reported that recurrence of NSCLC after surgery usually means that further curative treatment is unlikely. It explained that the only way to tell if surgery has been curative is to wait, and this results in continual anxiety for people with lung cancer and their families and carers. The company proposed adjuvant pembrolizumab for NSCLC in adults with a high risk of recurrence after complete resection and platinum-based adjuvant chemotherapy, only if their tumours have a PD-L1 TPS less than 50%. The committee

understood that there are no other immunotherapy treatment options available at this point in the treatment pathway. The patient organisation submission stated that there is an ongoing need to develop additional treatments that would reduce the risk of recurrence. The committee concluded that there was an unmet need for a treatment that reduces the risk of recurrence after complete resection.

Comparators

3.2 The company compared adjuvant pembrolizumab with active monitoring. The final scope for this evaluation also included:

- platinum doublet chemotherapy
- perioperative durvalumab (durvalumab with chemotherapy before surgery [neoadjuvant] then alone after surgery [adjuvant]; subject to NICE appraisal)
- adjuvant osimertinib (subject to NICE appraisal) and
- adjuvant atezolizumab (subject to NICE appraisal).

The company explained that people eligible for adjuvant pembrolizumab would have had adjuvant chemotherapy and so platinum doublet chemotherapy was not a relevant comparator. It stated that atezolizumab, osimertinib and perioperative durvalumab were not recommended in routine practice. It stated that even if they were, osimertinib and perioperative durvalumab would not be used in the same population as adjuvant pembrolizumab. It also stated that because of the proposed population of people with a PDL1 TPS less than 50% (see [section 3.1](#)), atezolizumab would not be considered a relevant comparator. So, the company had decided there was no alternative to pembrolizumab for this population and the relevant comparator was active monitoring. The committee agreed that adjuvant pembrolizumab would be used after platinum doublet chemotherapy, as specified in its marketing authorisation, so platinum doublet chemotherapy is not an appropriate comparator. Durvalumab would be suitable for a different population

(people with resectable NSCLC) and is not routinely commissioned. Osimertinib and atezolizumab are also not routinely commissioned because they are only available within the CDF. So, the committee agreed that active monitoring was the relevant comparator for pembrolizumab.

Clinical effectiveness

Clinical trial evidence

3.3 The clinical evidence for adjuvant pembrolizumab came from KEYNOTE 091, a phase 3 randomised controlled trial. This compared adjuvant pembrolizumab (200 mg every 3 weeks for 1 year) with placebo in adults with stage 1B (tumour size equal to or more than 4 cm), 2 or 3A NSCLC after complete surgical resection and with or without adjuvant chemotherapy. The primary outcome measure was disease-free survival (DFS). A key secondary outcome measure was overall survival (OS). The trial stratification factors included the use of previous adjuvant chemotherapy and PD L1 status (with TPS less than 1%; between 1% and 49%; and more than 50%). The population in the marketing authorisation is adults who had previous adjuvant chemotherapy (see [section 2.1](#)). In line with its proposed population, the company provided post hoc subgroup results from a subpopulation of adults who had previous adjuvant chemotherapy and whose tumours had a PD L1 TPS less than 50%. The company provided results from an interim analysis with a data cutoff date of January 2023 for the:

- previous adjuvant chemotherapy population (full licensed population; pembrolizumab n=506; placebo n=504)
- PD-L1 TPS less than 50% population (the company's proposed population; pembrolizumab n=363; placebo n=363).

Adjuvant pembrolizumab was associated with a statistically significant improvement in DFS compared with placebo for both populations. The greatest benefit was seen in the PD-L1 TPS less than 50% population:

- previous adjuvant chemotherapy population (full licensed population): hazard ratio 0.76 (95% confidence interval [CI] 0.64 to 0.91)
- PD-L1 TPS less than 50% population: hazard ratio 0.72 (95% CI 0.58 to 0.89).

Adjuvant pembrolizumab was also associated with improved OS compared with placebo for both populations, and the greatest benefit was again seen in the PD-L1 TPS less than 50% population:

- previous adjuvant chemotherapy population (full licensed population): hazard ratio 0.79 (95% CI 0.62 to 1.01)
- PD-L1 TPS less than 50% population: hazard ratio 0.73 (95% CI 0.55 to 0.97).

PD-L1 TPS less than 50% subgroup data

3.4 At the first meeting, the company proposed pembrolizumab as an adjuvant treatment for adults with NSCLC after complete surgical resection and adjuvant chemotherapy, and whose tumours have a PD-L1 TPS less than 50% (see [section 3.1](#)). This was a narrower population than in the [NICE final scope](#) and the marketing authorisation (see [section 2.1](#)). The company explained that this positioning of pembrolizumab is consistent with the clinical trial results, in which this subpopulation had the greatest clinical benefit from adjuvant pembrolizumab compared with placebo (see [section 3.3](#)). But this subgroup was not prespecified in KEYNOTE-091, so the results were from a post hoc analysis. The company stated that this subpopulation has a large unmet need and could benefit most from an additional adjuvant option. The committee acknowledged that there are currently no other adjuvant treatment options for people whose tumours have a PD-L1 TPS less than 50% (see [section 3.1](#)), but also noted there are no routinely commissioned options for people whose tumours have a PD-L1 TPS of 50% or over. The EAG was concerned that the decision to focus on this subgroup was driven by the data rather than biological plausibility. Also, it noted that the data was

from a post hoc analysis and so could be at risk of bias and type 1 error. This means that the data could potentially overestimate the effectiveness of pembrolizumab compared with placebo in this subpopulation. It explained that the smaller sample size of this post hoc subgroup, which was a subpopulation of the prespecified population (see [section 3.3](#)), reduced the power of the analyses. The EAG advised that this can prevent reliable conclusions being drawn and increases the risk that the results are down to chance. The committee considered the clinical-effectiveness results. It questioned why the full licensed population was not the focus of this evaluation, given that adjuvant pembrolizumab was also more effective than placebo in this broader population. The company stated that people whose tumours have a PD-L1 TPS of 50% or more were excluded from the proposed population because there is uncertainty about whether pembrolizumab is more clinically and cost effective compared with atezolizumab. It explained that clinical feedback suggested that pembrolizumab is not expected to become the preferred treatment option over atezolizumab. Clinical expert submissions disagreed, explaining that they expect more people to have adjuvant pembrolizumab. This is because using pembrolizumab as an alternative to atezolizumab for people with a PD-L1 TPS of 50% or more could give people the option of a treatment that is given less frequently, every 6 weeks rather than every 3 to 4 weeks. The committee noted that atezolizumab was recommended through the CDF and is not established practice in the NHS, so it was not considered to be a relevant comparator (see [section 3.2](#)). It understood that after a period of managed access, NICE will review the technology to decide if it can be recommended for routine commissioning.

The committee noted there is an unmet need regardless of the PD-L1 status of people's NSCLC. It also noted that the KEYNOTE-091 results unexpectedly showed that adjuvant pembrolizumab was less effective in the PD-L1 TPS 50% or more subgroup than in the PD-L1 TPS less than

50% subgroup. The company's UK Clinical Advisory Board also noted that the KEYNOTE-091 results for the PD-L1 TPS 50% or more subgroup contradicted clinical expectations. This is because there is established evidence that PD-1 inhibitors, such as pembrolizumab, typically have a greater efficacy in the PD-L1 TPS 50% or more subgroup. This was supported by clinical experts who explained that KEYNOTE-091 was designed with this clinical expectation. This was why the PD-L1 TPS 50% or more subgroup was a stratification factor in the trial and the PD-L1 TPS less than 50% subgroup was not predefined. The EAG's clinical experts agreed with this and noted that the mechanism underpinning greater clinical benefits in the PD-L1 TPS less than 50% subgroup is not yet understood. The company suggested that the reason for these results was because the trial placebo arm in the PD-L1 TPS 50% or more subgroup performed better than expected. But the EAG highlighted that the company had not provided evidence to support this claim, and the placebo arm could have instead underperformed in the PD-L1 TPS less than 50% subgroup. The committee noted that the company's proposed positioning of adjuvant pembrolizumab was in a narrower population than that in the NICE final scope. Also, it considered that the results of the KEYNOTE-091 subgroups could not be clinically explained so could be because of chance. The committee was aware that the [NICE health technology evaluations manual section on analysis of data for patient subgroups](#) states that subgroups should be based on an expectation of differential clinical or cost effectiveness because of known, biologically plausible mechanisms, social characteristics or other clearly justified factors. The committee thought that the company's decision to focus on the PD-L1 TPS less than 50% subgroup was not based on a biologically plausible rationale. Instead, it had been driven by the findings in the clinical trial and potentially the impact this had on the cost effectiveness of pembrolizumab. The committee decided that, given the company and the clinical experts could not fully explain the results from the PD-L1 TPS subgroups, the clinical and cost effectiveness of adjuvant pembrolizumab

in the PD-L1 less than 50% subgroup remains uncertain. The committee was not presented with cost-effectiveness analyses in the licensed population at the first committee meeting. It decided that the justification to restrict pembrolizumab to the PD-L1 TPS less than 50% subgroup because of unmet need was weak when there are no routinely commissioned treatments available in the PD-L1 TPS 50% or more subgroup. It also thought that the evidence from the PD-L1 TPS less than 50% subgroup had relevant limitations, including the post hoc nature of the analysis. It thought that the findings could be a result of chance. The committee concluded that it would like to see an analysis using the full licensed population, in addition to any subgroups that are based on known biologically plausible mechanisms, social characteristics or other clearly justified factors.

Relevant population for decision making

3.5 During consultation on the draft guidance, the company provided updated analyses that modelled the cost effectiveness of adjuvant pembrolizumab in the full licensed population. But the company still felt that the PD-L1 TPS less than 50% subgroup was the most appropriate for decision making. It noted that the results in the PD-L1 50% or more subgroup were unexpected and explained this was because of overperformance of the placebo arm in this subgroup. The company submitted additional evidence that showed that median DFS in the placebo arm of the PD-L1 50% or more subgroup (57.82 months) was numerically higher than the PD-L1 less than 1% (34.5 months) and the PD-L1 1% to 49% (32.89 months) subgroups. It also compared these to the median DFS from the IMpower010 trial, which were 35.7 months for the PD-L1 50% and over subgroup and 31.4 months for the PD-L1 1% to 49% subgroup. The company stated that because the PD-L1 50% or more subgroup was relatively small (see [section 3.3](#)) it was more susceptible to sampling bias, which could have caused the unexpected results. The company noted that between 30% and 40% of people with completely resected NSCLC are cured at surgery, but that it is impossible to know who these people are. It

stated that there could be more people with cured NSCLC in the PD-L1 50% or more subgroup than expected, which would have led to overperformance of the placebo arm of this subgroup. The company stated that it was pursuing reimbursement in the PD-L1 less than 50% subgroup only, to increase the certainty of the results and their applicability to UK clinical practice. The EAG advised that if there were more people with cured NSCLC than expected in the PD-L1 50% or more subgroup, this would also benefit the pembrolizumab arm and would not necessarily explain why it had lower relative effectiveness than the PD-L1 less than 50% subgroup. It also noted that if there were more people with cured NSCLC than expected in the PD-L1 50% or more subgroup, this would likely mean that there were fewer people with cured NSCLC than expected in the PD-L1 less than 50%. This would result in underperformance of the placebo group in that subgroup. The committee acknowledged the company's additional evidence but decided it was implausible that only 1 subgroup of the trial would be affected by sampling bias.

In its response to draft guidance consultation, the professional organisation stated there was an unmet need for the PD-L1 less than 50% population because there were no other adjuvant immunotherapies in this population. The patient group draft guidance consultation response noted that in the PD-L1 50% or more population, pembrolizumab would be perceived by patients as having an advantage over atezolizumab because it allowed a 6-weekly dosing over a 4-weekly dosing regimen. The committee noted these submissions. Because atezolizumab is only available through the CDF, not routine commissioning, it is not a relevant comparator for this appraisal (see [section 3.2](#)).

The committee maintained its decision that there was an unmet need in the whole population (see [section 3.4](#)). It acknowledged there were uncertainties in the evidence for the PD-L1 less than 50% population (see [section 3.4](#)). It emphasised its decision that any subgroup analyses

should be based on biological plausibility or clearly justified factors, and that the company's decision to restrict to the PD-L1 less than 50% population was not based on either of these factors. The committee concluded that the full licensed population was appropriate to use for decision making.

Baseline age

3.6 The mean age of the KEYNOTE-091 overall trial population was 64.3 years, and this was also the starting age in the economic model at the first meeting. The EAG's clinical experts highlighted that this was younger than seen in NHS clinical practice. So, the EAG was concerned about the generalisability of the KEYNOTE-091 trial age to clinical practice and the potential impact of this on the cost-effectiveness results. The EAG's clinical experts also noted that fewer people are likely to be cured in an older population (see [section 3.11](#)). The EAG was also concerned that age may be a treatment effect modifier. The company highlighted that the treatment effect of pembrolizumab did not differ across age groups in the PD-L1 TPS less than 50% DFS subgroup analysis, the company's proposed population. Clinical experts supported this, explaining that they had not seen age appear as an independent prognostic factor in lung cancer. In its base case, the EAG's starting age was 68 years based on registry data from people with NSCLC who had surgery in England in 2012. The company stated that people having adjuvant pembrolizumab must be fit enough to have surgery and to complete adjuvant chemotherapy, so would likely be younger than the average person diagnosed with NSCLC. This was supported by clinical experts who explained that the NHS targeted lung health check programme would likely result in more cases being diagnosed at earlier stages and in younger people. They advised that it was reasonable to expect the average age of people with NSCLC to decrease over time. The committee noted the evidence and decided it was plausible that the mean age would change with time. But it was aware that decision making should be based on data from current clinical practice. The NHS CDF clinical

lead informed the committee that the mean age of people having adjuvant atezolizumab through the CDF in NHS practice is 67 years (see the [NICE technology appraisal guidance on atezolizumab](#)). They expected this to be generalisable to the proposed population in this appraisal. The committee noted that the age of people in KEYNOTE-091 may reflect a younger and fitter population than in NHS clinical practice. It decided that the modelling should reflect NHS clinical practice. At the first meeting, the committee concluded that KEYNOTE-091 had a lower mean age than the potential population in current NHS practice, and that 67 years was the most appropriate age to be modelled. For the second meeting, the company and EAG base cases used a starting age of 67 years in the model. The committee concluded that this was appropriate for decision making.

Economic model

Company's modelling approach

3.7 The company developed a Markov state-transition model with 4 health states to model the cost effectiveness of adjuvant pembrolizumab compared with active monitoring. The health states were DFS, local-regional recurrence (LR), distant metastases (DM), and death. People enter the model in the DFS health state and move to the LR or DM health states, depending on the type of recurrence event they have. From LR, people could move to DM, and people could move to the death state from any other health state. Pembrolizumab was modelled to increase the probability of a cure in the long term, rather than just delaying recurrence. The model included a cure assumption, which meant that a proportion of people in the DF health state at a given time point would be considered cured (see [section 3.11](#)). This meant people who had pembrolizumab remained in the DF health state longer than people who had active monitoring. So fewer people experienced transitions to the recurrence health states, LR and DM. This separation in DFS between treatment arms was expected to continue (see [section 3.10](#)) and translate roughly into the same improvements in OS. This model structure implied that DFS

was a surrogate outcome for OS. In the absence of robust KEYNOTE-091 data, external data was used to model recurrence transition probabilities from LR and DM. To better align the modelled OS with the observed OS from KEYNOTE-091, the company applied a multiplier (see [section 3.14](#)). The biggest driver of the modelled benefits of pembrolizumab compared with active monitoring was the improvement in OS. The committee noted that there were some uncertainties because the lifetime survival extrapolations were reliant on DFS and there was limited OS data available. It concluded that the structure of the model was acceptable for decision making.

DFS modelling

3.8 The transition rates of people moving from the DFS health state were determined using DFS data from KEYNOTE091. The company fitted individual parametric models to each of the 3 transitions from DFS (DFS to LR, DFS to DM and DFS to death) and to each treatment arm. The company determined the statistical and visual fit of the parametric models for each transition and treatment arm in combination. It also took into account the [NICE Decision Support Unit technical support document 14 \(TSD14\)](#). This states that if the proportional hazards assumption does not seem appropriate it is likely to be most sensible to fit separate parametric models of the same type. At the first meeting, the company applied the same parametric model to both pembrolizumab and active monitoring for the PD-L1 less than 50% subgroup.

The company applied an exponential model for the transition from DFS to death. The EAG noted that the company's model selection did not explicitly account for the treatment effect waning in the pembrolizumab arm of KEYNOTE091 (see [section 3.10](#)). To capture this, the EAG chose to select different parametric models for the pembrolizumab and active monitoring arm. TSD14 states that it may be appropriate to fit separate parametric models to individual treatment arms, but to do so would require substantial justification because different models allow very different

shaped distributions. The EAG advised that treatment effect waning was visible within the observed DFS data, and that selecting different parametric curves for each treatment arm may be a reasonable way to account for this. At the first meeting, for the transition from DFS to LR, the EAG modelled an exponential curve for pembrolizumab and a generalised gamma curve for the active monitoring arm. For the transition from DFS to DM, the EAG selected a log-normal curve for the pembrolizumab arm and a Gompertz curve for active monitoring. Like the company, the EAG used the exponential curve to model the transition from DFS to death in both treatment arms.

The company responded that stronger evidence is needed to use different models for each treatment arm. It highlighted that modelling pembrolizumab to only delay recurrence and have no curative advantage is contrary to clinical expectation (see [section 3.11](#)). The company stated that using an exponential curve, which assumes a constant risk of hazard, to model the transition from DFS to LR in the pembrolizumab arm is likely to be inappropriate. This is because the risk of recurrence decreases over time as more people are cured. It also noted that using a Gompertz curve, which assumes no risk of recurrence soon after follow up, is likely inappropriate to model the DFS to DM transition in the placebo arm. This is because there is well-established evidence that shows there are ultra-late recurrences in early NSCLC. Applying these curves to only one of the treatment arms was also challenged by the company. The EAG responded that these were the best fitting curves to the observed data and the cure period. The company suggested alternative DFS curves that it believed had good visual and statistical fit, clinically plausible projections and followed the TSD14 guidance. This was the generalised gamma curve to model transition from DFS to LR and the log-normal curve to model the DFS to DM transition. At the first meeting, the committee considered the 3 different DFS curve selections presented. It noted that the EAG modelled DFS rates were the closest to the observed DFS from KEYNOTE-091 in both the treatment arms. But it decided there was not

enough justification to deviate from the TSD14 guidance of using the same parametric function in both treatment arms. The committee considered the company's proposed curves but noted that not enough evidence had been presented to select any. It concluded that it would like to see DFS modelled using the full licensed population and the PD-L1 TPS less than 50% subgroup (see [section 3.4](#)). This modelling should ensure that the post-cure rate of recurrence aligns with NSCLC literature (see [section 3.11](#)) and treatment waning is captured appropriately (see [section 3.10](#)).

DFS modelling in the full licensed population

3.9 In response to draft guidance consultation, the company provided an updated model for the full licensed population. But it noted there was uncertainty around the most appropriate distributions to use to extrapolate transitions from DFS to the LR and DM health states. The company stated that in the full licensed population there was some evidence to support deviating from TSD14 (see [section 3.8](#)). So, it provided 1 option that fitted different distributions to the different arms of the model. It provided 3 possible combinations of curves for the 2 arms:

- log-normal for the transitions from DFS to LR and DFS to DM for both arms (the best parametric fit for the pembrolizumab arm)
- generalised gamma for DFS to LR and log-normal for DFS to DM for both arms (the best parametric fit for the placebo arm)
- log-normal for pembrolizumab DFS to LR, generalised gamma for placebo DFS to LR, and log-normal for DFS to DM for both arms (the best parametric fits for each arm).

Because the committee decided in the first meeting that there was not enough evidence to deviate from TSD14 guidance (see [section 3.8](#)), the EAG used the parametric models that gave the best parametric fit for placebo in its base case. The EAG noted that although this choice gives a better fit for placebo overall, it still gave a poor fit in the final years of the

placebo arm. In an alternative scenario, the EAG provided parametric models that used the best parametric fit for placebo, based on the mean square error of DFS. This was generalised gamma for DFS to LR and Gompertz for DFS to DM. But the EAG acknowledged that this curve selection resulted in a poor fit for pembrolizumab in the final years of treatment. The company stated the EAG's alternative scenario was inappropriate because it modelled no DM recurrences in the placebo arm and showed convergence and crossing of the DFS curves between arms, which the company considered clinically implausible. The committee retained its position from the first committee meeting that it had not seen enough evidence to deviate from TSD14 (see [section 3.8](#)). It noted that the 2 remaining options were both plausible, but that using the generalised gamma distribution for the DF to LR transitions for both arms and the log-normal distribution for DF to DM transitions for both arms (the best parametric fit for the placebo arm) appeared to be more appropriate. The committee concluded that it would use this in its decision making.

Treatment effect waning

3.10 In the first meeting, the company did not apply any explicit treatment effect waning in the model. This meant that the benefits of adjuvant pembrolizumab were assumed to be sustained throughout the lifetime horizon. It stated that time in the DFS health state was only determined by the cure assumption (see [section 3.11](#)), the background mortality rates and the parametric models selected (see [section 3.8](#)). The company explained that pembrolizumab is a PD-1 inhibitor that blocks the interaction between PD-1 receptors and PD-L1 proteins, helping immune cells to attack cancer cells. It explained that this mechanism of action (see the [summary of product characteristics for pembrolizumab](#)) supports a sustained treatment effect. It said that this has been observed in both KEYNOTE-091 and in long-term data from other pembrolizumab indications. The EAG disagreed with the company, stating that there was significant evidence of treatment effect waning in the pembrolizumab arm.

It noted that in the observed DFS data for the PD-L1 TPS less than 50%

subgroup in KEYNOTE-091, the treatment benefit of pembrolizumab compared with placebo declines at every timepoint from 18 months. The company stated that the difference in observed DFS between pembrolizumab and placebo in KEYNOTE-091 remains relatively consistent until year 4. After year 4 the DFS difference meaningfully narrows, but at this point around two-thirds of people in the trial are censored and there are a very small number of events (n=19). So, the company said that a conclusion of treatment effect waning based on very limited data would be inappropriate. The EAG acknowledged that the data available at 5 years is limited. But it noted there is no other information to inform modelling, and the company still considered there to be enough data at these timepoints to make extrapolation assumptions (see [section 3.9](#)). The clinical experts explained that it is biologically plausible that immunotherapies could increase the proportion of people cured, and that they would expect the DFS separation to continue between pembrolizumab and active monitoring. They added that this was supported by their experience in clinical practice, where many people had survived after having immunotherapy. But they noted that there would be some people for whom the treatment effect would stop after finishing treatment. The committee acknowledged that the observed DFS curves trended together towards the end of the Kaplan–Meier curves. But this was based on a small number of people left at risk, so it was not a reliable assessment on the effect of treatment waning. It decided that a waning of the benefits of adjuvant pembrolizumab was clinically plausible, but that this had not been fully explored within the modelling. Treatment effect waning could be captured in a model either explicitly, for example, by forcing a convergence of hazards over time, or implicitly by accounting for waning in the survival estimates through selected parametric survival models. The committee concluded that it would like to see treatment effect waning explored within the new analyses. In response to the draft guidance consultation, the company confirmed that any treatment effect waning would be implicitly captured in the curve selection and the cure

assumption (see [section 3.11](#)). But it provided an additional sensitivity analysis that explored explicit treatment effect waning at 5 to 7 years, by matching the pembrolizumab hazards to placebo. It noted that this had a small effect on the cost-effectiveness estimates. The committee decided that although treatment effect waning was plausible, the presence of a cure assumption meant that its effect would be limited (because from 7 years most people in the DFS state would be considered cured). The sensitivity analysis submitted by company supported this because it had minimal impact on the cost-effectiveness estimates. The committee concluded that no additional modelling of treatment effect waning was needed.

Assumptions around cure

3.11 At the first meeting, the company model contained a cure assumption. It assumed that in the DFS health state, the proportion of people cured would rise from 0% at 5 years to 95% at 7 years. People considered cured were assumed to have no risk of transitions out of the DFS state to LR or DM states. The EAG highlighted that the 95% reduction in risk of recurrence originates from the [NICE technology appraisal guidance on pertuzumab for adjuvant treatment of HER2-positive early stage breast cancer](#). It was used in addition to a 5-year cure point so that the model's long-term recurrence rate aligned with the breast cancer literature. The EAG explained that a 95% cure rate was specific to that appraisal and has not been clinically validated. It highlighted that there is substantial uncertainty about the exact cure point and cure rate for early NSCLC. The company responded that, in its base case at the first committee meeting, the proportion of modelled ultra-late recurrences (0.8%, derived from Sonada et al. 2019) was in line with the NSCLC literature when using this 95% rate, which suggests that it is appropriate. The EAG noted that the ultra-late recurrence rate in the EAG base case at the first meeting was lower than the rate from NSCLC literature (Sonada et al. 2019), and that the EAG base case would need a 75% cure rate to align with the literature. The committee noted that the modelled proportion of ultra-late

recurrences should align with the NSCLC literature. People who remain in the DFS health state in the model were assumed to have age- and sex-matched general population mortality. The EAG's clinical experts advised that even people who remain in the DFS health state will have a mortality rate 50% to 60% higher than the general population. They explained this is caused by lasting damage to the lungs from cancer and surgery. To represent the higher risk of mortality in the cured fraction, the EAG thought a standardised mortality ratio of 1.5 should be applied to general population mortality. The company noted that all-cause mortality at year 15 in the model is already about 1.5 times that of the general population, which the EAG accepted. The committee decided that the modelling of cure was broadly appropriate in the company's base case. But it noted that the all-cause mortality for people remaining in the DFS health state should align with the clinical opinion. It concluded that the appropriateness of applying a standardised mortality ratio and using a 95% cure rate should be reassessed in any new analyses presented (see [section 3.4](#)).

Updated modelling of cure rate

3.12 In its response to draft guidance consultation, the company noted that the mean age in the Sonada et al. (2019) study was 64, compared with the mean age in the NHS clinical practice population of 67 (see [section 3.6](#)). It stated that the 0.8% estimate from the Sonada study might be an overestimate because a younger population would live longer and have more chance of a late recurrence. The company also noted that the Sonada study was based on resections done 20 years ago, and that its generalisability to current practice may be limited. So, it stated that the appropriateness of the 0.8% figure was uncertain. But the company did include various scenarios exploring calibration of ultra-late recurrence to an external literature estimate (0.8%) for the different curve selection options it had provided for the full licensed population (see [section 3.9](#)). It noted that when the best fitting curves for placebo were applied to both arms for both transitions, an 89.3% cure rate would result in ultra-late

recurrences in the placebo arm of 0.8%. The EAG, which had used the best fitting curves for placebo in its base case for the full licensed population, confirmed that it preferred to model an 89% cure rate in its base case. The committee noted the company's statement on the uncertainty of the 0.8% ultra-late recurrence rate. But it had not seen any other evidence on the incidence of ultra-late recurrences and decided that the estimate from the Sonada study was appropriate to use as a guideline for modelling cure rate. The committee maintained its preference from the first committee meeting that ultra-late recurrence should be aligned with external literature estimates (see [section 3.11](#)). The committee had decided to use the best fitting placebo curves (see [section 3.9](#)). So, it concluded that a cure rate of 89% was uncertain but appropriate for decision making.

Updated modelling of mortality

3.13 The company updated the model to include a standardised mortality ratio (SMR). It preferred to use an SMR of 1.453, which it derived from published literature, rather than the EAG expert assumption of 1.5. The company stated that the SMR included double counting. This was because the figure it was derived from included people who died because of lung cancer, and death associated with recurrent lung cancer was already captured in the model. So, the company suggested that the SMR should be seen as an upper estimate. The EAG agreed that the SMR would include some double counting, but no other appropriate evidence-based value was available. The committee noted that mortality in the cured population would be higher than in the general population. It acknowledged that some of this mortality would be because of late recurrence and would already be captured in the model. But it decided there would still be some excess mortality beyond that captured in the model. This is potentially because of comorbidities commonly found in people with NSCLC who are more likely to smoke or have smoked, or lasting damage to the lungs from cancer or surgery (see [section 3.10](#)).

The committee concluded that an SMR should be modelled in some form.

It concluded the SMR of 1.453 should be included in the model, because it is the best available estimate. But it acknowledged that this might be an upper estimate of the level or mortality in the cured population, and took this into account when considering the acceptable incremental cost-effectiveness ratio (ICER) threshold (see [section 3.17](#)).

Validation and calibration of modelled OS

3.14 The transition rates for people who have disease recurrence were estimated based on external data. This is because appropriate data from KEYNOTE-091 to inform the transitions was unavailable. To estimate the recurrence transition from the LR and DM health states, NSCLC literature and data from trials of subsequent treatments was used, respectively. Recurrence transitions were modelled in both treatment arms with an exponential curve. The company explained that using external data to model the pembrolizumab arm resulted in significantly different OS estimates compared with the KEYNOTE-091 results and real-world data. To better match the OS curves to the observed OS data, the company applied a single multiplier to all recurrence transitions in both arms. The EAG highlighted that calibrating the modelled OS to the observed OS in this way relied on combining multiple assumptions. This led to significant uncertainty in the modelled transition rates. Without access to the data, it could not validate whether the exponential curves were a reasonable fit. The EAG advised it was unlikely that a single modifier and the same parametric distribution would be appropriate to use across each transition and for both treatment arms, particularly because it believed there was evidence of treatment effect waning (see [section 3.10](#)). At the first meeting, the EAG explained that within the time constraints, no alternative approach had been presented, which made the cost-effectiveness estimates uncertain and the direction of the bias of this calibration approach unclear. To resolve this, it noted that using a partitioned survival model structure or adapting the model structure to allow for time-dependent transitions in people with recurrence would allow different modelling methods to be tested. Also, it said that further investigation of

the KEYNOTE-091 individual patient-level data would be useful to help to inform transition rates. The committee noted that the OS data was not very mature (pembrolizumab 23%; placebo 30% of OS events occurred), which meant that lifetime extrapolation of this data was very uncertain. The clinical experts supported this, explaining that there was not enough OS data for them to estimate long-term survival. The committee considered the large amount of uncertainty in the modelled transition rates and was not satisfied with the OS calibration that had been applied. At the first meeting, the committee concluded that it would like to see additional analyses that explored these uncertainties with transitions at later lines. For example, analysis that validates the OS modelling assumptions, such as providing visual fits of post-calibration extrapolations over observed OS and DFS data, for intermediate transition. Also, an exploration of the uncertainties in the model structure. For example, using mixture cure models, making post-recurrence extrapolation cures time-dependent, and applying a modifiable risk ratio to the transition rates of LR to death and DM to death, to match modelled OS.

Further exploration of modelled OS

3.15 During draft guidance consultation, the company submitted graphs showing visual fits of modelled OS and OS from KEYNOTE-091, both with and without OS calibration. It stated that updating the model to include time-dependant transitions was not possible within the time constraints. But it presented updated analyses that explored the impact of modelling cure in the LR health state. It noted that modelling cure in the LR health state did not have a large effect on the cost-effectiveness estimates, and suggested this was because the incidence of LR was similar between pembrolizumab and active monitoring. The EAG noted that the company scenario analyses that explored LR cure reduced the original uncertainty of not modelling time-dependent transitions out of LR. But the EAG stated that the certainty around DM was less clear, noting that small differences between models does not necessarily mean a small impact on the cost-effectiveness results. It also stated that curve choice for the KEYNOTE-

407 extrapolation, which was used to model distant metastases for squamous NSCLC and represented around 25% of people in the DM health state, had a significant effect on long-term outcomes. The company responded that there was no evidence to suggest that the distributions used in the DM health state were inappropriate. It also noted that in scenarios where the DM health state was not adjusted for, the cost effectiveness of pembrolizumab improved. The committee was satisfied with the exploration of transitions from the LR and DM health states. It decided that absence of time-dependent transitions was unlikely to have a large impact on cost-effectiveness estimates. The EAG's position from the first meeting was that using a single multiplier to calibrate modelled OS to trial observed OS relied on multiple assumptions and was associated with uncertainty (see [section 3.14](#)). The committee acknowledged this and concluded that the calibration of OS was associated with uncertainty but, in the absence of alternative modelling, was acceptable for decision making.

Severity

3.16 The company did not make a case to apply the severity modifier. NICE's methods on conditions with a high degree of severity did not apply.

Cost-effectiveness estimates

Acceptable ICER

3.17 [NICE's manual on health technology evaluations](#) notes that, above a most plausible ICER of £20,000 per quality-adjusted life year (QALY) gained, judgements about the acceptability of a technology as an effective use of NHS resources will take into account the degree of certainty around the ICER. The committee will be more cautious about recommending a technology if it is less certain about the ICERs presented. But it will also take into account other aspects including uncaptured health benefits. The committee decided that the most appropriate population was the full

licensed population ([section 3.5](#)). It noted the remaining uncertainty in the model, which arose from:

- the limited DFS evidence after 4 years, which meant that the duration of treatment effect and the most appropriate modelling of DFS is unknown (see [section 3.10](#) and [section 3.9](#))
- the way in which the assumption of cure was modelled and what mortality cured people would have (see [section 3.12](#))
- the reliance on a surrogate relationship between DFS and OS to model OS, and the recurrence transition rates used to generate OS, which is based on multiple combined assumptions (see [section 3.15](#)).

So, the committee concluded that an acceptable ICER would be around the middle of the range that NICE considers a cost-effective use of NHS resources (£20,000 to £30,000 per QALY gained).

Committee's preferred cost-effectiveness estimates

3.18 The committee's preferences for the cost-effectiveness modelling of adjuvant pembrolizumab in the full licensed population included:

- considering active monitoring as the only relevant comparator (see [section 3.2](#))
- using a starting age of 67 in the model (see [section 3.6](#))
- using generalised gamma distribution for DF to LR transitions and log-normal for DF to DM transitions for both arms (see [section 3.9](#))
- including no additional modelling of treatment effect waning (see [section 3.10](#))
- applying a cure rate of 0% at 5 years, rising to 89% at 7 years (see [section 3.12](#))
- applying an SMR of 1.453 to general population mortality to reflect higher mortality in the disease-free state (see [section 3.13](#))
- calibrating modelled OS to trial data by modifying all recurrent transitions (see [section 3.15](#)).

The committee concluded that the most plausible ICER when considering its preferred assumptions was within the range it considered a cost-effective use of NHS resources. The exact ICERs are confidential and cannot be reported here because of confidential discounts for technologies included in the modelling.

Other factors

Equality

3.19 The committee did not identify any equality issues.

Conclusion

Recommendation

3.20 The committee took into account the key uncertainties in the model and its preferred assumptions. The committee concluded that pembrolizumab is a cost-effective treatment option. So, pembrolizumab is recommended.

4 Implementation

4.1 Section 7 of the [National Institute for Health and Care Excellence \(Constitution and Functions\) and the Health and Social Care Information Centre \(Functions\) Regulations 2013](#) requires integrated care boards, NHS England and, with respect to their public health functions, local authorities to comply with the recommendations in this evaluation within 90 days of its date of publication.

4.2 Chapter 2 of [Appraisal and funding of cancer drugs from July 2016 \(including the new Cancer Drugs Fund\) – A new deal for patients, taxpayers and industry](#) states that for those drugs with a draft recommendation for routine commissioning, interim funding will be available (from the overall Cancer Drugs Fund budget) from the point of marketing authorisation, or from release of positive draft guidance, whichever is later. Interim funding will end 90 days after positive final guidance is published (or 30 days in the case of drugs with an Early

Access to Medicines Scheme designation or cost comparison evaluation), at which point funding will switch to routine commissioning budgets. The [NHS England Cancer Drugs Fund list](#) provides up-to-date information on all cancer treatments recommended by NICE since 2016. This includes whether they have received a marketing authorisation and been launched in the UK.

- 4.3 The Welsh ministers have issued directions to the NHS in Wales on implementing NICE technology appraisal guidance. When a NICE technology appraisal guidance recommends the use of a drug or treatment, or other technology, the NHS in Wales must usually provide funding and resources for it within 2 months of the first publication of the final draft guidance.
- 4.4 When NICE recommends a treatment ‘as an option’, the NHS must make sure it is available within the period set out in the paragraphs above. This means that, if a patient has non-small-cell lung cancer (NSCLC) with a high risk of recurrence after complete resection and platinum-based chemotherapy and the healthcare professional responsible for their care thinks that pembrolizumab is the right treatment, it should be available for use, in line with NICE’s recommendations.

5 Evaluation committee members and NICE project team

Evaluation committee members

The 4 technology appraisal committees are standing advisory committees of NICE. This topic was considered by [committee D](#).

Committee members are asked to declare any interests in the technology being evaluated. If it is considered there is a conflict of interest, the member is excluded from participating further in that evaluation.

The [minutes of each evaluation committee meeting](#), which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Chair

Raju Reddy

Vice Chair, technology appraisal committee D evaluation committee

NICE project team

Each evaluation is assigned to a team consisting of 1 or more health technology analysts (who act as technical leads for the evaluation), a technical adviser, a project manager and an associate director.

Lauren Elston and Cara Gibbons

Technical leads

Samuel Slayen and Adam Brooke

Technical adviser

Celia Mayers

Project manager

Ian Watson and Jasdeep Hayre

Associate director

ISBN: **[to be added at publication]**