NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Final appraisal document

Pembrolizumab for treating locally advanced or metastatic urothelial carcinoma after platinum-containing chemotherapy

1 Recommendations

1.1 Pembrolizumab is not recommended, within its marketing authorisation, for treating locally advanced or metastatic urothelial carcinoma in adults who have had platinum-containing chemotherapy.

1.2 This recommendation is not intended to affect treatment with pembrolizumab that was started in the Cancer Drugs Fund before this guidance was published. For those people, pembrolizumab will be funded by the company until they and their NHS clinician consider it appropriate to stop.

Why the committee made these recommendations

This appraisal reviews the additional evidence collected as part of the Cancer Drugs Fund managed access agreement for pembrolizumab for previously treated locally advanced or metastatic urothelial carcinoma after platinum-containing chemotherapy (NICE technology appraisal 519). Once final guidance is published, pembrolizumab will no longer be available in the Cancer Drugs Fund for this indication for people starting treatment, but people already taking it will be able to continue.

Treatment for previously treated locally advanced or metastatic urothelial carcinoma includes docetaxel or paclitaxel. Clinical trial evidence shows that pembrolizumab significantly improves overall survival compared with docetaxel and paclitaxel. Some evidence has also been collected through the Cancer Drugs Fund.

Issue date: March 2021

Atezolizumab is now also a possible treatment. But it was not established clinical practice in the NHS at the time of the original appraisal, so is not included in the scope for this review.

If an active treatment is not tolerated or people choose not to have it, best supportive care is given. No clinical or cost-effectiveness evidence was available for pembrolizumab compared with best supportive care.

Pembrolizumab meets NICE's criteria to be considered a life-extending treatment at the end of life. The most likely cost-effectiveness estimate for pembrolizumab is uncertain. This is because it is not clear which model of overall survival is most appropriate or how long the treatment benefit of pembrolizumab should continue. Even when pembrolizumab is offered with its agreed discount, the most plausible cost-effectiveness estimate remains higher than what NICE normally considers acceptable for end-of-life treatments. Therefore, pembrolizumab is not recommended.

2 Information about pembrolizumab

Marketing authorisation indication

1.3 Pembrolizumab (Keytruda, Merck Sharp & Dohme) has a marketing authorisation for 'the treatment of locally advanced or metastatic urothelial carcinoma in adults who have received prior platinum-containing chemotherapy'.

Dosage in the marketing authorisation

1.4 The dosage schedule is available in the summary of product characteristics.

Price

1.5 £2,630 per 100 mg vial (excluding VAT; company submission).

1.6 The company has a commercial arrangement. This makes pembrolizumab available to the NHS with a discount and it would have also applied to this indication if the technology had been recommended. The size of the discount is commercial in confidence. It is the company's responsibility to let relevant NHS organisations know details of the discount.

1.7 While available in the Cancer Drugs Fund (see NICE technology appraisal 519), pembrolizumab has a commercial arrangement (managed access agreement including a commercial access agreement).

3 Committee discussion

The appraisal committee considered evidence submitted by Merck Sharp & Dohme and a review of this submission by the evidence review group (ERG), and the technical report developed through engagement with stakeholders. See the committee papers for full details of the evidence.

The committee recognised that there were remaining areas of uncertainty in the analyses presented (see technical report, table 2, page 37), and took these into account in its decision making. It discussed issues 1 to 5 from the technical report, which were not resolved after technical engagement:

- choice of extrapolation for progression-free survival
- treatment switching
- choice of extrapolation curve and cut-off point for overall survival
- treatment effect duration
- PD-L1 expression subgroups.

The condition

Locally advanced or metastatic urothelial carcinoma substantially decreases quality of life

3.1 Urothelial carcinoma causes a number of symptoms, including haematuria (blood in the urine) and increased frequency, urgency and pain associated with urination. Surgical treatments such as urostomy can have a substantial effect on quality of life and restrict daily activities. The patient experts explained that chemotherapy is associated with unpleasant side effects such as fatigue, nausea and vomiting and puts people at a greater risk of infection. The committee was aware that many people with locally advanced or metastatic urothelial carcinoma are older and may have comorbidities, which can affect treatment decisions. It recognised that locally advanced or metastatic urothelial carcinoma has a significant impact on quality of life.

Current treatments and comparators

Paclitaxel, docetaxel and best supportive care are the relevant comparators for this appraisal

3.2 The committee was aware that the treatment pathway for locally advanced or metastatic urothelial carcinoma had changed since the original appraisal of pembrolizumab for this indication. This is because of NICE's technology appraisal guidance on atezolizumab for treating locally advanced or metastatic urothelial carcinoma after platinum-containing chemotherapy (from now, TA525). Atezolizumab was not established clinical practice in the NHS when the final scope for the original appraisal of pembrolizumab was issued. In a review of a drug funded by the Cancer Drugs Fund, no changes to the final scope of the original appraisal are allowed, so atezolizumab could not be included as a comparator (see section 6.25 of the NICE guide to the processes of technology appraisal).

At the time of the original appraisal of pembrolizumab, first-line treatment for locally advanced or metastatic disease was usually a platinum-containing chemotherapy regimen. For people who were not well enough or chose not to have this, best supportive care was offered. Retreatment with a first-line chemotherapy was also included in the scope for the original appraisal of pembrolizumab. However, it was not established clinical practice then, because retreatment was used before a second-line treatment option was available. Also, most clinicians would have used a taxane (paclitaxel and docetaxel). The committee agreed that treatment options for people with disease progression after platinum-containing chemotherapy at that time included docetaxel, paclitaxel or best supportive care. The committee concluded for the original appraisal that the most relevant comparators were paclitaxel, docetaxel and best supportive care.

The KEYNOTE-045 post-hoc subgroup results are most appropriate for decision making

3.3 The clinical-effectiveness evidence for pembrolizumab came from KEYNOTE-045, an open-label, randomised controlled trial. It included people with disease progression or recurrence of urothelial cancer after treatment with a platinum-containing regimen (cisplatin or carboplatin). The comparator arm in the trial was the investigator's choice of paclitaxel, docetaxel, or vinflunine. The company recognised that vinflunine is not used in clinical practice in the UK, and did a post-hoc subgroup analysis. This included:

- 188 people randomised to have pembrolizumab
- 182 people randomised to have the investigator's choice of paclitaxel or docetaxel (referred to as the ‘UK standard of care [UK SoC]’ control arm).

The committee concluded that the trial was good quality and the results were informative for decision making. It was aware that using post-hoc subgroup analyses introduced the risk of bias, and that excluding the vinflunine data reduces the statistical power of the trial. But the committee concluded that the post-hoc subgroup best reflects UK clinical practice and is the most appropriate evidence for decision making.

Pembrolizumab improves overall survival compared with docetaxel or paclitaxel

3.4 In KEYNOTE-045, progression-free survival and overall survival were co-primary end points. In the latest data cut of KEYNOTE-045, the median overall survival for pembrolizumab was 10.1 months (95% confidence interval [CI] 7.6 to 12.9) compared with 6.2 months (95% CI 5.2 to 7.4) for the UK SoC arm with a hazard ratio of 0.64 (95% CI 0.49 to 0.81). This suggests that pembrolizumab improves overall survival compared with docetaxel or paclitaxel. However, the median progression-free survival for pembrolizumab was 2.1 months (95% CI 2.0 to 2.2) compared with 3.3 months (95% CI 2.3 to 3.5) in the UK SoC arm, with a hazard ratio of 0.95 (95% CI 0.76 to 1.19). The committee concluded that pembrolizumab improves overall survival but does not appear to improve progression-free survival. The additional clinical data collected by Public Health England as part of the Systemic Anti-Cancer Therapy dataset while pembrolizumab was in the Cancer Drugs Fund did not contribute to this review.

The 2-stage method for subsequent immunotherapy in KEYNOTE-045 is appropriate in the original appraisal

3.5 If their disease progressed, people in the trial could have subsequent anti-PD-L1 or PD-1 treatment. This included atezolizumab, avelumab, durvalumab, nivolumab, and pembrolizumab. The company adjusted overall survival in the UK SoC arm to account for these treatments using the 2-stage method to adjust for treatment switching. The 2-stage method

used an acceleration factor (a ratio of the survivor function for the pembrolizumab and UK SoC treatment arms). This was to shrink the survival time of patients who had UK SoC, were eligible for subsequent therapy, and who then had anti-PD-L1 or PD-1 therapy. The ERG believed that the 2-stage method had disadvantages, but overall was the most appropriate. The committee concluded that the company's 2-stage method was appropriate for decision making in the original appraisal.

New KEYNOTE-045 data shows that the 2-stage method may not be appropriate, and the unadjusted method should also be taken into account

3.6 The November 2018 data cut from KEYNOTE-045 showed that the acceleration factor was larger and applied to more people in the trial. This meant the 2-stage adjustment had a greater influence on overall survival than it did in the original appraisal. The acceleration factor was 5.37 (95% CI 3.23 to 10.09) (based on 25 patients) after the November 2018 data cut, compared with 3.86 (95% CI 1.79 to 11.68) (based on 14 patients) using previous data. The ERG considered that both the 2-stage adjusted analyses and analyses without this adjustment for treatment switching should be carefully considered. It advised that the true overall survival benefit would be somewhere between the result of the 2 approaches. Using an approach without the adjustment might overestimate survival time in the UK SoC arm, but the 2-stage method might underestimate survival time in this arm too much. The ERG advised that the main uncertainties with the 2-stage adjustment were:

- The wide confidence interval around the acceleration factor showed a high degree of uncertainty.
- The adjustment method assumed an average adjustment for all people switching to anti-PD-L1 or PD-1 therapy. The ERG considered it unlikely that all patients who switched benefited equally from the anti-

PD-L1 or PD-1 treatment. This was because evidence from KEYNOTE-045 suggested pembrolizumab had less benefit than UK SoC for the first 3 months of follow up, and because immunotherapies have not been shown to benefit everyone.

- With the adjustment, the benefit would have been the same as if patients had anti-PD-L1 or PD-1 therapy earlier in their disease pathway. The KEYNOTE-045 trial data did not support this.
- There was potential for selection bias related to switching, and unmeasured prognostic factors could affect the data.

In response to the appraisal consultation document, the company advised that it considered the updated acceleration factor to be more reliable than the original acceleration factor. This was because it was calculated from a larger sample size and the confidence intervals were narrower and within the range of the originally calculated confidence intervals. The ERG stated that the main concern was not the size of the acceleration factor, but that the increased size meant the adjustment had more influence and so the existing uncertainties associated with the 2-stage method were more important. With the most up-to-date data from November 2018, 40 people on the UK SoC arm of the trial switched to an anti-PD-L1 or PD-1 treatment. The acceleration factor was calculated from the 25 people who switched when progression of their disease was documented. The acceleration factor was not applied to the overall survival time of 15 patients who switched at different times. The ERG stated that the company had not provided an established rule for switching. In response to consultation, the company provided a sensitivity analysis applying the acceleration factor to all 40 patients. In this, the hazard ratio for pembrolizumab compared with UK SoC was 0.55. However, the calculation of the acceleration factor was not adjusted to include these 15 patients. The committee considered that using the 2-stage adjustment for treatment switching likely underestimated the incremental cost-

effectiveness ratios (ICERs) but using no adjustment would overestimate the ICERs. It concluded that the true overall survival benefit was probably between that seen with an adjustment for treatment switching and that without an adjustment. The committee considered this issue further after an appeal (see sections 3.23 and 3.26).

PD-L1 positive subgroups are not clinically distinct

3.7 The company defined PD-L1 expression in KEYNOTE-045 by combined proportion score, which includes PD-L1 expression in both the solid tumour and the infiltrating immune cells. The company did not present an analysis showing the interaction between treatment effect and PD-L1 status, or results for the PD-L1 negative subgroup using data from the November 2018 cut-off. The committee agreed there was inherent uncertainty when considering estimates of effectiveness based on any subgroup data. The clinical expert explained that PD-L1 is not a predictive biomarker for pembrolizumab after platinum-containing chemotherapy, but it is more relevant for pembrolizumab for people when cisplatin is unsuitable. This is reflected in the marketing authorisation for pembrolizumab in the first-line indication for people when cisplatin is unsuitable, because it specifies PD-L1 expression through combined proportion score level. The clinical expert advised that diagnostic tissue samples for combined proportion score testing are taken before first-line treatment, and combined proportion score may change after platinum-based chemotherapy. This means combined proportion score and PD-L1 expression are not predictive biomarkers in this post-chemotherapy population. The committee agreed that PD-L1 positive subgroups were not clinically distinct subgroups for this indication. It concluded to not consider PD-L1 subgroups in its decision making.

Comparison with best supportive care

No evidence is available comparing pembrolizumab with best supportive care

3.8 The committee considered best supportive care as a relevant comparator, because a few people would have best supportive care if an active treatment was not tolerated or they chose not to have it (see section 3.2). There was no direct trial evidence comparing pembrolizumab with best supportive care. The company did not consider best supportive care a relevant comparator and, in the original appraisal, did not present any new clinical or cost-effectiveness evidence comparing pembrolizumab with best supportive care. Therefore, the committee concluded it was unable to make a recommendation on this, and agreed not to consider it further.

Adverse events

Pembrolizumab is well tolerated in clinical practice

3.9 Pembrolizumab is associated with some rare but unpleasant, and potentially serious, adverse events that are specific to immunotherapy. The committee understood that pembrolizumab was well tolerated and that patients considered it to have fewer severe adverse events than chemotherapy. The patient experts explained that, although pembrolizumab does have side effects, these are typically less than for chemotherapy for this indication. They suggested that pembrolizumab did not interfere with everyday activities as much. The committee concluded that pembrolizumab was well tolerated.

Assumptions used in the economic model

A 2-year stopping rule for pembrolizumab is appropriate

3.10 In the KEYNOTE-045 protocol, the maximum pembrolizumab treatment duration was 2 years from the first dose, when treatment must be

stopped. This was not reflected in the summary of product characteristics, which states that treatment should continue until disease progression or unacceptable toxicity. For pembrolizumab for other indications, and in TA525, a 2-year stopping rule was applied. The committee noted that the 2-year stopping rule was included in company's economic model, and concluded that it was appropriate.

A Weibull curve is the most appropriate to model progression-free survival in both treatment arms

3.11 In the original appraisal, the committee concluded that the Weibull curve for progression-free survival was appropriate. The committee noted that, for the review, the company still extrapolated progression-free survival from 21 weeks, but preferred a log-normal curve for the pembrolizumab arm. This was based on statistical and visual fit to the KEYNOTE-045 data, and then was used for the UK SoC arm to be consistent. The ERG considered it appropriate to extrapolate from 21 weeks, but only found the Weibull curve to consistently be among the best fitting curves for both the pembrolizumab and the UK SoC arms. This was according to the Akaike information criterion and the Bayesian information criterion. The ERG explained that NICE's technical support document 14 advises that when parametric models are fitted separately to individual treatment arms, the same extrapolation should be used for both arms. Otherwise, substantial justification would be needed to use different extrapolation models. The committee considered the Weibull curve to fit well to the almost-complete data for the UK SoC arm, and also to the 2 to 3-year progression-free survival data for pembrolizumab (the benefit is very uncertain beyond that). The Weibull curve was most consistent with the Kaplan−Meier data seen at 2 and 3 years in both arms of the KEYNOTE-045 trial, and was also a good visual fit. The committee concluded that the Weibull curve was the most appropriate curve to model progression-free survival and that it should be used for both the pembrolizumab and UK SoC arms.

A piecewise model is appropriate to model overall survival, and the best time to switch to a parametric curve is at 24 weeks

3.12 The company used a piecewise approach to model overall survival, in which Kaplan−Meier data are used first before switching to a parametric curve. This is because the cumulative hazard plot showed that the hazards crossed and therefore the proportional hazards assumption did not hold. The company incorporated switching to a parametric curve at week 24 in its base-case analysis because the cumulative hazard curves started separating from week 24. The committee agreed that the company's piecewise model was appropriate to model overall survival, and the best time for switching to a parametric curve was at 24 weeks.

The long-term effect of a stopping rule on the duration of treatment effect is unknown for immunotherapies, but a lifetime treatment effect is implausible

3.13 A 2-year stopping rule was appropriate for pembrolizumab (see section 3.10). The duration of continued treatment effect after implementation of a stopping rule is an area of uncertainty for all immunotherapies. Before this review, there were no data from KEYNOTE-045 on the effect of implementing the stopping rule, because the longest follow up was only 20.8 months. In the original appraisal, the committee concluded that a lifetime treatment effect was implausible. While a small number of patients could have 'immune memory' after the 2-year stopping point for treatment with pembrolizumab, this was uncertain. The clinical expert explained that the long-term effect of stopping immunotherapy at 2 years was still unknown for any disease.

Evidence of treatment effect duration from other pembrolizumab trials is not appropriate for decision making

3.14 The company highlighted that data supporting a long-term survival benefit was available across the pembrolizumab clinical study programme,

particularly from KEYNOTE-001 (melanoma, non-small-cell lung cancer), KEYNOTE-006 (melanoma) and KEYNOTE-024 (non-small-cell lung cancer). The committee was aware that melanoma and non-small-cell lung cancer trials for pembrolizumab had some of the strongest evidence for a sustained response in a small number of patients. However, it recognised that the evidence suggests that treatment effect duration varies in different types of cancers. It therefore agreed that the results from those trials were not generalisable to urothelial carcinoma.

There is no strong evidence to support the 5-year duration of treatment effect from the start of pembrolizumab treatment in the company's base case

3.15 For this review, the company used a 5-year treatment effect duration from the start of treatment with pembrolizumab in its base case, and 3 years and 10 years of treatment effect from the start of treatment in its scenario analyses. It supported its choice of a 5-year treatment effect duration in its base case by showing that the hazard ratio for overall survival for pembrolizumab compared with the UK SoC arm (using its preferred 2-stage adjustment, see section 3.5) had improved with additional follow-up data (median follow up 40.9 months, range 36.6 to 48.9 months). The comparison with the data from the original appraisal cannot be shown here as the hazard ratio is academic in confidence. The company explained that this trend was seen with the full trial population in the comparator arm of KEYNOTE-045 and when data for the UK SoC arm (unadjusted for treatment switching) was used. The company considered that a 2-year or 3-year cap on the duration of treatment effect from the start of treatment was inappropriate. This was because any longer-term benefit of pembrolizumab would not be taken into consideration, and extrapolation in the pembrolizumab arm did not fit well to the Kaplan−Meier overall survival data from the November 2018 data cut-off. The company indicated that with its preferred log-logistic curve for

extrapolation of overall survival (see section 3.20), 4.5% of people having pembrolizumab were modelled to still be alive 10 years after starting treatment. Around 50% of patients in KEYNOTE-045 stopped pembrolizumab 6 months after starting treatment. The clinical expert found it plausible that 5% to 10% of people having pembrolizumab might survive to 10 years after starting treatment (with a 2-year stopping rule). A patient expert and the Cancer Drugs Fund clinical lead agreed that there was uncertainty about how long people might survive after having pembrolizumab. This is because people with urothelial cancer tend to be older, with other comorbidities, so those people whose cancer responds to treatment could die from another cause before 10 years after starting treatment. The committee agreed that there was no strong evidence to support a 5-year or longer treatment effect, and no more than 5% of people treated with pembrolizumab might be alive after 10 years.

Based on the available evidence, a 3-year duration of treatment effect from the start of pembrolizumab is plausible

3.16 The ERG suggested that the improved hazard ratio for overall survival for pembrolizumab with the extended follow up could be explained by greater data completeness (patients in the trial progressing or dying in the longer follow-up period). The ERG preferred to use a 3-year duration of treatment effect in its exploratory base case, because it thought there was reasonably robust evidence of an effect up until 2 years from starting treatment, but limited support for an effect beyond 3 years. This was because after 3 years, there was only 1 death in the unadjusted (see section 3.5) UK SoC arm and none in the adjusted population. Although the ERG accepted that there was some evidence of sustained response for pembrolizumab, it also considered that the same was true for the UK SoC arm, with no evidence to suggest the hazard rate for long-term response was different across treatment arms after 2 years. The clinical expert advised that the sustained response from pembrolizumab was

greater than that for the UK SoC arm. They stated that there was a small group of people who had pembrolizumab supporting at least a 3-year duration of treatment effect from the start of treatment. The clinical expert explained that this was not the case for people who had chemotherapy, because very few people survive beyond 2 years. The committee considered that there was robust evidence to support a 3-year treatment effect after starting pembrolizumab (2 years of treatment plus 1 year of follow up). It concluded that, although the effect duration was uncertain, based on the available evidence a 3-year duration treatment from the start of pembrolizumab was plausible.

The company's new scenario analyses on duration of treatment effect are not appropriate for the model

3.17 In response to consultation, the company highlighted that 38.5% of people in the pembrolizumab arm had a best overall response of disease control. It presented several scenario analyses in which 38.5% of people continued to benefit from pembrolizumab for their lifetime, while the rest had the same benefit as the UK SoC arm after either 3 or 5 years. The ERG highlighted that the company had assumed the same level of response to pembrolizumab for people whose disease responded and people whose disease did not respond for the first 3 or 5 years of the model. It considered that the 2 groups would be likely to have quite different survival outcomes. The committee considered that it was arbitrary to split the pembrolizumab arm at 3 or 5 years and that a split at baseline with a different statistical analysis may have been more plausible. It also noted that the analysis was not applied to the UK SoC arm although there were people in the UK SoC arm whose disease also became stable. The committee concluded that the company's new scenario analyses were not appropriate for the model.

A 3-year to 5-year duration of treatment effect from the start of pembrolizumab treatment could be plausible

3.18 In response to consultation, the company also presented a summary of response rates from KEYNOTE-045. It highlighted that the median duration of response for responders was 29.7 months in the pembrolizumab arm and 4.4 months in the UK SoC arm. The 36-month overall survival rate was 20.7% in the pembrolizumab arm and 11.0% in the UK SoC arm. The proportion of responses lasting 24 months or more was 56.8% in the pembrolizumab arm and 28.3% in the UK SoC arm. The company also stated that the trial was not designed to show a treatment benefit beyond 3 years. At consultation, professional groups also highlighted these figures and stated that they were consistent with more positive long-term survival estimates than those previously assumed by the committee. The committee agreed that the Kaplan–Meier evidence did not suggest a long-term difference in hazard rates between the 2 treatment arms. It considered that there was robust evidence to support a 3-year treatment effect after starting pembrolizumab (see section 3.16). However, it also considered that the new figures suggested the relative treatment effect of pembrolizumab might continue beyond 3 years. The committee recalled that in the TA525, analyses with a treatment effect cap at 3 years after stopping were taken into account in its decision making but there was not enough evidence to support a specific duration of benefit. The committee agreed that the treatment effect duration was uncertain. It concluded that a 3-year to 5-year treatment effect from start of pembrolizumab treatment could be plausible. The committee considered this issue further after an appeal (see sections 3.23 and 3.24).

The costs of pembrolizumab are likely underestimated in the model

3.19 The NHS England commissioning expert highlighted that in KEYNOTE-045, people in the pembrolizumab arm who stopped taking pembrolizumab because they had a complete response or after the 2-year

stopping rule could restart pembrolizumab for up to 1 year if their disease progressed. The company explained that 10 people out of 188 in the pembrolizumab arm had retreatment with pembrolizumab and that the costs of this were not included in the model. It explained that it was difficult to separate out the benefit these people may have had with retreatment from the overall benefit of taking pembrolizumab. The committee concluded that although only a small proportion of patients had retreatment, the costs of pembrolizumab were likely underestimated in the model. The committee considered this issue further after an appeal (see sections 3.23 and 3.25).

There are 3 plausible overall survival extrapolation curves

3.20 In its base case, the company preferred the log-logistic extrapolation for overall survival. This choice was based on statistical and visual fit to the updated overall survival data from KEYNOTE-045 (see section 3.4). The company highlighted that the log-logistic curve gave a 3.2% 5-year survival rate for the UK SoC arm, consistent with the 2% to 3% figure given by the ERG's clinical expert in the original appraisal. The ERG preferred a log-logistic curve in its exploratory base case. But, it also considered the log-normal and generalised gamma plausible if some patients experienced the long-term survival benefit for pembrolizumab suggested by the company (with generalised gamma being the most optimistic). If no patients experienced this long-term survival benefit, then the ERG advised that the Weibull extrapolation would be plausible. The ERG explained that the company's preferred curve and anticipated long-tailed survival profile for pembrolizumab in the long term were plausible, but unsupported by evidence (see section 3.15). The committee acknowledged that there were a number of plausible overall survival extrapolation curves. Because a small number of people having pembrolizumab may survive to 10 years after starting treatment (see section 3.15), the committee agreed that the Weibull extrapolation would

penalise overall survival too harshly. But, the log-logistic, log-normal and generalised gamma were plausible if there were any survivors at 10 years. However, there was a high degree of uncertainty around long-term overall survival for pembrolizumab and all immunotherapies at 10 years because of a lack of data. So, the committee concluded that log-logistic, log-normal and generalised gamma were all plausible, and that all 3 should be taken into account in decision making.

The company's utility value estimates are appropriate

3.21 EQ-5D data were collected directly in KEYNOTE-045; these data are the preferred measure of health-related quality of life in adults. In the company's base case, vinflunine data was not included in the utility estimates because vinflunine is not used in UK clinical practice and is not included in the survival data (see sections 3.3 and 3.4). The company based the utility values on progression state and used the most recent age-related disutility algorithm. It also pooled the utility estimates across treatment arms. The committee agreed with the utility values estimates used in the company's economic model.

The cost-effectiveness estimates before the appeal

The most plausible ICER for pembrolizumab compared with docetaxel and paclitaxel is likely to be over £50,000 per QALY gained

3.22 The company's base-case deterministic ICER for pembrolizumab was £47,123 per quality-adjusted life year (QALY) gained compared with docetaxel or paclitaxel. This was based on the following assumptions: log-normal extrapolation for progression-free survival from 21 weeks; log-logistic extrapolation for overall survival from 24 weeks; a 5-year treatment effect duration from the start of treatment with pembrolizumab; 2-stage adjustment for treatment switching applied to the UK SoC arm. The committee noted that without the 2-stage adjustment for switching

(see section 3.5) the company's ICER increased to £56,422 per QALY gained. The ERG changed the company's base case to use a Weibull extrapolation for progression-free survival from 21 weeks, which was the committee's preferred assumption (see section 3.11). This increased the ICER for pembrolizumab to £48,518 per QALY gained, and to £58,850 per QALY gained without the 2-stage adjustment for treatment switching (both ICERs including a 5-year treatment duration effect). The ERG then also included a 3-year treatment effect duration, which the committee agreed was plausible (see section 3.18). The ICER for pembrolizumab increased to £53,678 per QALY gained with the 2-stage adjustment for treatment switching and to £65,469 per QALY gained without the 2-stage adjustment. Considering all 3 plausible options for the extrapolation of overall survival (log-logistic, log-normal and generalised gamma, see section 3.20), with the 2-stage adjustment for treatment switching, the Weibull extrapolation for progression-free survival and the 3-year treatment effect duration, the ICER ranged from £53,678 to £58,705 per QALY gained. The equivalent ICERs without the 2-stage adjustment ranged from £61,653 to £70,520 per QALY gained. The committee agreed the most plausible ICERs were somewhere between those with the 2-stage adjustment for treatment switching in the UK SoC arm, and those without the adjustment (see section 3.5). It also agreed that the ICER of £48,518 per QALY gained was at the lowest end of the range of plausible ICERs, but it was unlikely to be the most plausible because it was based on the most optimistic of the committee's preferred assumptions. When taking into account the uncertainty about the 2-stage adjustment, the uncertainty around the plausible treatment effect duration (3 to 5 years, see section 3.18) and the 3 plausible overall survival extrapolations (see section 3.20), the committee noted that the ICER could be as high as £70,520 per QALY gained. This was also unlikely to be the most plausible ICER because it was based on the most pessimistic of the committee's preferred assumptions. The committee also considered that the costs of

pembrolizumab could be underestimated in the model (see section 3.19), and that increasing the costs of pembrolizumab would increase the ICERs. Considering all these factors, the committee concluded that the most plausible ICER for pembrolizumab compared with docetaxel and paclitaxel was likely to be over £50,000 per QALY gained.

After the appeal

3.23 At the third appraisal committee meeting, the committee considered the appeal panel's decision to uphold 3 appeal points and refer these back to the appraisal committee for further consideration. These were:

- The committee needs to clearly explain its rationale for accepting a different approach to the duration of treatment effect than TA525 (see section 3.24).
- The committee should allow the company an opportunity to respond to the issue of retreatment costs (see section 3.25).
- The committee should consider a range of acceleration factors for the 2-stage method to adjust for treatment switching. Also, it should reconsider whether it is appropriate to give equal weight to analyses that did not adjust for treatment switching (see section 3.26).

The committee considered the company's updated analyses including a revised patient access scheme.

Differences in the clinical and economic evidence between this appraisal and TA525 mean it is appropriate to consider different treatment effect durations

3.24 The committee considered the first upheld appeal point (see section 3.23). It discussed the reasoning behind its previous conclusion to consider 3-year and 5-year treatment effect durations after starting pembrolizumab for decision making (see section 3.18). In TA525 all analyses that varied the treatment effect duration, from a lifetime effect to a 3-year effect after

stopping atezolizumab, had ICERs that were comfortably within the range normally considered cost effective for end-of-life technologies. The exact ICERs are confidential and cannot be reported here. The committee noted several other differences between the 2 appraisals:

- The 5-year treatment effect duration used in TA525 was not supported by robust evidence because the IMvigor211 trial had a maximum follow up of 25 months. However, extended follow-up data from KEYNOTE-045 were available. Those data suggested that the treatment benefit with pembrolizumab was unlikely to be sustained after 3 years (see section 3.16).
- Pembrolizumab treatment was only given for 2 years in KEYNOTE-045 but there was no treatment cap for atezolizumab in IMvigor211.
- In IMvigor211, patients continued taking atezolizumab until unmanageable toxicity or lack of clinical efficacy. This means that some people continued taking atezolizumab after their disease progressed. So, any treatment benefit may have lasted for longer than if treatment was stopped after disease progression (as in KEYNOTE-045).
- In KEYNOTE-045, 10 patients had retreatment with pembrolizumab after disease progression (see sections 3.19 and 3.25). These patients being offered a second course of pembrolizumab suggests that a long-term treatment effect after their initial course was not expected. In TA525, there was no evidence of retreatment with atezolizumab.
- In the model, the proportion of patients alive at 2 years still having treatment was lower for pembrolizumab than for atezolizumab.
- The company for atezolizumab did not provide analyses assuming a 3-year treatment effect duration from starting treatment.

The committee carefully considered these differences. It reiterated that there was no robust evidence to support a 5-year treatment effect, but acknowledged that it could be plausible (see section 3.18). In TA525, although it had not seen analyses assuming a 3-year treatment effect

duration from starting treatment, all ICERs were comfortably cost effective. This meant the committee was confident a 3-year treatment effect analysis would also have had a cost-effective ICER. It concluded that its rationale for considering analyses using 3-year and 5-year treatment effect durations from the start of pembrolizumab treatment was reasonable, based primarily on the difference in cost-effectiveness estimates between this appraisal and TA525, and supported by the differences in the evidence.

The cost of retreatment should be included at 3 years

3.25 The committee considered the second upheld appeal point (see section 3.23). After the appeal, the company submitted scenario analyses that included the costs of pembrolizumab retreatment for the 10 patients that had it. The company and ERG had different preferences about when to apply this cost in the model. The committee agreed that the ERG's preference (3 years) was more consistent with the data, but acknowledged the timing had negligible impact on cost-effectiveness results. The company could have provided analyses that removed potential survival benefit from retreatment instead of adding the costs, but it did not do so. The company advised that pembrolizumab retreatment has uncertain clinical benefit and does not reflect clinical practice in the NHS in England. So, it considered that its preferred analysis without the costs remained appropriate for decision making. The committee found it inconsistent to include the potential benefits of retreatment without the costs, so both should either be included or excluded. In the absence of an analysis removing the benefits of retreatment, it concluded that the costs should be applied at 3 years.

Unadjusted analyses are not suitable for decision making

3.26 The committee considered the third upheld appeal point (see section 3.23). It previously concluded that analyses that did not attempt to

adjust for treatment switching method should be taken into account (see section 3.6). After the appeal, both the company and ERG agreed that the unadjusted analyses were not appropriate for decision making. The committee agreed that the unadjusted analyses would be less robust than the 2-stage adjustment method. It concluded that unadjusted analyses were not suitable for decision making.

The ERG's analysis of post-progression survival times is suitable for decision making

3.27 Having concluded that analyses based on the 2-stage adjustment for treatment switching were appropriate for decision making, the committee discussed the acceleration factor used in the adjustment. The company's base case remained unchanged and used the point estimate of 5.37, but it provided a scenario analysis using the lower bound of the 95% confidence interval (3.23). The company submitted 3 additional analyses exploring the effect of different acceleration factors. The ERG described the limitations of those analyses:

- Including recensoring led to much less follow up and considerably less information on the control arm. This made it unlikely to be useful for decision making.
- Using an acceleration factor of 5.32 included people who had vinflunine, which is not licensed in England.
- Applying an acceleration factor of 5.37 (calculated based on 25 patients who switched after disease progression) to all 40 patients that switched, including 15 who switched at different times, was discussed previously (see section 3.6). The company did not provide more information on the characteristics of these patients.

The committee agreed that only the 5.37 and 3.23 acceleration factors were relevant to the decision, because the company's additional analyses had important limitations. At the clarification stage, the ERG asked the

company to provide the patient-level data and code needed to reproduce and explore the acceleration factor. The company provided the code but not the patient-level data. So, the ERG approximated progression and survival data using outputs from the model, to examine how different acceleration factors affected post-progression survival. The ERG's analysis predicted how long patients who switched treatment would have lived for if they had not switched treatment (the counterfactual). The company considered that the ERG's comparison of post-progression survival estimates was flawed, because it did not adjust the full standard care arm for the 15 patients who switched to an active treatment at different times. The committee would have liked to have seen a comparison of the characteristics of those 15 patients with the 25 who switched after disease progression, but the company did not provide those data. Therefore, it agreed that it was appropriate to consider the ERG's analyses in its decision making.

An acceleration factor of 5.37 is not plausible, and although 3.23 is more appropriate the most plausible value is very uncertain and may be lower

3.28 Using the company's preferred acceleration factor of 5.37, the ERG's analysis (see section 3.27) predicted that patients who switched would otherwise have had shorter post-progression survival than the average patient in the standard care arm. The exact data are confidential and cannot be reported here. Using the lower bound acceleration factor of 3.23, the ERG predicted that patients who switched would otherwise have had similar post-progression survival to the average patient in the standard care arm. It advised that this suggests the company's preferred acceleration factor (5.37) was adjusting survival on the standard care arm too much. This was because it attributed too much post-progression survival benefit to the effect of the new treatment, and too little to potential confounding factors. The company provided analyses that adjusted for several potential confounders such as age, gender and ECOG

performance status, but did not provide enough information to allow the ERG to validate these analyses. Therefore, the ERG's preferred analysis used the lower bound acceleration factor (3.23), which was closer to original acceleration factor applied in the original appraisal (3.86). The committee considered that patients who were offered and accepted a treatment switch were likely to have a relatively good prognosis and post-progression survival compared with the average patient having standard care. Therefore, it agreed that the company's preferred acceleration factor (5.37) produced clinically implausible results in the ERG's analysis, while the ERG's preferred acceleration factor (3.23) produced more plausible results. It acknowledged that 3.23 is closer to the value that was accepted in the original appraisal (3.86). The committee agreed that 5.37 adjusted survival on the standard care arm too much. It recalled the limitations of the 2-stage method (see section 3.6), and noted that the ERG had not been provided with the data to validate the company's 2-stage adjustment in detail. This meant the point estimate acceleration factor (5.37) and its lower bound (3.23) were both subject to the same methodological uncertainties. It also noted that 3.23 was an arbitrary value to use, presented only because it is the confidence interval's lower bound. Therefore, the committee agreed that neither value was robust, but the most plausible acceleration factor is likely to be closer to 3.23 than 5.37, and it could plausibly be even lower than 3.23. It concluded that it would consider analyses using both acceleration factors in its decision making,

but would be mindful that 3.23 was likely to be more plausible than 5.37 and the most plausible value could be even lower.

End of life

Life expectancy for people with urothelial carcinoma is less than 24 months

3.29 The committee considered the advice about life-extending treatments for people with a short life expectancy in NICE's guide to the methods of technology appraisal. For people with locally advanced or metastatic disease who have had platinum-containing chemotherapy, data from the company's model and from the literature showed that median overall survival was much less than 24 months for people having treatment with UK standard care. The clinical experts also agreed that they would expect people with locally advanced or metastatic urothelial carcinoma to live for less than 24 months. The committee concluded that the short life expectancy criterion was met.

Pembrolizumab extends life by at least 3 months, and meets the criteria for end-of-life treatments

3.30 The median overall survival for pembrolizumab in KEYNOTE-045 using the November 2018 cut-off was 10.1 months (95% CI 7.6 to 12.9) compared with 6.2 months (95% CI 5.2 to 7.4) for UK SoC (using a 2-stage method for adjustment). The committee concluded that pembrolizumab would extend life by more than 3 months, and therefore met the end-of-life criteria.

Cost-effectiveness estimates after the appeal

The most plausible ICER for pembrolizumab compared with docetaxel and paclitaxel is likely to be over £50,000 per QALY gained

3.31 After the appeal, the company's base-case deterministic ICER was £43,181 per QALY gained with the revised patient access scheme. The committee noted that the company's preferred assumptions had not changed after the appeal (see section 3.22). The committee took into account its preferred assumptions of:

- Considering both the 3-year and 5-year treatment effect durations from the start of pembrolizumab (see section 3.24).
- Adding retreatment costs at 3 years (see section 3.25).
- Considering analyses using the 2-stage method with acceleration factors of 3.23 and 5.37 to adjust for treatment switching, noting that 3.23 was more likely to be plausible than 5.37 (see sections 3.27 and 3.28).

The committee found the log-logistic, log-normal and generalised gamma overall survival functions all plausible (see section 3.20). Therefore, the ICERs considered for decision making ranged from £44,903 to £58,323 per QALY gained. The committee noted that the higher ICERs in this range were associated with an acceleration factor of 3.23, which it reiterated was more plausible than 5.37 and the most appropriate acceleration factor could be even lower. It concluded that the most plausible ICER was likely to be over £50,000 per QALY gained. It also agreed that most ICERs considered would need to be comfortably below £50,000 per QALY gained for it to be confident that pembrolizumab was cost effective, given the substantial uncertainty in the value of the acceleration factor and treatment effect duration.

Cancer Drugs Fund

Pembrolizumab cannot be recommended in the Cancer Drugs Fund

3.32 The aim of a Cancer Drugs Fund guidance review is to decide whether or not the drug can be recommended for routine use. Pembrolizumab for locally advanced or metastatic urothelial carcinoma in adults who have had platinum-containing chemotherapy will not remain in the Cancer Drugs Fund once the guidance review has been completed (see NICE's guide to the processes of technology appraisal).

Conclusion

Pembrolizumab is not recommended for routine use

3.33 The committee considered that the most plausible ICER was above the range that NICE normally considers a cost-effective use of NHS resources for a life-extending treatment at the end of life. It agreed that there is uncertainty surrounding the acceleration factor estimates (more likely to be 3.23 than 5.37, but may be even lower), and therefore the cost-effectiveness results. So, most ICERs in the range considered for decision making should be substantially below £50,000 per QALY gained (that is, the maximum weight of 1.7 applied to the normal range of maximum acceptable ICERs). Based on the range of ICERs considered in decision making, it concluded not to recommend pembrolizumab for treating locally advanced or metastatic urothelial carcinoma in adults who have had platinum-containing chemotherapy.

Other factors

3.34 No equality issues were identified.

3.35 The company did not highlight any additional benefits that had not been captured in the QALY calculations.

4 Review of guidance

4.1 The guidance on this technology will be considered for review 3 years after publication of the guidance. The guidance executive will decide whether the technology should be reviewed based on information gathered by NICE, and in consultation with consultees and commentators.

Lindsay Smith
Vice chair, appraisal committee
March 2021

5 Appraisal committee members and NICE project team

Appraisal 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 to be appraised. If it is considered there is a conflict of interest, the member is excluded from participating further in that appraisal.

The minutes of each appraisal committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

NICE project team

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

Amy Crossley, Kirsty Pitt, Abitha Senthinathan
Technical leads

Nicola Hay, Lucy Beggs, Jamie Elvidge

Technical advisers

Kate Moore

Project manager

ISBN: [to be added at publication]