



Australian Government

Medical Services Advisory Committee

## Public Summary Document

### ***Application No. 1522 – Programmed Cell Death Ligand 1 (PD-L1) immunohistochemistry testing for access to pembrolizumab as first-line therapy for patients with recurrent (not amendable to local treatment) or metastatic head and neck squamous cell carcinoma – codependent***

**Applicant:** Merck, Sharp and Dohme (Australia) Pty Limited

**Date of MSAC consideration:** MSAC 80<sup>th</sup> Meeting, 26-27 November 2020

Context for decision: MSAC makes its advice in accordance with its Terms of Reference, [visit the MSAC website](#)

#### **1. Purpose of application**

The integrated codependent submission (or applicant-developed assessment report (ADAR)) was received from Merck Sharp & Dohme (Australia) Pty Limited by the Department of Health. The submission requested:

- Medicare Benefits Schedule (MBS) listing of programmed death ligand 1 (PD-L1) testing using the 22C3 antibody assay for the evaluation of combined positive score (CPS)  $\geq 1$  to determine eligibility for access to pembrolizumab treatment in patients with previously untreated recurrent or metastatic (R/M) squamous cell carcinoma (SCC) of the oral cavity, pharynx or larynx; and
- Pharmaceutical Benefits Scheme (PBS) Authority Required listing for treatment with pembrolizumab in patients with previously untreated R/M SCC of the oral cavity, pharynx or larynx in patients who have evidence of Combined Positive Score (CPS)  $\geq 1$ .

#### **2. MSAC's advice to the Minister**

After considering the strength of the available evidence in relation to comparative safety, clinical effectiveness and cost-effectiveness, MSAC did not support the creation of a new Medicare Benefits Schedule (MBS) item for PD-L1 testing in patients with recurrent or metastatic head and neck squamous cell carcinoma (HNSCC) in alignment with the Pharmaceutical Benefits Advisory Committee (PBAC) decision to not recommend pembrolizumab in this codependent submission. MSAC considered the evidence presented did not adequately support the claim of codependency.

### **Consumer summary**

This application was from Merck, Sharp & Dohme (Australia) Pty Limited for public funding via the Medicare Benefits Schedule (MBS) for programmed cell death ligand 1 (PD-L1) immunohistochemistry testing to inform decisions about a patient's suitability for pembrolizumab as first-line treatment of recurrent or metastatic head and neck squamous cell carcinoma.

This application was a codependent application, which means that the applicant also applied to the Pharmaceutical Benefits Advisory Committee (PBAC) to list pembrolizumab on the Pharmaceutical Benefits Scheme (PBS) for those patients shown to be suitable by this immunohistochemistry testing. The PBAC did not recommend the extension of the PBS listing for pembrolizumab to include these patients.

Squamous cell carcinoma of the head or neck starts in the cells inside the mouth, nose or throat. Recurrent means the cancer has come back after being treated, and metastatic means the cancer has spread to other parts of the body.

Immunohistochemistry testing in tumours is a way for doctors to know whether a certain protein is in some tumours and how much of it is there (called 'expression'). Sometimes, if a patient's tumour is expressing certain proteins, some medicines work better, and testing is needed to show that patients are suitable for these medicines. In this application, the PD-L1 test result is reported as a combined positive score (CPS). The applicant stated that those patients with a CPS of 1 or more could be suitable for treatment with pembrolizumab – either on its own or in combination with other therapies.

The Medical Services Advisory Committee (MSAC) considered that there is not enough evidence to suggest that a patient's CPS score is linked to pembrolizumab being a better treatment.

### **MSAC's advice to the Commonwealth Minister for Health**

MSAC did not support the requested MBS funding of programmed cell death ligand 1 (PD-L1) immunohistochemistry testing because the codependent medicine (pembrolizumab) was not recommended by the PBAC as requested, and there was no clear evidence that pembrolizumab works better for those patients with a PD-L1 combined positive score (CPS) of one or more as stated by the applicant.

### **3. Summary of consideration and rationale for MSAC's advice**

MSAC noted that Application 1522 was for a new MBS item for an immunohistochemical (IHC) assay for the evaluation of programmed death ligand 1 (PD-L1) expression in patients with previously untreated recurrent or metastatic (R/M) head and neck squamous cell carcinoma (HNSCC). MSAC also accepted that the HNSCC population had been focussed to SCC of the oral cavity, pharynx and larynx.

The application proposed that the test would report a combined positive score (CPS) to help determine eligibility for pembrolizumab (whether as monotherapy or combination therapy), which was proposed for inclusion on the Pharmaceutical Benefits Scheme (PBS) for those patients who also have a PD-L1 CPS  $\geq 1$ . MSAC noted that the Pharmaceutical Benefits Advisory Committee (PBAC) had not recommended the listing of pembrolizumab for first-line treatment of R/M HNSCC in patients with a PD-L1 CPS score  $\geq 1$  at its November 2020 meeting.

MSAC considered that the biological rationale for the proposed codependence was weak. Reflecting this lack of a cohesive rationale, pembrolizumab is approved by the Therapeutic

Goods Administration (TGA) for the treatment of several different tumour types with varying requirements of PD-L1 positivity, with some indications requiring other biomarkers and some indications not requiring any evidence of any biomarker at all. Notably, the submission to PBAC for second-line pembrolizumab monotherapy for recurrent or metastatic HNSCC (after failure of platinum-based chemotherapy) was not limited to patients whose tumours express PD-L1 ([Pembrolizumab July 2018 PBAC Public Summary Document \[PSD\]](#)). This was based on the 20 March 2017 TGA approval of this second-line indication. Alongside its approval of the first-line HNSCC indication based on the KN048 trial, the TGA subsequently amended the second-line HNSCC indication (based on the KN040 trial) to also require second-line HNSCC patients to have a PD-L1 CPS  $\geq 1$ . MSAC also noted that two main PD-L1 expression scoring systems were CPS scoring which assesses tumour cells and immune cells and tumour positive score (TPS) that assesses tumour cells only.

MSAC noted that there is no reference standard for the clinical validity of PD-L1 testing. Rather, MSAC noted that Emancipator 2020 and Cohen 2019 reported the analytical performance (sensitivity and specificity) of the PharmDX 22C3 antibody used in the KN048 trial using response rates to pembrolizumab as a reference standard for the clinical utility of PD-L1 testing. These were retrospective analyses of R/M HNSCC patients who received second-line pembrolizumab selected from KN012 and KN055 (Emancipator 2020) and KN040 (Cohen 2019). MSAC noted the high sensitivity and low specificity reported for the test at the threshold of CPS  $\geq 1$ , and agreed with the ESCs and the Commentary that the apparently high sensitivity might be due to the test classifying over 80% of patients as being CPS  $\geq 1$  rather than the test accurately identifying patients who will respond to pembrolizumab. The Commentary estimated that 20.3% and 17.4% of the total population responded to pembrolizumab in Emancipator 2020 and Cohen 2019, respectively. MSAC agreed with the Commentary that this indicated there was a poor correlation between the proposed PD-L1 CPS positivity threshold and extent of response to pembrolizumab.

MSAC noted that the main clinical evidence for supporting PD-L1 testing to help determine pembrolizumab eligibility was from the KN048 trial. KN048 recruited patients irrespective of PD-L1 status, and stratification by PD-L1 status was initially based on a tumour positive score (TPS)  $\geq 50\%$ , not CPS  $\geq 1$ . After patient enrolment, but before data analysis, the trial protocol was amended from stratification of PD-L1 status from TPS  $\geq 50$  to CPS  $\geq 1$ . MSAC agreed with the Commentary and the TGA clinical evaluator that the randomised stratification by TPS status in KN048 no longer holds and so the comparisons on the basis of CPS (whether  $\geq 1$  or the alternative threshold explored of  $\geq 20$ ) were effectively non-randomised, and therefore had a higher risk of bias. MSAC also noted the baseline characteristics by CPS subgroup were imbalanced – for example, there were more women and fewer patients with larynx cancer in the pembrolizumab + chemotherapy arm. These characteristics were expected to be independent predictors of variation in prognosis and the extent of effectiveness of pembrolizumab.

MSAC also noted that KN048 further artificially enriched cohorts with poor justification, as a protocol amendment required that the final 180 patients enrolled exhibit ‘strongly positive’ PD-L1 expression (i.e. TPS  $\geq 50\%$ ). As a result, 85.3% of KN048 patients had CPS  $\geq 1$  and the proportion of patients with CPS  $\geq 20$  was also disproportionately increased. MSAC noted the PD-L1 CPS  $> 1$  prevalence of 45.6% in de Ruiter 2019. By comparison, only 22.5% of all patients in the trial were TPS positive. MSAC considered that this also resulted in limited external validity of the trial results, which affected the economic evaluation and likely overestimated the financial estimates.

From the KN048 trial, and in the intention-to-treat population irrespective of PD-L1 status, MSAC accepted that pembrolizumab in combination with chemotherapy was more effective

than the comparator arm, but concluded that greater effectiveness was not convincingly demonstrated for pembrolizumab as monotherapy (see Table 4). This affected the interpretation of the PD-L1-based results, which were conditional on the results of the overall results in the planned statistical analysis. MSAC noted that pembrolizumab monotherapy appeared worse in terms of overall survival for patients with CPS <1, but considered the result was not convincing because of the trial design and conduct problems, which also resulted in the subgroup of patients with CPS <1 being too small to conclude that pembrolizumab had no effect in this subgroup. MSAC queried whether the postulated codependency between PD-L1 status and the clinical benefit from pembrolizumab differed when it was used as monotherapy compared with its use in combination with chemotherapy.

MSAC concluded that the trial's problems due to the change in the definitions of PD-L1 assessment, enrichment of patients with TPS  $\geq 50\%$ , and complex statistical methodology contributed to a complicated trial that was difficult to understand and interpret, and did not fully inform whether PD-L1 CPS testing identified patients most likely to benefit from pembrolizumab treatment. These difficulties are illustrated by the different conclusions drawn by the TGA and the European Medicines Agency compared with the United States' Food and Drug Administration.

MSAC noted the results of the indirect comparisons of pembrolizumab (as monotherapy or with chemotherapy) from KN048 against its comparator of chemotherapy (carboplatin or cisplatin and 5-FU) from the EXTREME trial using the cetuximab + chemotherapy arms of the two trials as the common reference, and the failure to fully account for the clinical effects of second line (2L) nivolumab in the comparator arm.

MSAC noted there was poor concordance between 22C3 antibody used to test PD-L1 positivity in the KN048 trial and other PD-L1 IHC antibodies (SP263, SP142 and 28-8). Additionally MSAC noted that there may be substantial inter observer variability in assays using the 22C3 antibody. The study by de Meulenaere 2018 assessed PD-L1 expression using the 22C3 antibody at the  $\geq 1\%$  level and reported a correlation coefficient of agreement of  $r=0.621$  with only 65% of samples classified consistently by the four observers. MSAC noted that de Meulenaere 2018 did not use CPS scoring as raised in the applicant's pre-ESC and pre-MSAC responses. However, MSAC considered Meulenaere 2018 assessed inter observer variability in a research setting and variability would be expected to be greater in clinical practice whether based on CPS or not.

MSAC noted that there were several outstanding issues relating to the sample used for PD-L1 testing. MSAC noted the ESC's advice that there is evidence of discordant PD-L1 status between resection and core biopsy specimens. MSAC noted the 22C3 antibody has not been validated for use with fine needle aspirates (FNA) as raised in the pre-ESC response. Further to this, MSAC advised that FNA would not be appropriate for CPS testing because CPS testing requires staining of adjacent immune cells which may not be appropriately captured in a FNA.

MSAC considered necessary a mechanism to train pathologists on the interpretation of PD-L1 CPS testing in HNSCC and validation of their competency is required. MSAC noted advice from the National Pathology Accreditation Advisory Council (NPAAC) that PD-L1 CPS testing requires a robust quality assurance framework and few laboratories are currently performing this test.

Overall, MSAC concluded that there was substantial uncertainty regarding the practicalities and accuracy of implementing PD-L1 testing to achieve the estimated extent of pembrolizumab effectiveness in Australian patients with a reported CPS  $\geq 1$ .

MSAC considered there were several issues in the economic model. MSAC considered the economic model to be incorrect in that it assumed that the 22C3 antibody assay would produce no false positive or false negative results (that is, the model assumed 100% sensitivity and 100% specificity). MSAC also considered that the economic model overestimated the prevalence of PD-L1 positivity. MSAC noted the pre-MSAC argument for using a 7.5-year time horizon. The pre-MSAC response stated that a 5-year time horizon did not capture the sustained benefits of pembrolizumab in 15–20% of patients. MSAC agreed with the Commentary and the ESCs that the number of patients remaining at risk is below 50% after 12 months in the KN048 trial, and that survival estimates after this was informed by small patient numbers and therefore unreliable. MSAC also noted that using a time-dependent hazard ratio (HR) from fractional polynomial network meta-analysis instead of a constant HR from the Bucher indirect comparison favoured pembrolizumab at a 7.5-year time horizon, but favoured standard of care at a 5-year time horizon. MSAC maintained that a 5-year time horizon was appropriate. MSAC agreed with the Commentary that the submission's forecast that 80% of pembrolizumab use would be monotherapy may not be reasonable given that pembrolizumab with chemotherapy appears to be more efficacious with regard to improvement in overall survival.

MSAC also considered that the standard of care in the economic evaluation inappropriately excluded the efficacy of second-line nivolumab, but nivolumab was included in the costs. MSAC disagreed with the pre-MSAC response, which argued that the standard of care overall survival (OS) curve reasonably accounts for most benefits of second-line nivolumab in a subpopulation who progressed on platinum-based initial therapy within six months. MSAC agreed with the ESCs and the Commentary, which demonstrated a significantly underestimated OS in the comparator arm, resulting in an underestimated incremental cost-effectiveness ratio (ICER). MSAC also noted that the approach to extrapolation was not justified and was likely implausible with the benefit of pembrolizumab increasing over time.

MSAC considered the evidence presented did not adequately support the claim of codependency. MSAC advised that any future integrated codependent resubmission to both PBAC and MSAC should strengthen the evidence for this claim of codependence, provide further evidence about appropriate sample types for PD-L1 testing, and address the outstanding concerns relating to the economic model and financial estimates.

MSAC recalled a similar application ([Application No. 1440.1](#)) for PD-L1 testing in non-small cell lung cancer (NSCLC) was considered in November 2017, March 2018 and July 2018. MSAC acknowledged similarities between the applications, but considered any precedent from Application 1440.1 to not be completely relevant. Although apparently similar in that the PBS listing of nivolumab as second-line treatment of NSCLC irrespective of PD-L1 status changed the clinical utility consequences of poor PD-L1 test performance, the NSCLC application was based on a stronger trial with a stronger evidence base, the cut-off point between positive and negative PD-L1 results was more persuasively developed (refer to [MSAC Application 1440.1 PSD, July 2018](#)).

#### **4. Background**

IHC testing for the evaluation of PD-L1 expression in patients diagnosed with R/M SCC of the oral cavity, pharynx and larynx has not been previously considered by MSAC. MBS item 72814 was listed on 1 November 2018 for PD-L1 testing of NSCLC patients for eligibility for pembrolizumab. By comparison, Application 1522 is for HNSCC and is restricted to R/M patients.

PASC considered the PICO Confirmation for this application at its April 2018 meeting, followed by a second consideration at its April 2020 meeting.

## 5. Prerequisites to implementation of any funding advice

The PD-L1 IHC 22C3 PharmDX kit is currently approved for use in patients with NSCLC. The submission stated that when the HNSCC indication for pembrolizumab is approved by the TGA, the Australian instructions for the 22C3 PharmDX use will be updated to include the statement:

“PD-L1 IHC 22C3 PharmDX is indicated as an aid in identifying HNSCC (Head and neck squamous cell carcinoma) patients for treatment with KEYTRUDA® (pembrolizumab) at PD-L1 expression level CPS  $\geq 1$ .”

The TGA classifies all in vitro diagnostic (IVD) companion diagnostics as class 3 IVDs<sup>1</sup>.

The National Pathology Accreditation Advisory Council (NPAAC) advised that MSAC consider the following key questions/concerns:

- NPAAC confines its observations to issues related to the quality framework for testing (including accreditation and the existence of external quality assurance programs) and likely access to testing.
- Quality framework issue. Previous NPAAC comments to MSAC on Application 1570 “PD-L1 (Programmed death ligand 1) immunohistochemistry (IHC) testing for access to atezolizumab as first-line therapy for patients with locally advanced or metastatic triple-negative breast cancer (TNBC)” apply to this testing also:
  - It is critical that biomarker testing is offered in a robust quality assurance framework to protect patients against false positive or negative results; and
  - A different antibody clone of PD-L1 is utilized in an immunohistochemistry stain for lung tumours.
- PD-L1 testing for HSNCC is already being performed in a very few laboratories in Australia. The technical aspects of the staining are standard but the varying interpretation, including across tumour types, requires experience. Roche (the applicant to MSAC for Application 1570) has provided an on line module to train pathologists using PD-L1 testing in the reporting of triple negative breast cancer in which pathologists reporting the stain were required to undergo special on-line training and validation prior to accepting cases. In that disease context, pathologists were also double reading the stains to standardize the interpretation. There is an Australian EQA program on interpretation available. There is an existing EQA for the technical performance of the stain offered from the UK/IQNPath (UK+Europe).

## 6. Proposal for public funding

The requested MBS item is shown in Table 1.

---

<sup>1</sup> <https://www.tga.gov.au/publication/ivd-companion-diagnostics>

**Table 1 Requested MBS listing**

Category 6 – Pathology Services
Immunohistochemical examination by immunoperoxidase or other labelled antibody techniques using the 22C3 PD-L1 antibody of tumour material from a patient diagnosed with recurrent or metastatic squamous cell carcinoma of the oral cavity, pharynx or larynx incurable by local therapies to determine if the requirements relating to programmed cell death ligand 1 (PD-L1) status for access to pembrolizumab under the Pharmaceutical Benefits Scheme (PBS) are fulfilled.
Fee: \$74.50 75% = \$55.90 85% = \$63.35

Source: p31 of the submission

The Commentary noted that the requested MBS item descriptor in the submission differed to the descriptor in the [Ratified PICO Confirmation](#) in two ways:

- The proposed MBS item descriptor in the submission specifies that the 22C3 antibody is to be used, whereas the descriptor in the [Ratified PICO Confirmation](#) did not specify the antibody to be used. The justification for this change was other antibodies for PD-L1 testing (e.g. SP263) had weak concordance with 22C3 based on two studies by de Ruiter 2019 and Scott 2018. At  $CPS \geq 1$ , de Ruiter showed 87% positivity for SP263 on the Ventana platform, but only 45.6% for the 22C3 PharmDx kit. Similarly, Scott 2018 found that the overall positive agreement at  $CPS \geq 1$  between the 22C3 and the SP263 tests was only 75%, indicative of moderate agreement; and
- The condition has been changed from any R/M head and neck SCC to specifically R/M SCC of the oral cavity, pharynx or larynx. No justification for this change was provided explicitly but it was likely changed to align the MBS item descriptor to the requested PBS restriction, which would be consistent with the current PBS restriction for second line (2L) nivolumab.

The Commentary suggested that, to be consistent with its recent related considerations, MSAC may wish to consider also supporting the inclusion of the following Explanatory notes:

- The PD-L1 test will be required once only per patient.
- For this tested population, a PD-L1 combined positive score (CPS)  $\geq 1$  is required to fulfil the requirements for access to pembrolizumab under the PBS.

The proposed MBS fee was consistent with the fee for current PD-L1 testing for NSCLC (MBS item 72814).

The submission proposed that PD-L1 expression in HNSCC be measured with the CPS algorithm. CPS measures the number of PD-L1 stained cells, including tumour cells, lymphocytes, and macrophages, divided by the total number of viable tumour cells, multiplied by 100. Despite its derivation, the CPS is expressed as a value and not a percentage, and although the result can exceed 100, the maximum score is defined as CPS 100. A threshold of  $CPS \geq 1$  was nominated as the relevant threshold. This is consistent with the TGA's decision to register pembrolizumab for the following indication:

KEYTRUDA® (pembrolizumab), as monotherapy or in combination with platinum and 5-fluorouracil (5-FU) chemotherapy, is indicated for the first-line treatment of patients with metastatic or unresectable recurrent HNSCC, and whose tumours express PD-L1 [Combined Positive Score (CPS)  $\geq 1$ ] as determined by a validated test.

The submission expects that the majority (95%) of 22C3 antibody assay tests in Australia will be laboratory developed tests (LDTs) conducted on the Ventana BenchMark XT platform.

While this is different to the 22C3 PharmDX kit used in the pembrolizumab clinical trials which requires the Autostainer Link 48 platform, the Commentary noted that there is a high level of concordance between the 22C3 antibody protocols between the two platforms.

The Commentary queried whether the item descriptor should specify the sample to be tested. In KN048, the tissue sample could be obtained by either core biopsy or excision. The Commentary noted that there were two studies showed discrepancies between biopsy samples and resection samples (de Meulenaere 2018) and between core biopsy and excised or resected samples (Paintal 2019, refer to Comparative Effectiveness – Sample Type). The pre-ESC Response considered that limiting to either core biopsy or excision would be consistent with tissue samples used in the KN048 trial and also highlighted that the product information for the 22C3 antibody states that fine needle aspirates have not been validated.

While the need for quality assurance program was mentioned in the submission (p30), no further details were provided. The Commentary highlighted that the Royal College of Pathologists of Australasia (RCPA) noted in the [Ratified PICO Confirmation \(p17\)](#), that there is a need for quality assurance programs and accurate training in the diagnostic evaluation of PD-L1 in the current context. There are challenges due to variability and ambiguity in interpretation as well as the availability of multiple PD-L1 antibodies and variation in the affinity and staining intensity of the tumour cells and immune infiltrate. The pre-ESC response stated that CPS training with Australian pathologists is underway which should reduce concerns regarding intra and inter observer variability with R/M HNSCC PD-L1 testing. Four pathologists have been trained as “train the trainers”, with plans for 24 pathologists to be trained by the end of December 2020.

## **7. Summary of public consultation feedback/consumer Issues**

No consultation feedback was received.

## **8. Proposed intervention’s place in clinical management**

### **Description of proposed intervention**

The proposed medical service is an immunohistochemical (IHC) test for evaluation of PD-L1 expression to determine eligibility for treatment with pembrolizumab in patients with R/M SCC of the oral cavity, pharynx and larynx incurable by local therapies who have not had prior systemic therapy administered in the recurrent or metastatic setting. Tumour material obtained via a resected biopsy or core biopsy as part of usual care to confirm disease progression would be used for immunohistochemical testing using the 22C3 antibody as part of the PharmDX kit or as a laboratory developed test (LDT). The testing would be done by a pathologist alongside other immunohistochemical tests which are done routinely.

The CPS algorithm was nominated to be used to score PD-L1 IHC staining. CPS measures the number of PD-L1 stained cells, including tumour cells, lymphocytes, and macrophages, divided by the total number of viable tumour cells, multiplied by 100. KN048 (the Keynote-048 trial of pembrolizumab presented in the ADAR), used the 22C3 PharmDX kit which requires the Autostainer Link 48 platform.

## Description of medical condition(s)

The proposed population for PD-L1 testing was R/M SCC of the oral cavity, pharynx and larynx.

Currently, patients with R/M SCC of the oral cavity, pharynx and larynx do not undergo PD-L1 testing. Patients currently receive standard of care (SoC), defined as first-line (1L) carboplatin or cisplatin and 5-FU (chemotherapy) followed by second-line (2L) nivolumab. The submission estimated the proportion of patients who would receive nivolumab to be 66.7% based on PBS script data, but the Commentary noted that this could not be verified.

Figure 1 and Figure 2 present the current and proposed treatment algorithms, respectively, from the submission. The Commentary considered the current and proposed algorithms presented in the submission were similar to the algorithms presented in the [Ratified PICO Confirmation](#).

It was proposed that patients diagnosed with R/M SCC of the oral cavity, pharynx or larynx would have their PD-L1 status (either CPS <1 or CPS ≥1) determined using IHC and the most recent biopsy sample. The submission claimed that a sample biopsy is currently routinely conducted in the majority of recurrent patients to confirm recurrence. For patients who are unable to be re-biopsied or who have insufficient fresh tissue for testing, archival tissue from their biopsy at the initial diagnosis would be used for PD-L1 testing.

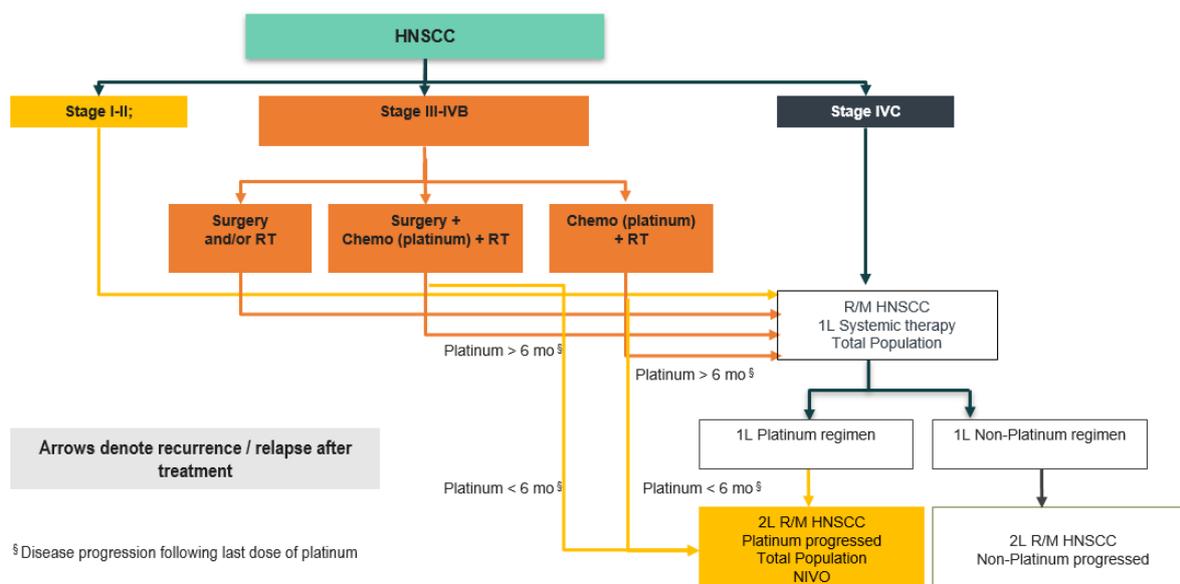
Patients whose tumour is PD-L1 positive (CPS ≥1) would receive either:

- first line (1L) pembrolizumab plus chemotherapy and second line (2L) methotrexate or taxane monotherapy, or
- 1L pembrolizumab monotherapy and 2L platinum plus fluorouracil (5-FU) or platinum single agent chemotherapy or taxane monotherapy.

Patients whose tumour is PD-L1 negative (CPS <1) would receive SoC (chemotherapy), followed by 2L nivolumab.

The Commentary noted that, although the submission presented information stating that the prevalence of CPS ≥1 was similar between archival tissue (n=159) and fresh tissue (n=442) in KN048, it did not provide evidence that the archival or fresh tissue samples were accurately classified. The Commentary considered that an intra-tumour comparison at two different time points (e.g. first sample from tumour block tested as fresh tissue, then a second test on same block tested few months later to once 'fresh tissue' becomes 'archival tissue' and then concordance evaluated) would have been required to support a conclusion that fresh and archival tissue would be interchangeable. The Commentary noted that there is some evidence that PD-L1 expression is dynamic in that it can change over time, and can also differ between primary and metastatic tumours.

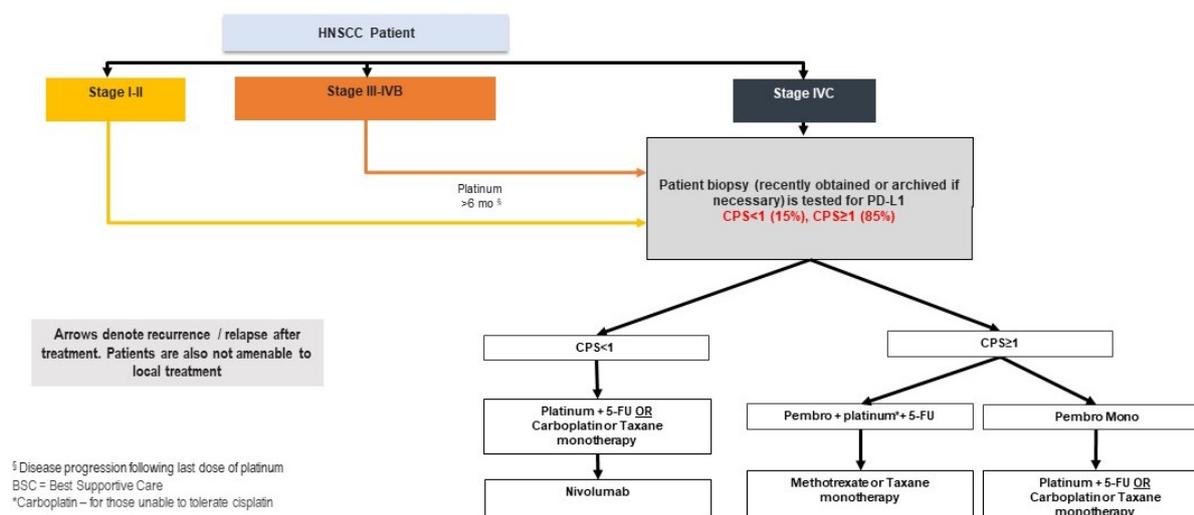
The submission suggested the PD-L1 testing would become a part of the diagnostic workup in patients with newly diagnosed metastatic HNSCC.



**Figure 1 Current treatment algorithm presented by the submission**

Source: Figure 1.2-1, p 9 of the submission

1L Cetuximab monotherapy is PBS listed for patients with stage III, IVa and IVb HNSCC who are contraindicated to or unable to take cisplatin.



**Figure 2 Proposed treatment algorithm presented by the submission**

Source: Figure 1.2-2, p 10 of the submission

## 9. Comparator

No PD-L1 testing was the appropriate nominated comparator for PD-L1 testing.

The submission nominated SoC as the comparator for pembrolizumab treatment. SoC was defined as 1L carboplatin or cisplatin and 5-FU (chemotherapy), followed by 2L nivolumab (for an estimated 66.7% of patients was nominated as the comparator for pembrolizumab. The Commentary considered this was the appropriate comparator for pembrolizumab.

## 10. Comparative safety

The PD-L1 testing strategy (22C3-CPS) used to assess PD-L1 CPS positivity in the KN048 trial was considered the evidentiary standard against which the percent agreement with other testing strategies were measured.

The submission presented a linked evidence approach to support its claim that targeting of patients with R/M HNSCC of the oral cavity, pharynx and larynx, whose tumours have PD-L1 CPS  $\geq 1$  with pembrolizumab will identify patients, who may derive the most benefit from immunotherapy (Table 2).

**Table 2 Summary of the linked evidence approach**

Study type	Type of evidence supplied	Extent of evidence supplied	Overall risk of bias in clinical trials
Accuracy and performance of the test (analytical validity)	No reference standard was identified. Emancipator 2020 included a portion of patients enrolled in KN012 and KN055 (n=252) and aimed to determine whether CPS or TPS is the preferred PD-L1 scoring method in advanced HNSCC, and Cohen 2019 (n=475) did a similar study with the patients in KN040.	<input type="checkbox"/> k=2 n=727	Full QUADAS-2 assessment in Table 2.10-2, p115-116 of the submission. Overall risk of bias was moderate, as there were issues with patient population being different from the requested population, with no information about blinding of assessment and the use of response rates for reference standard.
Prognostic evidence	Four systematic reviews – Tang 2020 (n=1729), Troiano 2019 (n=1060), Yang 2018 (n=3105) and Peng 2017 (n=1777) – which examined the relationship between PD-L1 and survival in HNSCC were identified by the submission.	<input type="checkbox"/> k=3 n=5,894	Risk of bias not assessed by submission. Significant overlap in the included systematic reviews.
Change in patient management	Not explicitly assessed, but CPS thresholds based on KN048 results.	<input type="checkbox"/> k=0 n=0	NA
Treatment effectiveness			
Predictive effect (treatment effect variation)	Based on KN048 subgroups by CPS	<input type="checkbox"/> k=1 n=47	Risk of bias possibly high as the protocol was changed from measuring PD-L1 via TPS to CPS
Treatment effect (enriched)	Based on indirect comparison of KN048 subgroups by CPS and ITT population in EXTREME	<input type="checkbox"/> k=2 n=919	Risk of bias possibly high due to indirect comparison

CPS = combined positive score; HNSCC = head and neck squamous cell carcinoma; k=number of studies, n=number of patients; NA = not applicable; PD-L1 = programmed death ligand 1; QUADAS-2 = Quality Assessment of Diagnostic Accuracy Studies 2; TPS = tumour proportion score

Source: Table 2.1, p96 of the Commentary

### *Adverse events from testing*

As PD-L1 testing is intended to be done on tissue previously biopsied for confirmation or diagnosis of recurrence or metastasis, it was not anticipated that there will be any additional adverse events from testing.

### *Adverse events from changes in clinical management*

The submission presented safety data from:

- patients treated with pembrolizumab  $\pm$  chemotherapy and cetuximab plus chemotherapy from KN048; and
- patients treated with cetuximab plus chemotherapy and chemotherapy alone from EXTREME.

However, the submission did not present relevant comparative safety between pembrolizumab  $\pm$  chemotherapy with SoC. The submission presented an indirect comparison of adverse events between pembrolizumab  $\pm$  chemotherapy with chemotherapy alone. The

Commentary did not consider this comparison appropriate or informative comparison as grade 3-5 drug related adverse events in patients treated with pembrolizumab ± chemotherapy in KN048 were compared to all cause grade 3-4 adverse events in patients treated with chemotherapy alone in EXTREME. No treatment related grade 3-4 adverse events were reported for EXTREME, therefore a meaningful comparison could not be conducted. Additionally, there was no consideration of safety compared with 2L nivolumab.

A naive comparison of all cause grade 3-5 adverse events in KN048 and EXTREME was conducted during the evaluation and is presented in Table 3.

**Table 3 Naïve indirect comparison of all cause grade 3-5 adverse events pembrolizumab monotherapy and pembrolizumab plus chemotherapy with chemotherapy alone**

	KN048		EXTREME	Pembro mono vs chemo alone		Pembro + chemo vs chemo alone	
	Pembro mono	Pembro + chemo	Chemo alone	RR (95%CI)	RD% (95%CI)	RR (95%CI)	RD% (95%CI)
All grade 3-5 AE	164/300 (54.7)	235/276 (85.1)	164/220 (76)	0.73 (0.64, 0.83)	-20 (-28, -12)	1.14 (1.04, 1.25)	11 (3, 18)
<i>All grade 3-5 AE (corrected in Pre-ESC response)</i>	164/300 (54.7)	235/276 (85.1)	171/215 (79.5) <sup>a</sup>	0.69 (0.61, 0.78)	-25 (-32, -17)	1.07 (0.98, 1.16)	6 (-1, 13)
Drug-related grade 3-5 AE	52/300 (17.0)	198/276 (71.7)	NR	NC	NC	NC	NC

AE = adverse events, pembro = pembrolizumab, mono = monotherapy, chemo = chemotherapy, RR = relative risk, RD = risk difference, NR = not reported, NC = not calculable

Text in italics indicate values updated from the pre-ESC response and recalculated by the evaluation. Figures in bold indicate statistically significant differences.

Source: Table MSAC.4, p15 of the Commentary

<sup>a</sup> Includes 7 deaths reported in Vermorken 2008 (Grade 5)

The Commentary concluded that, using all cause grade 3-5 adverse events (as treatment related adverse events were not reported in EXTREME), pembrolizumab plus chemotherapy had a higher incidence of grade 3-5 adverse events than chemotherapy alone, which was plausible and expected as:

- the PBAC has previously considered that cetuximab plus chemotherapy was inferior in safety compared to chemotherapy alone (para 6.23, p14 Cetuximab PSD March 2018), and given that KN048 showed that pembrolizumab plus chemotherapy was comparable to cetuximab plus chemotherapy in safety, it would be logical to expect that pembrolizumab plus chemotherapy was also inferior in safety to chemotherapy alone; and
- the addition of pembrolizumab to chemotherapy should result in an increased benefit as well as adverse events.

The pre-ESC response claimed the analysis in the Commentary excluded 7 reported deaths and had a different denominator to the EXTREME trial publication (Vermorken 2008). The pre-ESC response claimed that based on its recalculated relative risk, there was no statistically significant differences between pembrolizumab with chemotherapy with chemotherapy alone.

The Commentary considered the conclusion that pembrolizumab monotherapy was superior in safety to chemotherapy was reasonable, however the magnitude of safety benefit was

likely overestimated by the submission as the same incidences of adverse events (all cause instead of treatment related) were not considered. Additionally, the Commentary highlighted that the safety profile of pembrolizumab monotherapy (mainly immune related such as hypothyroidism or pneumonitis) was also significantly different to the profile of chemotherapy alone (more commonly associated with blood chemistry such as neutropenia and anaemia).

The Commentary highlighted that false positives from CPS testing would result in patients who should be treated with SoC being exposed to pembrolizumab. In patients using pembrolizumab monotherapy instead of SoC, there may be benefits in safety as the submission claims that pembrolizumab monotherapy is superior in safety compared to SoC. In patients using pembrolizumab plus chemotherapy, patients would experience additional adverse events associated with pembrolizumab, but in the proportion of patients treated with 2L nivolumab, the adverse events associated with pembrolizumab may simply replace the adverse events from nivolumab as both medicines belong to the same class (anti PD-1/PD-L1 agents). Additionally, false-positive patients would be denied treatment with PBS-subsidised 2L nivolumab as the current restrictions restricts use to patients who have not received prior treatment with a programmed cell death-1 (PD-1) inhibitor for this condition.

False negatives from CPS testing would result in patients who would benefit from pembrolizumab therapy being treated with SoC instead. In patients who would have been treated with pembrolizumab monotherapy, a false negative would expose patients to adverse events associated with chemotherapy (such as neutropenia and anaemia) rather than immune related adverse events (such as hypothyroidism and pneumonitis).

## **11. Comparative effectiveness**

### *Prognostic evidence*

The submission claimed that there was no conclusive evidence that there is any relationship between PD-L1 expression and prognosis in HNSCC based on a summary of four meta-analyses which considered the relationship between PD-L1 and prognosis in patients with HNSCC. The Commentary considered that this may not be a reasonable conclusion. The Commentary highlighted that Yang 2018, the largest of the systematic reviews by number of patient and studies, reported that:

- the only study in pharyngeal SCC (Vassilakopoulou 2016, n=260), high PD-L1 expression (>59th percentile in automated quantitative protein analysis) was associated with longer OS in multivariate analysis (HR=0.570, 95% CI: 0.333, 0.973, P=0.039); and
- patients with positive PD-L1 expression had improved PFS compared to patients with negative PD-L1 expression (HR=0.71, 95% CI: 0.55, 0.93, P=0.01) across six studies.

### *Predictive evidence*

The overall survival of patients by PD-L1 CPS status in KN048 is reported in Table 4. CPS  $\geq 1$  was proposed as the threshold for eligibility, as analysis of data from the phase 1b study KN012 demonstrated that the CPS scoring was more sensitive than TPS scoring and was more predictive of improved overall survival (OS) and progression free survival (PFS).

**Table 4 OS results for ITT and subgroup by PD-L1 CPS status in KN048**

Population	Pembrolizumab		Cetuximab + chemotherapy		Hazard ratio (95%CI)	P-value	Test for interaction p-value (I <sup>2</sup> )
	Deaths, n/N (%)	Median OS (95% CI)	Deaths, n/N (%)	Median OS (95% CI)			
<b>Pembrolizumab monotherapy</b>							
ITT	237/301 (78.7)	11.5 (10.3, 13.4)	264/300 (88.0)	10.7 (9.3, 11.7)	0.83 (0.70, 0.99) <sup>a</sup>	0.01985 <sup>b</sup>	NA
CPS <1	40/44 (90.9)	7.9 (4.7, 13.6)	35/45 (77.8)	11.3 (9.1, 15.9)	1.72 (1.06, 2.79)	0.029	<b>0.002 (89.80%)</b>
CPS ≥1	197/257 (76.7)	12.3 (10.8, 14.3)	229/255 (89.8)	10.3 (9.0, 11.5)	<b>0.74 (0.61, 0.90)</b>	<b>0.003</b>	
CPS <20	142/167 (85.0)	10.3 (8.4, 12.1)	153/175 (87.4)	10.3 (9.1, 12.2)	1.05 (0.84, 1.32)	0.667	<b>0.002 (89.86%)</b>
CPS ≥20	94/133 (70.7)	14.8 (11.5, 20.6)	108/122 (88.5)	10.7 (8.8, 12.8)	<b>0.58 (0.44, 0.78)</b>	<b>&lt;0.001</b>	
<b>Pembrolizumab plus chemotherapy</b>							
ITT	213/281 (75.8)	13.0 (10.9, 14.7)	247/278 (88.8)	10.7 (9.3, 11.7)	<b>0.72 (0.60, 0.87)<sup>a</sup></b>	<b>0.00025</b>	NA
CPS <1	36/39 (92.3)	11.3 (9.5, 14.0)	34/43 (79.1)	10.7 (8.5, 15.9)	1.18 (0.73, 1.90)	0.498	<b>0.025 (80.18%)</b>
CPS ≥1	177/242 (73.1)	13.6 (10.7, 15.5)	213/235 (90.6)	10.4 (9.1, 11.7)	<b>0.65 (0.53, 0.80)</b>	<b>&lt;0.001</b>	
CPS <20	128/154 (83.1)	11.8 (10.4, 14.0)	146/165 (88.5)	10.2 (8.9, 12.1)	0.83 (0.65, 1.05)	0.123	0.109 (61.18%)
CPS ≥20	84/126 (66.7)	14.7 (10.3, 19.3)	98/110 (89.1)	11.0 (9.2, 13.0)	<b>0.60 (0.45, 0.82)</b>	<b>&lt;0.001</b>	

<sup>a</sup> Based on Cox regression model with Efron's method of tie handling with treatment as a covariate stratified by ECOG, HPV status and PD-L1 status. In case the event count in any stratum is <5, stratification factors will be eliminated in the order of ECOG->HPV status->PD-L1 status until event count in every stratum is ≥5

<sup>b</sup> p-value boundary for statistical significance was 0.0059

Text in bold indicates statistically significant results

CPS = combined positive score, ITT = intention to treat, NA = not applicable, OS = overall survival, PD-L1 = programmed death ligand 1  
Source: Table 1, p9; Table PBAC.9, p43 of the Commentary

The Commentary considered that there is evidence from KN048 that patients with CPS ≥1 who were treated with pembrolizumab monotherapy had a statistically significantly lower risk of death compared to patients treated with cetuximab plus chemotherapy, but such a relationship was not observed in the ITT population.

Patients treated with pembrolizumab plus chemotherapy had a statistically significantly lower risk of death compared to patients treated with cetuximab plus chemotherapy in both the CPS ≥1 and ITT populations. Consequently, the Commentary considered that there may not be a valid reason to restrict use of pembrolizumab + chemotherapy to patients with CPS ≥1.

The Commentary also highlighted several factors that may affect the interpretation of PD-L1 status and response to pembrolizumab:

- the CPS <1 subgroup in KN048 was small (15% of the total trial population) and therefore lacked statistical power;
- the evidence between CPS ≥1 and response to pembrolizumab based on clinical trials may be circumstantial as the submission did not provide a biological explanation as to why CPS but not TPS was correlated to response to pembrolizumab;
- KN048 enrolment was stratified by TPS and not CPS, therefore any results based on CPS may be subject to bias; and

- cetuximab plus chemotherapy is not the standard of care in Australia, therefore the results from KN048 cannot be directly applied to the Australian population.

### *Comparative analytical performance*

No reference standard for PD-L1 testing in R/M HNSCC is available. However, two studies, Emancipator 2020 and Cohen 2019, evaluated the sensitivity and specificity of the PharmDX 22C3 antibody in pembrolizumab clinical trials using response rates as the reference standard (Table 5).

**Table 5 Sensitivity and specificity of PharmDX 22C3 assay versus response rate**

Cut off	Emancipator 2020			Cohen 2019		
	Sensitivity	Specificity	Prevalence (N=252)	Sensitivity	Specificity	Prevalence (N=244)
CPS $\geq$ 1	0.922	0.189	NR (83.3%)	0.94	0.23	195 (80%)
TPS $\geq$ 1%	0.745	0.323	NR (69.0%)	0.64	0.44	140 (57%)
CPS $\geq$ 20	0.667	0.542	NR (50.0%)	0.58	0.64	96 (39%)
TPS $\geq$ 20%	0.627	0.592	NR (45.2%)	0.56	0.64	94 (39%)
CPS $\geq$ 50	0.431	0.741	NR (29.4%)	0.50	0.78	64 (26%)
TPS $\geq$ 50%	0.412	0.756	NR (27.8%)	0.47	0.77	65 (27%)

Source: Table MSAC.9, p21 of the Commentary

CPS = combined positive score; TPS = tumour proportion score

The submission claimed that the analytical validity of the PD-L1 test is supported by the high sensitivity (>90%) at CPS  $\geq$ 1, meaning that the majority of patients who respond are detected.

The Commentary considered that this may be due to be a result of the high prevalence as estimated by the clinical trials in the submission (>80% were considered to be CPS  $\geq$ 1) while the overall rate of response in immunotherapy monotherapy was approximately 20% (reconstructed during the evaluation using results from Emancipator 2020 and Cohen 2019). The Commentary considered that this indicated that there was likely a poor correlation between PD-L1 positivity and response.

The Commentary considered that both Emancipator 2020 and Cohen 2019 also showed that the test had a low positive predictive value for CPS  $\geq$ 1 (22.4% in Emancipator 2020 and 20.5% in Cohen 2019). This indicated that the high sensitivity is due to the test classifying a large proportion of patients as CPS  $\geq$ 1, rather than the test being accurate at predicting the patients who may be responders. This was also evident in the low specificity of the test, which showed the test was not particularly accurate at identifying patients who were non-responders.

### *Concordance of PD-L1 IHC tests*

The submission claimed there is high concordance between the 22C3 PharmDX kit and the 22C3 LDT, the likely test to be implemented in the Australian setting. Vainer 2019, one of three studies presented in the submission, demonstrated that the concordance between the 22C3 PharmDX assay used in KN048 and 22C3 LDT on the Ventana Benchmark XT platform, which can be replicated by pathologists in Australia as most laboratories have Ventana platforms, was high, with 96% overall percentage agreement. This suggests that a 22C3 LDT on the Ventana platform would be the preferred approach to PD-L1 testing in HNSCC in Australia. The Commentary considered that this was reasonable, however, noted it may be inappropriate to assume 100% sensitivity and specificity in the economic model

presented, given that an alternate assay, the 22C3 LDT will be more commonly used in Australia.

The Commentary concluded that, based on the studies presented in the submission (de Ruiter 2019, Scott 2018 and Vainer 2019) and another included during the evaluation (de Meulenaere 2018), that the concordance in HNSCC between 22C3 and the alternative antibodies SP263 or SP142 was not high, though the concordance between 22C3 and the alternative antibody 28-8 appears to be higher. The Commentary queried whether the item descriptor should be limited to the 22C3 antibody assay.

The Commentary highlighted that the submission has not proven that 22C3 is better at predicting response to pembrolizumab compared to the other antibodies, simply that the classification of PD-L1 status using the various antibodies do not have a high degree of agreement.

#### *Observer variability*

The submission claimed that there was robust intra and inter observer variability based on three studies (overall per cent agreement  $\geq 88\%$ ). The Commentary also presented data from De Meulenaere 2018 as it provided data on inter observer variation for the 22C3 assay between four observers across 99 samples. At the  $\geq 1\%$  cut off, the correlation coefficient of agreement between all four was reported as  $r=0.621$ , with only 65% of samples being classified consistently by all four observers as PD-L1 positive or negative. The Commentary considered that this provided further evidence that there is variability between observers which may be relevant to clinical practice.

The pre-ESC response considered that the findings from De Meulenaere 2018 were not reasonable to assess inter-observer reliability for CPS scoring using the 22C3 antibody as the study used tumour cell scoring at the  $\geq 1\%$  level, not CPS scoring.

#### *Prevalence*

The submission estimated the prevalence of CPS  $\geq 1$  in patients with R/M HNSCC, of the oral cavity, pharynx or larynx, to be around 85.2% based on the proportion of patients in KN048. The TGA evaluator (TGA CER p38-39) also noted that the proportion of patients with CPS  $< 1$  was likely to be higher in the Australian population than in the KN048 population, particularly as KN048 had artificially enriched cohorts with protocol amendment 01 specifying that the final 180 patients enrolled must exhibit 'strongly positive' PD-L1 expression (i.e. TPS  $\geq 50\%$ ). The Commentary noted that if the prevalence of CPS  $\geq 1$  be lower than expected, the number of patients treated with pembrolizumab will be lower than expected, and the cost per test per positive patient would increase.

#### *Change in clinical management in practice*

It was anticipated that 85% of all patients with R/M HNSCC of the oral cavity, pharynx and larynx, ECOG PS 0-1 and CPS  $\geq 1$  will be treated with pembrolizumab  $\pm$  chemotherapy. The Commentary considered there may be potential for leakage into the CPS  $< 1$  population, as there is evidence that pembrolizumab + chemotherapy is more effective in KN048 in all patients irrespective of PD-L1 status, *however this would be outside the proposed TGA indication which is limited to CPS  $\geq 1$* . The submission assumed that 80% of patients will be treated with pembrolizumab monotherapy and 20% will be treated with pembrolizumab plus chemotherapy, based on expert opinion. The Commentary highlighted that this may not be reasonable given that pembrolizumab plus chemotherapy appears to be more efficacious with regard to improvement in overall survival.

### *Tissue sample*

The Commentary highlighted that there were other uncertainties with regards to PD-L1 testing using the 22C3 antibody in R/M HNSCC.

- In KN048, all enrolled patients must have had excisional or core biopsy available for PD-L1 testing. However, there appears to be variation in results depending on whether excisional or core biopsy samples were used. In Paintal 2019, who obtained samples from 20 patients with HNSCC who had undergone fine needle aspiration (FNA) (n=10) or core biopsy of a metastatic lesion (n=10), followed by an excisional biopsy (n=9) or resection (n=11), the estimated sensitivity and specificity of core biopsy compared to excised or resected tissue was 77% and 100% respectively, with the core biopsy identifying seven of the nine positive samples and one of the one negative samples as classified by the excised or resected tissue. Additionally, de Meulenaere 2018 also reported that there was poor agreement for the 22C3 antibody between biopsy samples and resection samples (Cohen's  $\kappa=0.175$ ) from 15 samples in oropharyngeal SCC. This demonstrates that there may have been heterogeneity within the KN048 PD-L1 classifications depending on the type of tissue samples tested, which creates uncertainty within the PD-L1 classification presented in KN048. The MSAC may wish to consider whether the MBS item number should clearly specify which samples (i.e. resection and/or core biopsy) need to be used.
- It was unclear if PD-L1 status identified with archival tissue would result in the same classification as using fresh tissue. The only evidence provided by the submission was that the proportion of CPS  $\geq 1$  using fresh tissue was similar to using archival tissue in KN048 (86% vs 83%), which was insufficient evidence to conclude that using fresh and archival tissue would result in the same classification in the same patient.

### *Clinical effectiveness of pembrolizumab*

The submission relied on an indirect comparison of the results from KN048 with the EXTREME trial (n=442), which compared cetuximab plus chemotherapy with chemotherapy alone in patients with R/M HNSCC of the oral cavity, pharynx and larynx, using cetuximab plus chemotherapy as the common comparator.

## **12. Economic evaluation**

The submission presented two modelled cost-effectiveness analyses that separately compared (i) pembrolizumab monotherapy with SoC and (ii) pembrolizumab plus chemotherapy with SoC in a population of patients with previously untreated R/M HNSCC of the oral cavity, pharynx or larynx. Outputs from the model in all patients and in the CPS  $\geq 1$  subpopulation were estimated and then used to inform a weighted ICER based on test accuracy and prevalence of CPS  $\geq 1$ .

The Commentary considered that it was unreasonable to assume that there were no false positive or false negatives as:

- concordance between the 22C3 PharmDX kit used in the KN048 trial and the 22C3 LDT conducted on the Ventana BenchMark XT platform (most commonly used platform in Australia) was high but not perfect (96% in published literature); and
- there may be inter-observer disagreement in the 22C3 assay as reported by de Meulenaere 2018, indicating that there may be significant possibility of false positives and negatives in clinical practice.

Therefore the Commentary considered that the base case model was likely optimistic with regards to testing accuracy and underestimates the true ICER.

The Commentary also considered that the economic evaluation did not accurately portray the comparative efficacy between pembrolizumab ± chemotherapy with SoC as the efficacy of 2L nivolumab was not considered, though the costs of 2L nivolumab were inconsistently applied. Therefore, the economic evaluation as provided in the submission should be considered as incomplete and likely uninformative.

Table 6 presents the results of the economic evaluation.

**Table 6 Results of the economic evaluation in all patients including prevalence data**

Component	Pembrolizumab ± chemotherapy	SoC	Increment
Pembrolizumab monotherapy vs SoC			
Cost	\$redacted <sup>3</sup>	\$30,972	\$redacted
QALY	1.28	0.83	0.45
Incremental cost per QALY gained			\$redacted <sup>1</sup>
Pembrolizumab plus chemotherapy vs SoC			
Cost	\$redacted <sup>3</sup>	\$30,679	\$redacted
QALY	1.35	0.76	0.58
Incremental cost per QALY gained			\$redacted <sup>1</sup>
Weighted pembrolizumab monotherapy (80%) and pembrolizumab plus chemotherapy (20%) <sup>a</sup>			
Cost	\$redacted <sup>3</sup>	\$30,914	\$redacted
QALY	1.30	0.82	0.48
Weighted incremental cost per QALY gained			\$redacted <sup>1</sup>

<sup>a</sup> values calculated as weighted 80% of pembrolizumab monotherapy cost, life year and quality adjusted life year from pembrolizumab monotherapy model and 20% of pembrolizumab plus chemotherapy model

LY = life year, QALY = quality adjusted life year

Source: Section 3 Diag. Sens. Wkbk.xlsx using inputs from Section 3 workbook.xslm

The redacted values correspond to the following ranges:

<sup>1</sup> \$55,000 to < \$75,000/QALY gained.

The base case results from the economic model as well as key univariate sensitivity analyses around testing thresholds and accuracy of the test are presented in Table 7. The Commentary noted that the model was more sensitive to the changes in specificity than sensitivity of the PD-L1 test.

**Table 7 ICERs and considerations of various PD-L1/pembrolizumab funding scenarios**

Univariate analyses	Model/comparison	Incremental costs	Incremental effectiveness (QALY)	Incremental cost-effectiveness (\$/QALY)	Percent change from base case
<b>Base case</b>					
Pembrolizumab treated population only (CPS ≥1 only)	Pembro mono	\$redacted	0.53	\$redacted <sup>1</sup>	-
	Pembro Chemo	\$redacted	0.69	\$redacted <sup>1</sup>	-
	Weighted			\$redacted <sup>1</sup>	-
<b>Sensitivity analyses</b>					
No CPS testing in model (base case CPS ≥1)	Pembro mono	\$redacted	0.41	\$redacted <sup>1</sup>	+17.9%
	Pembro Chemo	\$redacted	0.58	\$redacted <sup>1</sup>	+13.1%
	Weighted			\$redacted <sup>1</sup>	+16.6%
CPS threshold increased to CPS ≥20 in model (base case CPS ≥1)	Pembro mono	\$redacted	0.74	\$redacted <sup>1</sup>	-13.0%
	Pembro Chemo	\$redacted	0.98	\$redacted <sup>2</sup>	-16.3%
	Weighted			\$redacted <sup>2</sup>	-13.8%
Whole R/M HNSCC population and testing in model (assumes pembro for CPS ≥1 population, SOC for CPS <1)	Pembro mono	\$redacted	0.45	\$redacted <sup>1</sup>	-
	Pembro Chemo	\$redacted	0.58	\$redacted <sup>1</sup>	-
	Weighted			\$redacted <sup>1</sup>	-
Use results from Cohen 2019 (94.0% sensitivity and 23% specificity) <sup>a</sup>	Pembro mono	\$redacted	0.39	\$redacted <sup>1</sup>	+12.7%
	Pembro Chemo	\$redacted	0.55	\$redacted <sup>1</sup>	+9.7%
	Weighted			\$redacted <sup>1</sup>	+11.9%
Sensitivity of PD-L1 test used in Australia assumed to be 75% of that used in KN048 (base case sensitivity 100%) <sup>b</sup>	Pembro mono	\$redacted	0.34	\$redacted <sup>1</sup>	+0.1%
	Pembro Chemo	\$redacted	0.44	\$redacted <sup>1</sup>	+0.1%
	Weighted			\$redacted <sup>1</sup>	+0.1%
Specificity of PD-L1 test used in Australia assumed to be 75% of that used in KN048 (base case specificity 100%) <sup>b</sup>	Pembro mono	\$redacted	0.44	\$redacted <sup>1</sup>	+4.2%
	Pembro Chemo	\$redacted	0.58	\$redacted <sup>1</sup>	+3.3%
	Weighted			\$redacted <sup>1</sup>	+4.0%

<sup>a</sup> Cohen 2019 evaluated the sensitivity and specificity of the PharmDX 22C3 antibody in the KN040 trial (n=475) using response rates as the reference standard.

<sup>b</sup> Submission results could not be duplicated during the evaluation.

NB. Weighted assuming 20% patients receive pembrolizumab + chemotherapy and 80% receive pembrolizumab monotherapy.

Abbreviations: QALY = quality-adjusted life-year; SOC = standard of care, CPS = combined positive score, R/M HNSCC = recurrent or metastatic head and neck squamous cell carcinoma, PD-L1 = programmed death ligand 1

Text in italics indicate values calculated or extracted or corrected during the evaluation.

Source: Table MSAC.5, pp17-18; Table MSAC.15, p28; Table 3.7.2, p187 of the Commentary

The redacted values correspond to the following ranges:

<sup>1</sup> \$55,000 to < \$75,000/QALY gained

<sup>2</sup> \$45,000 to < \$55,000/QALY gained

### 13. Financial/budgetary impacts

The submission used an epidemiological approach to estimate the expected cost to the MBS of listing the PD-L1 test. The submission assumed that PD-L1 testing will be conducted as a routine part of the diagnostic workup, all newly diagnosed patients with HNSCC, specifically of the oral cavity, pharynx and larynx, were considered to have one PD-L1 test. This was revised to reflect the proposed MBS item descriptor which was limited to patients with R/M SCC of the oral cavity, pharynx and larynx will have their biopsy tested for PD-L1 CPS status. The estimated use of the PD-L1 testing and the costs to the MBS are presented in Table 8.

**Table 8 Estimated use and financial implications**

		Year 1	Year 2	Year 3	Year 4	Year 5	Year 6
<b>Estimated extent of use of PD-L1 testing</b>							
	Projected SCC of the oral cavity, pharynx, larynx cancers exclude nasopharyngeal	redacted <sup>1</sup>					
A	R/M SCC of the oral cavity, pharynx, larynx cancers (excludes nasopharyngeal)	redacted <sup>1</sup>					
B	Cost of CPS testing (\$59.60 × A) <sup>a</sup>	\$redacted <sup>2</sup>					
<b>Estimated financial implications of administration costs</b>							
C	Total administration costs of pembrolizumab ± chemotherapy	\$redacted <sup>2</sup>					
D	Total administration costs offset from medicines not used	\$redacted <sup>2</sup>					
E	Net administration cost (C – D)	\$redacted <sup>3</sup>					
<b>Net financial implications</b>							
F	Net cost to MBS (B + E)	\$redacted <sup>3</sup>					

<sup>a</sup> \$59.60 calculated as 80% benefit of \$74.50

Source: Table MSAC.16, p28 of the Commentary.

Note: Estimates were revised during the evaluation using information from Section 4 Workbook (1L HNSCC).xlsx

The redacted values correspond to the following ranges:

<sup>1</sup> 500 to < 5,000

<sup>2</sup> \$0 to < \$10 million

<sup>3</sup> net cost saving

The Commentary considered the main uncertainty with the MBS financial estimates related to the estimated usage of pembrolizumab (which is dependent on the prevalence of CPS ≥1 and assumed uptake rate, among other factors), as well as the proportion of 2L nivolumab offset by the listing of pembrolizumab.

#### 14. Key issues from the ESCs for MSAC

ESCs key issue	ESCs advice to MSAC
Rationale for codependency	The ESCs advised that the rationale for codependency was not strong for PD-L1 expression (CPS ≥1) and overall survival outcomes with pembrolizumab. The ESCs noted that the Therapeutic Goods Administration (TGA) indication for pembrolizumab was for the first-line treatment of patients with metastatic or unresectable recurrent HNSCC, and whose tumours express PD-L1 (CPS ≥1). However, the ESCs highlighted that the KN048 trial demonstrated that pembrolizumab with chemotherapy had a statistically significant overall survival benefits in the ITT analysis that included patients with CPS <1. While the CPS <1 subgroup appeared to gain no benefit from treatment with pembrolizumab plus chemotherapy in KN048, the sample size was small and therefore lacked statistical power to demonstrate a difference.
Item descriptor	The ESCs noted that the proposed item descriptor was for squamous cell carcinoma (SCC) of the oral cavity, pharynx or larynx while the TGA indication for pembrolizumab is for head and neck squamous cell carcinoma (HNSCC). The ESCs noted that the requested item descriptor was consistent with the existing PBS listing for nivolumab and the requested restriction for pembrolizumab.  The ESCs supported limiting PD-L1 testing to the 22C3 antibody in the item descriptor as there was poor concordance between 22C3 and other antibodies used in PD-L1 testing.  MSAC may wish to consider whether the item descriptor should clearly specify the sample type (resection and/or core biopsy), noting there is some evidence there may be differences in PD-L1 status between resection and core biopsy specimens. The pre-ESC Response also highlighted that the product information for the 22C3 antibody states that fine needle aspirates have not been validated and considered that limiting to either core biopsy or excision would be consistent with tissue samples used in the KN048 trial.
Inter-operator variability	The ESCs considered that there may be substantial inter-operator variability in the classification of PD-L1 status using the 22C3 assay based on the findings of de Meulenaere 2018. The ESCs noted the pre-ESC response's concerns that de Meulenaere 2018 assessed tumour cell scoring at the ≥1%

ESCs key issue	ESCs advice to MSAC
	threshold, rather than CPS scoring, and therefore not relevant to support inter-observer variability with CPS scoring at the $\geq 1$ level. However, this still represented indirect evidence as the submission did not provide a biological explanation why CPS rather than TPS was correlated to pembrolizumab response as CPS was selected based on the results of a non-randomised Phase 1b pembrolizumab study.
Economic issues	The ESCs also considered the assumption that the 22C3 antibody assay has 100% sensitivity and 100% specificity as the evidentiary standard in the economic model was implausible due to issues with inter-observer variability and imperfect concordance between the testing platform used in the KN048 trial and the Ventana BenchMark XT platform most commonly used in Australia. The ESCs highlighted that test performance parameters from Cohen 2019 (94.0% sensitivity and 23% specificity) increased the weighted ICER by 11.9%.

## ESCs discussion

The ESCs noted that the proposed MBS item descriptor is for the PD-L1 CPS testing of squamous cell carcinoma (SCC) of the oral cavity, pharynx or larynx while the TGA indication for pembrolizumab is for head and neck squamous cell carcinoma (HNSCC). The ESCs noted that the requested item descriptor was consistent with the existing PBS listing for nivolumab and the requested restriction for pembrolizumab and considered this appeared appropriate.

The ESCs considered the submission's request to limit PD-L1 testing to the 22C3 antibody in the MBS item descriptor was appropriate as there was poor concordance between this and other available antibodies used in PD-L1 testing.

The ESCs suggested that the proposed MBS item descriptor should specify that PD-L1 CPS testing should occur once per cancer recurrence or metastasis, rather than once per patient as suggested by the Commentary.

The ESCs considered that it may be appropriate for the MBS item descriptor to specify the sample type that should be used for testing. The ESCs noted that there was some evidence to suggest poor agreement for the 22C3 antibody between biopsy samples and resection samples and therefore there could be variation in PD-L1 CPS status depending on the type of sample tested. The ESCs noted that KN048 trial tested excisional or core biopsy samples and agreed with the Commentary that there may have been heterogeneity within the PD-L1 classifications in KN048 depending on the type of tissue samples tested. The ESCs considered this creates uncertainty within the PD-L1 classification presented in KN048. The ESCs noted that the pre-ESC Response stated that limiting to either core biopsy or excision would be consistent with tissue samples used in the KN048 trial and also highlighted that the product information for the 22C3 antibody states that fine needle aspirates have not been validated. The ESCs considered that any additional evidence regarding agreement in PD-L1 CPS status between biopsy and resection samples may be informative for MSAC.

The ESCs agreed with the submission that a fresh sample was preferred, though positivity and negativity rates for PD-L1 were similar for archival and fresh tissue in KN048. The ESCs agreed with the Commentary that the evidence presented in the submission suggested that the prevalence of CPS  $\geq 1$  was similar between fresh and archival tissue, but did provide evidence that the archival or fresh tissue samples were accurately classified and therefore did not fully support the claim that fresh and archival tissue would be interchangeable. Additionally, the ESCs noted that there is some evidence that PD-L1 expression is dynamic and can vary between the primary tumour and metastases.

The ESCs noted that the TGA-approved indication for pembrolizumab was for the first line treatment of patients with metastatic or unresectable recurrent HNSCC, and whose tumours express PD-L1 (CPS  $\geq 1$ ). However, the ESCs considered that the rationale for codependency was not strong. The ESCs noted that the test for interaction demonstrated significant differences in treatment effect by PD-L1 positivity. However, the ESCs highlighted the following factors raised in the Commentary weakened the argument for codependency.

- The KN048 trial demonstrated that pembrolizumab with chemotherapy had a statistically significant overall survival benefit in the ITT analysis that included patients with CPS  $< 1$ . While the CPS  $< 1$  subgroup appeared to gain no benefit from treatment with pembrolizumab plus chemotherapy in KN048, the sample size was small (15% of the total trial population) and therefore lacked statistical power.
- The submission did not provide a biological explanation why CPS but not TPS was correlated to response to pembrolizumab in R/M HNSCC. The CPS  $\geq 1$  threshold was used based on the findings of KN012, a non-randomised open label phase 1b study of pembrolizumab in R/M HNSCC.

Overall, based on the subgroup analyses of KN048 in R/M HNSCC, the ESCs accepted that pembrolizumab had a larger effect in patients with a PD-L1 CPS  $\geq 1$ , but the subgroup of patients with CPS  $< 1$  was too small to conclude that pembrolizumab had no effect in this subgroup.

The ESCs noted that the key clinical data presented in the submission was based on the KN048 trial which compared pembrolizumab  $\pm$  chemotherapy with cetuximab + chemotherapy. Evidence for the treatment effect of pembrolizumab  $\pm$  chemotherapy compared with chemotherapy alone was based on an indirect comparison with the chemotherapy arm in the EXTREME trial. The ESCs noted that this comparison did not inform the intended comparison for the economic evaluation of 1L pembrolizumab  $\pm$  platinum based chemotherapy followed by second line (2L) chemotherapy versus 1L chemotherapy followed by 2L nivolumab.

The ESCs noted that PD-L1 CPS testing using the PD-L1 IHC 22C3 PharmDX kit used in KN048 was considered by the submission to be the evidentiary standard. The ESCs noted that 22C3 antibody assay tests in Australia will be laboratory developed tests (LDTs) conducted on the Ventana BenchMark XT platform. The ESCs noted that the concordance between this platform and the platform used for the evidentiary standard is high, but imperfect at 96%.

The ESCs considered that there was substantial inter-observer variability PD-L1 status using the 22C3 assay. The ESCs noted the pre-ESC response argued that de Meulenaere 2018 (quoted by the Commentary) assessed tumour cell scoring (TPS) at the  $\geq 1\%$  threshold, rather than CPS scoring, and therefore findings were not relevant to inter-observer variability with CPS scoring at the  $\geq 1$  level. However, in the absence of a clear biological rationale for choosing CPS over TPS (as CPS was selected based on the results of a non-randomised Phase 1b pembrolizumab study), the ESCs considered that these results still represented indirect evidence.

The ESCs noted that the submission claimed analytical validity against a reference standard of response rates to pembrolizumab was supported by high sensitivity. The ESCs noted that the high sensitivity may be due to the test classifying a large proportion of patients as CPS  $\geq 1$  ( $> 80\%$  of patients) rather than the test being accurate at predicting patients who are likely to respond to pembrolizumab, consistent with the low specificity of PD-L1 CPS  $\geq 1$  testing.

The ESCs noted that the economic evaluation in the PD-L1 CPS  $\geq 1$  population was presented as a base case, however the ESCs considered that as an integrated codependent submission the economic evaluation in the whole R/M HNSCC population was the appropriate base case and the sensitives around this are the most meaningful.

The ESCs also considered the assumption that the 22C3 antibody assay has 100% sensitivity and 100% specificity as the evidentiary standard in the economic model was implausible given the aforementioned issues with inter-observer variability and imperfect concordance between the testing platform used in the KN048 trial and the Ventana BenchMark XT platform most commonly used in Australia. The ESCs highlighted that test performance parameters from Cohen 2019 (94.0% sensitivity and 23% specificity) increased the weighted ICER by 11.9% and the base case model is optimistic with regards to testing accuracy.

The ESCs noted that the ICERs in the CPS  $\geq 1$  subgroup and the ITT population were similar due to the high prevalence of PD-L1 positivity and the assumption of 100% sensitivity and 100% specificity. The ESCs noted that chemotherapy was dominant in the CPS  $< 1$  population therefore an increase in false positives will increase the ICER.

The ESCs noted that the economic model used the ITT results and the CPS  $\geq 1$  subgroup to back-calculate the ICER for the CPS  $< 1$  subgroup. The ESCs noted that this method resulted in implausible scenarios where the efficacy of the Standard of Care (SoC) varied by CPS status, though the pre-ESC response argued that the incremental survival estimate measured between arms was equivalent regardless of differences between the SoC arm in the PD-L1 CPS subpopulations.

The ESCs considered that the current model may be inadequate to support decision making due to the exclusion of benefits expected from patients who would be treated with 2L nivolumab, which resulted in a significant underestimate of overall survival in the comparator arm and therefore underestimated the ICER. In addition, the ESCs also highlighted a number of issues that increased uncertainty in the modelled cost-effectiveness, and may favour pembrolizumab, including: use of a time horizon (7.5 years) that may be optimistic; the use of the inverse hazard ratio from a fractional polynomial network meta-analysis rather than the Bucher analysis in the base case; use of utility values that were higher than previously considered by PBAC and a lack of transparency due to custom Visual Basic for Applications (VBA) macros and custom functions.

The submission estimated net costs to the MBS based on additional costs from PD-L1 testing and costs of administration of pembrolizumab and chemotherapy. The ESCs noted that there were a number of areas of uncertainty in financial impact estimates presented. The ESCs agreed with the Commentary and the TGA Clinical Evaluation Report (CER, p38-39) that the prevalence of PD-L1 CPS positivity (85.2%) from the KN048 trial may be higher than the prevalence for the Australian population. The ESCs (p2) noted that the pre-ESC response claimed that the prevalence of PD-L1 CPS positivity at the  $\geq 1$  threshold has been between 79% and 85% across its trials for R/M HNSCC. The ESCs also noted that the prevalence estimates were uncertain and these would impact on changes in MBS costs.

The ESCs noted advice from the Royal College of Pathologists of Australasia (RCPA) that there is a need for quality assurance programs and accurate training in the diagnostic evaluation of PD-L1 ([1522 Ratified PICO confirmation, p17](#)). The ESCs noted that the pre-ESC response stated that CPS training with Australian pathologists is underway with four pathologists having been trained with plans for 24 pathologists to be trained by the end of December 2020. The ESCs queried whether this was sufficient to provide access to patients throughout Australia.

## **15. Other significant factors**

Nil.

## **16. Applicant comments on MSAC's Public Summary Document**

MSD is disappointed with the outcome for this application due to concerns about codependency, and the lack of acceptance of the biomarker enrichment data in the submission. MSD will be working on a resubmission to address the issues raised by MSAC to ensure patients with recurrent or metastatic SCCHN will be able to receive targeted treatment in this setting. MSD believes the lack of alignment between PBAC and MSAC processes continue to create challenges for precision medicine in Australia, acknowledging that other countries, such as the UK, France and the USA have already approved and incorporated PD-L1 testing into their H&N testing/treatment paradigm.

## **17. Further information on MSAC**

MSAC Terms of Reference and other information are available on the MSAC Website: [visit the MSAC website](#)