



Australian Government

Medical Services Advisory Committee

Public Summary Document

Application No. 1414 – PD-L1 testing for access to pembrolizumab for the treatment of locally advanced or metastatic NSCLC

Applicant: Merck, Sharp & Dohme (Australia)

Date of MSAC consideration: MSAC 68th Meeting, 24-25 November 2016

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

1. Purpose of application and links to other applications

The codependent application requested:

- Medicare Benefits Schedule (MBS) listing of an immunohistochemistry (IHC) test for evaluation of Programmed Cell Death 1 Ligand (PD-L1) expression to determine eligibility for treatment with pembrolizumab in patients with locally advanced or metastatic (Stage IIIb/IV) non-small cell lung cancer (NSCLC); and
- Pharmaceutical Benefits Scheme (PBS) Section 100 (Efficient Funding of Chemotherapy) Authority Required listing for the treatment of locally advanced or metastatic NSCLC in patients who have evidence of high expression of PD-L1, defined as a tumour proportion score (TPS) $\geq 50\%$.

2. MSAC's advice to the Minister

After considering the strength of the available evidence in relation to the comparative safety, clinical effectiveness and cost effectiveness, MSAC did not support public funding of PD-L1 IHC as a companion diagnostic test for selecting patients with NSCLC for treatment with pembrolizumab. MSAC considered that PD-L1 IHC as a companion diagnostic test has insufficient evidence of analytical validity (and documented poor reproducibility), weak evidence of clinical validity (lacks ability to predict response to therapy) and weak evidence of clinical utility (insufficient information to guide treatment).

MSAC noted that the Royal College of Pathologists of Australasia (RCPA) does not currently endorse the use of PD-L1 testing as a biomarker to exclude patients from therapy, but has acknowledged that higher levels of PD-L1 expression are associated with increased likelihood of response in some types of tumours.

MSAC advised that any resubmission would need to be critiqued and considered by ESC.

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

The application to list PD-L1 IHC testing in the MBS was part of an integrated codependent submission, which also requested that PBAC consider listing of pembrolizumab in the PBS to treat NSCLC. The proposed PBS criteria included a requirement that the patient must have evidence of high expression of PD-L1, defined as a tumour proportion score (TPS) $\geq 50\%$.

There are around 12,000 new diagnoses of lung cancer in Australia each year with NSCLC accounting for approximately 85% of all lung cancer cases. While the identification of oncogenic targets (epidermal growth factor receptor (*EGFR*) mutation and anaplastic lymphoma kinase (*ALK*) gene rearrangement) and the development of agents targeting these mutations have improved outcomes for some NSCLC patients, the majority have no option other than untargeted chemotherapy.

Pembrolizumab belongs to a new class of immunotherapy working on the PD-1 pathway that may help these patients. PD-L1 is preferentially expressed on the surface of NSCLC tumour cells and binds to PD-1 receptors on T-cells to switch off the immune response. Antibodies that bind to PD-1 on the T-cells (receptor binding; e.g. pembrolizumab and nivolumab) or PD-L1 or PD-L2 on the tumour cells (ligand binding; e.g. durvalumab and atezolizumab) disrupt this pathway, allowing the T-cells to recognise the tumour cells and initiate activated death of the tumour cell. The rationale for testing levels of PD-L1 expression is that it may predict variation in the extent of clinical response to pembrolizumab treatment.

MSAC noted that the proposed clinical algorithm involves biomarker tests for *EGFR* and *ALK* undertaken on patients with non-squamous or not otherwise specified NSCLC histology. If the tumour is *EGFR* wildtype, *ALK* gene rearrangement negative or the patient has squamous histology then a platinum doublet or pemetrexed would be prescribed as first-line therapy. If disease progression occurs, then PD-L1 testing would be undertaken, with those assessed as positive (with a TPS $> 50\%$) then treated with pembrolizumab. The applicant stated in its Pre-MSAC response that the intention was that PD-L1 IHC testing would be performed at initial diagnosis of NSCLC, at the same time as *EGFR* mutation and *ALK* IHC testing.

MSAC noted that, unlike many other companion tests, PD-L1 has a wide range of expression and hence the results reported are not dichotomous and are challenging to quantify. In this application, a sample was considered to be positive if the TPS is $\geq 50\%$. MSAC noted that the stability of PD-L1 as a biomarker varied before and after treatment and across different stages of disease making the identification of patients likely to benefit from PD-L1 agents challenging. MSAC also noted that PD-L1 expression is inducible and may vary during the course of disease.

In considering the evidence provided to support the analytical validity of the test, MSAC noted that there is currently no PD-L1 test reference standard. MSAC noted that its Protocol Advisory Sub-Committee (PASC) had recommended that, in the absence of a test reference standard, the Clinical Trial Assay (CTA) be used as the evidentiary standard. The CTA was used to screen for eligibility to the Keynote 010 trial (KN-010), which provided the main clinical evidence for the submission. PASC also requested that an assessment of comparative assay performance for alternative PD-L1 tests should be presented.

While the applicant did not directly request MBS listing of a specific PD-L1 IHC test, the evidence presented to support the MBS listing almost exclusively referred to a proprietary test kit (Dako PD-L1 IHC 22C3 pharmDx test; hereafter the Dako 22C3 test). When reviewing data relating to the evidentiary standard, MSAC noted that the differences between

the CTA and the Dako 22C3 test are minor and that they are essentially the same staining system. MSAC was therefore concerned that the CTA and Dako 22C3 test categorised the same samples differently in a significant number of samples included in the unpublished Dako KN-010 Retest report. **Redacted**

MSAC considered that these factors are likely to occur routinely in diagnostic laboratories and hence may reflect application of the test in practice.

Redacted

MSAC noted that four different assays using four different antibodies have been used in clinical trials to determine the level of PD-L1 and/or PD-1 expression in NSCLC tumour cells and/or immune cells. MSAC was concerned that the findings of several different concordance studies (e.g. the unpublished BluePrint study, which compared the concordance of these four assays), revealed that different stains give rise to different results in a significant number of samples. In addition, it is likely that these antibodies will be provided to pathology laboratories for use in in-house diagnostics tests for PD-L1 IHC, creating further options for testing. MSAC considered that, if approved, PD-L1 testing would therefore require a quality assurance program to standardise testing in diagnostic laboratories. The quality assurance program would need to address interpretation of the test results for PD-L1 positivity using the other assays/antibodies likely to be available. MSAC noted that such a program is not yet established. Correspondence from the RCPA indicated that the College does not currently endorse the use of PD-L1 testing as a biomarker to exclude patients from therapy, but has acknowledged that higher levels of PD-L1 expression are associated with increased likelihood of response in some types of tumours. MSAC considered this would make establishment of an inter-laboratory quality program to manage these performance issues unlikely until test interpretation issues can be resolved.

In considering the clinical validity of the test, MSAC noted that no gold standard was available to measure immunosuppression by the PD-1 pathway. To link the performance of the test with the outcome of interest, the submission calibrated the percentage of cells with PD-L1 protein against the response to pembrolizumab in the NSCLC cohorts of the Keynote 001 (KN-001) Phase 1 single-arm study. MSAC noted that such trial-specific calibration may not be generalisable to clinical practice, and noted that this initial calibration was not subsequently validated beyond this small study population.

The TPS threshold of 50% used in the proposed clinical algorithm was determined using Receiver Operating Characteristic (ROC) curves from the KN-001 NSCLC cohorts for PD-L1 positivity as the closest point to the optimum of all true positives and no false positives (i.e. “by maximising Youden’s index”). MSAC considered this to be a simplistic approach because it did not consider the trade-off between false positives and true positives, which should reflect the differing downstream consequences in terms of under- versus over-treatment. In addition, a visual inspection of the area under the ROC curve presented in the Pre-MSAC response suggested a poorly performing test overall. However, MSAC noted data from the KN-001 trial, which indicated that the overall response rate in patients with a TPS of 25–49% was twice that of those with a TPS <1%, suggesting patients with 25–49% of tumour cells expressing PD-L1 may still benefit from pembrolizumab treatment. MSAC agreed with the advice from the joint meeting of the Evaluation Sub-Committee of MSAC and the Economics Sub-Committee of PBAC (the ESCs) that the selection of a TPS threshold of 50% for effectiveness may be arbitrary and that a proportion of patients with a lower TPS score may still benefit from pembrolizumab treatment.

In considering the clinical utility of the test, MSAC questioned whether PD-L1 testing is more accurate than other measures of PD-1 dependence (e.g. microsatellite instability or tumour infiltrating lymphocytes) with this type of information not being provided in the submission. MSAC noted that pembrolizumab is currently listed in the PBS for late-stage melanoma without any requirement for companion testing to determine eligibility for treatment. MSAC also noted that, in a study comparing nivolumab with docetaxel in NSCLC (Bahmer J et al, 2015), the expression of PD-L1 was neither prognostic nor predictive of benefit.

MSAC noted the applicant's proposal that, rather than undertake PD-L1 testing using newly obtained tissue from a new biopsy, testing should occur at initial diagnosis to avoid a delay in starting second-line pembrolizumab treatment. MSAC acknowledged that this proposal has benefit for the patient and the system. However, MSAC was concerned that studies undertaken by Ilie M et al 2015 and Kitazono S et al 2015 demonstrated poor reproducibility between biopsy and resection specimens.

The positive predictive value and the negative predictive value are dependent on the prevalence of responsive tumours and positive tests: that is, the same test will have different predictive accuracy if the prevalence of the condition varies within the populations.

Redacted

MSAC was concerned that the claim of codependency was not adequately supported by the data provided. MSAC noted that the KN-010 trial only included patients who had a PD-L1 positive tumour (classified as either weakly 1–49% or strongly $\geq 50\%$ positive). MSAC also noted that the pooled PD-L1 data provided from this trial also did not allow partitioning out of the baseline prognostic effect of the test and its predictive power.

Finally, MSAC considered that the economic evaluation was confounded by the poor performance of the test, the patient population included in the model (any patient with NSCLC), the timing of the biopsy and questions around the cost of the test and the cost consequences of false positives and negatives.

MSAC foreshadowed that an integrated codependent resubmission would need to address the following issues for MSAC:

- Analytical validity
 - Define gold standard or reference standard
 - Confirm real-world reproducibility between assays (taking into account different antibodies or platforms), laboratories, pathologists and tumour material (for example, site – biopsy or resection, metastases or primary, pre/post treatment, patient characteristics)
 - Provide guidance on how doctors should interpret results obtained with a different PD-L1 assay (other than the Dako 223C Market Ready Assay) to establish eligibility for pembrolizumab in NSCLC
- Clinical validity
 - Calibrate against a stable functional biological metric, and address the issue of PD-L1 inducibility
 - Unacceptably low sensitivity and specificity for tumour response
 - Document the clinical validity in key trial/s
 - Clarify the biological basis for PD-L1 testing in NSCLC as compared to other indications such as melanoma that do not require a test for pembrolizumab eligibility

- Clinical utility
 - Demonstrate the stability of the result from a biopsy obtained at initial diagnosis/presentation to inform a treatment decision for potentially later stage disease
 - Clarify any significance of histotypes on clinical utility of PD-L1 expression (in particular, in patients with squamous histology)
 - Compare the performance of PD-L1 as a biomarker vs other measures (for example, other biomarkers of programmed cell death)
 - Assess clinical utility in key trial/s (in particular, to address the paucity of data in PD-L1 expression negative patients)
 - Better justify the chosen PD-L1 expression TPS threshold of 50% (for inclusion in the PBS restriction)
- Support from professional bodies, in particular the RCPA, in developing a quality control program (QAP) including methodology and interpretation across different testing platforms.

4. Background

MSAC has not previously considered PD-L1 IHC testing for access to pembrolizumab for the treatment of locally advanced or metastatic NSCLC.

5. Prerequisites to implementation of any funding advice

The PD-L1 IHC 22C3 pharmDx test was listed by the Therapeutic Goods Administration in the Australian Register of Therapeutic Goods (ARTG) on 17 November 2016.

A Quality Assurance Program will be required, so that equivalency can be achieved between different tests, different antibodies, and different ways of reporting.

6. Proposal for public funding

The proposed MBS item descriptor (Table 1) was consistent with the descriptor in the Consultation Protocol for MSAC Application 1414, except that the eligible patient population was specified as ‘a patient diagnosed with non-small cell lung cancer’ instead of being unspecified. However, this item descriptor has not been restricted to NSCLC patients with Stage IIIB/IV disease, so testing could occur at an earlier stage of disease.

The proposed MBS item descriptor does not specify a testing platform, consistent with MBS policy. However, the proposed MBS item descriptor is broad enough to allow testing of any PD-L1 test available, not just the Dako PD-L1 IHC 22C3 pharmDx test. MSAC considered that a broad MBS item descriptor is not supported by the data the applicant has supplied to date.

Table 1: Proposed MBS item

Category 6 – Pathology Services	
MBS item number	
	Immunohistochemical examination by immunoperoxidase or other labelled antibody techniques using the PD-L1 antibody of tumour material from a patient diagnosed with non-small cell lung cancer to determine if the requirements relating to programmed death ligand 1 (PD-L1) status for access to pembrolizumab under the Pharmaceutical Benefits Scheme (PBS) are fulfilled.
	Fee: \$redacted Benefit: 75% = \$redacted 85% = \$redacted

The application requested that the PD-L1 IHC test be a pathologist-determinable test and that an amendment be made to Note P.1.2 “Services Where Request Not Required” to include the

above item number. This is consistent with other IHC tests (including of *ALK* in NSCLC) and with *EGFR* mutation testing of NSCLC patients, which are pathologist-determinable, but is inconsistent with *ALK* gene rearrangement testing, which is not pathologist-determinable.

The application requested an MBS fee of \$redacted, which is aligned with MBS item number 72848 for human epidermal growth factor receptor 2 (HER2) IHC testing; both tests require the counting of cells and assessing the intensity of staining.

The item descriptor for *EGFR* testing specifies the requestor: 'requested by, or on the advice of, a specialist or consultant physician.' No requestor was specified in the proposed item descriptor.

7. Summary of Public Consultation Feedback/Consumer Issues

PASC received two responses from peak bodies and two responses from consumer/caregivers.

Issues raised in the responses were:

- The use of PD-L1 assessed by IHC as a suboptimal biomarker, and the difficulties of implementing safe and reliable PD-L1 IHC testing in the routine environment
- That patients negative for PD-L1 can still respond and derive benefit from pembrolizumab (test enriches for, but does not perfectly predict, likelihood of response to PD-1 blockade)
- That the expression of the marker is inducible and could change over time
- There are problems with heterogeneity of expression across the section
- Possible necessity for a second biopsy to perform this testing
- IHC assays have different thresholds and scoring algorithms, there are current international efforts to try to harmonise assays.

8. Proposed intervention's place in clinical management

The application indicated that testing for PD-L1 expression should occur at diagnosis, although pembrolizumab is only an option when disease progression occurs subsequent to treatment with platinum-based chemotherapy. The application anticipated that the addition of PD-L1 testing will identify the cohort of these patients (PD-L1 TPS \geq 50%) who will be eligible to receive pembrolizumab therapy. However, there is some uncertainty about PD-L1 expression levels over time, especially as they may change after exposure to treatment and with disease progression or evolution.

Two alternative scenarios for testing on progression after platinum-based chemotherapy using either archival or fresh biopsy tissue were proposed in the MSAC Consultation Protocol. Both of these scenarios would cause a delay in treatment and having a fresh biopsy has a negative consequence for the patient. However, as PD-L1 expression levels may increase after further treatment, testing of fresh biopsy tissue if available in patients who were initially below the TPS threshold for access to pembrolizumab would enable patients whose PD-L1 expression levels increased to above the TPS threshold during previous treatments to receive pembrolizumab.

The ESCs noted that, although testing of fresh biopsy tissue in patients who were initially above the TPS threshold for access to pembrolizumab would be less likely to occur, it is possible that the PD-L1 expression in such patients may have subsequently decreased, resulting in a false positive test result and inappropriate treatment with pembrolizumab.

Testing for PD-L1 positivity at the point of diagnosis of advanced NSCLC would be in addition to *EGFR* and *ALK* testing (if *EGFR* negative), using the same biopsy sample. The addition of PD-L1 testing should not require an additional biopsy as only a single 4–5 micron section is required for the test. In its Pre-MSAC response, the applicant expressed a preference for PD-L1 IHC testing to be performed at initial diagnosis of NSCLC

9. Comparator

As there is currently no reference standard for PD-L1 testing, PASC recommended that the evidentiary standard should be the Clinical Trial Assay that was used to determine patient eligibility for the KN-010 trial. PASC also requested that an assessment of comparative assay performance for alternative PD-L1 tests should be presented.

The evidence base used to determine prognosis based on PD-L1 status and for accuracy of the PD-L1 test is summarised in Table 2.

Table 2: Evidence for the test performance

Prognostic evidence	Comparison of outcomes in patients receiving usual care conditioned on the presence or absence of the biomarker	☒ k=4 SRs k=10 studies, n=2,776
Comparative analytical performance	Diagnostic accuracy of the PD-L1 IHC 22C3 pharmDx test in NSCLC compared with the nominated evidentiary standard	☒ k=3 n=158
	Inter- and intra-rater reliability	☒ k=3 n=248
	Concordance between the PD-L1 IHC 22C3 pharmDx test and other existing PD-L1 tests	☒ k=5 n=709

^a reference standard available; ^b reference standard not available

NSCLC = non-small cell lung cancer; PD-L1 = programmed death ligand 1; SR = systematic review

10. Comparative safety

PD-L1 testing would be performed on tissue sections taken from a biopsy specimen obtained as part of standard diagnostic work-up and would not incur any direct risks to patients. The main risk to the patient would occur if an additional biopsy is required in order to obtain tissue to perform the IHC test. This could result in complications such as pneumothorax and haemorrhage. However, IHC only uses one 4 micron section compared to approximately 50 microns required for *EGFR* testing, and is unlikely that a re-biopsy would be required for the PD-L1 test alone. Hence, the addition of the PD-L1 biomarker to the testing protocol at diagnosis would be unlikely to increase the overall re-biopsy rate, unless a re-biopsy were to be obtained due to concerns with PD-L1 expression changes following platinum-based chemotherapy or between an initial biopsy and a patient subsequently progressing to developing more advanced disease and thus potentially eligible for pembrolizumab.

Redacted

11. Comparative effectiveness

Prognostic evidence

Four meta-analyses conducted including Asian patients found that those with PD-L1-positive NSCLC had a worse prognosis than those who had PD-L1-negative tumours. Meta-analysis of three studies conducted during the evaluation showed that PD-L1 positivity may be associated with longer survival in Caucasian patients. This contradictory result may be due to the many inconsistencies between the studies, such as differing stages of disease and variability in the testing approach, rather than a difference in prognostic effect due to ethnicity. Further evidence is needed to verify if there is a difference.

Comparative analytical performance

According to standard analytical validity metrics, the Market Ready Assay PD-L1 IHC 22C3 pharmDx test was accurate compared to the Clinical Trial Assay evidentiary standard (Table 3).

Table 3: Diagnostic accuracy of the PD-L1 IHC 22C3 pharmDx test compared to the Clinical Trial Assay

	Sensitivity	Specificity	PPV	NPV
≥1% TPS *	98.8% [95% CI 93.6, 99.8]	97.3% [95% CI 90.7, 99.3]	97.6% [95% CI 91.8, 99.4]	98.6% [95% CI 92.6, 99.8]
≥50% TPS	100% [95% CI 90.8, 100]	99.1% [95% CI 94.3, 99.9]	98.0% [95% CI 87.8, 99.9]	100% [95% CI 95.8, 100]

CI = confidence interval; NPV = negative predictive value; PD-L1 = programmed death ligand 1; PPV = positive predictive value

* Calculated during the evaluation

Source: Table BT.6.i.3 of the commentary

To date, four different assays using four different antibodies have been used in clinical trials to determine the level of PD-L1 and/or PD-1 expression in NSCLC tumour cells and/or immune cells (Table 4).

Table 4: Concordance (% agreement) between the four PD-L1 IHC tests used in clinical trials

	Dako 28-8		Ventana SP142		Ventana SP263	
Dako 22C3	≥1% TPS	95%	≥1% TPS	83%	≥1% TPS	90%
	≥50% TPS	93%	≥50% TPS	91%	≥50% TPS	85%
Dako 28-8			≥1% TPS	81%	≥1% TPS	91%
			≥50% TPS	91%	≥50% TPS	82%
Ventana SP142					≥1% TPS	76%
					≥50% TPS	81%

IHC = immunohistochemistry; PD-L1 = programmed death ligand 1

Source: Scheel et al. (2016)

The Dako PD-L1 IHC 22C3 pharmDx test (developed to support pembrolizumab) and the Dako 28-8 pharmDx assay (developed to support nivolumab) demonstrated high concordance, with 93–95% overall agreement at both a 1% and a 50% TPS threshold. In contrast, the Ventana SP142 assay (developed to support atezolizumab) consistently labelled fewer tumour cells, and the Ventana SP263 assay (developed to support durvalumab) consistently labelled more than the other assays. Thus, the Ventana SP142 and SP263 assays would need to be carefully calibrated so that the TPS thresholds used by the Dako assays and the two Ventana assays identify the same population for treatment with anti PD-1/PD-L1 antibodies. Thus, PD-L1 testing would require a Quality Assurance Program (QAP) to standardise PD-L1 testing and reporting in diagnostic laboratories. MSAC noted that these concordance data alone were insufficient to establish whether the different PD-L1 assays could be used interchangeably without first resolving the outstanding issues from the Dako KN-010 Retest report comparing the Clinical Trial Assay and the Market Ready Assay for Dako 22C3, and other clinical validity and clinical utility issues.

Prevalence

The prevalence of PD-L1 above a specific TPS threshold may differ if different tests, thresholds and/or antibodies are used. **Redacted.** The prevalence of PD-L1 expression in all patients screened for KN-001 (N=502) was 71% at TPS ≥1% and 30.4% at TPS ≥50% (not reported for KN-010). Information regarding the total number of patients screened for both trials was not provided and thus the proportion of total patients screened for the trials who were Australians could not be determined. The reliability of the estimates in the Australian patients screened for KN-001 and KN-010 trials and their external validity to the Australian setting remain uncertain.

The submission estimated 10,399 incident NSCLC cases in Year 1, increasing to 12,093 in Year 5. Of these, 55% were assumed to have Stage IIIB/IV disease at diagnosis. The number of NSCLC cases that progress to Stage IIIB/IV disease each year is 3,833 in Year 1 increasing to 4,457 in Year 5. Of incident and progressed cases, 60% were estimated to receive first-line treatment, and these patients were assumed to be eligible for PD-L1 testing and, following progression, later-line treatment with pembrolizumab. Prevalent cases of Stage IIIB/IV disease are also considered eligible for PD-L1 testing in Year 1 (Table 5).

Table 5: Numbers of patients and estimates used in the economic and financial analyses

Annual estimates	Submission base case
Number of patients tested	11,519 in Year 1 (to capture incident and prevalent cases), to 6,651 in Year 5 (incident cases)
Prevalence	redacted%
Number of test positive patients	redacted
Number of patients treated	redacted
Number of tests per patient treated	redacted
Test cost per patient treated	\$redacted

Source: *Constructed during evaluation*

Claim of co-dependence

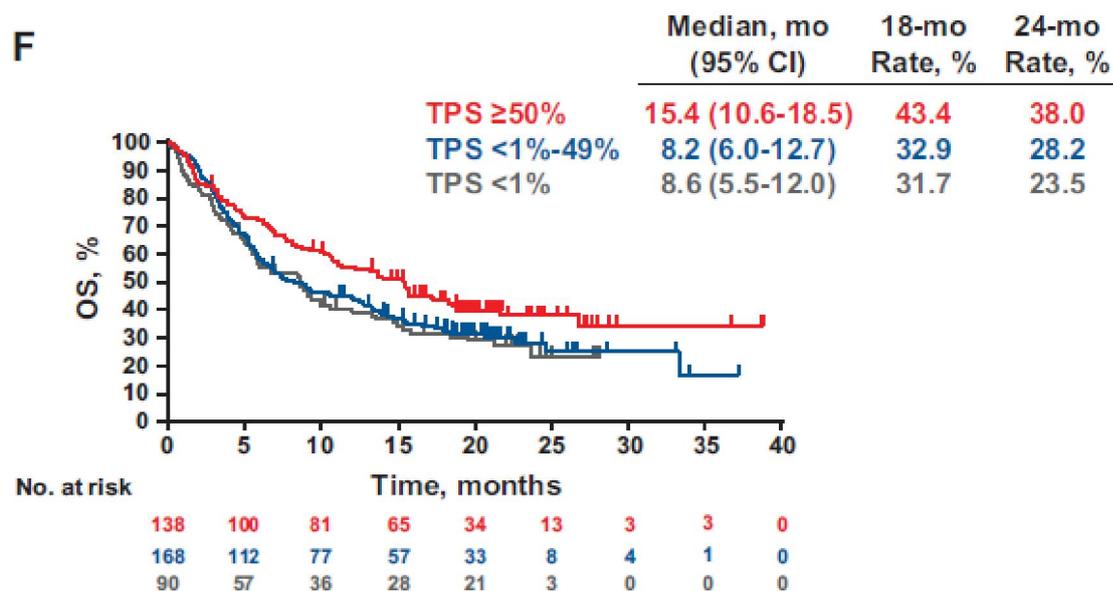
The submission presented a claim that treatment guided by PD-L1 status, where PD-L1 strong positives (i.e. TPS $\geq 50\%$) are treated with pembrolizumab, and PD-L1 negatives (i.e. TPS $< 1\%$) and weakly positives (i.e. TPS 1–49%) are treated with docetaxel or pemetrexed, results in improved outcomes versus the comparator, which is no testing and docetaxel or pemetrexed treatment.

Evidence to support the treatment effect modification by PD-L1 status from the key trial provided (KN-010) was limited to a comparison of strongly positive (TPS $\geq 50\%$) vs weakly positive (TPS 1–49%).

Further evidence beyond Trial KN-010 of treatment effect variation

Study KN-001 was a Phase 1 single-arm study which showed (Figure 1) that there was increasing OS associated with pembrolizumab in treating NSCLC as PD-L1 TPS increased (from $< 1\%$ [median OS = 8.6 months; 18-month OS rate of 32%], 1%–49% [median OS = 8.2 months; 18-month OS rate of 33%] and $\geq 50\%$ [median OS = 15.4 months; 18-month OS rate of 43%]). In this study, only 90 patients had both NSCLC and TPS $< 1\%$.

Figure 1: (KN-001): Kaplan Meier curves of overall survival in previously treated NSCLC patients by PD-L1 expression levels TPS $\geq 50\%$, 1%-49%^β, and $< 1\%$



^β The title of the figure in the abstract by Hui et al indicated the middle PD-L1 stratum was 1–49% rather than $< 1\%$ -49%. The former appeared more plausible and informative for mutually exclusive subsets.

CI = confidence interval; PD-L1 = programmed death ligand 1; TPS = tumour proportion score.

Source: Figure 2, Panel B from Hui et al (2016)¹.

This represented weak evidence to support the prior claim of codependence between being PD-L1 positive (or not) and the effectiveness of pembrolizumab.

Evidence of a treatment effect variation of anti-PD1 agents in NSCLC, by PD-L1 expression status, was also published in a recent meta-analysis by Abdel-Rahman et al (2016). The meta-analysis assessed the correlation between PD-L1 levels and effectiveness outcomes of PD-1/PD-L1 inhibitors (nivolumab, pembrolizumab, atezolizumab, durvalumab and avelumab) in advanced NSCLC. The benefit from PD-1 inhibitors versus docetaxel in second-line treatment of NSCLC appeared to be limited to the PD-L1 $> 1\%$ population.

The submission presented data in the translation section for nivolumab, by PD-L1 expression, in patients with non-squamous (CheckMate 057 trial) and squamous (CheckMate 017 trial) histology. These trials have been previously considered by the PBAC (nivolumab for non-squamous and squamous NSCLC, March 2016 PSDs). The data presented supported the conclusion of treatment effect modification by PD-L1 status in patients with non-squamous histology, but not in patients with squamous histology. The submission raised several issues with the CheckMate 017 data:

- the biological plausibility of a treatment effect variation is equally applicable in squamous as it is in non-squamous (as it relates to the biological effect that the PD-1 antibody has on the ability of the tumour to evade the immune system);
- a trend, although not significant, was suggestive of a treatment effect variation by PD-L1 expression;

¹ Abstract presentation by Hui et al. Long-Term Overall Survival For Patients With Advanced NSCLC Enrolled In the KEYNOTE-001 Study of Pembrolizumab. American Society of Clinical Oncology (ASCO) Annual Meeting June 3-7, 2016, Chicago, Illinois.

- a substantial proportion of patients (17%) had unquantifiable PD-L1 expression and these patients showed the biggest difference favouring nivolumab; and
- higher PD-L1 expression thresholds were associated with fewer patient numbers.

The literature has reported inconsistencies in the treatment effect variation of PD-L1 expression on the effectiveness of PD-1 /PD-L1 inhibitors, across squamous and non-squamous histology subtypes. A meta-analysis by Gandini et al (2016) showed that PD-L1 expression correlated with the clinical response to antibodies targeting the PD-1/PD-L1 axis in non-squamous NSCLC and not in squamous NSCLC (odds ratio (OR) of summary objective response rate for PD-L1 positive vs negative: squamous NSCLC: OR = 1.49; 95% CI 0.48, 4.64; non-squamous NSCLC: OR = 3.78; 95% CI 1.54, 9.24).

Nivolumab data in squamous NSCLC from CheckMate-017 (Brahmer (2015)) indicated the benefit of nivolumab over docetaxel was independent from PD-L1 expression regardless of the threshold used. In contrast, there was a convincing treatment effect variation with PD-L1 expression in the nivolumab versus docetaxel non-squamous trial (CheckMate-057, Borghaei (2015) and Horn (2015)). Overall, there are inadequate robust data currently available to support a confident position in this regard.

12. Economic evaluation

The application presented a modelled economic evaluation, cost-utility analysis and cost-effectiveness analysis (life-years gained).

13. Financial/budgetary impacts

Test cost/patient tested: The proposed MBS fee for PD-L1 testing is \$74.50, benchmarked against the fee for a similar item, MBS 72848.

An epidemiological approach was used in the application to estimate the number of patients eligible for PD-L1 testing treatment each year, over a five-year period. This was based on data presented in the applicant-commissioned ONCOSight report that estimated the number of Stage IIIB/IV NSCLC patients treated with first-line therapy. A summary of the approach used in the commissioned report is presented in Table 6.

Table 6: Estimated number of patients who would be tested

	Year 0	Year 1	Year 2	Year 3	Year 4	Year 5
No. prevalent Stage IIIB/IV cases	redacted	redacted	redacted	redacted	redacted	redacted
No. incident NSCLC cases	redacted	redacted	redacted	redacted	redacted	redacted
No. incident Stage IIIB/IV NSCLC cases (55%)	redacted	redacted	redacted	redacted	redacted	redacted
No. earlier cases that progress to Stage IIIB/IV	redacted	redacted	redacted	redacted	redacted	redacted
Total patients uptake PD-L1 testing (prevalent year prior + incident cases)	redacted	redacted	redacted	redacted	redacted	redacted

NSCLC = non-small cell lung cancer

Source: ONCOSight report

Assumptions used to generate these estimates this number may not be reasonable as:

- the proportion of patients who have Stage IIIB/IV disease at diagnosis may be an underestimate (55%, compared to 65%);

- the number of patients who progress to Stage IIIB/IV disease was not adequately described.

The application estimated that, if PD-L1 testing and pembrolizumab were both listed, the net cost to the MBS would be approximately \$redacted over the first five years (see further detail in Table 7).

Table 7: Cost of PD-L1 testing and pembrolizumab treatment to the MBS

	Year 1	Year 2	Year 3	Year 4	Year 5
Cost of PD-L1 testing	\$redacted	\$redacted	\$redacted	\$redacted	\$redacted
Administration & management costs	\$redacted	\$redacted	\$redacted	\$redacted	\$redacted
Total cost to the MBS	\$redacted	\$redacted	\$redacted	\$redacted	\$redacted

These are likely to be underestimates as the submission did not consider patients who progress to Stage IIIB/IV disease to be eligible for PD-L1 testing. Furthermore, only costs of PD-L1 testing and those associated with chemotherapy administration and disease management were included for the costs to the MBS. Costs associated with re-biopsy or re-testing of unevaluable samples have not been considered.

14. Key issues from ESC for MSAC

Clinical issues

The joint meeting of the Evaluation Sub-Committee of MSAC and the Economics Sub-Committee of PBAC (the ESCs) noted the extent of difficulties posed by implementation of a PD-L1 expression test as a criterion for access to a drug via the Pharmaceutical Benefits Scheme. These included issues of reproducibility, robustness and equivalence of results across different assays, and that PD-L1 expression had been shown to be inducible and dynamic, both of which would make identifying patients likely to benefit from PD-L1 agents a challenging issue. The ESCs also noted concerns about the validity of any particular PD-L1 expression threshold as a predictor of treatment response.

The proposed Medical Benefits Schedule (MBS) item descriptor did not restrict testing to non-small cell lung cancer (NSCLC) patients with Stage IIIB/IV disease, so testing could occur at earlier stages of disease. The Pre-Sub-Committee response (PSCR) provided to the ESCs confirmed that the proposed MBS item descriptor is intended to allow testing in all NSCLC patients at the initial diagnostic workup, regardless of stage of disease. The ESCs considered that it would be usual to include the stage of cancer in an MBS item descriptor for an item of this type, particularly to have the test conducted close in time to the decision to treat if the biomarker result changes over time and/or following treatment. The ESCs did not consider the example of the *EGFR* test item (73337) for access to PBS-funded tyrosine kinase inhibitors (also for NSCLC) to be a relevant comparison given that reflex testing at diagnosis or presentation of those patients was already established in clinical practice prior to listing of that test item and that this had been an outcome of a stakeholder meeting to discuss these and other problems around that codependent combination. However, the ESCs were also concerned about the implications of re-biopsying some patients that would be imposed by a stage-specific disease test requirement. The ESCs noted that the amount of tissue required for a PD-L1 IHC was nonetheless considerably smaller than for an *EGFR* test and that the patient's archival tissue block obtained at diagnosis would most likely be sufficient if biomarker stability is not a concern. To better inform this consideration, the ESCs requested that in its Pre-PBAC/pre-MSAC response, the applicant address the stability of the PD-L1 expression biomarker, for example over time, in response to likely treatments, and across primary tumours and metastases.

Although the proposed MBS item descriptor did not specify a named test, the submission almost exclusively referred to the Dako 22C3 test. The Instructions accompanying this test define a specimen as positive if programmed death ligand 1 (PD-L1) expression, reported as a tumour proportion score (TPS) is $\geq 1\%$. This does not match the requested PBS restriction for pembrolizumab requiring PD-L1 expression TPS $\geq 50\%$. The PSCR clarified that information is provided on both TPS thresholds in these instructions.

It was noted that different drugs in this class may have different affinities for the PD-1/PD-L1 epitope and so optimal clinical benefit may occur at different PD-L1 expression levels.

The applicant was asked to describe the preferred course of action for a treating physician in receipt of a test result obtained with an alternative assay. The PSCR addressed this by suggesting that the MBS item be restricted to use of the 22C3 antibody.

The ESCs considered that the nominated 50% TPS threshold was essentially arbitrary and would likely deny access to patients having lower PD-L1 expression levels who would benefit from treatment with pembrolizumab. It appears that a 50% TPS threshold has no biological basis and instead has been proposed as a means to narrow the population to support the cost-effectiveness analysis for the drug. The ESCs noted that the number of patients in the clinical data exceeding the 50% TPS threshold was small both as a proportion of total patients recruited and of those who responded. In addition, the KN-010 trial excluded patients with $< 1\%$ TPS, an exclusion criterion that limited the value of this trial for an evaluation of PD-L1 expression as a predictor of drug response.

The ESCs acknowledged the advice from the RCPA “Based on current data we do not endorse the use of PD-L1 as a biomarker to exclude patients from therapy but acknowledge that higher levels of PD-L1 expression are associated with increased likelihood of response in some types of tumours”.

The use of alternative tests to determine PD-L1 expression may identify different patient populations at the same TPS threshold because of the different staining conditions and/or different anti-PD-L1 antibodies used (with different affinities for the PD-L1 epitope).

While there is a high level of concordance between the Dako 22C3 test (used in pembrolizumab studies) and the Dako 28-8 test (used in nivolumab studies), the concordance between these tests and the Ventana SP124 test (atezolizumab) and the Ventana SP264 test (durvalumab) is lower. The data suggests that the Ventana SP124 test consistently stains less tumour cells than the other tests and the Ventana SP264 test consistently stains more.

PD-L1 testing requires a Quality Assurance Program (QAP) to standardise PD-L1 testing and reporting in diagnostic laboratories. It is unclear if a QAP is in development. The PSCR stated that the applicant is currently working with the RCPAQAP on implementation. The ESCs however noted the concerns raised and also the Department’s advice that development of a QAP would be complex in these circumstances and thus unlikely to be achievable in a timeframe consistent with usual implementation of an MBS item.

Economic issues

The submission stated that the implications of false positive results were included in the economic analyses, but this was not the case. The base case was respecified during the evaluation.

The ESCs noted that no rationale for the proposed MBS item fee of \$**redacted** was provided, and that basing the pricing for a single antibody IHC test (PD-L1 test) on the

specific IHC test for ER/PR/HER2 status (item 72848) had not been justified. The ESCs considered that the proposed fee per test was probably over-estimated given that the fee for IHC test items is driven by the type and number of antibodies used per specimen:

- MBS #72846 (general IHC item)(any 1-3 antibodies): \$59.60
- MBS #72847 (any 4-6 antibodies): \$89.40
- MBS #72848 (specific antibodies for ER/PR/HER2): \$74.50
- MBS #72849 (any 7-10 antibodies): \$104.30
- MBS # 72850 (any 11 or more): \$119.20

There is no reason for a PD-L1 test fee to be higher than the general item (72846).

Financial issues

The number of tests was likely to be underestimated:

- The submission only considered patients who were diagnosed at Stage IIIB/IV as being eligible for testing, and did not consider use in patients who progress to Stage IIIB/IV from an earlier stage (as allowed within the proposed MBS item descriptor).
- The proportion of cases that are Stage IIIB/IV at diagnosis (55%) may have been underestimated in the submission as the literature indicates a higher proportion (65%).

The approach used in the submission to estimate the number of patients treated with pembrolizumab did not take into account the accuracy of the test. This means that there was no allowance for inappropriate treatment (i.e. false positive patients) with pembrolizumab.

15. Other significant factors

Nil.

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

MSD is disappointed with the outcome, particularly given the test is registered, available for use, and included in various guidelines overseas. We will endeavour to work through the issues with MSAC so that patients whose tumours are strongly PD-L1 positive can get access to pembrolizumab as soon as possible.

17. Further information on MSAC

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