Assessment Run C1 2017

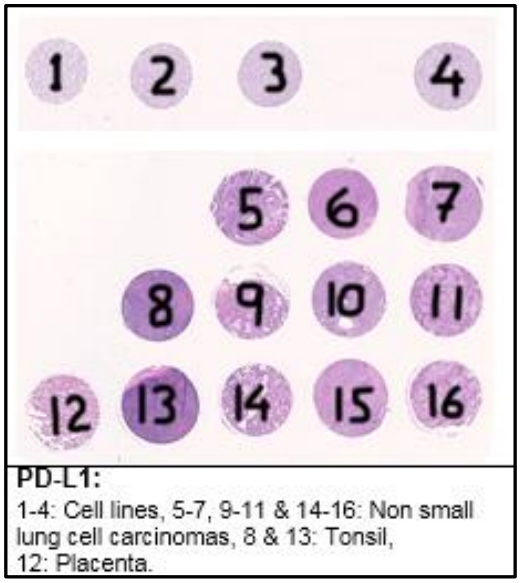
PD-L1

The first assessment in this new NordiQC Companion module C1 focused on the accuracy of the PD-L1 IHC assays performed by the participating laboratories to identify patients with NSCLC to be treated with immune therapy as either first line treatment (Keytruda®) or second line treatment (Keytruda® and Opdivo®). The PD-L1 expression levels in the circulated material used for the assessment were characterized by two CE IVD / FDA approved companion and complementary IHC assays, 28-8 pharmDX, SK005 Dako/Agilent and 22C3 pharmDX, SK006 Dako/Agilent. The associated cut-off values and interpretation guidelines were used accordingly to these two PD-L1 IHC assays.

Material

The slides to be stained for PD-L1 comprised the following 16 materials

	PD-L1 IHC TPS score*	**Eglible for treatment****
Cell line controls***		
1. Cell line 1	NA	NA
2. Cell line 2	NA	NA
3. Cell line 3	NA	NA
4. Cell line 4	NA	NA
Tissue controls		
8. Tonsil	NA	NA
13. Tonsil	NA	NA
12. Placenta	NA	NA
NSCLC		
5. NSCLC	No <1%	No
9. NSCLC	No <1%	No
14. NSCLC	No <1%	No
10. NSCLC	Low 1-49%	Yes
15. NSCLC	Low 1-49%	Yes
6. NSCLC	High ≥50%	Yes
7. NSCLC	High ≥50%	Yes
11. NSCLC	High ≥50%	Yes
16. NSCLC	High ≥50%	Yes



PD-L1:
1-4: Cell lines, 5-7, 9-11 & 14-16: Non small lung cell carcinomas, 8 & 13: Tonsil, 12: Placenta.

* Tumour proportion score (TPS) determined by PD-L1 IHC 28-8, SK005 & 22C3, SK006 Dako performed in NordiQC reference lab.
** Using present recommendations for cut-off value of TPS of 1-49% and ≥ 50% for second line (Keytruda® and Opdivo®) and first line treatment (Keytruda®), respectively.
*** Cell lines, Horizon Discovery, prod. Id HD788.

All tissues were fixed in 10% neutral buffered formalin.

The participating laboratories were asked to perform the PD-L1 IHC assay accordingly to the protocol being used in the laboratory and also to interpret the PD-L1 expression level and submit these scores to NordiQC. This allowed both an assessment of the technical performance / analytical accuracy of the PD-L1 IHC assays but also information on the reproducibility and concordance of the interpretation of PD-L1 expression among the laboratories.

Participation

Number of laboratories registered for PD-L1 IHC C1	71
Number of laboratories returning PD-L1 IHC	68 (96%)
Number of laboratories returning PD-L1 scoring sheet	65 (96%)

PD-L1 IHC, Technical assessment
Criteria for assessing a staining as **Optimal** included:
The staining is considered perfect or close to perfect in all of the included tissues.
TPS is concordant to the NordiQC reference data is obtained in all 9 NSCLC.

Criteria for assessing a staining as **Good** included:
The staining is considered acceptable in all of the included tissues. However, the protocol may be optimized to ensure the best staining intensity, counter staining, morphology and signal-to-noise ratio.
TPS is still concordant to the NordiQC reference data in all 9 NSCLC.

Criteria for assessing a staining as **Borderline** included:
The staining is considered insufficient, e.g., because of a generally too weak staining, a false negative staining or a false positive staining reaction of one of the included tissues. The protocol should be optimized.
TPS is **not** found concordant to the NordiQC reference data in all 9 NSCLC.

Criteria for assessing a staining as **Poor** included:
The staining is considered very insufficient e.g., because of a false negative or a false positive staining reaction of more of the included tissues.
An optimization of the protocol is urgently needed.
TPS is **not** found concordant to the NordiQC reference data in all 9 NSCLC.

PD-L1 IHC, Interpretation
All participating laboratories were asked to submit a scoring sheet with their interpretation of the tumour proportion score (TPS) in the nine lung NSCLC. Results were compared to NordiQC data from the reference laboratory to analyze scoring consensus.
Results: 68 laboratories participated in this assessment and 50% achieved a sufficient mark. Assessment marks for IHC PD-L1 assays and PD-L1 antibodies are summarized in table 1.

Table 1. **Assessment marks for IHC assays and antibodies run C1, PD-L1 IHC**

CE-IVD / FDA approved PD-L1 assays	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
22C3 pharmDX, **SK006**	12	Dako/Agilent	10	1	0	1	92%	92%
22C3 pharmDX, **SK006**[4]	2	Dako/Agilent	0	0	1	1	-	-
28-8 pharmDX, **SK005**	7	Dako/Agilent	3	3	1	0	86%	86%
SP263, 790-4905	16	Ventana/Roche	9	2	2	3	69%	77%
SP142, 740-4859	1	Ventana/Roche	0	0	0	1	-	-
Antibodies[3] for laboratory developed PD-L1 assays, conc. antibody	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
mAb clone **22C3**	13	Dako/Agilent	1	1	4	7	15%	-
mAb clone **E1L3N**	8	Cell Signaling	1	1	1	5	25%	-
mAb **CAL10**	1	Biocare	0	0	1	0	-	-
rmAb clone **28-8**	6	Abcam	0	1	1	4	17%	-
rmAb clone **ZR3**	1	Zeta Corporation	1	0	0	0	-	-
Antibodies for laboratory developed PD-L1 assays, RTU	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
rmAb clone **SP142**	1	Spring Biosystems	0	0	0	1	-	-
Total	68		25	9	11	23	-	-
Proportion			37%	13%	16%	34%	50%	-

1) Proportion of sufficient stains (optimal or good).
2) Proportion of sufficient stains with optimal protocol settings only, see below.
3) mAb: mouse monoclonal antibody, rmAb: rabbit monoclonal antibody.
4) RTU system developed for the Agilent/Dako`s semi-automated systems (Autostainer Link48) but used by laboratories on different platforms (Ventana Benchmark and Dako Omnis).

Detailed Analysis

CE IVD / FDA approved assays
PD-L1 IHC 22C3 pharmDx (SK006, Dako): 10 of 12 (83%) protocols were assessed as optimal. Protocols with optimal results were based on heat induced epitope retrieval (HIER) in EnVision™ Flex target retrieval solution low pH 6.1 (SK006) at 95-99°C for 20 min. in PT Link and 30 min. incubation of the primary Ab, linker and polymer. Using these protocol settings, 11 of 12 (92%) laboratories produced a sufficient staining result (optimal or good).

PD-L1 IHC 28-8 pharmDx (SK005, Dako): 3 of 7 (43%) protocols were assessed as optimal. Protocols with optimal results were based on HIER in EnVision™ Flex target retrieval solution low pH 6.1 (SK005) at 97°C for 20 min. in PT Link and 30 min. incubation of the primary Ab, linker and polymer. Using these protocol settings, 6 of 7 (86%) laboratories produced a sufficient staining result.

SP263 (790-4905, Ventana): 9 of 16 (56%) protocols were assessed as optimal. Protocols with optimal results were typically based on HIER in Cell Conditioning 1 (CC1), efficient heating time 56-64 min. in BenchMark Ultra, 16 min. incubation of the primary Ab OptiView as detection kit. Using these protocol settings, 10 of 13 (77%) laboratories produced a sufficient staining result.

Concentrated antibodies for laboratory developed (LD) assays
mAb **22C3**: One protocol with an optimal result was based on HIER using Target Retrieval Solution (TRS) Low pH 6.1 (Dako) at 97°C for 50 min. on-board, Omnis (Dako). The mAb clone 22C3 was diluted 1:30, incubated for 30 min. at 22°C. EnVision Flex+ (Dako) was used as detection system and an incubation time of 10 min. in linker and 20 min. in polymer. The protocol was performed on Omnis, Dako

mAb **E1L3N**: One protocol with an optimal result was based on HIER using Bond Epitope Retrieval Solution 2 pH 9 (Leica) at 100°C for 30 min. on-board, Bond III (Leica). The mAb clone E1L3N was diluted 1:1,100, incubated for 30 min. at room temp. Refine (Leica) was used as detection system and an incubation time of 8 min. in linker and 8 min. in polymer. The protocol was performed on Bond III, Leica.

rmAb **ZR3**: One protocol with an optimal result was based on HIER in PT-module using Tris-EDTA/EGTA pH 9 at 98°C for 20 min. The rmAb clone ZR3 was diluted 1:200, incubated for 30 min. at room temp. Biosite HistoPlus (Nordic Biosite) was used as detection system and an incubation time of 30 min. in polymer. The protocol was performed on Autostainer, Thermo.

Comments
In this first NordiQC run for PD-L1 in the companion module C1, a pass rate of 50% was observed. Insufficient PD-L1 IHC staining results were most frequently characterized by a reduced proportion of cells being PD-L1 positive compared to the level expected and defined by the two PD-L1 IHC pharmDx assays, SK005 and SK006, Dako/Agilent. Consequently, incorrect and too low TPS in one or more of the NSCLC was observed and in particular the tissue cores no. 6, 10 and 16 were challenging. The NSCLC tissue cores no. 6 and 16 were expected to be TPS High ≥50%, but only demonstrated a weak to moderate staining intensity characterized by SK005 and SK006. In the insufficient results these two tumours typically were categorized as TPS low. This pattern was seen in 68% of the insufficient results.
In 15% of the insufficient results an increased proportion of PD-L1 positive cells were observed compared to the level defined by the reference assays. This typically resulted in a TPS low 1-49% in core no. 14 expected to be TPS negative. The remaining insufficient results were characterized by poor signal-to-noise ratio and/or impaired morphology.

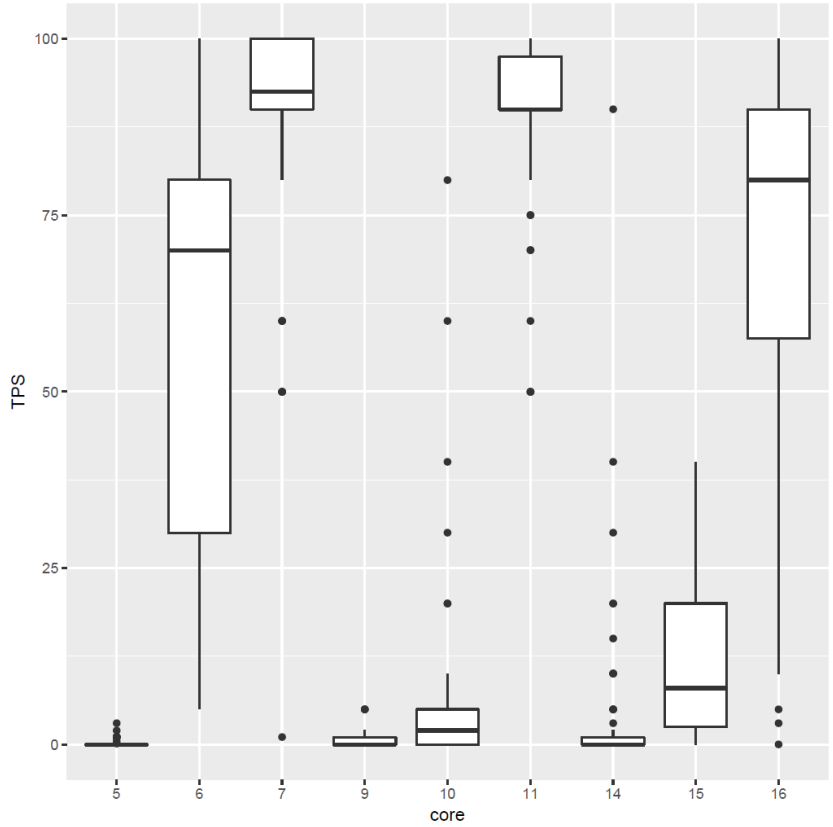
The Dako/Agilent 22C3 pharmDx assay SK006, applied by protocol settings in compliance with the vendor recommendations, was most successful with an overall pass rate of 92%. Two laboratories used the kit off-label on other staining platforms (BenchMark, Ventana and Omnis, Dako) and both produced an insufficient result. In this context, it must be emphasized that off-label use of approved assays is problematic since the intended use is violated and require an extended and challenging internal validation. Also the Dako/Agilent 28-8 pharmDx assay 28-8, was found successful with a pass rate of 86%. However a reduced proportion of optimal results was seen, which was caused by excessive cytoplasmic staining reaction compromising the interpretation. All protocols used for 28-8 (SK005) were performed in compliance with the vendor recommendations.
The Ventana PD-L1 IHC assay 790-4905, SP263 was the most widely used assay for PD-L1 and provided an overall pass rate of 77%, when based on protocol settings in compliance with the vendor recommendations.

Grouped together, the three approved PD-L1 IHC assays, 22C3 SK006 Dako, 28-8 SK005 Dako and SP263 790-4905 Ventana gave a pass rate of 80% (28 of 35 protocols performed by the intended settings).

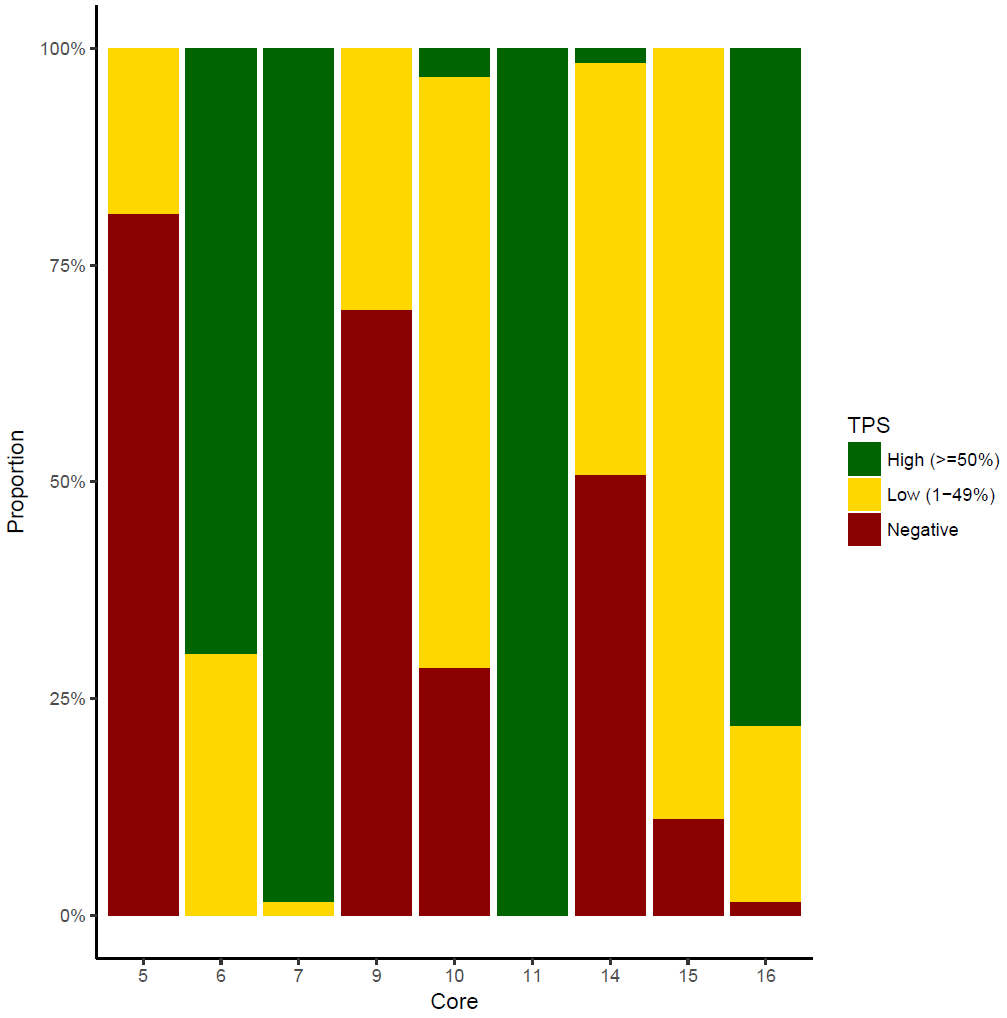
Laboratory developed (LD) assays were used by 44% (30 of 68) of the participants and for this group a pass rate of 20% (6 of 30) was observed. mAb clone 22C3 and rmAb 28-8 were used by 19 participants within a LD assay and a significantly reduced pass rate of 16% compared to the corresponding pharmDx assays SK005 and SK006 was seen. This very low pass rate for LD assays for PD-L1 IHC clearly indicates the challenge to identify and validate a reliable protocol to give concordant results to the vendor validated assays.
As this was the first assessment for PD-L1 IHC, data to identify reliable and recommended LD assays are limited. However clones 22C3, E1L3N and ZR3 could all be used to set-up a protocol providing an optimal staining result. For LD assays performed on BenchMark, Ventana the use of amplification kit in combination with OptiView as detection system seemed to be less successful. From the experience of other NordiQC assessments, this tyramide based IHC system seems to give a binary staining result being either negative or strongly positive. This has shown to be very successful for ALK demonstration in lung adenocarcinoma, but seems to be less adequate to demonstrate a dynamic range of expression levels which is essential for PD-L1 interpretation and establishment of TPS using present guidelines. In this assessment, LD assays based on OptiView + amplification kit gave the expected staining result in tissue core no. 7 and 11 characterized as TPS High with an intense expression level on the membranes of the tumour cells. Nevertheless, simultaneously an insufficient result in the tissue core no. 6 and 16 was observed. These two NSCLC were by the two pharmDx assays SK005 and SK006 expected to be TPS high, but only displayed a weak to moderate membrane staining reaction, which in the majority of LD assays based on Optiview + amplification could not be demonstrated and consequently the TPS changed from high to low.

PD-L1 interpretation and scoring consensus:
Participants were asked to evaluate the percentage of PD-L1 positive tumour cells in each of the nine NSCLC included in the assessment material. The overall interpretation of PD-L1 expression and consensus rates of the participants are shown in figures 1 and 2.

Figure 1.



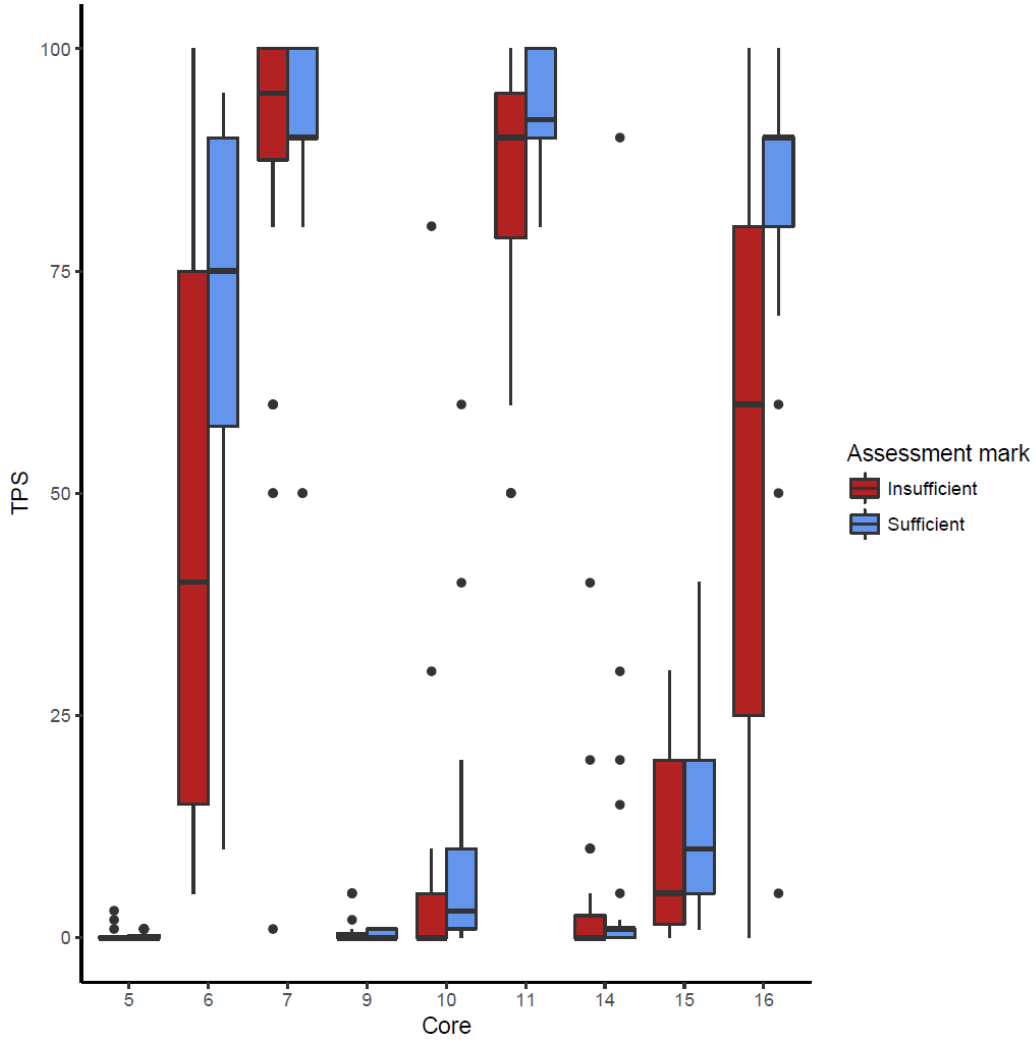
NordiQC PD-L1 run C1: participants interpretation of percentage of positive tumour cells
Figure 2.



NordiQC PD-L1 run C1: participant interpretation of PD-L1 TPS – impact on treatment

As seen in figures 1 and 2, low consensus rates for PD-L1 interpretation by the participants was observed. This finding is not surprising and can be related both to the subjectivity and complexity to interpret the TPS, but also to the low pass rate and poor accuracy of the PD-L1 IHC assays used by the participants.
Concerning the complexity to score PD-L1 it was thus revealed that 20% of the participants classified tumour core no. 5 as TPS low. This tumour was consistently TPS negative (both by NordiQC reference protocols and all participants), but included immune cells/macrophages, which was interpreted as tumour cells by many participants – see photo Fig. 2a.
When stratifying for the assessment marks, analysis revealed that scores among participants that had received an insufficient mark ("Borderline" or "Poor") reported lower TPS scores than laboratories that had received a sufficient mark (see figure 3). However, this difference was not statistical significant and showed some overlap in scores between the two groups.

Figure 3.



NordiQC PD-L1 run C1: interpretation concordance for labs with sufficient results versus insufficient results

Controls

In this assessment both normal tissues and cell lines were included in order to evaluate their potential as controls for PD-L1.

Tonsil and placenta were used as positive and negative tissue controls. In this assessment, tonsil was found to be superior to placenta, as tonsil displayed a range of PD-L1 expression, whereas placenta showed a more binary expression level with cells being either negative or positive. Using PD-L1 IHC 28-8, SK005 and 22C3, SK006 Dako/Agilent and obtaining an optimal staining result, tonsil displayed the following reaction pattern: No staining reaction in the vast majority of lymphocytes including mantle zone and germinal centre B-cells, a weak to moderate, typically punctuated membranous staining reaction of the majority of germinal centre macrophages and finally a moderate to strong staining reaction of the majority of epithelial crypt cells. SP263, 790-4905 Ventana/Roche provided similar staining pattern, but with an increased number of immune cells (see photo Fig. 1b).

However, it was observed that a fully acceptable staining pattern in tonsil could be obtained together with insufficient and false negative result in the NSCLC. This scenario was observed by the using Optiview + amplification within a LD assay typically based on mAb 22C3. This underlines the need to identify more reliable positive tissue controls for PD-L1 and/or improve the interpretation criteria for a sufficient staining reaction in tonsil e.g. more accurately specify number and intensity of cells expected to be demonstrated.

In this light, four different cell lines (Horizon Discovery, Cambridge UK) were included in this assessment, primarily to evaluate if this material in combination with digital image analysis can be used to evaluate staining quality for PD-L1 and potentially be used as standard reference material for the validation of the precision of PD-L1 IHC assays. Subsequent analysis will be performed by NordiQC and published at a later stage.

Conclusion

This was the first NordiQC assessment of PD-L1. 68 laboratories participated and a pass rate of 50% was observed. Companion diagnostic PD-L1 assays showed a significant superior pass rate compared to LD assays. The PD-L1 22C3 IHC assay, SK006 was most successful, giving a pass rate of 92%. Overall LD assays provided a worrying pass rate of 20%.

Insufficient results were typically characterized by a reduced TPS compared to the level expected and caused by imprecise calibration of the LD assay typically using tyramide based detection systems (OptiView+ amplification, Ventana).

Tonsil is at present the preferred choice as positive and negative tissue control for PD-L1. The majority of epithelial crypt cells must show a moderate to strong staining reaction, while the germinal centre macrophages must show a weak to moderate membranous staining reaction. No staining must be seen in the vast majority of lymphocytes.

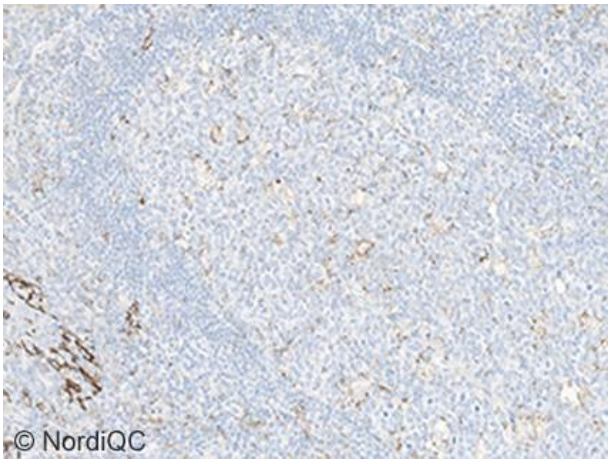


Fig. 1a
Optimal PD-L1 staining result of tonsil using the pharmDx IHC PD-L1 assay, SK006, Dako/Agilent based on the mAb clone 22C3.
Crypt epithelial cells show a moderate to strong, staining reaction, while the majority of germinal centre macrophages show a weak to moderate membranous staining reaction. The vast majority of lymphoid cells are negative.
Also compare with Figs. 2a – 5a, same protocol.

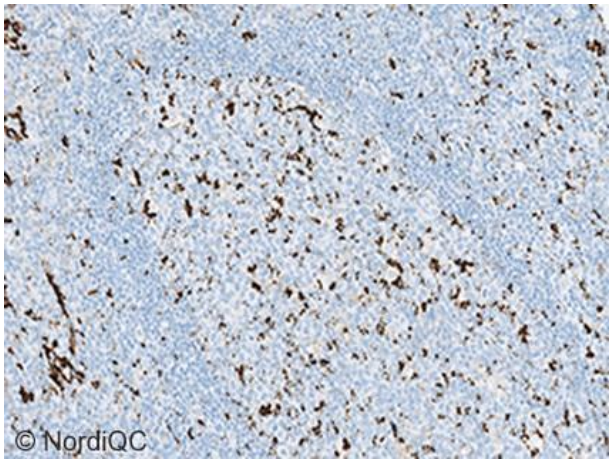


Fig. 1b
PD-L1 staining result of tonsil using the mAb clone 22C3 as concentrated format (M3653, Dako/Agilent) within a laboratory developed assay. The assay was based on OptiView + amplification and performed on BenchMark Ultra (Ventana/Roche)
A moderate to strong staining reaction is seen in both crypt epithelial cells and in germinal centre macrophages. However the assay was not balanced to demonstrate the range of weak to strong PD-L1 expression in tumour cells in the NSCLC used in the assessment material.
Compare with Figs. 2b – 5b, same protocol.

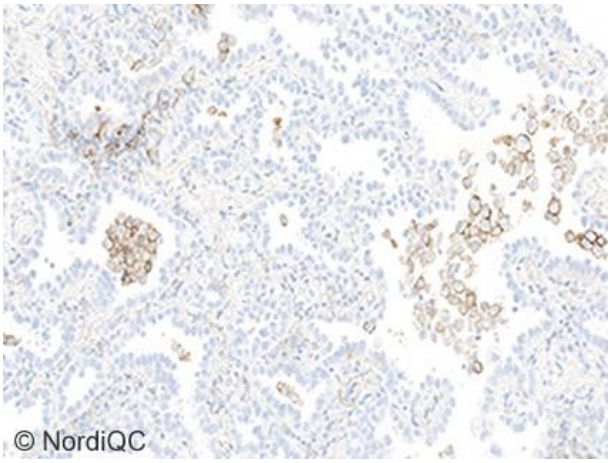


Fig. 2a
Optimal PD-L1 staining result of the NSCLC tissue core no. 14 using same protocol as in Fig. 1a.
The neoplastic cells are all negative and only macrophages and dispersed lymphocytes show a distinct membranous staining reaction serving as internal positive tissue control.
The tumour is categorized as TPS negative.

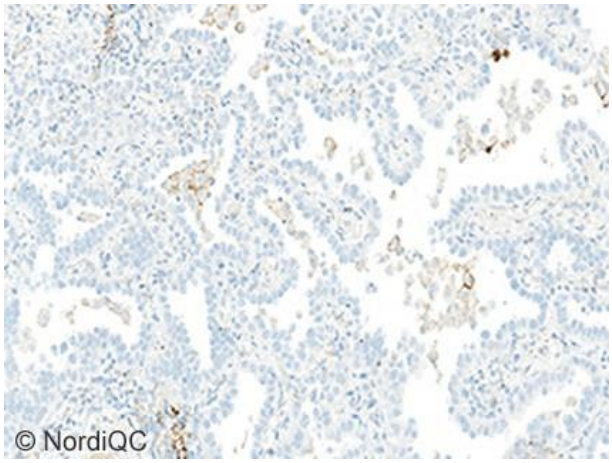


Fig. 2b
PD-L1 staining result of the NSCLC tissue core no. 14 using same protocol as in Fig. 1b.
The neoplastic cells are all negative and only macrophages and dispersed lymphocytes are demonstrated.
The tumour is categorized as TPS negative.
However also compare with Figs. 3 – 5b

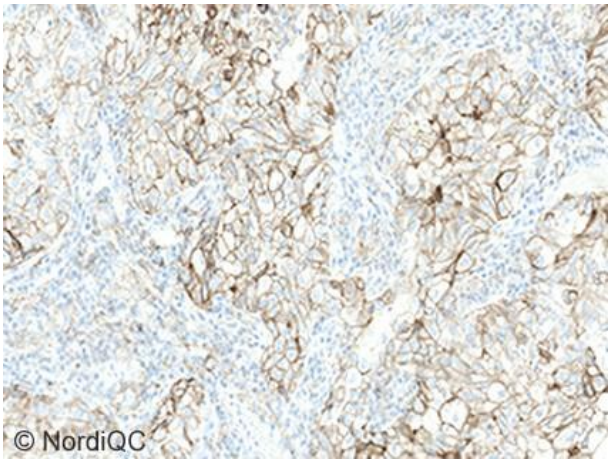


Fig. 3a
Optimal PD-L1 staining result of the NSCLC tissue core no. 11 using same protocol as in Figs. 1a and 2a.
The vast majority of neoplastic cells show a moderate and distinct membranous staining reaction.
The tumour is categorized as TPS high and thus eligible for 1’ line immune therapy.

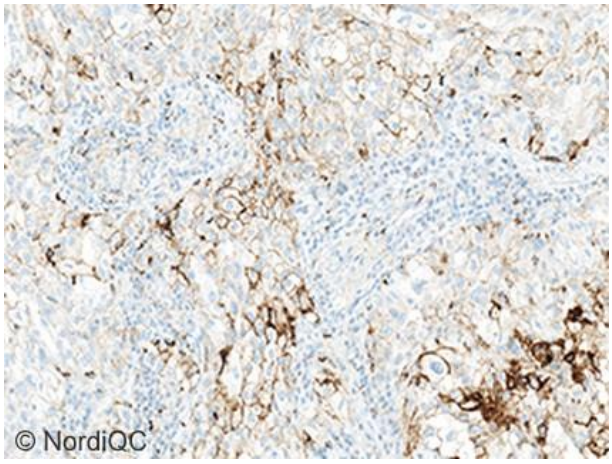


Fig. 3b
PD-L1 staining result of the NSCLC tissue core no. 11 using same protocol as in Figs. 1b and 2b.
The majority of neoplastic cells show a moderate to strong membranous staining reaction. The reaction is very granular compared to the pattern seen in Fig. 3a caused by the use of amplification.
The tumour is categorized as TPS high and thus eligible for 1’ line immune therapy.

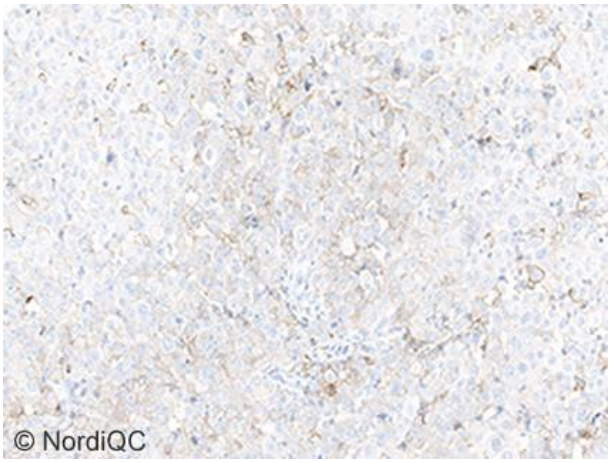


Fig. 4a
Optimal PD-L1 staining result of the NSCLC tissue core no. 6 using same protocol as in Figs. 1a - 3a.
The majority of neoplastic cells show a weak to moderate membranous staining reaction - in some areas only very weak and partial membranous staining reaction.
The tumour is categorized as TPS high and thus eligible for 1’ line immune therapy.

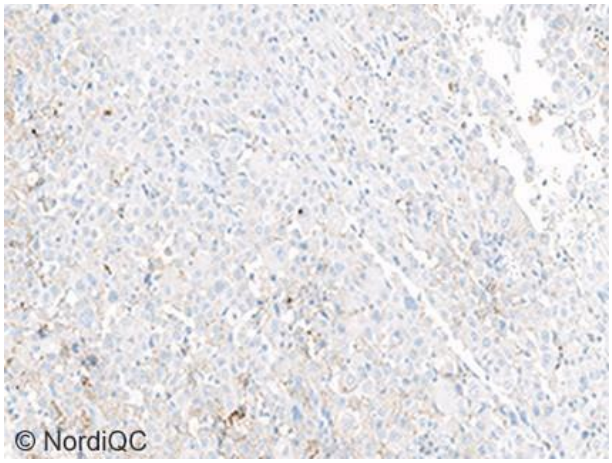


Fig. 4b
Insufficient PD-L1 staining result of the NSCLC tissue core no. 6 using same protocol as in Figs. 1b - 3b.
Only few cells show a distinct membranous staining reaction, whereas the vast majority are negative. Few immune cells show a strong staining reaction. The assay has not been calibrated to demonstrate the weak PD-L1 expression in the tumour cells.
The tumour is consequently categorized as TPS low, and thus only eligible for 2’ line immune therapy.

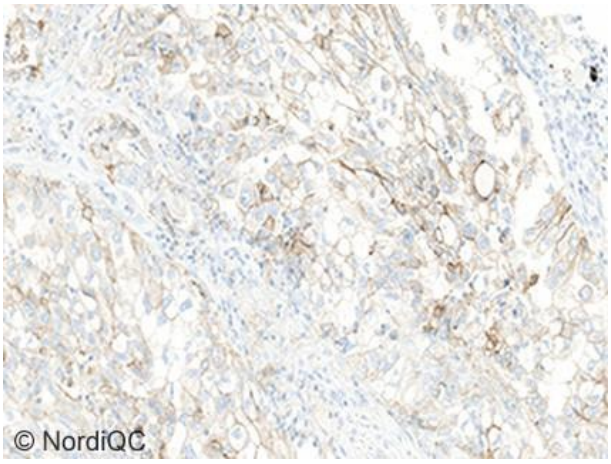


Fig. 5a
Optimal PD-L1 staining result of the NSCLC tissue core no. 16 using same protocol as in Figs. 1a - 4a.
The majority of neoplastic cells show a weak to moderate membranous staining reaction - in some areas only very weak and partial membranous staining reaction.
The tumour is categorized as TPS high and thus eligible for 1’ line immune therapy.

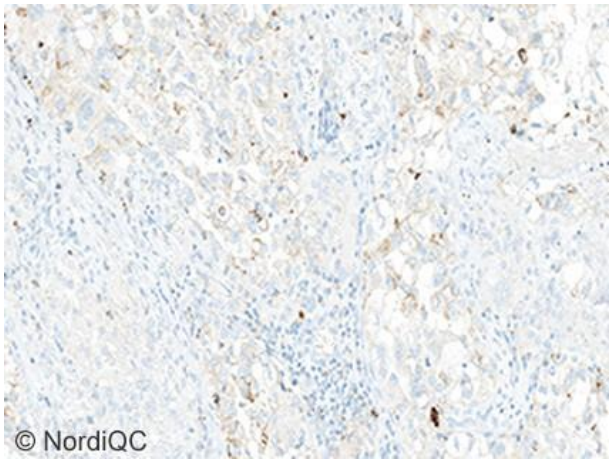


Fig. 5b
Insufficient PD-L1 staining result of the NSCLC tissue core no. 16 using same protocol as in Figs. 1b - 4b.
Only few cells show a distinct membranous staining reaction, whereas the vast majority are negative. Few immune cells show a strong staining reaction. The assay has not been calibrated to demonstrate the weak PD-L1 expression in the tumour cells.
The tumour is consequently categorized as TPS low and thus only eligible for 2’ line immune therapy.

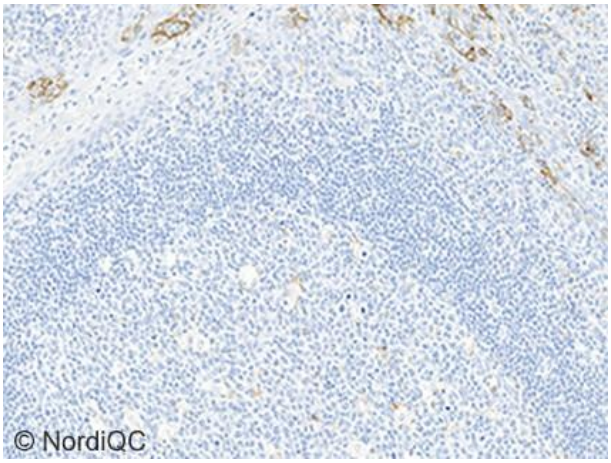


Fig. 6a
Insufficient PD-L1 staining result of tonsil using mAb clone E1L3N within a laboratory developed assay. The assay was based on a protocol providing a too low analytical sensitivity.
Only crypt epithelial cells show a moderate and distinct staining reaction, while germinal centre macrophages are negative.
Also compare with Fig. 6b, same protocol.

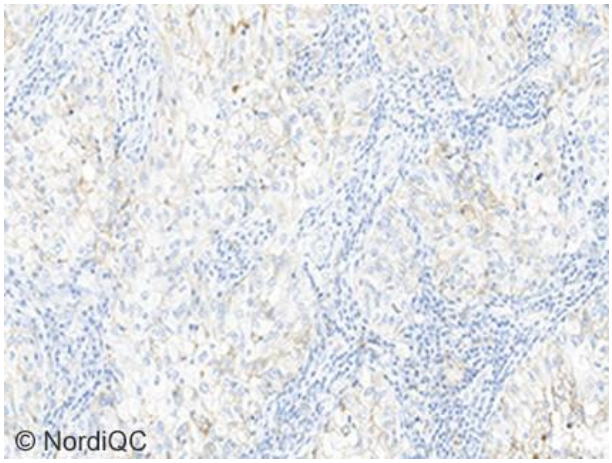


Fig. 6b
Insufficient PD-L1 staining result of the NSCLC tissue core no. 16 using same protocol as in Fig. 6a.
Only few cells show a distinct membranous staining reaction, whereas the vast majority are negative. The assay has not been calibrated to demonstrate the weak PD-L1 expression in the tumour cells.
The tumour is consequently categorized as TPS low and thus only eligible for 2' line immune therapy.

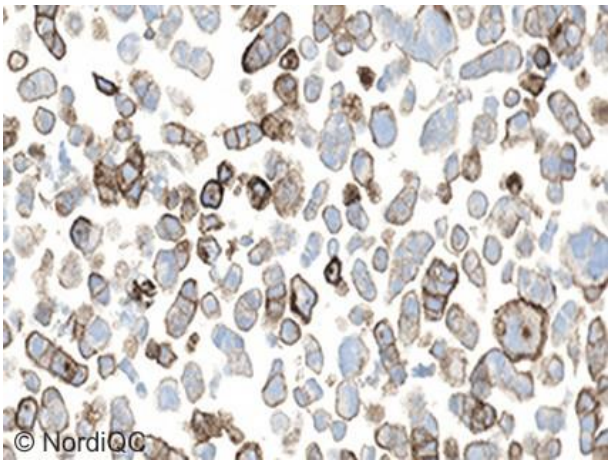


Fig. 7a
Optimal PD-L1 staining result of the cell line (Horizon Discovery) core no. 2 using the pharmDx IHC PD-L1 assay, SK006, Dako/Agilent based on the mAb clone 22C3.
Virtually all cells show an intermediate PD-L1 expression. Subsequent digital image analysis will be performed to evaluate if a defined reference score (number of cells and intensity) can be identified to monitor the accuracy and precision of PD-L1 IHC assays.

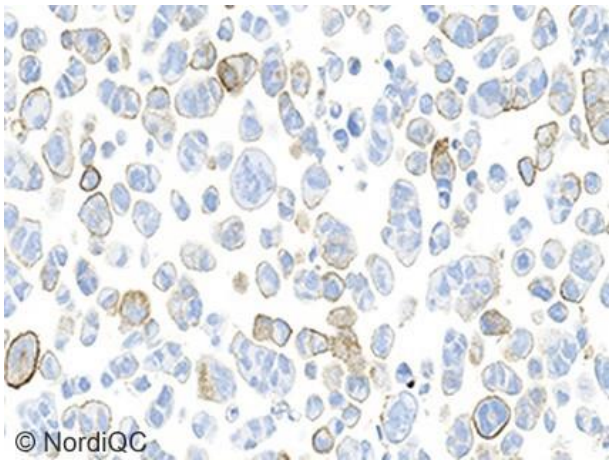


Fig. 7b
Insufficient PD-L1 staining result of the cell line (Horizon Discovery) core no. 2 mAb clone E1L3N within a laboratory developed assay - same protocol as used in Figs. 6a and 6b.
A significantly reduced proportion and intensity of the cells demonstrated is seen.

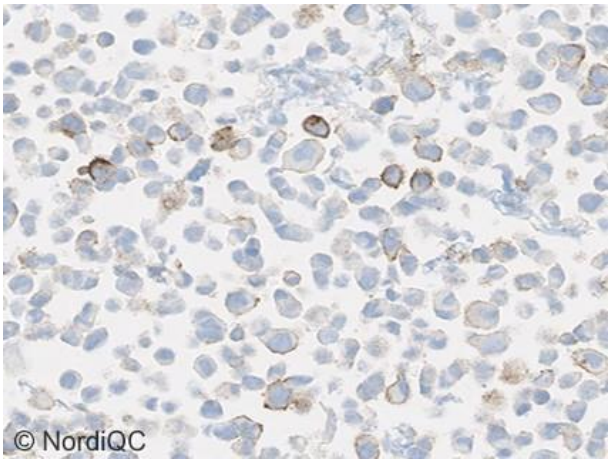


Fig. 8a
Optimal PD-L1 staining result of the cell line (Horizon Discovery) core no. 3 using the pharmDx IHC PD-L1 assay, SK006, Dako/Agilent based on the mAb clone 22C3.
The majority of cells show a low PD-L1 expression. Subsequent digital image analysis will be performed to evaluate if a defined reference score (number of cells and intensity) can be identified to monitor the accuracy and precision of PD-L1 IHC assays.

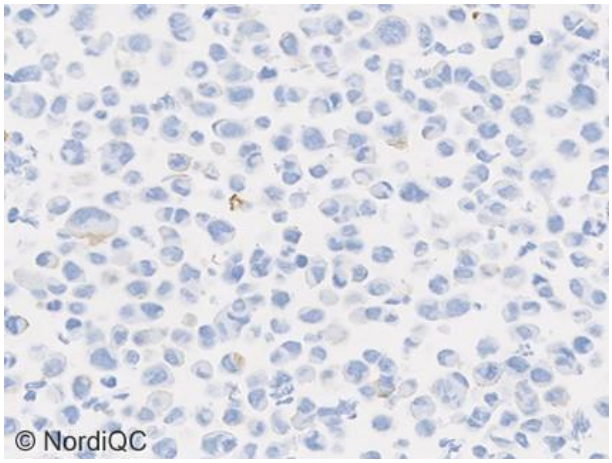


Fig. 8b
Insufficient PD-L1 staining result of the cell line (Horizon Discovery) core no. 3 mAb clone E1L3N within a laboratory developed assay - same protocol as used in Figs. 6a and 6b.
A significantly reduced proportion and intensity of the cells demonstrated is seen.

SN/LE/JLL/RR 20.04.2017