Assessment Run C6 2019
PD-L1 KEYTRUDA®

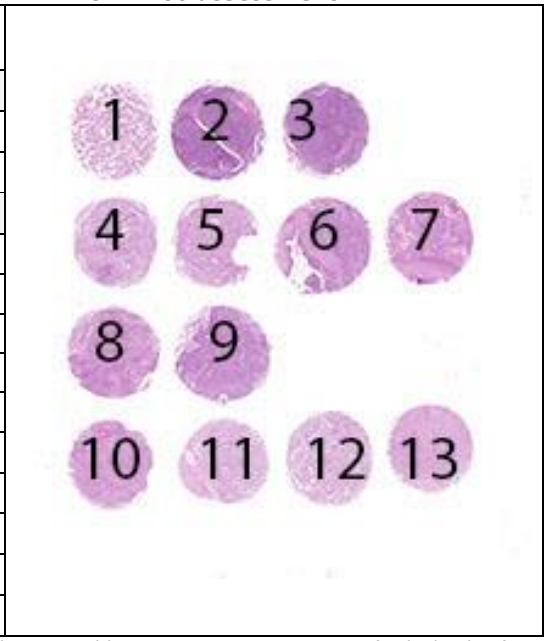
Purpose

This was the sixth assessment for PD-L1 in the NordiQC Companion module and it was modified compared to the previous five assessments. This assessment PD-L1 KEYTRUDA® primarily focused on the evaluation of the analytical accuracy of the IHC assays performed by the NordiQC participants to identify patients with NSCLCs and urothelial carcinomas to be treated with KEYTRUDA® as immune therapy. PD-L1 22C3 pharmDx, SK006 (Dako/Agilent) was used as reference standard method, and accuracy was evaluated in the carcinomas with the dynamic and critical relevant expression levels of PD-L1 characterized and evaluated by TPS and CPS. The obtained score in NordiQC is indicative of the performance of the IHC tests but due to the limited number and composition of samples internal validation and extended quality control e.g. regularly measuring the PD-L1 results is needed.

Material

Table 1. **Content of the TMA used for the NordiQC PD-L1 KEYTRUDA® C6 assessment**

	PD-L1 IHC TPS/CPS score*
Tissue controls	
1. Placenta	See section for controls
2-3. Tonsil	See section for controls
Carcinomas	
4. NSCLC	TPS: No; <1%
5. NSCLC	TPS: No; <1%
6. NSCLC	Excluded**
7. NSCLC	TPS: Low; ≥1-49%
8. NSCLC	TPS: High; ≥50%
9. NSCLC	TPS: High; ≥50%
10. Urothelial carcinoma	CPS: <10
11. Urothelial carcinoma	CPS: ≥10
12. Urothelial carcinoma	CPS: ≥10
13. Urothelial carcinoma	CPS: ≥10



* Tumour proportion score (TPS) and combined positive score (CPS) determined by PD-L1 IHC 22C3, SK006 (Dako/Agilent) performed in NordiQC reference lab.
** Excluded due to a combined feature of severe heterogeneity and levels with too few vital neoplastic cells

All tissues were fixed in 10% neutral buffered formalin.

The participating laboratories were asked to perform the PD-L1 IHC assay for treatment with KEYTRUDA®, interpret the PD-L1 expression level using the TPS and CPS scoring system and submit the stained slides and scores to NordiQC. This allowed assessment of the technical performance (analytical accuracy) of the PD-L1 KEYTRUDA® assays and provided information on the reproducibility and concordance of the PD-L1 interpretation results among the laboratories.

PD-L1 KEYTRUDA®, Technical assessment
In order to account for heterogeneity of PD-L1 expression in the individual tumour cores included in the tissue TMA blocks, reference slides were made through out the blocks. Every twentieth slide throughout the blocks were stained for PD-L1 using the CE IVD / FDA approved 22C3 pharmDx kit SK006 (Dako/Agilent), and also by the CE IVD approved assay (NSCLC, KEYTRUDA®) SP263 790-4905 (Ventana/Roche) in a NordiQC reference laboratory. During the assessment, TPS and CPS categories for each tissue core on the submitted slides were compared to the level in the nearest reference slide of 22C3 pharmDX SK006 (Dako).

Criteria for assessing a staining as Optimal include:
The staining is considered perfect or close to perfect in all of the included tissues.
TPS/CPS is concordant to the NordiQC reference data in all 9 carcinomas.

Criteria for assessing a staining as Good include:
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/CPS is still concordant to the NordiQC reference data in all 9 carcinomas.

Criteria for assessing a staining as Borderline include:
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/CPS is **not** concordant to the NordiQC reference data in 1 of 9 carcinomas.

Criteria for assessing a staining as Poor include:
The staining is considered very insufficient e.g., because of a false negative or a false positive staining reaction staining of more of the included tissues.
An optimization of the protocol is urgently needed.
TPS/CPS is **not** concordant to the NordiQC reference data in more than 1 of 9 carcinomas.

A staining can also be assessed as **borderline/poor** in case the interpretation and scoring is significantly hampered e.g. by impaired morphology, excessive background reaction etc.

Participation

Number of laboratories registered for PD-L1 KEYTRUDA IHC C6	188
Number of laboratories returning PD-L1 KEYTRUDA IHC	183 (97%)*
Number of laboratories returning PD-L1 scoring sheet	162 (89%)

*One laboratory was excluded due to contamination with a CD45 like antibody, which compromised the assessment of the staining. Data is not included in the analyze below.

Results: 182 laboratories participated in this assessment and 75% achieved a sufficient mark. Assessment marks for IHC PD-L1 assays and PD-L1 antibodies are summarized in Table 3 (see next page).

Table 2. **Proportion of sufficient results for PD-L1 (lung)/KETRUDA in the six NordiQC runs performed**

	C1 2017	C2 2018	C3 2018	C4 2018	C5 2019	C6 2019
Participants, n=	68	145	146	163	176	182
Sufficient results	50%	84%	91%	86%	80%	75%

Performance history
This was the sixth NordiQC assessment of PD-L1 (lung)/KEYTRUDA®. A slightly reduced pass rate was obtained in C6 (see Table 2) compared to the latest assessments. The number of new participants seems to be consistently increasing.

Conclusion
In this sixth NordiQC run for PD-L1 (lung)/KEYTRUDA® in the companion module C6, an overall pass rate of 75% was achieved. Insufficient PD-L1 IHC staining results were most frequently characterized by a reduced proportion of PD-L1 positive cells compared to the level expected as defined by the PD-L1 IHC pharmDx assay, SK006 (Dako/Agilent). This resulted in a too low TPS/CPS in one or more of the carcinomas.
Several companion diagnostic assays and laboratory developed (LD) assays based on clones as 22C3, E1L3N, CAL10, ZR3 and BSR90 could provide optimal results. The companion diagnostic PD-L1 IHC assays, 22C3 SK006/GE006 and 28-8 SK005 from Dako/Agilent and SP263 from Ventana/Roche performed in concordance to the product guidelines, provided a high proportion of sufficient results and

was superior to LD assays. Within LD assays, irrespectively of Ab clone being used, meticulous calibration and validation of the assay is required.

Table 3. **Assessment marks for IHC assays and antibodies run C6, PD-L1 KEYTRUDA®**

CE-IVD / FDA approved PD-L1 assays	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
rmAb clone SP263, **740-4907**[3]	6	Ventana/Roche	6	-	-	-	100%	100%
rmAb clone SP263, **740-4907**[4]	2	Ventana/Roche	1	-	1	-	-	-
rmAb clone SP263, **741-4905**[5]	30	Ventana/Roche	17	6	4	3	77%	81%
rmAb clone SP263, **790-4905**[6]	27	Ventana/Roche	19	7	1	-	96%	96%
mAb clone 22C3 pharmDX, **SK006**[7]	22	Dako/Agilent	8	9	4	1	77%	76%
mAb clone 22C3 pharmDX, **SK006**[4]	9	Dako/Agilent	2	3	1	3	56%	-
mAb clone 22C3 pharmDX, **GE006**[8]	11	Dako/Agilent	8	3	-	-	100%	100%
mAb clone 22C3 pharmDX, **GE006**[4]	2	Dako/Agilent	-	-	1	1	-	-
rmAb clone 28-8 pharmDX, **SK005**[9]	3	Dako/Agilent	2	1	-	-	-	-
Antibodies[10] for laboratory developed PD-L1 assays, concentrated antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
mAb clone **22C3**	50	Dako/Agilent	15	18	14	3	66%	76%
mAb clone **E1L3N**	6	Cell Signaling	1	1	2	2	33%	-
rmAb **CAL10**	2 3	Biocare Zytomed Systems	1	2	1	1	60%	-
rmAb clone **ZR3**	1 1 1	Zeta Corporation Nordic Biosite Gene Tech	2	-	1	-	-	-
rmAb clone **QR1**	1	Diagomics	-	1	-	-	-	-
rmAb **BSR90**	1	Nordic Biosite	-	1	-	-	-	-
rmAb clone **SP142**	1	Spring Biosystems	1	-	-	-	-	-
Ready-To-Use antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
rmAb clone 73-10, **PA0832**	1	Leica Biosystems	-	1	-	-	-	-
rmAb clone CAL10, **API 3171**	1	Biocare	-	-	-	1	-	-
mAb clone MX070C, **MAB-0854**	1	Maixin	1	-	-	-	-	-
Total	182		84	53	30	15		
Proportion			46%	29%	17%	8%	75%	

1) Proportion of sufficient stains (optimal or good).
2) Proportion of sufficient stains with optimal protocol settings only, see below.
3) Approved for IMFINZI in urothelial cancers in US.
4) RTU system used on a different platform than it was developed for.
5) Approved for IMFINZI, KEYTRUDA and OPDIVO in NSCLC in EU.
6) Analytical claim
7) Approved for KEYTRUDA in NSCLC in EU/US.
8) Approved for KEYTRUDA in NSCLC in EU
9) Approved for OPDIVO in NSCLC in EU/US
10) mAb: mouse monoclonal antibody, rmAb: rabbit monoclonal antibody.

Detailed Analysis
CE IVD / FDA approved assays

SP263 (740-4907, Ventana): 6 of 6 (100%) protocols were assessed as optimal. The protocols were based on Heat Induced Epitope Retrieval (HIER) in Cell Conditioning 1 (CC1), efficient heating time 32-64 min., 16 min. incubation of the primary Ab, OptiView with or without OptiView Amplification as detection system and performed on BenchMark Ultra.

SP263 (741-4905, Ventana): 17 of 30 (57%) protocols were assessed as optimal. Protocols with optimal results were based on HIER in CC1, efficient heating time 52-64 min., 16 min. incubation of the primary Ab, OptiView as detection system and performed on BenchMark Ultra. Using these protocols settings, 21 of 26 (81%) laboratories produced a sufficient staining result (optimal or good).

SP263 (790-4905, Ventana): 19 of 27 (70%) protocols were assessed as optimal. Protocols with optimal results were typically based on HIER in CC1, efficient heating time 52-64 min., 16 min. incubation of the primary Ab, OptiView with or without OptiView Amplification as detection system and performed on BenchMark Ultra. Using these protocol settings, 23 of 24 (96%) laboratories produced a sufficient staining result.

PD-L1 IHC 22C3 pharmDx (SK006, Dako): 8 of 22 (36%) protocols were assessed as optimal. Protocols with optimal results were typically based on HIER using EnVision™ Flex Target Retrieval Solution (TRS) low pH 6.1 (SK006) at 95-99°C for 20-25 min. (PT Link), 30 min. incubation of the primary Ab, EnVision Flex+ as the detection system and performed on Autostainer Link 48. Using these protocol settings, 16 of 21 (76%) laboratories produced a sufficient staining result.

PD-L1 IHC 22C3 pharmDx (GE006, Dako): 8 of 11 (73%) protocols were assessed as optimal. Protocols with optimal results were based on HIER using EnVision™ Flex Target Retrieval Solution (TRS) low pH 6.1 (GV805) at 95-99°C for 40 min. (PT Link), 40 min. incubation of the primary Ab, EnVision Flex+ as the detection system and performed on Omnis. Using these protocol settings, 11 of 11 (100%) laboratories produced a sufficient staining result.

PD-L1 IHC 28-8 pharmDx (SK005, Dako): 2 of 3 protocols were assessed as optimal. Protocols with optimal results were based on HIER using EnVision™ Flex TRS low pH 6.1 at 97°C for 20 min. (PT Link), 30 min. incubation of the primary Ab, EnVison Flex+ as the detection system and performed on Autostainer Link 48. Using these protocol settings, all three laboratories produced a sufficient staining result.

Table 4 summarizes the proportion of sufficient and optimal marks for the most commonly used CE IVD / FDA approved assays. The performance was evaluated both as "true" plug-and-play systems performed strictly accordingly to the vendor recommendations and by laboratory modified systems changing basal protocol settings. Only protocols performed on the specific IHC stainer device are included.

Table 4. **Comparison of pass rates for vendor recommended and laboratory modified protocols**

CDx assay*	**Vendor recommended protocol settings****		**Laboratory modified protocol settings*****	
	Sufficient	Optimal	Sufficient	Optimal
Ventana BenchMark Ultra rmAb SP263, **740-4907**	3/3	3/3	3/3	3/3
Ventana BenchMark XT, GX, Ultra rmAb SP263, **741-4905**	19/22 (86%)	16/22 (73%)	4/8 (50%)	1/8 (13%)
Ventana BenchMark XT, GX, Ultra rmAb SP263, **790-4905**	14/14 (100%)	13/14 (93%)	12/13 (92%)	6/13 (46%)
Dako Autostainer Link 48+ mAb 22C3 pharmDX, **SK006**	14/19 (74%)	7/19 (37%)	3/3	1/3
Dako Omnis mAb 22C3 pharmDX, **GE006**	11/11 (100%)	8/11 (73%)	-	-
Dako Autostainer Link 48+ rmAb 28-8 pharmDX, **SK005**	3/3	2/3	-	-

*See Table 3 for examples of approval and indications.
**Protocol settings recommended by vendor – Retrieval method and duration, Ab incubation times, detection kit, IHC stainer/equipment.
***Modifications in one or more of above mentioned parameters. Only protocols performed on the specified vendor IHC stainer are included.

Concentrated antibodies for laboratory developed (LD) assays
mAb **22C3**: 15 of 50 (30%) protocols were assessed as optimal of which 7 were stained on the Benchmark (Ventana) and six on the Omnis (Dako) platform.
On BenchMark XT/Ultra (Ventana), the protocols providing optimal results were typically based on a titre of 1:40-50, primary Ab incubation time of 32-64 min., HIER in CC1 (efficient heating time 48-64 min.) and OptiView as detection system. Using these protocol settings, 12 of 15 (80%) laboratories produced a sufficient staining result. Four laboratories applied OptiView Amplification to the protocol settings listed above. 2 of 4 produced a sufficient staining result.
On Omnis (Dako), the protocols providing optimal results were typically based on a titre of 1:20-30 of the primary Ab, incubation time of 30-60 min., HIER in TRS low pH 6.1 (Dako) at 97°C (efficient heating time 30-50 min.) and EnVision FLEX+ as detection system. Using these protocol settings, 8 of 9 (89%) laboratories produced a sufficient staining result.
In total, 28 of 37 (76%) laboratories used optimal protocol settings and obtained a sufficient staining result.

mAb **E1L3N**: One protocol provided an optimal result. The protocol was based on HIER using an alkaline-buffer at 95°C for 30 min. The mAb clone E1L3N was diluted 1:100, incubated for 30 min. at room temp. and visualized using a 3-layer detection system. Using these protocol settings, 2 of 2 laboratories obtained a sufficient staining result.

rmAb **CAL10**: One protocol was assessed as optimal. The protocol was based on a titre of 1:50, primary Ab incubation time of 15 min., HIER in an alkaline buffer for 30 min. and visualized using a 3-layer detection system. Using these protocol settings, 2 of 3 laboratories obtained a sufficient staining result.

rmAb **ZR3**: Two protocols were assessed as optimal. The protocols were based on a titre of 1:150, primary Ab incubation time of 20-50 min., HIER in an alkaline buffer for 15-40 min. and visualized using a 2- or 3-layer detection system. Using these protocol settings, 2 of 2 laboratories obtained a sufficient staining result.

rmAb **SP142**: One protocol was assessed as optimal. The protocol was based on a titre of 1:400, primary Ab incubation time of 30 min., HIER in a non-alkaline buffer for 25 min. and visualized using a 3-layer detection system. Only one laboratory used rmAb SP142.

Table 5. **Optimal results for PD-L1 for the most commonly used antibodies as concentrates on the four main IHC systems***

Concentrated antibodies	**Ventana/Roche BenchMark GX/XT/Ultra**		**Dako/Agilent Autostainer**		**Dako/Agilent Omnis**		**Leica Bond III/Max**	
	CC1 pH 8.5	CC2 pH 6.0	TRS pH 9.0	TRS pH 6.1	TRS High pH	TRS Low pH	BERS2 pH 9.0	BERS1 pH 6.0
mAb clone **22C3**	7/25	-	1/3	0/4	1/1	5/8	0/3	0/1
mAb clone **E1L3N**	-	-	-	-	-	-	1/2	-
rmAb clone **CAL10**	0/1	-	-	-	1/1	-	-	-
rmAb clone **ZR3**	-	-	-	-	1/1	-	-	-

*Antibody concentration applied as listed above, HIER buffers and detection kits used as provided by the vendors of the respective platforms.
**number of optimal results/number of laboratories using this buffer

Block construction and assessment challenges
The tissue micro array (TMA) block constructed for this PD-L1 run consisted of 6 NSCLCs, 4 urothelial carcinomas 2 tonsils and 1 placenta. The NSCLCs were selected so the slides cut from the block would contain 2 NSCLCs being TPS negative (<1% PD-L1 positive tumour cells), 2 NSCLCs of each group: TPS low (≥1-49%) and TPS high (≥50%). The urothelial carcinomas were selected to contain 1 carcinoma with CPS <10 and 3 with CPS≥10, one with PD-L1 expression primarily in immune cells, one with PD-L1 expression primarily in tumour cells and one with PD-L1 expression in both tumour cells and immune cells, respectively. Reference slides throughout the block were stained using the companion diagnostic assays 22C3 pharmDX SK006 (Dako). In total, four identical blocks were constructed and cut slides were sent to the participants. Tissue core no. 6 was excluded from the assessment due to too few tumour cells in some slides and excessive heterogeneity.

Reviewing the reference slides from the blocks, heterogenic expression of PD-L1 was seen in one of the tumor cores. In tissue core no. 8, initially scored as TPS high (≥50%), slides with a TPS low ≥1-49% was identified. During the assessment, TPS and CPS categories for each tissue core on the submitted slides were compared to the level in the nearest reference slides of 22C3 pharmDX SK006 (Dako). Heterogeneity in PD-L1 expression is well known in NSCLCs and the assessment in this sense emulated clinical settings.

Comments

In this sixth NordiQC assessment for PD-L1 for (lung)/KEYTRUDA®, the prevalent feature of an insufficient staining result was a too weak or false negative staining result, being observed in 82% (37 of 45) of the insufficient staining results. As shown in Table 6, a false negative staining result has been the most common reason for insufficient staining results in all NordiQC PD-L1 (lung)/KEYTRUDA® assessments with an average occurrence of 75%. In this run, 11% (5 of 45) of the insufficient results were caused by a false positive staining result. Poor-signal-to-noise ratio, excessive cytoplasmic staining reaction interfering the interpretation or technical issues were observed in the remaining 7% of the insufficient results (3 of 45). 29% (53 of 182) of the participants obtained a score as Good. In 60% of these (32 of 53), this was due to a general weak staining result or a reduced TPS, but with no change in the TPS-category. In the remaining 40%, excessive background staining, impaired morphology or cytoplasmic staining reaction interfering the interpretation was observed. No obvious reason for this observation could be identified.

Table 6. **Characteristics of insufficient results in the last six NordiQC PD-L1 (lung)/KEYTRUDA® runs.**

	False Negative (TPS changes from high to low or low to negative)	**False Positive (TPS changes from negative to low or low to high)**	**Other Cause "Technical"**
C1 (50% insufficient)	68%	15%	17%
C2 (16% insufficient)	68%	23%	5%
C3 (9% insufficient)	100%	-	-
C4 (14% insufficient)	74%	4%	22%
C5 (20% insufficient)	59%	18%	23%
C6 (25% insufficient)	82%	11%	7%
Average (22% insufficient)	75%	12%	13%

Tissue cores no. 4, 7 and 8 were the most challenging tissues to obtain optimal staining for the laboratories in this assessment and required an accurate and carefully calibrated protocol. The majority of false negative results were especially seen in core 8, changing the TPS category from high to low, which would change the status for first line immune therapy using the present guidelines for NSCLCs as e.g. in EU. Virtually all insufficient results were related to incorrect TPS categories in one or more of the NSCLCs, whereas the CPS scores only were affected in the urothelial carcinomas in a few cases. PD-L1 IHC demonstration was in this assessment more successful in urothelial carcinomas versus NSCLCs. No plausible reasons for this difference could be identified, but the expression levels in the materials used for the assessments in combination with the different cut-off values for the two entities might have favoured PD-L1 demonstration in urothelial carcinomas. In order to evaluate IHC accuracy NordiQC strives to include neoplasias with PD-L1 levels close to the critical and clinical relevant thresholds for positivity focusing on both intensity, proportion and subtypes of cells to be scored. In this assessment NSCLCs might thus have been more challenging compared to the urothelial carcinomas

The Ventana PD-L1 IHC assays 740-4907/741-4905/790-4905, based on the SP263 clone, were the most widely used assays for demonstration of PD-L1 and provided a pass rate of 87% (55 of 63). Applying protocol settings in compliance with the vendor recommendations, the pass rate was 92% (36 of 39). In comparison, protocols based on laboratory modified protocol settings obtained a pass rate of 79% (19 of 24). A common modification was addition of OptiView Amplification kit based on tyramide, giving a stronger and more granular staining reaction that could interfere the interpretation. It is well known from previous assessments in NordiQC, both for other epitopes (general module) and previous PD-L1 runs, that IHC assays based on tyramide amplification can be challenging as especially low-level expressing tissue structures may be negative and, if not carefully calibrated, can cause false positive staining result in structures expected to be negative.

The Dako/Agilent 22C3 pharmDx assay SK006 provided an overall pass rate of 77% (17 of 22). When using the recommend protocol settings from Dako, the pass rate was 74% (14 of 19). Three laboratories modified the protocols - all obtained a sufficient result.
Nine laboratories applied the SK006 RTU product on another stainer platform than the Dako Autostainer Link 48, providing a pass rate of 56% (5 of 9).

All 11 laboratories using the Dako/Agilent 22C3 pharmDx assay GE006 for Omnis used the recommended protocol settings and all obtained optimal results.
It was thus observed that the recently launched PD-L1 22C3 GE006 assay for Omnis was more successful compared to 22C3 pharmDx SK006. No data concerning protocol settings submitted to NordiQC could explain the difference observed.

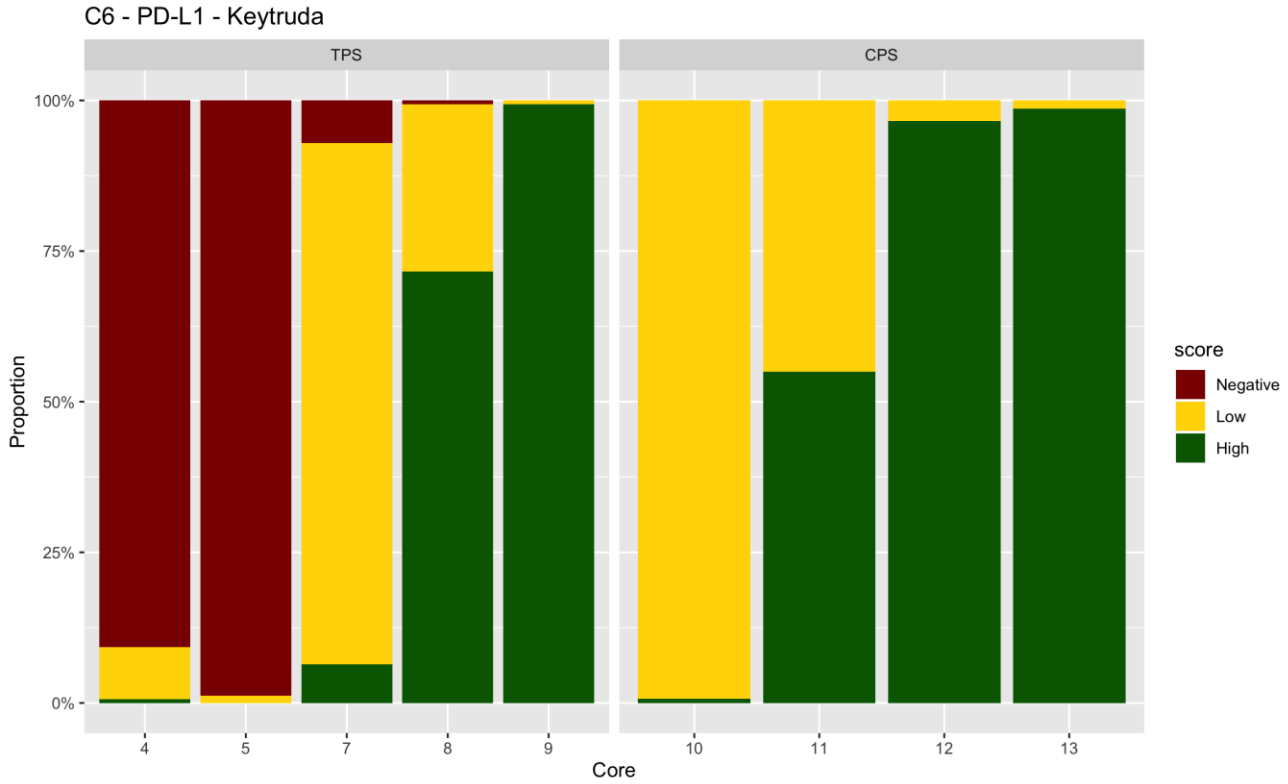
The pharmDx SK005 28-8 (Dako/Agilent) was used by three laboratories. All used the recommended protocol settings with sufficient results.

Grouped together, and using vendor recommended protocol settings, the CE IVD approved PD-L1 IHC assays provided a pass rate of 89% (64 of 72).

Laboratory developed (LD) assays either based on a concentrated Ab, a "non-companion diagnostic approved" RTU format, or an approved companion diagnostic assay not used strictly accordingly to vendor recommendations, were used by 60% (110 of 182) of the participants. For this group a pass rate of 66% (73 of 110) was observed.

The mAb clone 22C3 was the most widely used concentrated Ab within a LD assay (n=50) and the pass rate was 66% (33 of 50). This is slightly reduced to the C5 run, where 74% of the LD assays based on this clone provided a sufficient result. In this run, 66% (33 of 50) of the LD assays based on mAb clone 22C3 as concentrate provided a sufficient result, 30% (15 of 50) optimal, compared to 85% (28 of 33), 50% (15 of 30) optimal, using the corresponding companion diagnostic assays SK006 and GE006 with the recommended protocol settings – see Table 3.
26 laboratories used the mAb clone 22C3 on the Ventana BenchMark stainer platform and obtained a pass rate of 69% (18 of 26). 18 protocols were based on OptiView as detection system providing a pass rate of 78% (14 of 18) and 33% optimal (6 of 18) optimal. 6 protocols based on 22C3 in combination with OptiView and Amplification provided a pass rate of 50% (3 of 6), but only one optimal. When using OptiView with Amplification, an excessive aberrant granular staining pattern was seen. This pattern was accepted when it did not significantly compromise the interpretation. However, in one case the interpretation was hampered and the protocols was evaluated as insufficient.
9 laboratories used the mAb clone 22C3 as concentrate on the Dako Omnis stainer platform and obtained a pass rate of 89% (8 of 9) and 67% optimal (6 of 9).

PD-L1 interpretation and scoring consensus:
Participants were asked to score each of the cores using either tumour proportion score (TPS) for the NSCLC or combined positive score (CPS) for the urothelial carcinomas.



Graph 1. **NordiQC PD-L1 run C6: Tumour Proportion scores (TPS) in NSCLCs (core no. 4-9) and Combined Positive Score (CPS) in urothelial carcinomas (core no. 10-13). Core no. 6 was excluded.**

As seen in Graph 1, a relatively high consensus rates were observed for the tissue core 4,5,9,10, 12 and 13, whereas the consensus rate were significantly lower in tissue core 7,8 and 11 (tissue core 6 was excluded from the assessment).

For the tissue core no. 8, tissue heterogeneity was observed throughout the blocks and the TPS category changed especially for one of the blocks used in this assessment between TPS low and TPS high. Tissue core no. 7 was technically challenging to stain – both false negative and false positive staining reactions were seen. This also seems to be reflected in the scoring results from the participants. Concerning tissue core no 11, which only had positive immune cells (and not tumour cells), a relative high number of laboratories scored this as CPS low. One possible explanation could be that some laboratories scored the core using the TPS scoring system and not the CPS scoring system, which also includes positive immune cells in the CPS score.

When stratifying for assessment marks, analysis indicated that participants that had received an insufficient mark (borderline or poor) for the technical assessment of their PD-L1 result also had a higher tendency to perform an incorrect read-out of TPS and/or CPS in the submitted slides.

Controls

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 dynamic and clinical relevant range of PD-L1 expression levels, whereas placenta virtually only contained cells (throphoblasts) with high level PD-L1 expression. Using PD-L1 IHC 28-8 (SK005, Dako/Agilent), 22C3 (GE006 and SK006, Dako/Agilent) or SP263 (790-4905/4907 and 741-4905, Ventana/Roche) 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, no staining reaction in superficial epithelial 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. It was observed that rmAb SP263 (790-4905/4907 and 741-4905, Ventana/Roche) typically provided a higher proportion of positive inter and intra-follicular immune cells compared to the Dako/Agilent PD-L1 assays (SK005, SK006 and GE006).

However, it was observed that occasionally a fully acceptable staining pattern in tonsil could be obtained together with an insufficient result in the carcinomas. This underlines the need to identify more reliable positive and negative (tissue) controls for PD-L1 and/or improve the interpretation criteria for a sufficient

staining reaction in tonsils e.g. more accurately specify number and intensity of cells expected to be demonstrated and correlated to the PD-L1 IHC test/clone used. Some clones, e.g. mAb clone CAL10 typically gave a stronger staining reaction in more germinal centre macrophages compared to mAb clone 22C3.

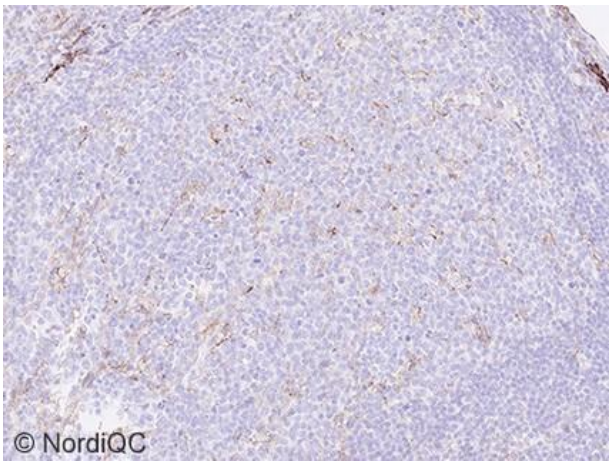


Fig. 1a
Optimal staining result of tonsil (germinal centre) using the 22C3 pharmDx IHC PD-L1, SK006, Dako/Agilent on Autostainer 48Link following the recommended protocol settings. Same protocol used in Figs. 2a-5a. The majority of germinal centre macrophages show an at least weak but distinct membranous staining reaction. The vast majority of lymphoid cells are negative.

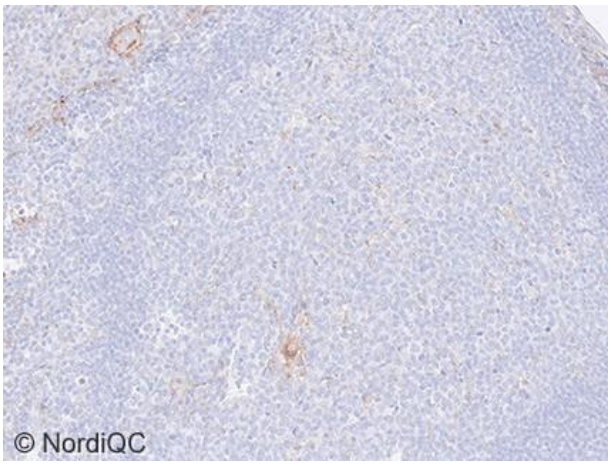


Fig. 1b
Staining result of tonsil (germinal centre) using the 22C3 pharmDx IHC PD-L1 SK006, Dako/Agilent on a BenchMark XT, Ventana. Same protocol in Fig 2b. Germinal centre macrophages show a weak and less distinct staining reaction compared to the optimal result in Fig. 1a.

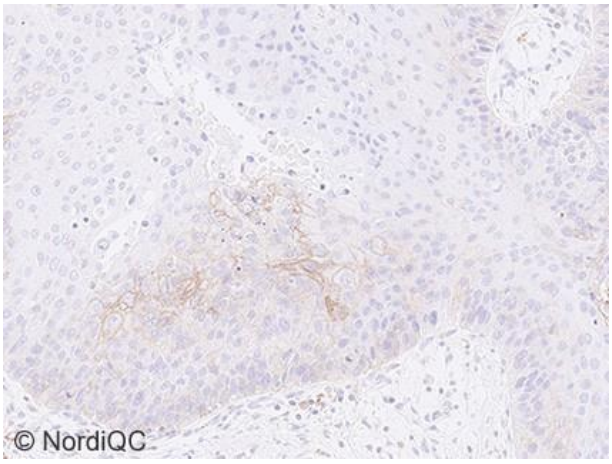


Fig. 2a
Optimal staining result in NSCLC tissue core no. 7, using same protocol as in Fig. 1a. Approximately, 10-15% of the neoplastic cells in the whole core (in TMA II) show a weak but distinct membranous staining reaction. The tumour was categorized as TPS low (1-49%) and thus eligible for second line immune therapy.

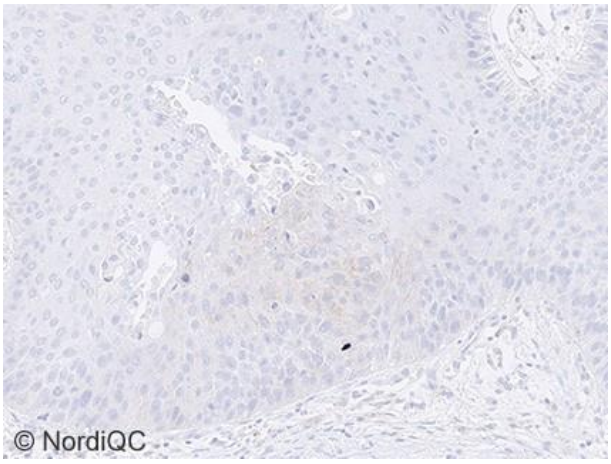


Fig. 2b
Insufficient staining result in NSCLC tissue core no. 7, using same protocol as in Fig. 1b.
Virtually all tumour cells are negative providing a TPS below 1%. Compare with optimal staining in Fig. 2a.

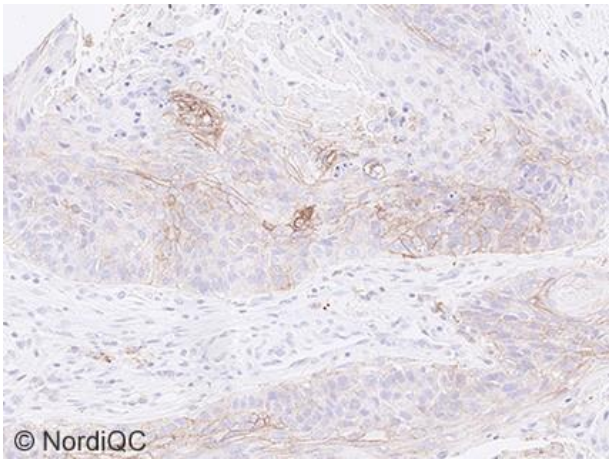


Fig. 3a
Optimal staining result in NSCLC tissue core no. 7, using same protocol as in Figs. 1a-2a. Approximately, 20-30% of the neoplastic cells in the whole core (in TMA I) show a weak but distinct membranous staining reaction. The tumour was categorized as TPS low (1-49%) and thus eligible for second line immune therapy.

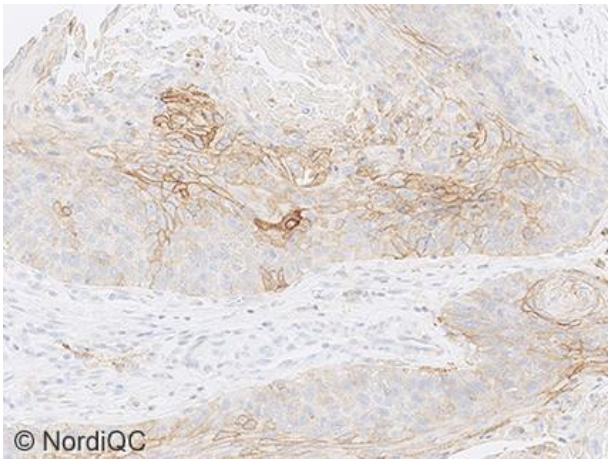


Fig. 3b
Insufficient staining result in NSCLC tissue core no. 7, using the SP263 assay 741-4905, Ventana, with modified protocol settings (decreased HIER time). Same protocol in Fig. 4b.
A significantly increased number of positive neoplastic cells is obtained changing the TPS category from low to high. Compare with optimal result in Fig. 3a, same area.

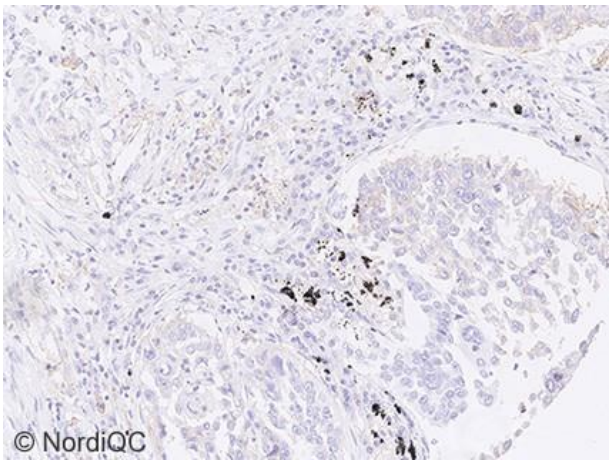


Fig. 4a
Optimal staining result in NSCLC tissue core no. 4, using same protocol as in Figs. 1a-3a. Only few scattered neoplastic cells show a weak membranous staining reaction. The tumour was categorized as negative.

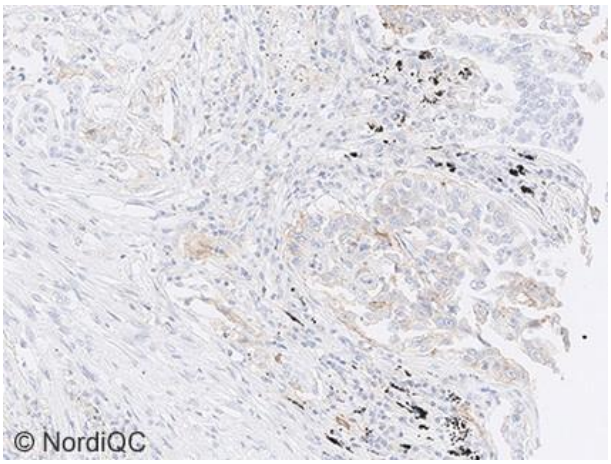


Fig. 4b
Insufficient staining result in NSCLC tissue core no. 4, using the same protocol as in Fig. 3b.
An increased number of positive neoplastic cells is observed changing the TPS category from negative to low. Compare with optimal result in Fig. 4a, same area.

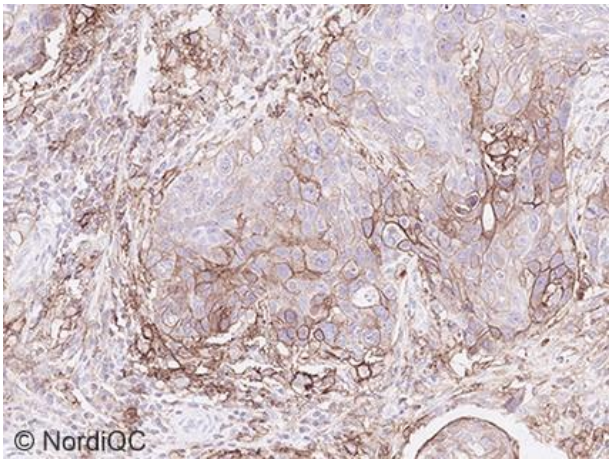


Fig. 5a
Optimal staining result in NSCLC tissue core no. 8, using same protocol as in Figs. 1a-4a. Approximately, 70-90% of the neoplastic cells in the whole core (in TMA I) show a weak to moderate, distinct membranous staining reaction. The tumour was categorized as TPS high (≥50%) and thus eligible for first line immune therapy.

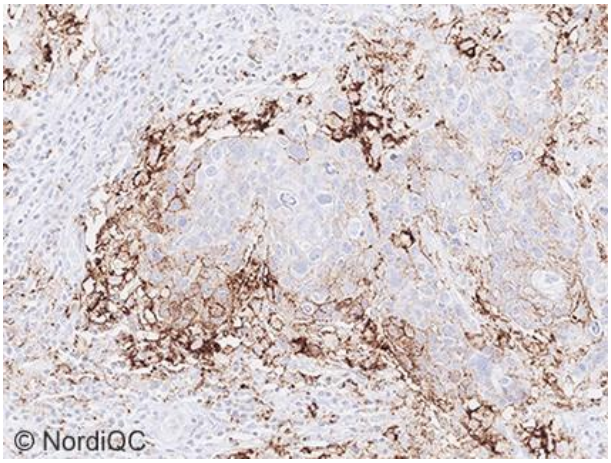


Fig. 5b
Insufficient staining result in NSCLC tissue core no. 8, using the mAb clone 22C3 as a concentrate in a laboratory developed assay. The protocol was based on HIER in CC1 (Ventana/Roche) and OptiView with Amplification as detection system (BenchMark Ultra, Ventana/Roche). Although the tumour was categorized as TPS high, the majority of neoplastic cells displays a granular reaction pattern (due to the amplification step), rather than the continuous and homogenous staining pattern seen in Fig. 5a. Cells at the stromal interface show an intense staining reaction and it is virtually impossible to identify the nuclear details and cell origin being IC cells or tumour cells while tumour cells central in the invasive area are unstained. If protocols are not carefully calibrated with this system, the interpretation of the reactions can be obscured due to granular deposit formed with tyramide based detection systems or the reactions can be too weak, risking that patients are positioned in wrong TPS categories.

HLK/RR/LE/SN 12.12.2019