

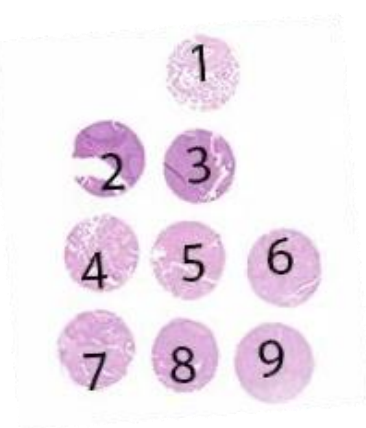
Purpose

This was the fourteenth assessment for PD-L1 in the NordiQC Companion Module. This assessment for PD-L1 TPS/CPS (KEYTRUDA®) primarily focused on the evaluation of the analytical accuracy of the IHC assays performed by the NordiQC participants to identify patients with non-small cell lung cancer (NSCLC) and triple negative breast carcinoma (TNBC) to be treated with KEYTRUDA® as immunotherapy. PD-L1 22C3 PharmDx (Dako/Agilent), was used as the reference standard method, and accuracy was evaluated in carcinomas with the dynamic and critical relevant expression levels of PD-L1 characterized by TPS and CPS. The scores obtained by NordiQC participants is indicative of the performance of the IHC tests but due to the limited number and composition of samples, additional internal validation/verification and extended quality control e.g. regularly measuring the PD-L1 results, is needed.

This was the fourth assessment for PD-L1 TPS/CPS comprising TNBCs being integrated in the material circulated at the expense of urothelial carcinomas (same cut-off's and scoring methods for the two entities).

Material

Table 1. Content of the TMA used for the NordiQC PD-L1 TPS/CPS (KEYTRUDA®) C14 assessment.

	PD-L1 IHC TPS/CPS score*	
Tissue controls		
1. Placenta	See section for controls	
2. Tonsil	See section for controls	
3. Tonsil	See section for controls	
Carcinomas		
4. NSCLC	TPS: No; <1%	
5. NSCLC	TPS: Low; 20-40%**	
6. NSCLC	TPS: High; 85-100%	
7. TNBC	CPS: <10 IC#	
8. TNBC	CPS: ≥10; 30-60 IC#	
9. TNBC	CPS: ≥10; 40-90 IC#	

* Tumour proportion score (TPS) and combined positive score (CPS) determined by PD-L1 IHC 22C3, pharmDx (Dako/Agilent) performed in NordiQC reference lab.

** The tumour showed heterogeneity in the different levels within and in between the TMA's used. In three of the eight TMA's used for the assessment, areas with TPS 50-60% were observed.

IC, Immune cells - TC; Tumour cells

All tissues were fixed in 10% neutral buffered formalin.

The participating laboratories were asked to perform their PD-L1 IHC assay for predicting likely response to KEYTRUDA® as a treatment option, evaluate the PD-L1 expression level using the TPS and CPS scoring system, and to submit their stained slides and scores to NordiQC. This allowed assessment of the technical performance (analytical accuracy) of the PD-L1 TPS/CPS assays and provided information on the reproducibility and concordance of the PD-L1 read-out results among the laboratories.

PD-L1 TPS/CPS, Technical assessment

In order to account for heterogeneity of PD-L1 expression in the individual tumour cores included in the tissue micro array (TMA) blocks, reference slides were made throughout the blocks. The PD-L1 expression levels were thus characterized in every twenty-fifth slide and during the assessment, TPS and CPS categories for each tissue core on the submitted slides from the participants were compared to the level in the nearest reference slide.

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 carcinomas.

Criteria for assessing a staining as Good include:

The staining is considered acceptable (correct PD-L1 TPS/CPS category) in all of the included tissues. PD-L1 expression in one or more tissues varies significantly from the expected TPS/CPS scores, but still in the correct category. The protocol may be optimized to ensure analytical accuracy. The technical quality may be improved for e.g. counter staining, morphology and signal-to-noise ratio.
TPS/CPS is still concordant to the NordiQC reference data obtained in all carcinomas.

Criteria for assessing a staining as Borderline include:

The staining is considered insufficient because of a false negative or false positive staining reaction in one of the included carcinomas. The protocol should be optimized.
TPS/CPS is **not** concordant to the NordiQC reference data in one of the 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 of more than one of the included carcinomas. Optimization of the protocol is urgently needed.
TPS/CPS is **not** concordant to the NordiQC reference data in two or more of the carcinomas.

An IHC result can also be assessed as **borderline/poor** related to technical artefacts, e.g. poor signal-to-noise ratio, excessive counterstaining, impaired morphology and/or excessive staining compromising the scoring.

Participation

Number of laboratories registered for PD-L1 KEYTRUDA IHC C14	262
Number of laboratories returning PD-L1 KEYTRUDA IHC slides	243 (93%)
Number of laboratories returning PD-L1 scoring sheet	216

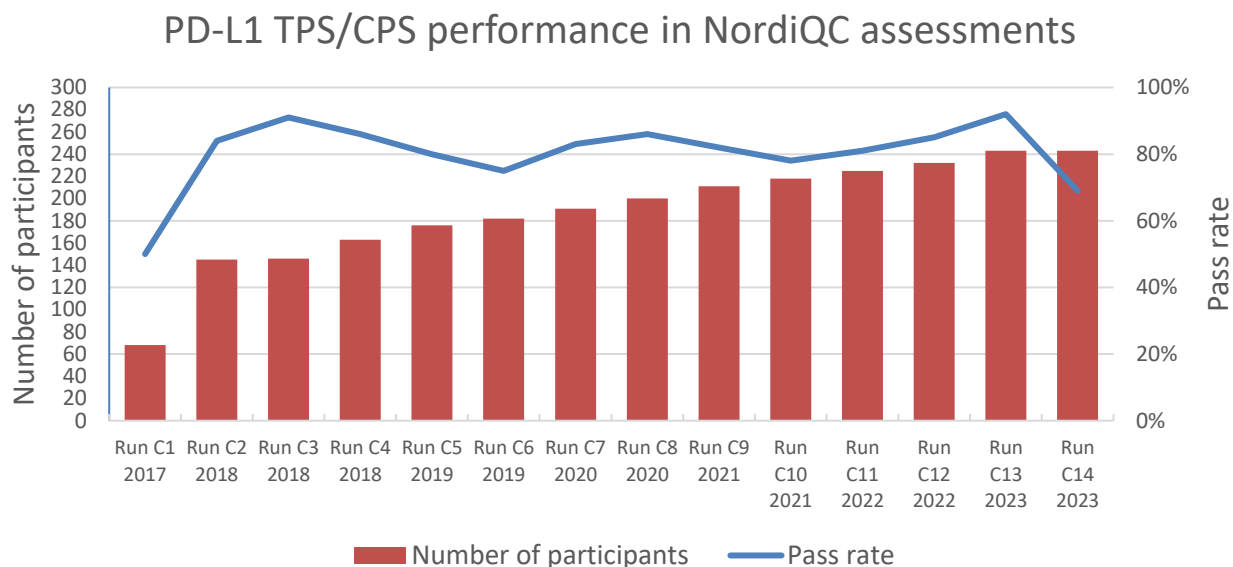
Results

243 laboratories participated in this assessment and returned slides. 69% of the participants achieved a sufficient mark. Assessment marks for IHC PD-L1 assays and PD-L1 antibodies are summarized in Table 2 (see page 4). All slides returned after the assessment were assessed and laboratories received advice if the result was insufficient, but the data was not included in this report.

Performance history

This was the fourteenth NordiQC assessment of PD-L1 TPS/CPS (KEYTRUDA®). Until C14 a relatively consistent pass rate has been observed with an upward trend seen. However, at C14 a decrease in pass rate was observed with the latest runs results shown in Graph 1 below. The number of new participants have recently been consistently increasing by about 3-5% in each run but has remained the same for runs C13 and C14.

Graph 1. Proportion of sufficient results for PD-L1 TPS/CPS (KEYTRUDA®) in the NordiQC runs performed.



Conclusion

This was the fourteenth NordiQC assessment of PD-L1 for TPS/CPS status with focus on NSCLCs and TNBCs. 243 laboratories participated and a pass rate of 69% was observed.

The PD-L1 IHC pharmDx assay, 22C3 GE006, Dako/Agilent applied in concordance to the vendor recommended guidelines, was most successful companion diagnostic assay providing a pass rate of 100%, with an optimal rate of 90%, being superior to the other companion diagnostic assays and LD assays based on concentrated Abs. Surprisingly and as yet unexplained due to unknown reasons, the Ventana/Roche PD-L1 IHC assays 741-4905 and 740-4907 for BenchMark (Ultra Plus/Ultra/XT/GX) based on the rmAb clone SP263 provided an overall low pass rate of 40% being inferior to the level expected and obtained in previous runs, which contributed to an overall reduced pass rate for C14.

In this run and similar to observations seen in previous NordiQC runs PD-L1 TPS/CPS, the insufficient PD-L1 IHC results were most frequently characterized by a reduced proportion of PD-L1 positive cells compared to the level expected and defined by the NordiQC reference standard method resulting in false negative results.

Table 2. Assessment marks for IHC assays and antibodies run C14, PD-L1 TPS/CPS (KEYTRUDA®)

CE-IVD / FDA approved PD-L1 assays	n	Vendor	Optimal	Good	Borderline	Poor	Suff. ¹	OR ²
rmAb clone SP263, 741-4905 (VRPS) ³	37	Ventana/Roche	-	16	20	1	43%	0%
rmAb clone SP263, 741-4905 (LMPS) ⁴	1	Ventana/Roche	-	-	1	-	-	-
rmAb clone SP263, 740-4907 (VRPS) ³	10	Ventana/Roche	1	2	7	-	30%	10%
mAb clone 22C3 pharmDX, SK006 (VRPS) ³	18	Dako/Agilent	11	4	2	1	83%	61%
mAb clone 22C3 pharmDX, SK006 (LMPS) ⁴	25	Dako/Agilent	15	7	3	-	88%	60%
mAb clone 22C3 pharmDX, GE006 (VRPS) ³	30	Dako/Agilent	27	3	-	-	100%	90%
mAb clone 22C3 pharmDX, GE006 (LMPS) ⁴	14	Dako/Agilent	8	5	1	-	93%	57%
rmAb clone 28-8 pharmDX, SK005 (VRPS) ³	3	Dako/Agilent	1	2	-	-	-	-
Antibodies ⁵ for laboratory developed PD-L1 assays, concentrated antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff. ¹	OR ²
mAb clone 22C3	53	Dako/Agilent	13	28	8	4	77%	25%
rmAb CAL10	2	Zytomed Systems	2	-	2	-	50%	50%
	2	Biocare Medical						
rmAb clone E1L3N	3	Cell Signaling	-	1	1	1	-	-
rmAb clone QR1	2	Quartett	-	2	-	-	-	-
rmAb clone 28-8	1	Dako/Agilent	-	1	-	-	-	-
rmAb clone SP142	1	Abcam	-	-	1	-	-	-
Ready-To-Use antibodies ⁶	n	Vendor	Optimal	Good	Borderline	Poor	Suff. ¹	OR ²
rmAb clone SP263, 790-4905⁶ (VRPS) ³	13	Ventana/Roche	-	1	11	1	8%	-
rmAb clone SP263, 790-4905⁶ (LMPS) ⁴	20	Ventana/Roche	-	11	8	1	55%	-
rmAb clone SP142, 740-4859⁴ (LMPS) ⁴	1	Ventana/Roche	-	-	1	-	-	-
rmAb clone 73-10 PA0832	3	Leica Biosystems	-	3	-	-	100%	-
mAb clone C9C9 CPM-0273	1	Celnovte	-	1	-	-	-	-
rmAb MX070C MAB-0854	1	Fuzhou Maixin	-	1	-	-	-	-
rmAb clone AC37 PA168 / AD80167	1	Abcarta	1	-	-	-	-	-
rmAb clone BP6099 I12052E	1	Biolynx	-	1	-	-	-	-
Total	243		79	89	66	9		
Proportion			32%	37%	27%	4%	69%	

1) Proportion of sufficient stains (optimal or good).

2) Proportion of optimal results.

3) Vendor recommended protocol settings – RTU product used in compliance to protocol settings, platform and package insert.

4) Laboratory modified protocol settings for a RTU product applied either on the vendor recommended platform(s) or other platforms.

5) mAb: mouse monoclonal antibody, rmAb: rabbit monoclonal antibody.

6) Ready-To-Use antibodies without predictive claim.

Detailed Analysis

CE IVD / FDA approved assays

SP263 (741-4905, Ventana/Roche): In total, 0 of 38 (0%) protocols were assessed as optimal. This product has a locked protocol on all BenchMark platforms and cannot be changed. The protocol is based on Heat Induced Epitope Retrieval (HIER) in Cell Conditioning 1 (CC1) at 100°C for 64 min., 16 min. incubation of primary Ab and OptiView as detection system. Using these protocols settings and applied on the BenchMark platform, 16 of 38 (42%) laboratories produced a sufficient staining result (good).

SP263 (740-4907, Ventana/Roche): In total, 1 of 10 (10%) protocols were assessed as optimal. This product has a locked protocol on BenchMark platform and cannot be changed. The protocol is based on HIER in CC1 at 100°C for 64 min., 16 min. incubation of primary Ab and OptiView as detection system. Using these protocols settings, 3 of 10 (30%) laboratories produced a sufficient staining result (optimal or good).

PD-L1 IHC 22C3 pharmDx (SK006, Dako/Agilent): In total, 11 of 18 (61%) protocols were assessed as optimal. Protocols with optimal results were typically based on the vendor recommended protocol settings based on HIER using EnVision™ FLEX Target Retrieval Solution (TRS) low pH 6.1 at 95-99°C for 20 min. in 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, 15 of 18 (83%) laboratories produced a sufficient staining result.

SK006 was frequently used by modified protocol settings e.g., electing for other platforms such as Ventana BenchMark or performed manually with an overall comparable performance as shown in Table 2.

PD-L1 IHC 22C3 pharmDx (GE006, Dako/Agilent): In total, 27 of 30 (90%) protocols were assessed as optimal. Protocols with optimal results were typically based on the vendor recommended protocol settings HIER using EnVision™ FLEX TRS low pH 6.1 (GV805) at 95-99°C for 40 min., 40 min. incubation of the primary Ab, EnVision™ FLEX+ as the detection system and performed on Omnis. Using these protocol settings, 30 of 30 (100%) laboratories produced a sufficient staining result.

Table 3 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 according to the vendor recommendations and by laboratory modified systems changing basal protocol settings. Only protocols performed on the specific automated IHC platform are included.

Table 3. **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 XT, GX, Ultra rmAb SP263, 741-4905	16/37 (43%)	0/37 (0%)	-	-
Ventana BenchMark Ultra rmAb SP263, 740-4907	3/10 (30%)	1/10 (10%)	-	-
Dako Autostainer Link 48+ mAb 22C3 pharmDX, SK006	15/18 (83%)	11/18 (61%)	12/13 (92%)	8/12 (67%)
Dako Omnis mAb 22C3 pharmDX, GE006	30/30 (100%)	27/30 (90%)	8/8 (100%)	6/8 (75%)
Dako Autostainer Link 48+ rmAb 28-8 pharmDX, SK005	3/3	1/3	-	-

*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 clone **22C3**: In total, 13 of 53 (25%) protocols were assessed as optimal of which 30 were stained on the BenchMark Ultra platform (Ventana/Roche), 1 on the BenchMark Ultra Plus platform (Ventana/Roche), 1 on BenchMark XT platform (Ventana/Roche), 11 on the Omnis platform (Dako/Agilent), 4 on Autostainer Link 48 (Dako/Agilent), 4 on BOND-III platform (Leica Biosystems), 1 on BOND Max platform (Leica Biosystems) and 1 manually.

On BenchMark Ultra, the protocols providing optimal results were based on a titre of 1:20-50 for mAb clone 22C3, incubation time of 60-120 min., HIER in CC1 for 48-64 min. and OptiView as the detection system. Using these protocol settings, 6 of 30 (20%) laboratories produced optimal staining results, and 22 of 30 (73%) laboratories produced sufficient staining results.

On BenchMark Ultra Plus, one protocol providing an optimal result being based on a titre of 1:50 for mAb clone 22C3, incubation time of 60 min., HIER in CC1 for 48 min. and OptiView as the detection system. Using these protocol settings, 1 of 1 laboratory produced optimal staining results.

On Omnis, the protocols providing optimal results for mAb clone 22C3 were based on a titre of 1:20-50 of the primary Ab, incubation time of 30-40 min., HIER in TRS low pH 6.1 at 97°C for 30-40 min. and EnVision™ FLEX+ as detection system. Using these protocol settings, 4 of 11 (36%) laboratories produced optimal results and 11 of 11 (100%) laboratories produced a sufficient staining result.

rmAb clone **CAL10**: Overall, 2 of 4 protocols were assessed as optimal.

The optimal protocols for this clone was based on a titre of 1:50 of the primary Ab, incubation time of 15-20 min., HIER in BERS2 (Leica Biosystems) pH 9 at 100°C for 20-30 min., Bond™ Refine as the detection system and performed on Bond III, Leica Biosystems. Using these protocol settings, 2 of 2 laboratories produced optimal staining results.

Table 4. **Optimal results for PD-L1 for the most commonly used antibody as concentrate on the four main IHC systems***

Concentrated antibodies	Ventana/Roche BenchMark Ultra		Dako/Agilent Autostainer		Dako/Agilent Omnis		Leica Biosystems BOND III	
	CC1 pH	CC2 pH	TRS pH	TRS pH	TRS High pH	TRS Low pH	BERS2 pH	BERS1 pH
	8.5	6.0	9.0	6.1			9.0	6.0
mAb clone 22C3	6/30** (20%)	-	1/4	1/4	-	4/11	2/4	-

*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 reference standards

The tissue micro array (TMA) blocks constructed for this PD-L1 run consisted of three NSCLCs, three TNBCs, two tonsils and one placenta. The NSCLCs were selected to comprise PD-L1 expression levels for each TPS category: TPS negative (<1% PD-L1 positive tumour cells), TPS low (≥1-49%) and TPS high (≥50%). The TNBCs were selected to comprise one carcinoma with CPS<10 and two carcinomas with CPS≥10 - one with PD-L1 expression primarily in immune cells and one with PD-L1 expression in both tumour cells and immune cells. Reference slides throughout the individual TMA blocks (interval at each twenty-fifth slide) were stained using the companion diagnostic assay 22C3 pharmDX (Dako/Agilent). In total, eight identical TMA blocks were constructed and eight of these used for this assessment.

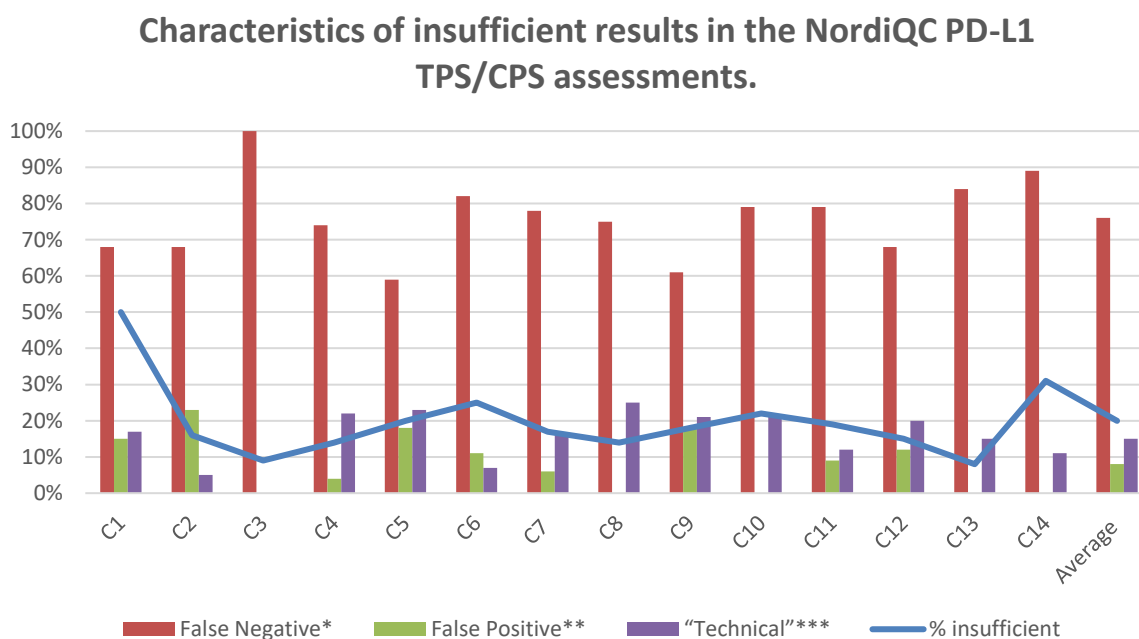
Reviewing the reference slides from the blocks, a heterogenic expression of PD-L1 was seen in one of the tumour cores. Of particular importance for the NSCLC, tissue core no. 5, focal areas with TPS 50-60% (TPS High) were observed and as such increased to the main level of 20-40% (TPS Low). During the assessment, TPS and CPS categories for each tissue core on the submitted slides were compared to the level in the nearest reference slides. Heterogeneity in PD-L1 expression is well known in NSCLCs and the assessment in this sense emulated clinical settings.

Comments

In this fourteenth NordiQC assessment for PD-L1 TPS/CPS (KEYTRUDA®), the prevalent feature of an insufficient staining result was a false negative staining result, being observed in 89% of the insufficient results. 11% of the insufficient results were caused by technical issues as poor-signal-to-noise ratio, excessive cytoplasmic staining reaction or a coarse and indistinct granular staining reaction compromising the scoring of the PD-L1 status in one or more of the carcinomas.

As shown in Graph. 2, a false negative staining result has been the most common reason for insufficient staining results in all NordiQC PD-L1 TPS/CPS (KEYTRUDA®) assessments with an average occurrence of 76%.

Graph 2. **Prevalence and characteristics of insufficient results**



* TPS changes from high to low or low to negative. And/or CPS changes from ≥ 10 to < 10 .

** TPS changes from negative to low or low to high. And/or CPS changes from < 10 to ≥ 10 .

*** Interpretation compromised e.g. by poor-signal-to noise ratio, poor morphology, excessive cytoplasmic staining reaction etc.

In this assessment and in concordance with previous runs the majority of insufficient results were related to incorrect TPS categories in one or more of the NSCLCs, whereas the CPS categories of the TNBCs only were affected in a few cases. This observation is similar to the results obtained and described in previous NordiQC PD-L1 TPS/CPS assessments with the combination of NSCLCs, TNBC's and urothelial carcinomas. However, there is a larger percentage of overall insufficient results for Run C14 compared to recent Runs.

In order to evaluate IHC accuracy NordiQC strives to include neoplasms with PD-L1 levels close to the critical and clinically relevant thresholds for positivity focusing on both intensity, proportion and subtypes of cells to be scored mimicking real-life diagnostics.

The NSCLC, tissue cores no. 5, characterized as TPS low respectively by the NordiQC reference standard method, was the most challenging to obtain an optimal result. In addition, the TNBC, tissue core no. 7, expected to show a CPS < 10 with a positive staining reaction primarily in the immune cells, also was found to be challenging.

37% (n=89 of the results submitted were marked as "Good". In 78% of these (69 of 89), this was due to a significantly reduced TPS/CPS level, but with no change of the TPS/CPS-category in any of the carcinomas and thus still an accurate PD-L1 status for treatment decision. Only in 3% (3 of 89) an increased TPS/CPS level was observed compared to the level expected, but again without any change in the TPS/CPS-category and PD-L1 status. In the remaining 19% (17 of 89) of the results assessed as "Good" these were characterized by poor signal-to-noise ratio, impaired morphology, too weak or excessive counterstaining and/or a coarse granular staining reaction compromising the evaluation of the membranous staining reaction. The latter only seen for protocols based on OptiView with amplification kit (Ventana/Roche).

The Ventana/Roche PD-L1 IHC assays 741-4905 and 740-4907 for BenchMark (Ultra Plus/Ultra/XT/GX) with predictive claims, based on the SP263 clone, were used by 19% (47 of 243) of the participants and in total provided an overall pass rate of 40% (19 of 47), with only 2% (1 of 47) being assessed as optimal when applied by protocol settings in compliance with vendor recommendations (see Table 3). The assays are locked for central protocol settings and based on HIER in CC1 for 64 min., incubation in primary Ab for 16 min. and use of OptiView as the detection system. Overall, the proportion of optimal results for SP263 has in the latest TPS/CPS assessments been reduced compared to results seen previously and related to general lower TPS/CPS scores in the carcinomas included in the assessment materials. The significantly reduced analytical sensitivity observed is in contrast to earlier assessment runs and many publications. E.g., the published comparative study by Noske et al.¹ "*Interassay and interobserver comparability study of four programmed death-ligand 1 (PD-L1) immunohistochemistry assays in triple-negative breast cancer*" indicating an enhanced proportion of PD-L1 positive cases using SP263 compared to 22C3 using CPS \geq 10 as cut-off. Same has been observed in several publications focusing on PD-L1 TPS in NSCLC as e.g. Tsao et al.², Torlakovic et al.³ and Kim et al.⁴. At present, no explanation for this discrepancy observed has been identified.

The Dako/Agilent 22C3 pharmDx assay GE006 for Dako Omnis was used by 12% (30 of 243) of the participants providing a pass rate of 100% (90% optimal) when applied by protocol settings in compliance with vendor recommendations (see Table 3).

Similar to the data generated in previous runs, it was observed that the PD-L1 22C3 GE006 assay for Omnis was more successful compared to the 22C3 pharmDx SK006 for Autostainer Link 48. The superior performance of GE006 might in part be related to a more consistent reproducibility of the 22C3 pharmDx assay on the fully automated Dako Omnis platform compared to the assay when applied on the semi-automated Autostainer Link 48. In this context it has to be emphasized that the 22C3 GE006 assay for Dako Omnis is by Dako/Agilent only validated for PD-L1 status and predictive claim in NSCLC with TPS as scoring system and at present not validated by Dako/Agilent for any indication with CPS as scoring system including TNBC.

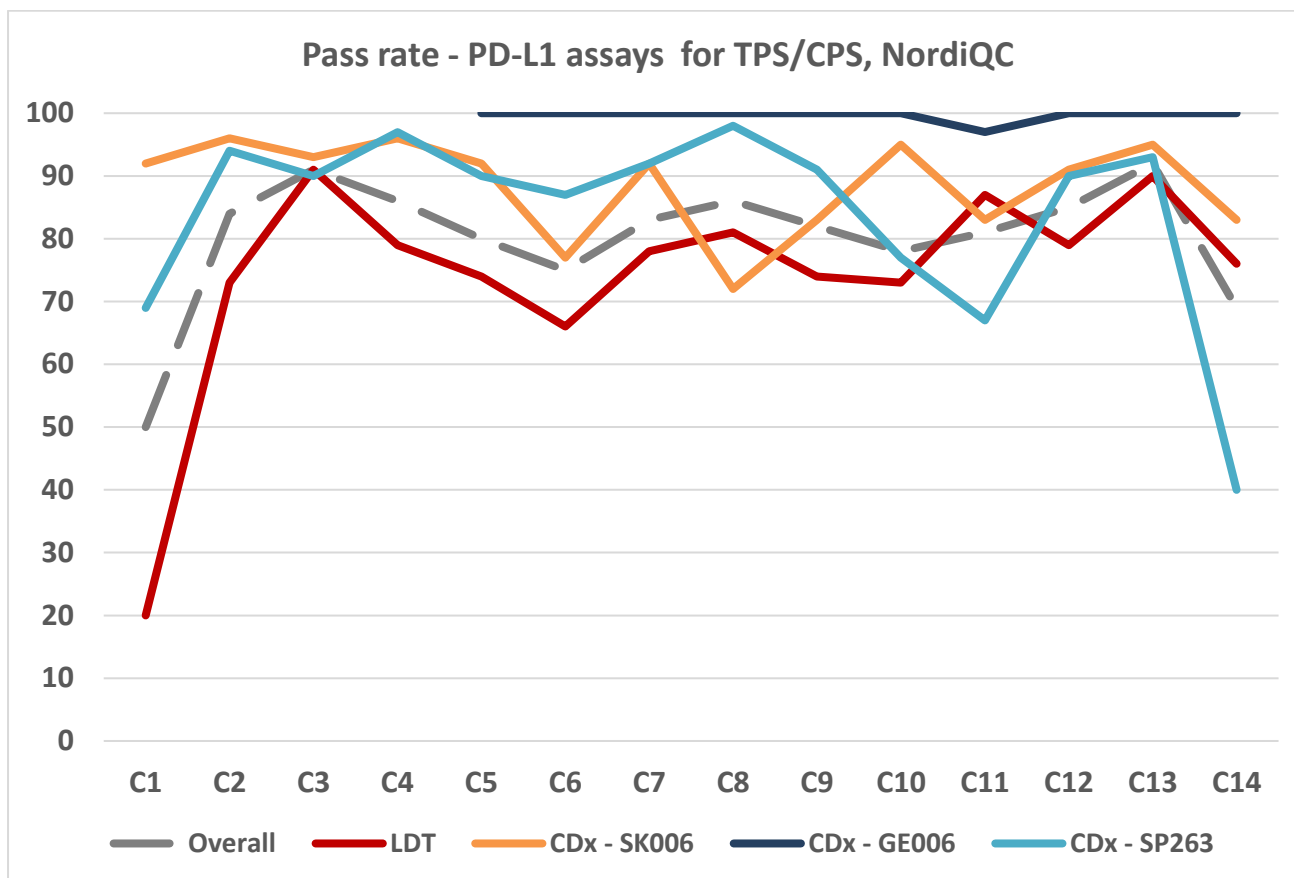
The Dako/Agilent 22C3 pharmDx assay SK006 for Autostainer Link 48 was used by 7% (18 of 243) of the participants and provided a pass rate of 83% (61% optimal) when applied by protocol settings in compliance with vendor recommendations (see Table 3). The 22C3 SK006 assay was frequently applied off-label (n=25), both on Autostainer 48 Link using modified protocol settings or on non-Autostainer Link 48 platforms as e.g. BenchMark Ultra (Ventana/Roche) and Omnis (Dako/Agilent), and as shown in Table 2 in this run with similar performance when applied as per recommendations or by modified off-label settings.

The Dako/Agilent pharmDx SK005 28-8 for Autostainer Link 48 was used by 3 laboratories. All used the recommended protocol settings with all being assessed as sufficient (2 =Good, 1= Optimal).

Laboratory developed (LD) assays either based on a concentrated Ab, a RTU Ab without any predictive claim or a companion diagnostic assay not used strictly accordingly to vendor recommendations were applied by 54% (132 of 243) of the participants, which is an increase compared to 52% in the previous two assessments – C13 and C12. For this group a pass rate of 76% was observed which is a decrease compared to the level of 90% as seen in the last assessment run – C13. Focusing on the performance of PD-L1 LD assays from C2-C14, excluding the initial run C1 and start-up phase to identify "best practice LD assays", the mean pass rate for LD assays has been 79% (range 66%-91%) compared to e.g., 100% for the 22C3 GE006 pharmDx (Dako/Agilent), 88% for 22C3 SK006 pharmDx (Dako/Agilent) and 84% for the SP263 assay (Ventana/Roche).

The performance of CDx and LD IHC assays for PD-L1 is summarized and shown in Graph 3 below.

Graph 3. Proportion of pass rates for PD-L1 TPS/CPS IHC assays in the NordiQC runs performed.



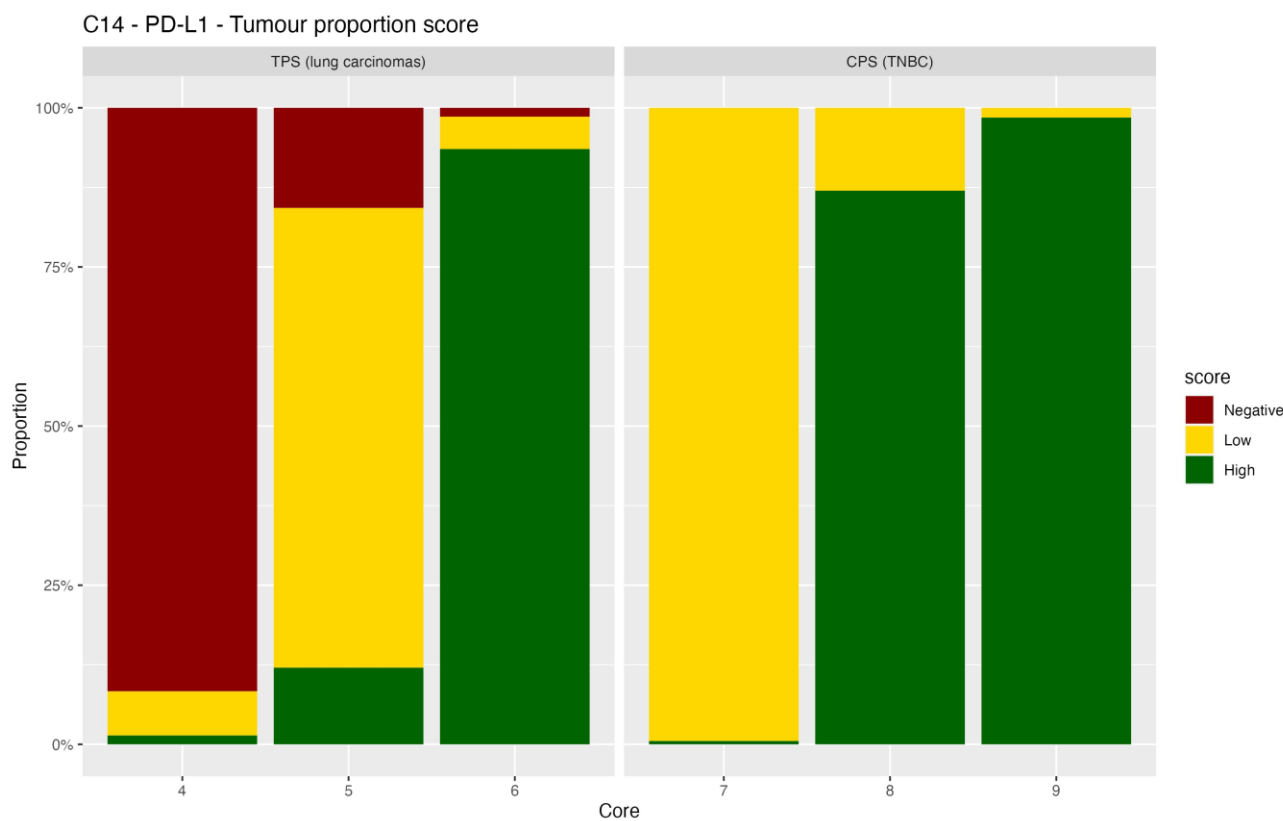
The mAb clone 22C3 was the most widely used concentrated Ab within a LD assay (n=53) providing a pass rate of 77% and an optimal rate of 25% which is decreased compared to run C13 (84%, 41%, respectively).

As described above for optimal protocol settings for mAb clone 22C3 as concentrated format, successful and interlaboratory reproducible settings have been identified for BenchMark (Ventana/Roche) and Omnis (Dako/Agilent) and these now seem to be widely consolidated within the laboratories providing a pass rate largely comparable to most of the companion diagnostic assays in this assessment as show in Graph 3 above.

As mentioned in previous reports the performance of mAb clone 22C3 on BOND III / BOND MAX (Leica Biosystems) has shown to be inferior, however, in run C13 there was a 100% sufficient pass rate, with 1 participant achieving an optimal result. However, this was not repeated in Run C14 as 0 participants achieved an optimal result. Cumulated data for runs C8 - C14 focusing on the performance of mAb clone 22C3 on the BOND platforms have shown a pass rate of 40% (10 of 25), with only 1 optimal result achieved. Only a small number of data observations generated so far and so conclusions are to be taken with caution. The recently launched Leica Biosystems PD-L1 IHC RTU assay based on rmAb clone 73-10 with intended use on Bond III, was used by 6 participants in C13 and 3 participants in C14 and when used by vendor recommended protocol settings obtaining a pass rate of 100%.

PD-L1 interpretation and scoring consensus:

Participants were asked to score each of the cores using either tumour proportion score (TPS) for the NSCLCs or combined positive score (CPS) for the TNBCs.



Graph 4. NordiQC PD-L1 run C14: Tumour Proportion scores (TPS) in NSCLCs (core no. 4-6) and Combined Positive Score (CPS) in TNBCs (core no. 7-9).

As seen in Graph 4, relatively high consensus rates were observed for the tissue cores no. 4, 6, 7 and, 9, whereas the consensus rate being significantly lower in tissue core no. 5 but also slightly reduced in tissue core no. 8.

Controls

Throughout all assessments for PD-L1 TPS/CPS tonsil and placenta have been used as positive and negative tissue controls and tonsil has been found to be superior to placenta, as tonsil typically display a dynamic and clinically relevant range of PD-L1 expression levels from weak, low to high, whereas placenta typically only contain cells (throphoblasts) with high level PD-L1 expression.

In tonsil, protocols with optimal results for PD-L1 TPS/CPS status typically provide the following reaction pattern:

A moderate to strong predominantly membranous staining reaction in the crypt epithelial cells, a weak to moderate, typically punctuated membranous staining reaction of the majority of germinal centre macrophages and scattered intra- and interfollicular lymphocytes and macrophages showing a coarse punctuated granular cytoplasmic staining reaction. No staining reaction in the vast majority of lymphocytes and normal stratified squamous epithelial cells.

It has been observed that different assays and/or clones for PD-L1 TPS/CPS status give different staining patterns in tonsil, which must be taken into account when evaluating the reaction pattern and to verify if the result is as expected. The rmAb clone SP263 (741-4905, 790-4905, 740-4907), Ventana/Roche) typically provide a higher proportion of positive inter- and intra-follicular immune cells compared to the Dako/Agilent 22C3 PD-L1 assays (SK006 and GE006). For other clones, e.g. mAb clone CAL10 and E1L3N typically a stronger staining reaction in more germinal centre macrophages were seen compared to mAb clone 22C3, when the clones still provided otherwise optimal and accurate results in the carcinomas. This emphasizes that the expected test performance characteristics in tonsil must be correlated to the PD-L1 IHC test/clone used both for the inter- and intra-PD-L1 IHC reproducibility evaluation.

1. Noske A, Wagner DC, Schwamborn K, et al. Interassay and interobserver comparability study of four programmed death-ligand 1 (PD-L1) immunohistochemistry assays in triple-negative breast cancer. *Breast*. 2021;60:238-244.
2. Ming Sound Tsao, Keith M. Kerr, Mark Kockx, et al. PD-L1 Immunohistochemistry Comparability Study in Real-Life Clinical Samples: Results of Blueprint Phase 2 Project. *Journal of Thoracic Oncology*. 2018;13(9):1302-1311
3. Torlakovic E, Lim HJ, Adam J, et al. "Interchangeability" of PD-L1 immunohistochemistry assays: a meta-analysis of diagnostic accuracy. *Mod Pathol*. 2020;33(1):4-17.
4. Kim SY, Kim T-E, Park CK, Yoon H-K, et al. Comprehensive Comparison of 22C3 and SP263 PD-L1 Expression in Non-Small-Cell Lung Cancer Using Routine Clinical and Conditioned Archives. *Cancers*. 2022; 14(13):3138.

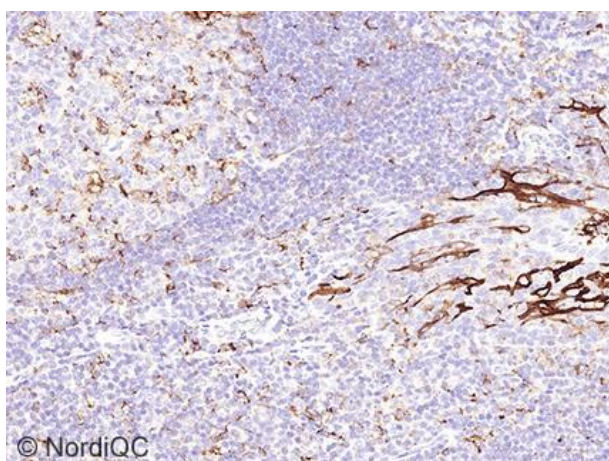


Fig. 1a
Optimal staining result of tonsil using the PD-L1 IHC 22C3 SK006 pharmDx kit, Dako/Agilent following the vendor recommended protocol settings.
A weak to moderate, but distinct punctuated membranous staining reaction of germinal centre macrophages and dispersed lymphocytes is seen. A moderate to strong staining reaction is seen in the reticulated crypt epithelial cells (right)
No staining reaction is seen in the vast majority of lymphocytes.
Also compare with Figs. 2a – 6a, same protocol.

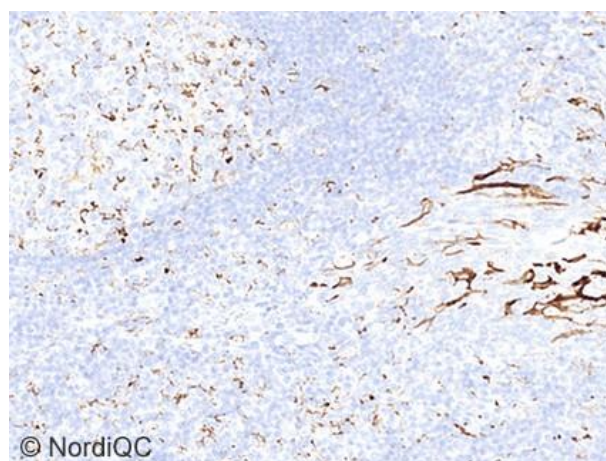


Fig. 1b
Staining result of tonsil, using the PD-L1 IHC SP263 assay 741-4905 for BenchMark Ultra, Ventana/Roche.
The result in tonsil is concordant to the expected pattern and expression levels as seen in Fig. 1a and obtained by the 22C3 pharmDx kit. However, overall, the protocol provided an insufficient result characterized by a too low level of analytical sensitivity – see Figs. 3b and 4b.
In this context it has to be emphasized that tonsil isolated cannot be used as quality control for PD-L1 IHC testing as no consistent information on low level demonstration of PD-L1 can be extracted.

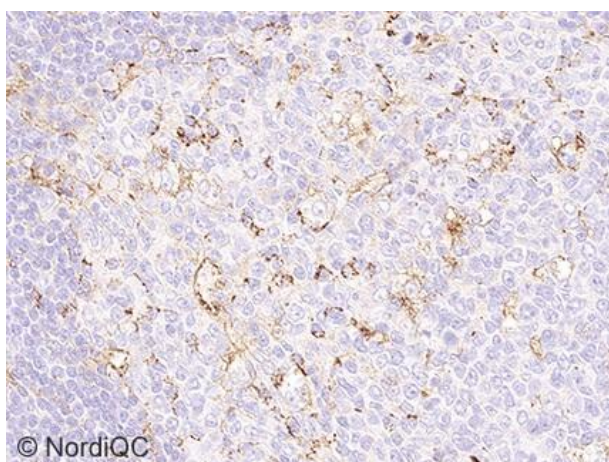


Fig. 2a
High-power field of tonsil using same protocol as in Fig. 1a. Germinal centre macrophages and dispersed T-cells are clearly demonstrated. Macrophages show a weak membranous staining reaction while T-cells a more granular punctuated membranous staining reaction.

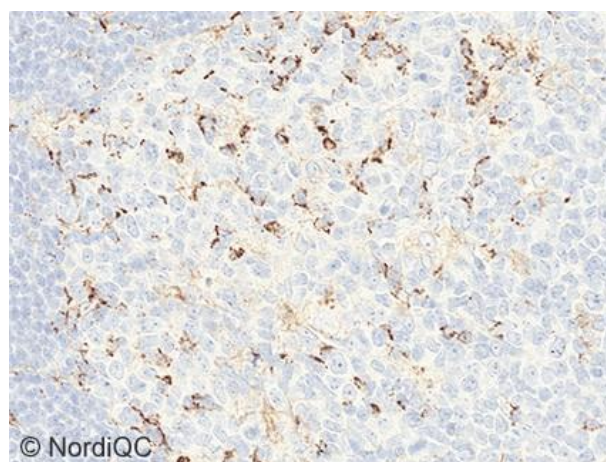


Fig. 2b
High-power field of tonsil using same protocol as in Fig. 1b. Mainly dispersed T-cells are demonstrated showing a granular punctuated membranous staining reaction. A reduced staining reaction is seen in the germinal centre macrophages.

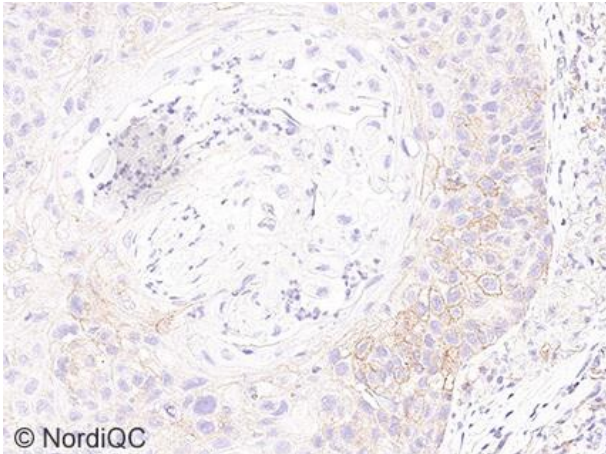


Fig. 3a
Optimal staining result of the NSCLC, tissue core no. 5, using the same protocol as in Figs. 1a and 2a. A weak to moderate, but distinct staining reaction is seen in dispersed tumour cells. The tumour was categorized as TPS Low ($\geq 1-49\%$) and thus eligible for second line immune therapy with KEYTRUDA® (different regional cut-offs occur).

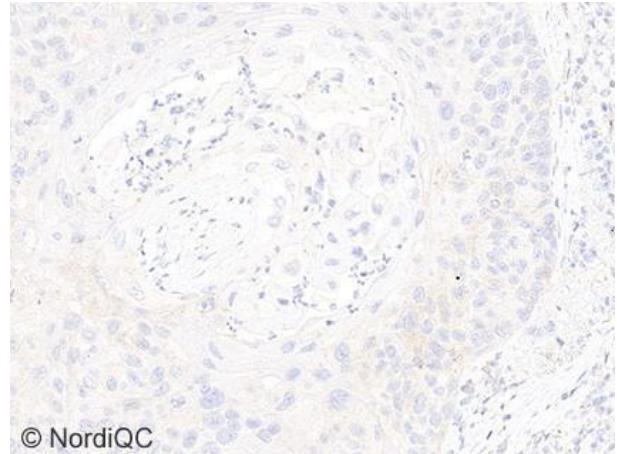


Fig. 3b
Insufficient staining result of the NSCLC, tissue core no. 5, using the same protocol as in Figs. 1b and 2b. Less than 1% of the tumour cells show a membranous staining reaction and the PD-L1 category being changed from the expected TPS Low to TPS Negative and the tumour not being eligible for immune therapy. Compare to the expected result as shown in Fig. 3a.

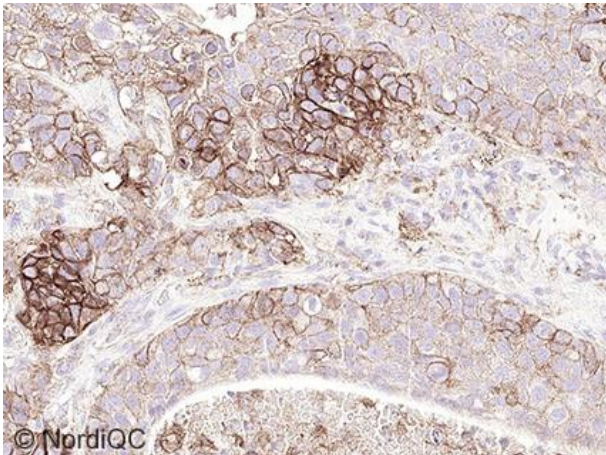


Fig. 4a
Optimal staining result of the NSCLC, tissue core no. 6, using the same protocol as in Figs. 1a - 3a. A weak to strong, distinct membranous staining reaction is seen in virtually all tumour cells. The tumour was categorized as TPS High ($\geq 50\%$) and thus eligible for first line immune therapy with KEYTRUDA® (different regional cut-offs occur).

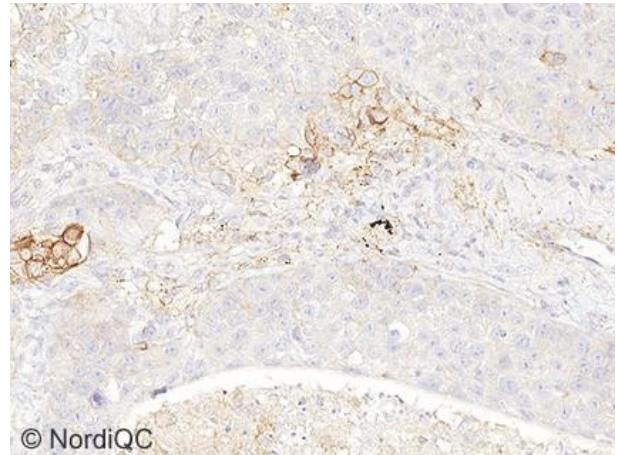


Fig. 4b
Insufficient staining result of the NSCLC, tissue core no. 6, using the same protocol as in Figs. 1b - 3b. Only dispersed tumour cells are demonstrated and significantly reduced compared to the level expected and overall the result categorized as TPS Low ($\geq 1-49\%$). The PD-L1 category consequently being changed from the expected TPS High to TPS Low and the tumour not being eligible for first line immune therapy. Compare to the expected result as shown in Fig. 4a.

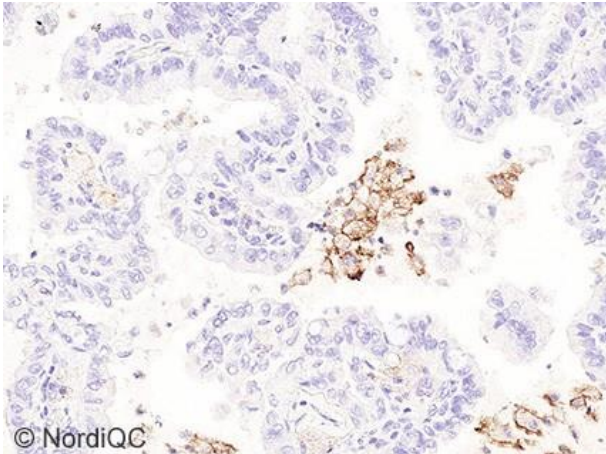


Fig. 5a
 Optimal staining result of the NSCLC, tissue core no. 4, using the same protocol as in Figs. 1a – 4a.
 A moderate, distinct membranous staining reaction is seen in the alveolar macrophages while the neoplastic cells are negative.
 The tumour was categorized as TPS Negative (<1%) and thus not eligible for immune therapy with KEYTRUDA® (different regional cut-offs occur).

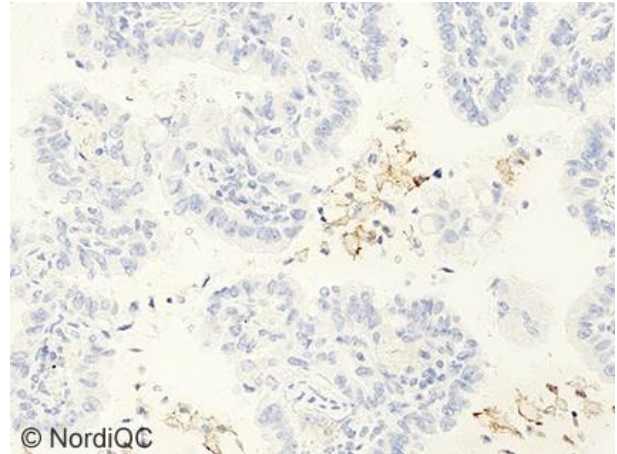


Fig. 5b
 Staining result of the NSCLC, tissue core no. 4, using the same protocol as in Figs. 1b – 4b.
 A weak, but distinct membranous staining reaction is seen in the alveolar macrophages while the neoplastic cells are negative.
 The tumour was as expected categorized as TPS Negative (<1%) and thus not eligible for immune therapy with KEYTRUDA® (different regional cut-offs occur).

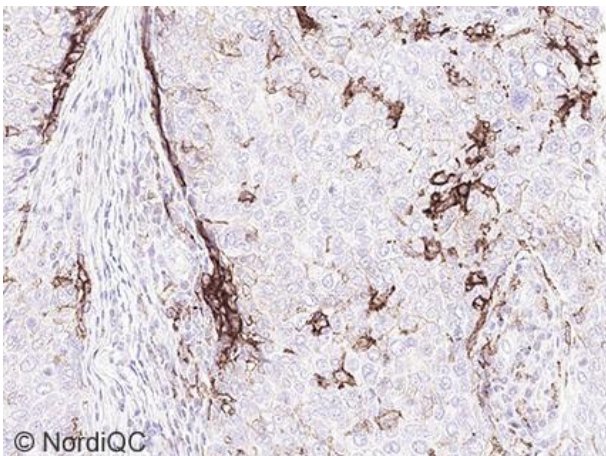


Fig. 6a
 Optimal staining result of the TNBC, tissue core no. 8, using the the 22C3 pharmDx, Agilent/ Dako following the vendor recommended protocol settings.
 A weak to strong and distinct membranous staining reaction is seen in many immune cells, while virtually all tumour cells are negative .
 The tumour was categorized as CPS≥10 and thus eligible for immune therapy with KEYTRUDA®.

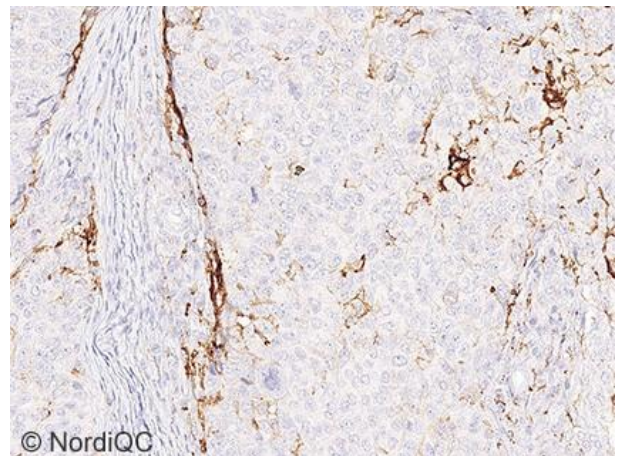


Fig. 6b
 Staining result of the TNBC, tissue core no. 8, using the SP263 PD-L1 IHC assay, Ventana/Roche following the vendor recommended protocol settings.
 Overall a slightly reduced intensity and proportion of immune cells is observed but the tumour still categorized as CPS≥10 and thus eligible for immune therapy with KEYTRUDA®.

JH/LE/SN 11.12.23