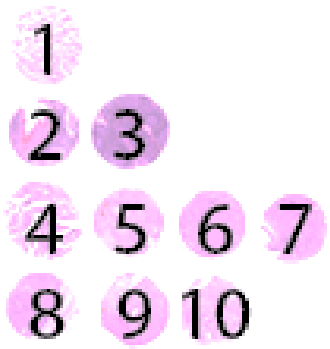
Assessment Run C12 2022
PD-L1 TPS/CPS

Purpose
This was the twelfth 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. 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 second 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®) C12 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: Low; 5-40%***
6. NSCLC	TPS: High; 80-90%
7. NSCLC	TPS: High; 100%
8. TNBC	CPS: <10
9. TNBC	CPS: ≥10; 30-40 IC#
10. TNBC	CPS: ≥10; 100 TC# + IC#



* Tumour proportion score (TPS) and combined positive score (CPS) determined by PD-L1 IHC 22C3, pharmDx (Dako/Agilent) performed in NordiQC reference lab.
** Focally in two of the seven TMA's used for the assessment, areas with TPS 1-2% were observed
*** The tumour showed heterogeneity in the different levels within and in between the TMA's used. Focally in one of the seven TMA's used for the assessment, areas with TPS 50% 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 <u>Optimal</u> include:
The staining is considered perfect or close to perfect in all of the included tissues.
<u>TPS/CPS is concordant to the NordiQC reference data in all carcinomas.</u>

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.

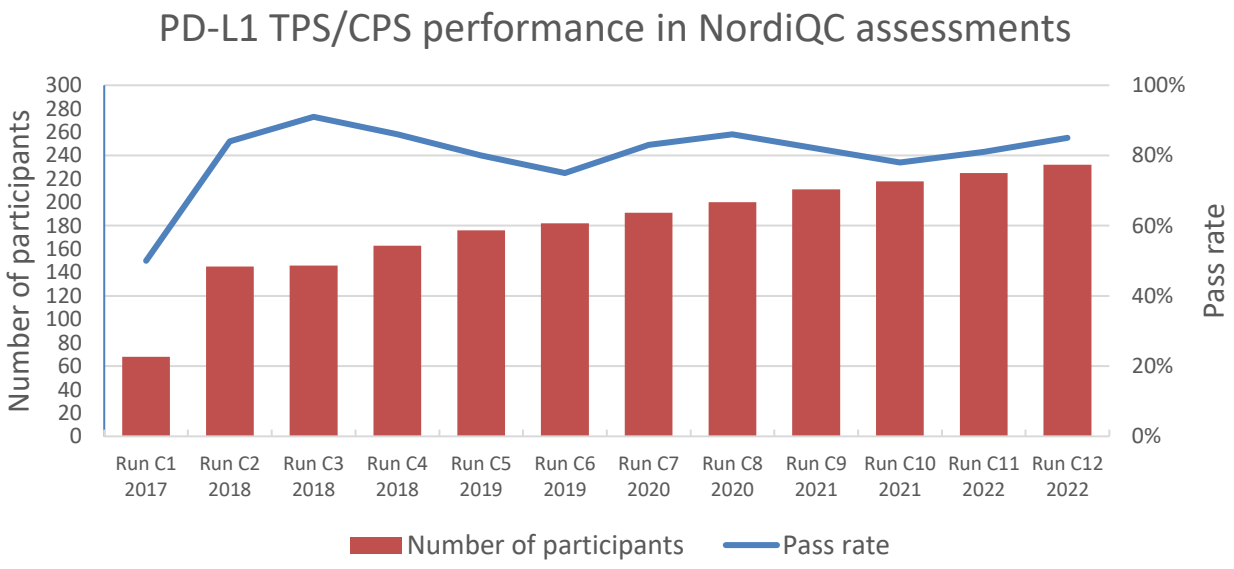
Participation

Number of laboratories registered for PD-L1 KEYTRUDA IHC C12	251
Number of laboratories returning PD-L1 KEYTRUDA IHC slides	232 (92%)
Number of laboratories returning PD-L1 scoring sheet	210

Results
232 laboratories participated in this assessment and returned slides. 85% 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 twelfth NordiQC assessment of PD-L1 TPS/CPS (KEYTRUDA®). A relatively consistent pass rate has been observed in the latest runs as shown in Graph 1 below. The number of new participants seems to be consistently increasing by about 3-5% in each run.

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



Conclusion
This was the twelfth NordiQC assessment of PD-L1 for TPS/CPS status with focus on NSCLCs and TNBCs. 232 laboratories participated and a pass rate of 85% was observed.

The PD-L1 IHC pharmDx assay, 22C3 GE006, Dako/Agilent applied in concordance to the vendor recommended guidelines, was most successful providing a pass rate of 100%, with an optimal rate of 79%, being superior to the other companion diagnostic assays and LD assays based on concentrated Abs or RTU systems without predictive claim.

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 methods resulting in false negative results.

Table 2. **Assessment marks for IHC assays and antibodies run C12, 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	10	25	1	1	95%	27%
rmAb clone SP263, **741-4905 (LRPS)[4]**	1	Ventana/Roche	1	-	-	-	-	-
rmAb clone SP263, **740-4907 (VRPS)[3]**	15	Ventana/Roche	3	9	3	-	80%	20%
rmAb clone SP142, **741-4860 (VRPS)[3]**	1	Ventana/Roche	-	1	-	-	-	-
rmAb clone SP142, **740-4859 (VRPS)[3]**	1	Ventana/Roche	-	1	-	-	-	-
mAb clone 22C3 pharmDX, **SK006 (VRPS)[3]**	22	Dako/Agilent	13	7	-	2	91%	59%
mAb clone 22C3 pharmDX, **SK006 (LMPS)[4]**	12	Dako/Agilent	3	7	1	1	83%	25%
mAb clone 22C3 pharmDX, **GE006 (VRPS)[3]**	28	Dako/Agilent	22	6	-	-	100%	79%
mAb clone 22C3 pharmDX, **GE006 (LMPS)[4]**	9	Dako/Agilent	5	2	-	2	78%	56%
rmAb clone 28-8 pharmDX, **SK005 (VRPS)[3]**	3	Dako/Agilent	1	1	1	-	-	-
Antibodies[5] for laboratory developed PD-L1 assays, concentrated antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
mAb clone **22C3**	39	Dako/Agilent	13	18	4	4	79%	33%
rmAb **CAL10**	3 2	Zytomed Systems Biocare Medical	1	1	1	2	40%	20%
rmAb clone **E1L3N**	3	Cell Signaling	1	1	-	1	-	-
rmAb clone **QR1**	3	Quartett	-	2	1	-	-	-
rmAb clone **ZR3**	2	Zeta Corporation	2	-	-	-	-	-
rmAB clone **SP142**	1	Abcam	1	-	-	-	-	-
rmAb clone **IHC411**	1	GenomeMe	1	-	-	-	-	-
rmAB clone **BP6099**	1	Biolynx	-	-	1	-	-	-
Ready-To-Use antibodies[6]	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
rmAb clone SP263, **790-4905[6] (VRPS)[3]**	16	Ventana/Roche	4	9	2	1	81%	25%
rmAb clone SP263, **790-4905[6] (LMPS)[4]**	23	Ventana/Roche	8	12	2	1	87%	35%
rmAB clone **73-10**	3	Leica Biosystems	1	1	1	-	-	-
rmAB **MX070C**	2	Fuzhou Maixin	2	-	-	-	-	-
rmAB clone **BP6099**	1	Biolynx	1	-	-	-	-	-
mAb clone C9C9 **CPM-0278**	1	Celnovte	-	1	-	-	-	-
rmAb clone AC37 **AD80167**	1	Abcarta	-	-	1	-	-	-
rmAb clone RM320 **8263-C010**	1	Sakura Finetek	1	-	-	-	-	-
Total	232		94	104	19	15		
Proportion			41%	45%	8%	6%	85%	

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, 10 of 37 (27%) 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 35 of 37 (95%) laboratories produced a sufficient staining result (optimal or good).

SP263 (740-4907, Ventana/Roche): In total, 3 of 15 (20%) protocols were assessed as optimal. This product has a locked protocol on BenchMark Ultra 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, 12 of 15 (80%) laboratories produced a sufficient staining result.

PD-L1 IHC 22C3 pharmDx (SK006, Dako/Agilent): In total, 13 of 22 (59%) 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, 20 of 22 (91%) 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 inferior performance as shown in Table 2.

PD-L1 IHC 22C3 pharmDx (GE006, Dako/Agilent): In total, 22 of 28 (79%) protocols were assessed as optimal. Protocols with optimal results were typically based on the vendor recommended protocol settings based on 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, 28 of 28 (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 accordingly 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**	35/37 (95%)	10/37 (27%)	-	1/1
Ventana BenchMark Ultra rmAb SP263, **740-4907**	12/15 (80%)	3/15 (20%)	-	-
Ventana Benchmark Ultra rmAb SP142, **741-4860**	2/2	0/2	-	-
Dako Autostainer Link 48+ mAb 22C3 pharmDX, **SK006**	20/22 (91%)	13/22 (59%)	10/12 (83%)	3/12 (25%)
Dako Omnis mAb 22C3 pharmDX, **GE006**	28/28 (100%)	22/28 (79%)	7/9 (78%)	5/9 (56%)
Dako Autostainer Link 48+ rmAb 28-8 pharmDX, **SK005**	2/2	0/2	-	-

*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**: In total, 13 of 39 (33%) protocols were assessed as optimal of which 24 were stained on the BenchMark Ultra platform (Ventana/Roche), 1 on BenchMark XT platform (Ventana/Roche), 7 on the Omnis platform (Dako/Agilent), 4 on Autostainer Link 48 (Dako/Agilent), 2 on BOND-III platform (Leica Biosystems) and 1 manually.
On BenchMark Ultra, the protocols providing optimal results for the mAb clone 22C3 were based on a titre of 1:20-50 of the primary Ab, incubation time of 32-80 min., HIER in CC1 for 48-80 min. and OptiView as the detection system. Using these protocol settings, 21 of 24 (88%) laboratories produced a sufficient staining result.

On Omnis, the protocols providing optimal results for the mAb clone 22C3 were based on a titre of 1:20-30 of the primary Ab, incubation time of 30-40 min., HIER in TRS low pH 6.1 at 97°C for 40-50 min. and EnVision™ FLEX+ as detection system. Using these protocol settings, 5 of 7 (71%) laboratories produced a sufficient staining result.

rmAb **E1L3N**: 1 of 3 protocols were assessed as optimal.
The protocol providing an optimal result was based on a titre of 1:200 of the primary Ab, incubation time of 30-40 min., Tris-EDTA / EGTA pH 9 at 98°C for 20 min. (water bath), BrightVision as the detection system and performed on a Dako Autostainer Link 48.

rmAb **ZR3**: 2 of 2 protocols were assessed as optimal.
One protocol was based on HIER using an alkaline buffer (BERS2, Leica Biosystems) at 100°C for 30 min. The rmAb clone ZR3 was diluted 1:100, incubated for 30 min. at room temp., visualized using a 2-layer detection system (Zeta Universal Polymer HRP) and performed on a Leica Bond III platform.
The other protocol was based on HIER in Tris/EDTA pH 9 in a water bath at 95°C for 20 min. The rmAb clone ZR3 was diluted 1:100 for 50 min. at room temp., visualized by GTVision (Gene Tech) and performed on a Gene Tech GeneStainer platform.

rmAb **CAL10**: 1 of 5 protocols were assessed as optimal
The optimal protocol was based on HIER using an alkaline buffer BERS2 (Leica Biosystems) at 99°C for 30 min. The rmAb clone CAL10 was diluted 1: 50, incubated for 15 min. at room temp., visualized by Bond™ Refine detection kit and performed on a Leica Biosystems BOND-III platform.

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 GX/XT/Ultra**		**Dako/Agilent Autostainer**		**Dako/Agilent Omnis**		**Leica Biosystems 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**	8/25** (32%)	-	-	2/4	0/1	2/6 (33%)	0/2	-

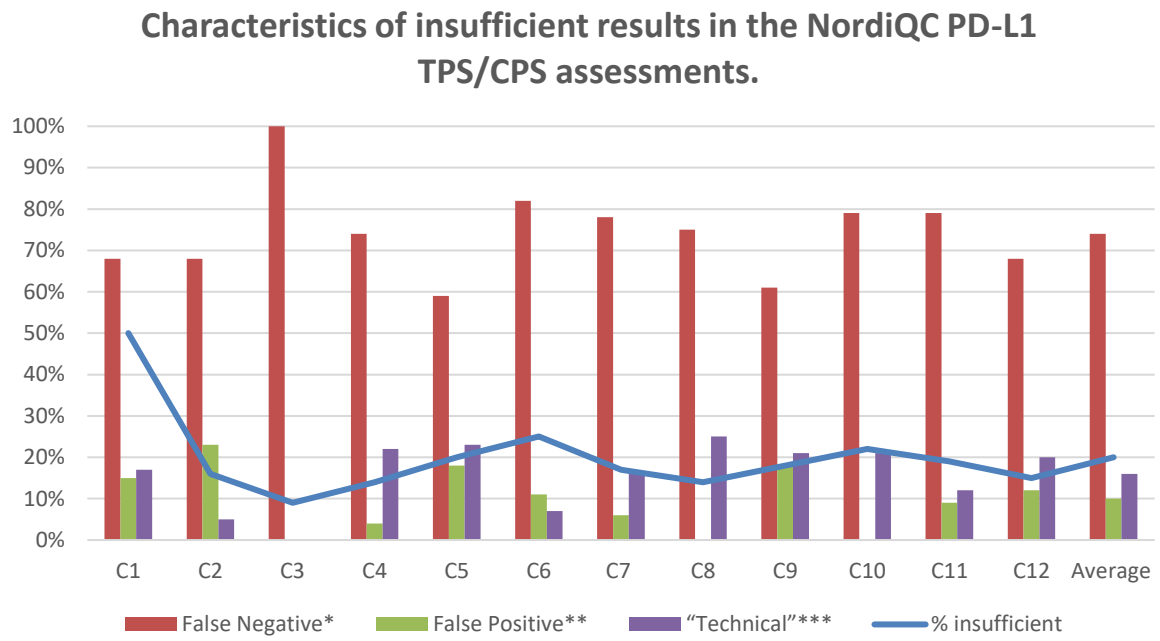
*Antibody concentration applied as listed above, HIER buffers and detection kits used as provided by the vendors of the respective platforms.
**number of optimal results/number of laboratories using this buffer.

Ready-To-Use antibodies for laboratory developed (LD) assays
rmAb **SP263** (790-4905, Ventana/Roche): In total, 12 of 39 (31%) protocols provided an optimal result. Protocols with optimal results were typically based on HIER in CC1 at 95-100°C, efficient heating time 40-64 min., incubation of the primary Ab of 36-37 min., OptiView as detection system and performed on BenchMark Ultra. Using a range of these protocols settings 33 of 38 (87%) laboratories produced a sufficient staining result. One participant achived a result scored as "Good" using HIER in an alkaline buffer BERS2 (Leica Biosystems) for 60 min., antibody incubation time of 15 min. at room temperature and visualisation by Bond™ Refine detection kit on the Leica Biosystems BOND-III platform.

Block construction and assessment reference standards
The tissue micro array (TMA) blocks constructed for this PD-L1 run consisted of four 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 seven of these used for this assessment. Reviewing the reference slides from the blocks, a heterogenic expression of PD-L1 was seen in two of the tumour cores. In the NSCLC, tissue core no. 4, areas with TPS 1-2% were observed. In the NSCLC, tissue core no. 5, focally in one of the seven TMA's used for the assessment, areas with TPS 50% were observed.
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 twelfth 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 68% of the insufficient results. 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 74%. 12% of the insufficient results were related to a false positive staining result while the remaining 20% 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.

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 was fully identical to the results obtained and described in previous NordiQC PD-L1 TPS/CPS assessments with the combination of NSCLCs, TNBC's and urothelial carcinomas.

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 two NSCLCs, tissue cores no. 5 and 6, characterized as TPS low and high, by the NordiQC reference standard methods, respectively, were most challenging to obtain an optimal result. Virtually all false negative results were as such seen in one or both of these NSCLCs, changing the TPS category compared to the level expected and defined by the CE IVD approved PD-L1 IHC assays used as the NordiQC reference standard methods. In addition, the TNBC, tissue core no. 9, expected to show a CPS≥10 (range 30-40) with a positive staining reaction primarily in the immune cells, also was found to be challenging and was typically also false negative when one or both of the two NSCLCs showed an inferior result.

In contrast, virtually all protocols provided the expected PD-L1 status in both the NSCLC, tissue core no. 7, characterized by NordiQC to show a strong membranous staining reaction in all tumour cells and the TNBC, tissue core no. 10, with CPS≥10 expressed in both immune cells and tumour cells. The false positive cases were only observed in the NSCLCs and seen in tissue core no. 4 and 5 changing the TPS status from negative to low and low to high, respectively.

45% (n=104) of the results submitted were marked as "Good". In 76% of these (79 of 104), 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. In only 1 of 104 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 24% (25 of 104) 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/XT/GX) with predictive claims, based on the SP263 clone, were used by 23% of the participants and in total provided an overall pass rate of 90%, 25% 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.

In contrast to the last PD-L1 TPS/CPS assessment run - C11, the two Ventana/Roche PD-L1 assays based on the rmAb clone SP263 provided an improved performance (pass rate) compared to the inferior level caused by a reduced analytical sensitivity observed in the previous run. However, the proportion of optimal results still being reduced compared to results seen previously and related to general lower TPS / CPS scores in the carcinomas included in the assessment materials. The reduced analytical sensitivity observed is in contrast to earlier assessment runs and many publications.

The recently 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" indicate an enhanced proportion of PD-L1 positive cases using SP263 compared to 22C3 using CPS≥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] The reduced analytical sensitivity especially observed in C11 and also C12 is in contrast to the data recently published by Sompuram et al.[4] showing that "the SP263 assay was the most sensitive PD-L1 approved assay" when comparing the levels for low limit of PD-L1 demonstration in IHC calibrators containing different levels and dynamic range of purified protein PD-L1 analytes.

The Dako/Agilent 22C3 pharmDx assay GE006 for Dako Omnis was used by 16% of the participants and in this assessment the most successful assay providing a pass rate of 100% (79% 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. Cumulated data for the latest 7 successive runs has shown a pass rate of 99% (143 of 144) for laboratories using GE006 by vendor recommended protocol settings. In comparison a pass rate of 85% (121 of 143) for laboratories using SK006 by vendor recommended protocol settings has been obtained.
The different pass rates observed have to be taken with caution due to relatively few data observations, but a clear trend so far has been observed in the latest six successive runs performed. 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 15% of the participants and provided a pass rate of 91% (59% optimal) when applied by protocol settings in compliance with vendor recommendations (see Table 3). The 22C3 SK006 assay was frequently (n=12 of 34 participants) applied off-label both on Autostainer Link 48 and a non-Autostainer Link 48 platform as e.g. BenchMark Ultra/GX/XT (Ventana/Roche) and as shown in Table 2 with inferior performance. In total 5 laboratories used the SK006 assay on BenchMark with a pass rate of 80%, 20% optimal, which is an improvement compared to run C11 data for the same (caution for few data points).

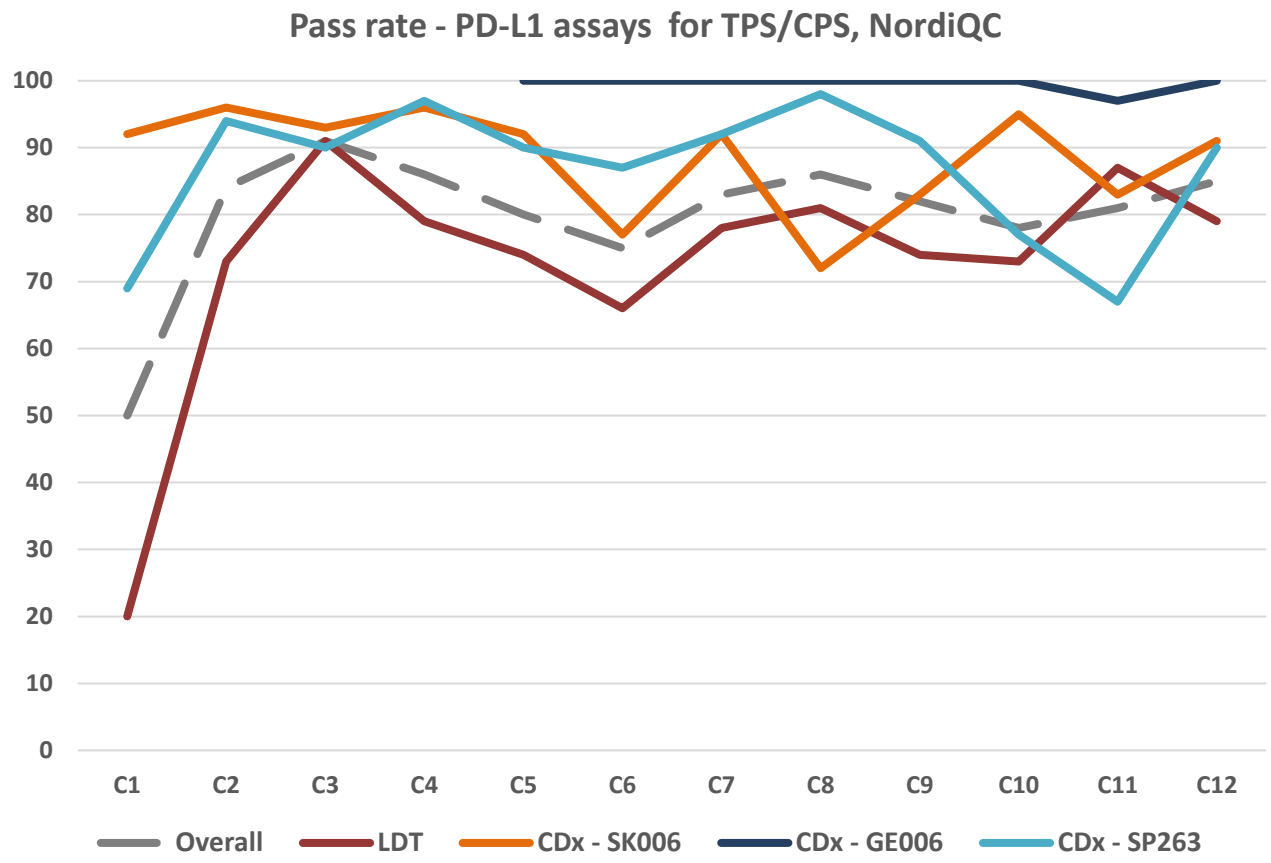
The Dako/Agilent pharmDx SK005 28-8 for Autostainer Link 48 was used by 3 laboratories. All used the recommended protocol settings with 2 results being assessed as sufficient (1 =Good, 1= Optimal) and 1 as Borderline.

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 47% (109 of 232) of the participants, which was increased compared to 37% in the previous assessment – C11. For this group a pass rate of 79% was observed being a reduced compared to the level seen in the last

assessment run. Focusing on the performance of PD-L1 LD assays from C2-C12, excluding the initial run C1 and start-up phase to identify "best practice LD assays", the mean pass rate for LD assays has been 77% (range 66%-91%) compared to e.g. 99% for the 22C3 GE006 pharmDx (Dako/Agilent) and 88% for both the SP263 assay (Ventana/Roche) and 22C3 SK006 pharmDx (Dako/Agilent).

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

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



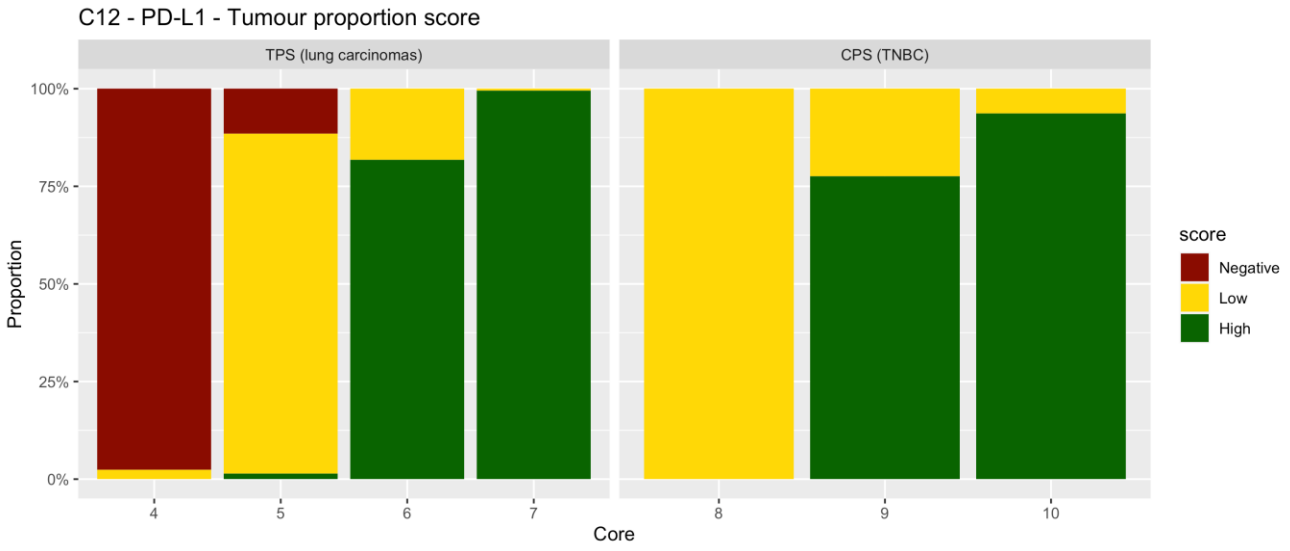
The mAb clone 22C3 was the most widely used concentrated Ab within a LD assay (n=39) providing a pass rate of 79%, 33% optimal which is reduced compared to run C11.

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 fully comparable and even superior to some 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. Cumulated data for runs C8 - C12 focusing on the performance of mAb clone 22C3 on the BOND platforms have shown a pass rate of 31% (5 of 16), no optimal, despite the clone 22C3 was applied by similar central protocol settings on BOND compared to both BenchMark and Omnis, but so far with limited success. Only a few data observations generated and conclusions to be taken with caution, but as mentioned the same trend has now been observed in 5 successive runs.
For the BOND platform, when using the rmAb CAL10 (conc.), rmAb 73-10 (RTU), rmAb SP263 (RTU) and rmAb ZR3 (conc.) with a range of protocols Optimal results were achieved in this run.

PD-L1 read-out 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 C12: Tumour Proportion scores (TPS) in NSCLCs (core no. 4-7) and Combined Positive Score (CPS) in TNBCs (core no. 8-10).**



As seen in Graph 4, relatively high consensus rates were observed for the tissue core 4, 7, 8 and 10, whereas the consensus rates were somewhat lower in tissue core 5, 6 and 9. This also correlates with the insufficient and false negative results typically being seen in these tissue cores.

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, 790-4905/4907, 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. Sompuram, Seshi R et al. Quantitative comparison of PD-L1 IHC assays against NIST standard reference material 1934." Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc. 2022;35(3): 326-332.

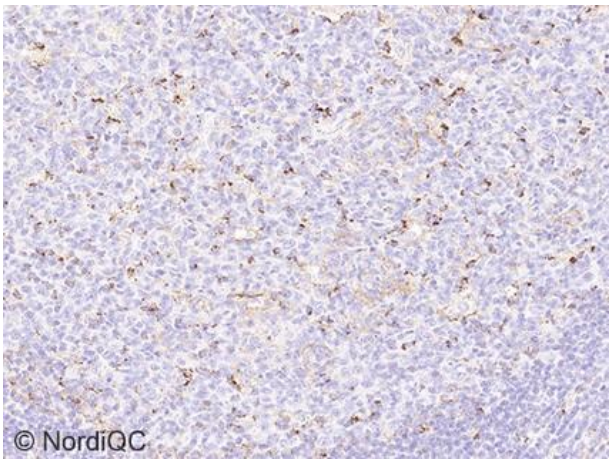


Fig. 1a
Optimal staining result of tonsil using the PD-L1 IHC 22C3 pharmDx kit, GE006 Dako/Agilent on the Omnis platform following the vendor recommended protocol settings.
A weak to moderate, but distinct punctuated membranous staining reaction of germinal centre macrophages and dispersed lymphocytes is seen.
No staining reaction is seen in the vast majority of lymphocytes.
Also compare with Figs. 2a – 6a, same protocol.

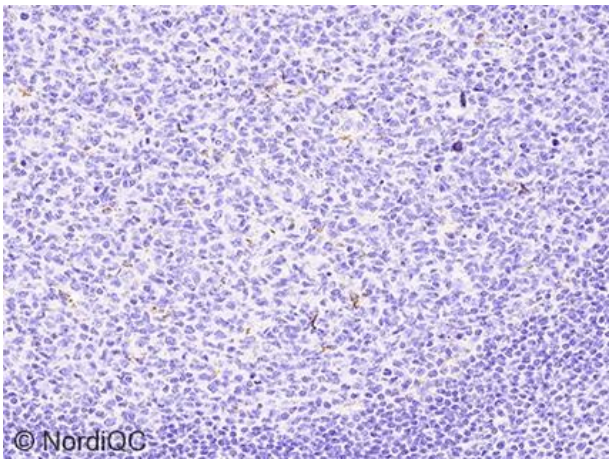


Fig. 1b
Insufficient staining result of tonsil, using the mAb clone 22C3 as concentrate within a laboratory developed assay on Dako Omnis providing a too low analytical sensitivity.
Only few lymphocytes and germinal centre macrophages show a weak membranous staining reaction.
Also compare with the insufficient results seen in two of the included carcinomas in the TMA, as shown in Figs. 4b and 5b.

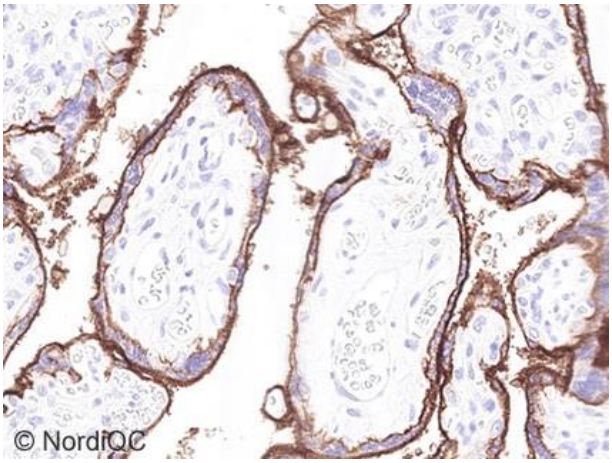


Fig. 2a
Optimal staining result of the placenta, using the same protocol as in Fig. 1a.
Virtually all trophoblasts show a strong membranous staining reaction, but also a weak to moderate staining reaction of the cytoplasmic compartment.

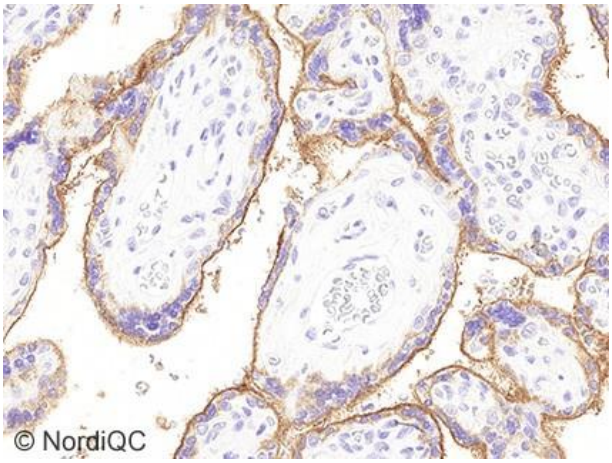


Fig. 2b
Staining result of the placenta, using the same protocol as in Fig. 1b.
The vast majority of trophoblasts show a moderate membranous staining reaction, whereas the cytoplasm only display a barely perceptible staining reaction.
Placenta was found to be a valuable supplemental tool to tonsil for the evaluation of level of analytical/technical sensitivity of the PD-L1 IHC test when focusing on the staining pattern in the trophpoblasts.
Also compare with Figs. 3b - 5b, same protocol.

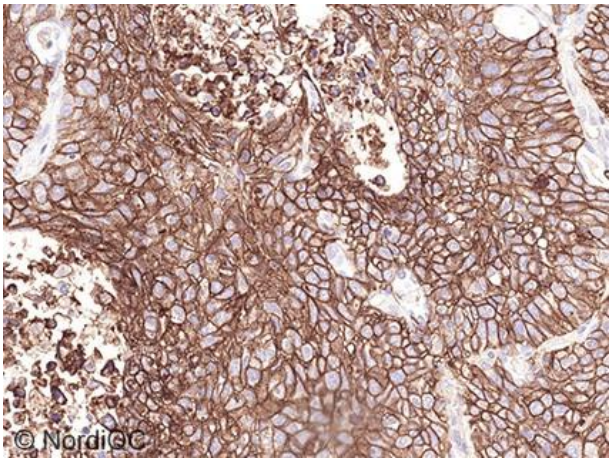


Fig. 3a
Optimal staining result of the NSCLC, tissue core no. 7, using the same protocol as in Figs. 1a and 2a.
A strong distinct membranous staining reaction is seen in virtually all tumour cells.
The tumour was categorized as TPS high (≥50%) and thus eligible for first line immune therapy with KEYTRUDA® (different regional cut-offs occur).

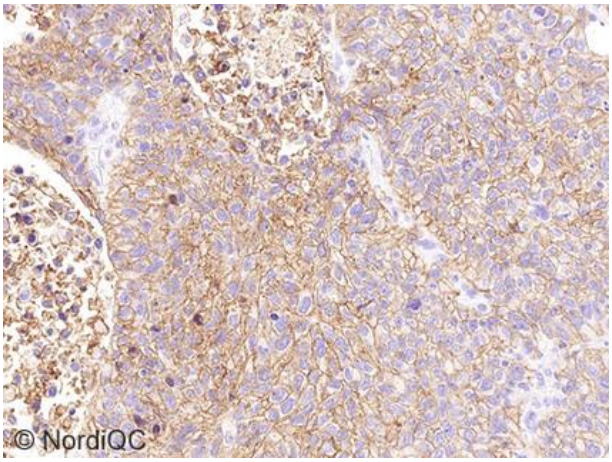


Fig. 3b
Staining result of the NSCLC, tissue core no. 7, using the same protocol as in Figs. 1b and 2b.
The vast majority of tumour cells show a weak to moderate membranous staining reaction.
The tumour was despite a reduced intensity of the tumour cells still categorized as TPS high (≥50%) and thus eligible for first line immune therapy with KEYTRUDA® (different regional cut-offs occur).
However, also compare with Figs. 4b and 5b, same protocol.

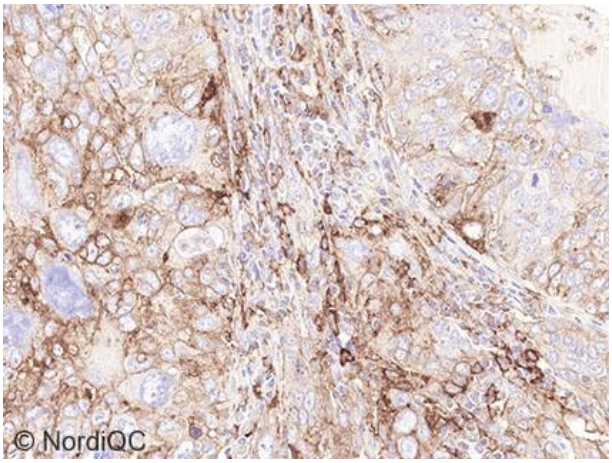


Fig. 4a
Optimal staining result of the NSCLC, tissue core no. 6, using the same protocol as in Figs. 1a - 3a.
A weak to moderate membranous staining reaction is seen in most tumour cells.
The tumour was categorized as TPS high (≥50%) and thus eligible for first line immune therapy with KEYTRUDA® (different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

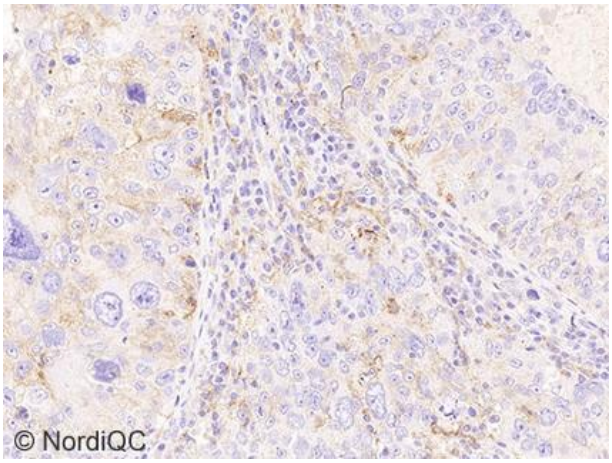


Fig. 4b
Insufficient staining result of the NSCLC, tissue core no. 6, using the same protocol as in Figs. 1b – 3b.
Only scattered tumour cells show a weak membranous staining reaction changing the TPS category from the expected high to low – same field as Fig. 4a.
The intensity of the membranes in this tumour was reduced compared to the NSCLC tissue core no. 7 (see Fig. 3a) and thus more "technically" challenging, but of diagnostic importance.
Compare to the expected result as shown in Fig. 4a.

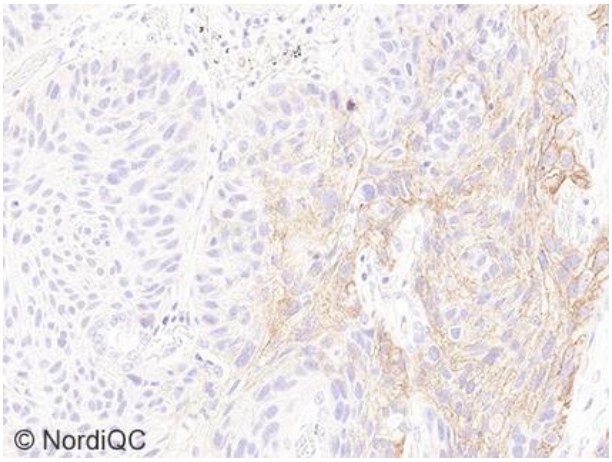


Fig. 5a
Optimal staining result of the NSCLC, tissue core no. 5, using the same protocol as in Figs. 1a - 4a.
A weak to moderate, but distinct staining reaction is seen in dispersed tumour cells.
The tumour was categorized as TPS low (≥1-49%) and thus eligible for second line immune therapy with KEYTRUDA® (different regional cut-offs occur).

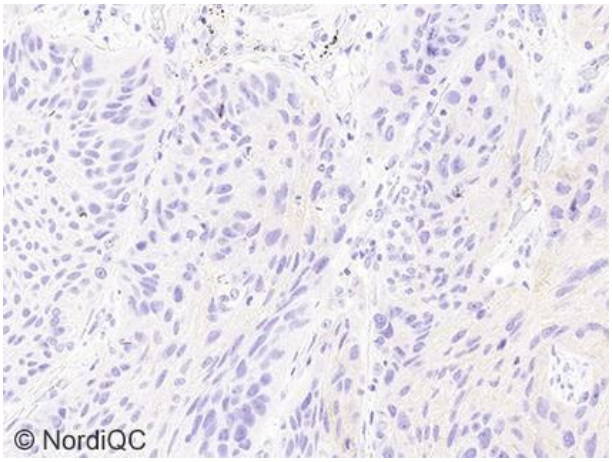


Fig. 5b
Insufficient staining result of the NSCLC, tissue core no. 5, using the same protocol as in Figs. 1b – 4b.
Virtually no staining reaction in the tumour cells is seen.
The PD-L1 category changed from the expected TPS low to negative and not being eligible for immune therapy.
Compare to the expected result as shown in Fig. 5a.

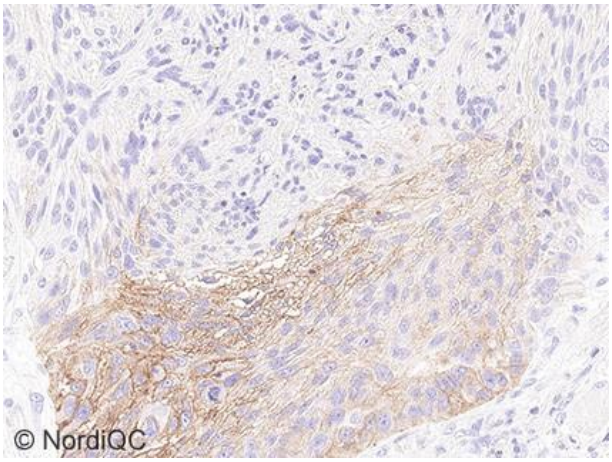


Fig. 6a
Optimal staining result of the NSCLC, tissue core no. 5, using the mAb clone 22C3 as concentrate within an optimally calibrated laboratorory developed assay and performed on BenchMark Ultra using OptiView as detection system.
A weak to moderate, but distinct staining reaction is seen in dispersed tumour cells.
The tumour was categorized as TPS low (≥1-49%) and thus eligible for second line immune therapy with KEYTRUDA® (different regional cut-offs occur).

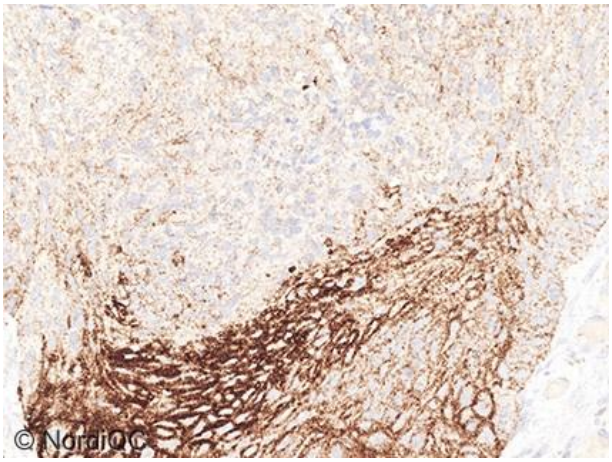


Fig. 6b
Insufficient staining result of the NSCLC, tissue core no. 5, using the 22C3 GE006 pharmDx by laboratory modified protocol settings on BenchMark Ultra.
A granular and extended membranous staining reaction for PD-L1 is seen in most tumour cells changing the PD-L1 category from TPS low (≥1-49%) to TPS high (≥50%).
The protocol was based on a detection system with tyramide amplification (OptiView + Amplification).

JH/LE/SN 13.12.22