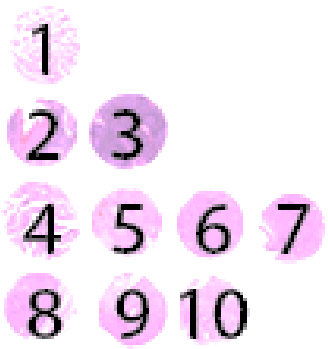
Assessment Run C13 2023
PD-L1 TPS/CPS

Purpose
This was the thirteenth 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 third 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®) C13 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; 20-40%**
6. NSCLC	TPS: High; 50-90%
7. NSCLC	TPS: High; 100%
8. TNBC	CPS: <10
9. TNBC	CPS: ≥10; 50-100 IC#
10. TNBC	CPS: ≥10; 50-100 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.
** The tumour showed heterogeneity in the different levels within and in between the TMA's used. In three of the six TMA's used for the assessment, areas with TPS 50-80% 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.

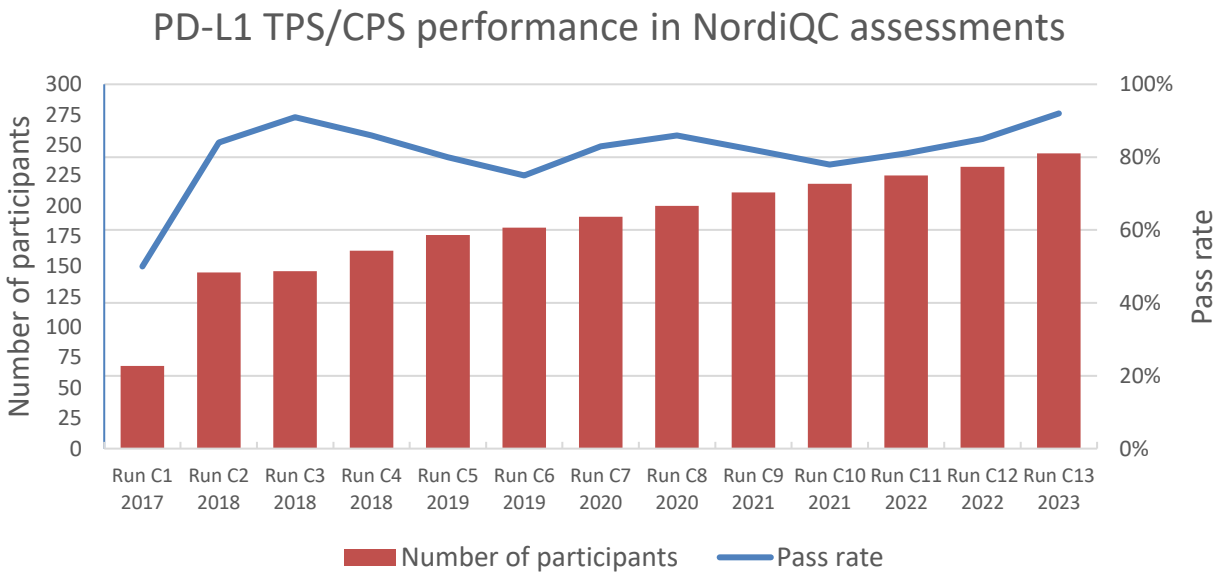
Participation

Number of laboratories registered for PD-L1 KEYTRUDA IHC C13	261
Number of laboratories returning PD-L1 KEYTRUDA IHC slides	243 (93%)
Number of laboratories returning PD-L1 scoring sheet	211

Results
243 laboratories participated in this assessment and returned slides. 92% 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 thirteenth NordiQC assessment of PD-L1 TPS/CPS (KEYTRUDA®). A relatively consistent pass rate has been observed with an upward trend seen 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 thirteenth NordiQC assessment of PD-L1 for TPS/CPS status with focus on NSCLCs and TNBCs. 243 laboratories participated and a pass rate of 92% 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 79%, being superior to the other companion diagnostic assays and LD assays based on concentrated Abs. Leica Biosystems RTU system for PD-L1 based on rmAb clone 73-10 without predictive claim showed an impressive performance giving a pass rate of 100%, 83% optimal.

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 C13, 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]**	41	Ventana/Roche	5	33	3	-	93%	12%
rmAb clone SP263, **741-4905 (LPMS)[4]**	2	Ventana/Roche	-	1	1	-	-	-
rmAb clone SP263, **740-4907 (VRPS)[3]**	12	Ventana/Roche	3	9	-	-	100%	25%
mAb clone 22C3 pharmDX, **SK006 (VRPS)[3]**	19	Dako/Agilent	14	4	-	1	95%	74%
mAb clone 22C3 pharmDX, **SK006 (LMPS)[4]**	20	Dako/Agilent	13	5	2	-	90%	65%
mAb clone 22C3 pharmDX, **GE006 (VRPS)[3]**	29	Dako/Agilent	23	6	-	-	100%	79%
mAb clone 22C3 pharmDX, **GE006 (LMPS)[4]**	18	Dako/Agilent	12	4	2	-	89%	67%
rmAb clone 28-8 pharmDX, **SK005 (VRPS)[3]**	3	Dako/Agilent	2	1	-	-	-	-
Antibodies[5] for laboratory developed PD-L1 assays, concentrated antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
mAb clone **22C3**	44	Dako/Agilent	18	19	7	-	84%	41%
rmAb **CAL10**	4 1	Zytomed Systems Biocare Medical	2	2	-	1	80%	40%
rmAb clone **E1L3N**	4	Cell Signaling	1	3	-	-	-	-
rmAb clone **QR1**	2	Quartett	2	-	-	-	-	-
rmAb clone **28-8**	1	Dako/Agilent	-	1	-	-	-	-
rmAb clone **ZR3**	1	Zeta Corporation	-	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, **790-4905[6] (VRPS)[3]**	13	Ventana/Roche	-	11	2	-	85%	-
rmAb clone SP263, **790-4905[6] (LMPS)[4]**	16	Ventana/Roche	1	15	-	-	100%	6%
rmAb clone 73-10 **PA0832**	6	Leica Biosystems	5	1	-	-	100%	83%
rmAb MX070C **MAB-0854**	2	Fuzhou Maixin	1	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	-	-	-	-	-
rmAb clone BP6099 **I12052E**	1	Biolynx	1	-	-	-	-	-
Total	243		106	118	17	2		
Proportion			44%	48%	7%	1%	92%	

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, 5 of 41 (12%) 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, 33 of 41 (93%) laboratories produced a sufficient staining result (optimal or good).

SP263 (740-4907, Ventana/Roche): In total, 3 of 12 (25%) 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, 12 of 12 (100%) laboratories produced a sufficient staining result.

PD-L1 IHC 22C3 pharmDx (SK006, Dako/Agilent): In total, 14 of 19 (74%) 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, 18 of 19 (95%) 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, 23 of 29 (79%) 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, 29 of 29 (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**	38/41 (93%)	5/41 (12%)	-	-
Ventana BenchMark Ultra rmAb SP263, **740-4907**	12/12 (100%)	3/12 (25%)	-	-
Dako Autostainer Link 48+ mAb 22C3 pharmDX, **SK006**	18/19 (95%)	14/19 (74%)	18/20 (90%)	13/20 (65%)
Dako Omnis mAb 22C3 pharmDX, **GE006**	39/29 (100%)	23/29 (79%)	16/18 (89%)	12/18 (67%)
Dako Autostainer Link 48+ rmAb 28-8 pharmDX, **SK005**	3/3	2/3	-	-

*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, 18 of 44 (41%) protocols were assessed as optimal of which 25 were stained on the BenchMark Ultra platform (Ventana/Roche), 1 on BenchMark XT platform (Ventana/Roche), 10 on the Omnis platform (Dako/Agilent), 4 on Autostainer Link 48 (Dako/Agilent), 2 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 1:20-40 for mAb clone 22C3, incubation time of 48-120 min., HIER in CC1 for 48-64 min. (and up to 80 min. on Discovery Ultra) and OptiView as the detection system. Using these protocol settings, 9 of 25 (36%) laboratories produced optimal staining results, and 19 of 25 (76%) laboratories produced sufficient 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 40-50 min. and EnVision™ FLEX+ as detection system. Using these protocol settings, 6 of 10 (60%) laboratories produced optimal results and 10 of 10 (100%) laboratories produced a sufficient staining result.

rmAb clone **E1L3N**: 1 of 4 protocols were assessed as optimal.
The protocol providing an optimal result was based on a titre of 1:100 of the primary Ab, incubation time of 30 min., HIER in Bond™ Epitope Retrieval Solution 2 (BERS2; Leica Biosystems) pH 9 at 100°C for 30 min. and Bond™ Refine as the detection system and performed on a Leica BOND-III.

rmAb clone **CAL10**: 2 of 5 protocols were assessed as optimal.
The optimal protocols for this clone was based on:

- Leica BOND-III - a titre of 1:50 of the primary Ab, incubation time of 20 min., HIER in BERS2 (Leica Biosystems) pH 9 at 100°C for 20 min. and Bond™ Refine as the detection system.
- IntelliPATH (Biocare) - a titre of 1:100 of the primary Ab, incubation time of 30 min., HIER in Citrate pH 6 (Zytomed Systems) at 110°C for 7 min. and ZytoChem Plus (HRP) Polymer Kit as the detection system.

rmAb clone **QR1**: 2 of 2 protocols were assessed as optimal.
The protocols providing optimal results were based on a titre of 1:100-150 of the primary Ab, incubation time of 20-60 min., HIER in Bond™ Epitope Retrieval Solution (Leica Biosystems) pH 9 at 100°C for 30 min. and Bond™ Refine as the detection system and performed on a Leica BOND-III and Leica BOND Max.

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**	9/25** (36%)	-	-	1/4 (25%)	1/1	5/9 (56%)	1/3	-

*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.

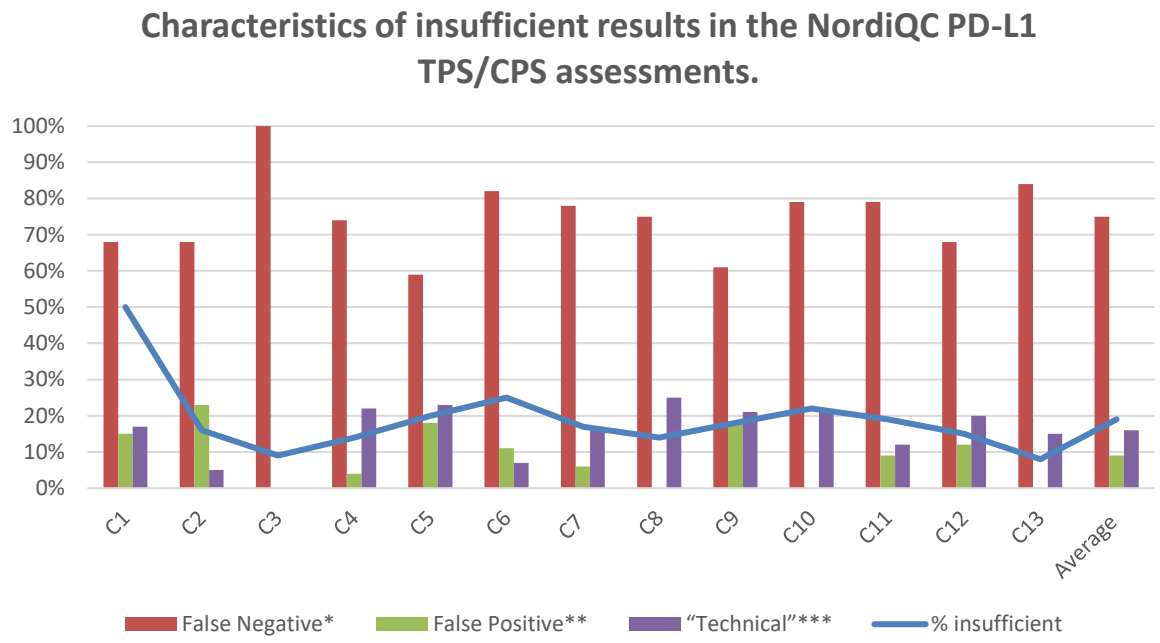
Ready-To-Use antibodies for laboratory developed (LD) assays
rmAb clone **73-10**, product no. PA0832, Leica Biosystems, Bond III
In total, 5 of 6 (83%) protocols were assessed as optimal. Protocols with optimal results were typically based on the vendor recommended protocol settings using HIER in Bond™ Epitope Retrieval Solution 1 (BERS1; Leica Biosystems) pH 6 at 100°C for 20 min. and Bond™ Refine as the detection system and performed on a Leica BOND-III. Using these protocol settings, 6 of 6 (100%) laboratories produced a sufficient staining result.

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 six of these used for this assessment.
Reviewing the reference slides from the blocks, a heterogenic expression of PD-L1 was seen in three of the tumor cores. Of particular importance for the NSCLC, tissue core no. 5, focal areas with TPS 50-80% (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 thirteenth 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 84% 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 75%. 0% of the insufficient results were related to a false positive staining result while the remaining 15% 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 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.

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 respectively by the NordiQC reference standard methods, was the most challenging to obtain an optimal result. In addition, the TNBC, tissue core no. 8, expected to show a CPS<10 with a positive staining reaction primarily in the immune cells, also was found to be challenging.

49% (n=118) of the results submitted were marked as "Good". In 73% of these (86 of 118), 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 4% (5 of 118) 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 23% (27 of 118) 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 22% of the participants and in total provided an overall pass rate of 94%, with 15% 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 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. 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≥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 18% of the participants 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. 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.

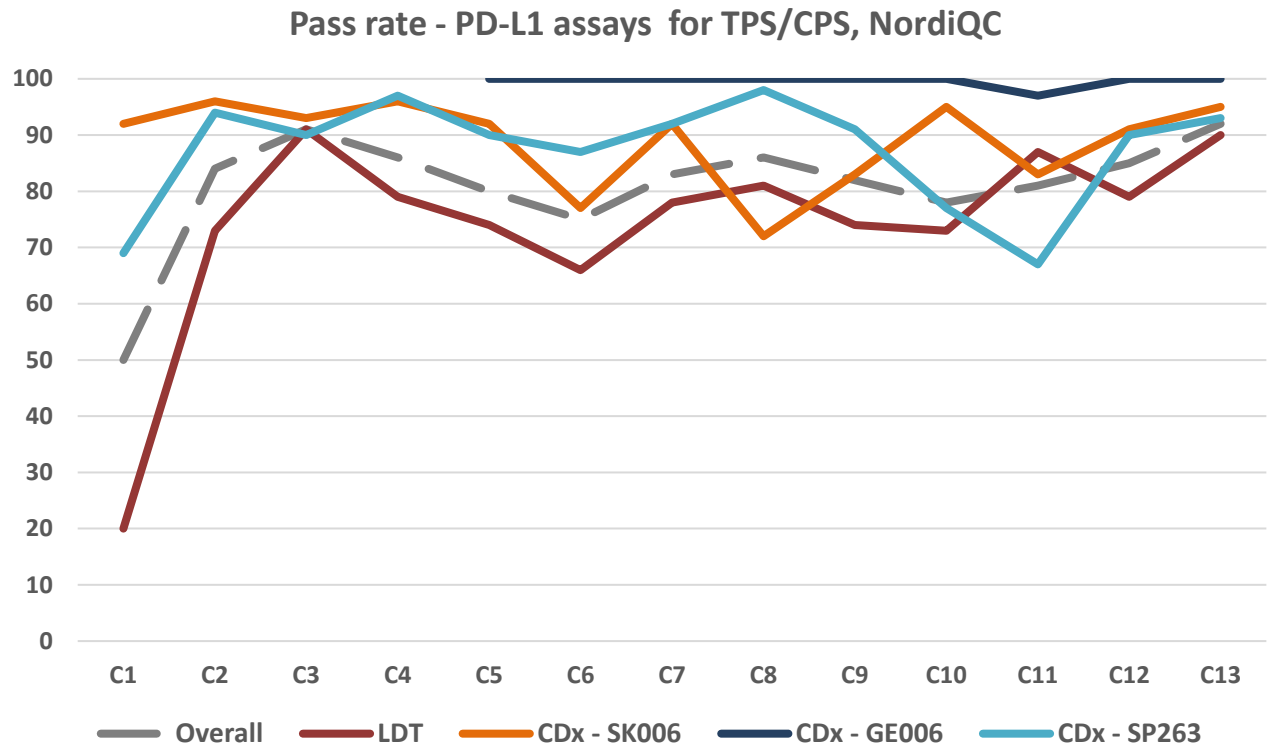
The Dako/Agilent 22C3 pharmDx assay SK006 for Autostainer Link 48 was used by 8% of the participants and provided a pass rate of 95% (74% optimal) when applied by protocol settings in compliance with vendor recommendations (see Table 3). The 22C3 SK006 assay was frequently applied off-label (n=20), both on Autostainer 48 Link using modified protocol settings or on non-Autostainer Link 48 platforms as e.g. BenchMark Ultra/GX/XT (Ventana/Roche) and Omnis (Dako/Agilent), and as shown in Table 2 with inferior performance.

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 (1 =Good, 2= 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 52% (126 of 243) of the participants, which was increased compared to 47% in the previous assessment – C12. For this group a pass rate of 90% was observed which is an increase compared to the level of 79% as seen in the last assessment run. Focusing on the performance of PD-L1 LD assays from C2-C13, 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) and 89% 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.

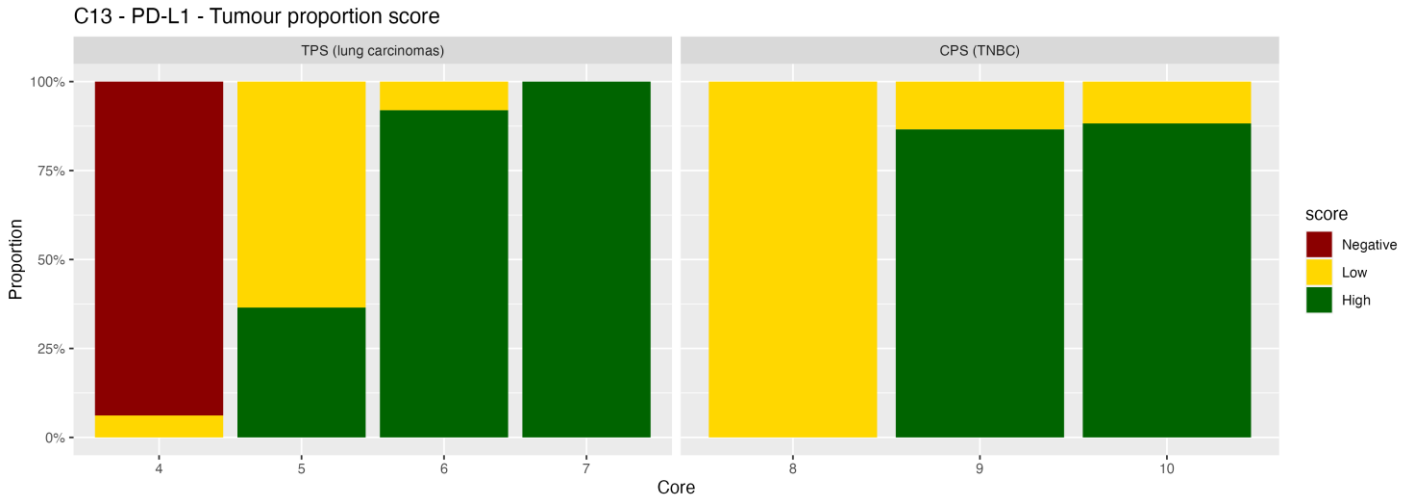
Graph 3. **Proportion of pass rates for PD-L1 TPS/CPS assays in the NordiQC runs performed.**



The mAb clone 22C3 was the most widely used concentrated Ab within a LD assay (n=44) providing a slightly improved pass rate of 84%, 41% optimal which is increased compared to run C12 (79%, 33%). 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 this run there was a 100% sufficient pass rate, with 1 participant achieving an optimal result. Cumulated data for runs C8 - C13 focusing on the performance of mAb clone 22C3 on the BOND platforms have shown a pass rate of 42% (8 of 19), 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 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 C13: 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 cores no. 4, 6, 7 and, 8, whereas the consensus rates being significantly lower in tissue core no. 5 but also slightly reduced in cores no.9 and 10.

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.

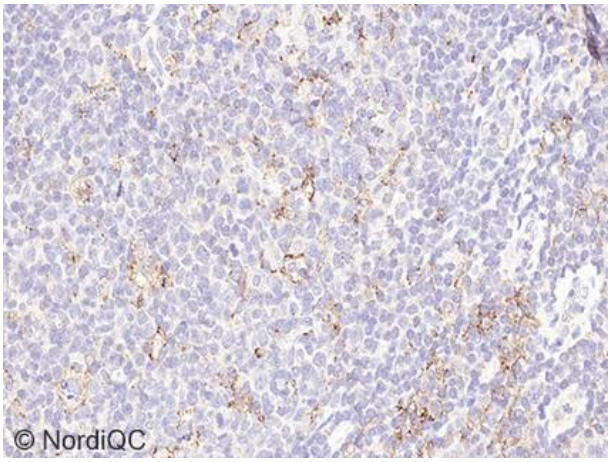


Fig. 1a
Optimal staining result of tonsil using the PD-L1 IHC 22C3 pharmDx kit, Dako/Agilent following the vendor recommended protocol settings.
A weak to moderate, but distinct punctuated membranous staining reaction of germinal centre macrophages and dispersed lymphocytes is seen.
No staining reaction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

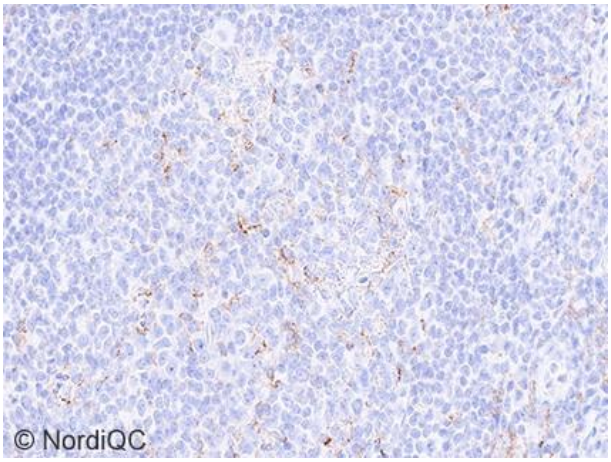


Fig. 1b
Staining result of tonsil, using the mAb clone 22C3 as concentrate on BenchMark Ultra, Ventana/Roche.
The result in tonsil is slightly inferior to the expected level as seen in Fig. 1a and obtained by the 22C3 pharmDx kit as the contours of cells demonstrated being diffuse.
Overall, the protocol provided an insufficient result characterized by a too low level of analytical sensitivity – see Figs. 3b and 4b.
In this context is 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.

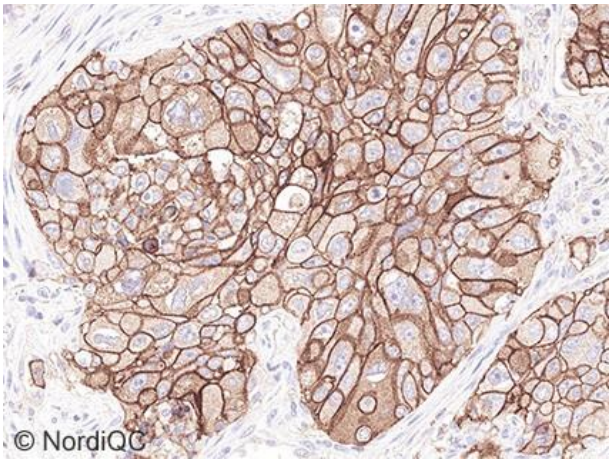


Fig. 2a
Optimal staining result of the NSCLC, tissue core no. 7, using the same protocol as in Fig. 1a.
A moderate to strong, distinct membranous staining reaction is seen in virtually all tumour cells.
The tumour was categorized as TPS High (≥50%) and thus eligible for first line immune therapy with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

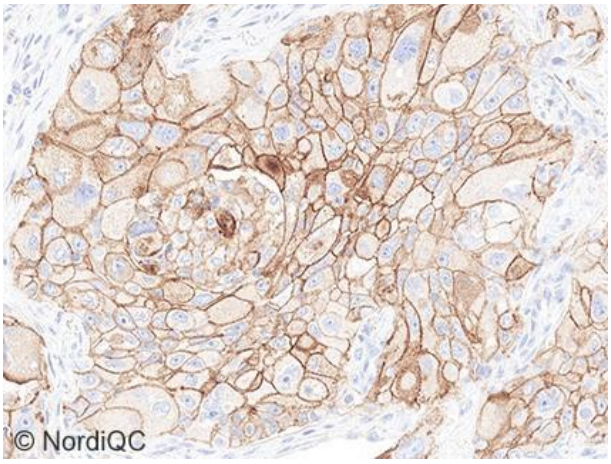


Fig. 2b
Staining result of the NSCLC, tissue core no. 7, using the same protocol as in Fig. 1b.
A moderate and distinct membranous staining reaction is seen in virtually all tumour cells.
Despite the intensity of the membrane reaction is reduced, the tumour is still categorized as TPS High (≥50%) and thus eligible for first line immune therapy with KEYTRUDA® (different regional cut-offs occur).
Also compare with Figs. 3b - 4b, same protocol.

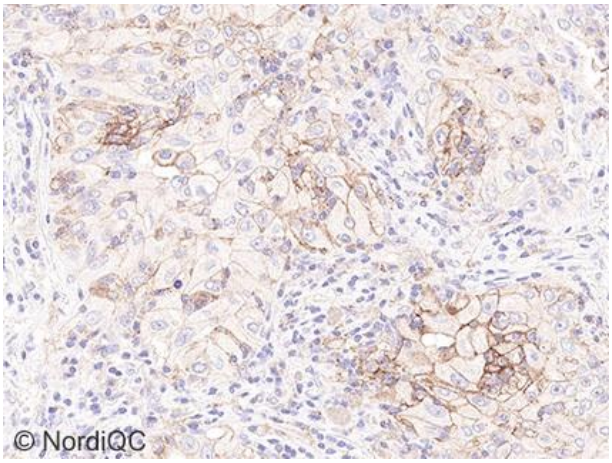


Fig. 3a
Optimal staining result of the NSCLC, tissue core no. 6, using the same protocol as in Figs. 1a and 2a.
A weak to moderate, but distinct staining reaction is seen in most tumour cells.
The tumour was categorized as TPS High (≥50%) and thus eligible for first line immune therapy with KEYTRUDA® (different regional cut-offs occur).

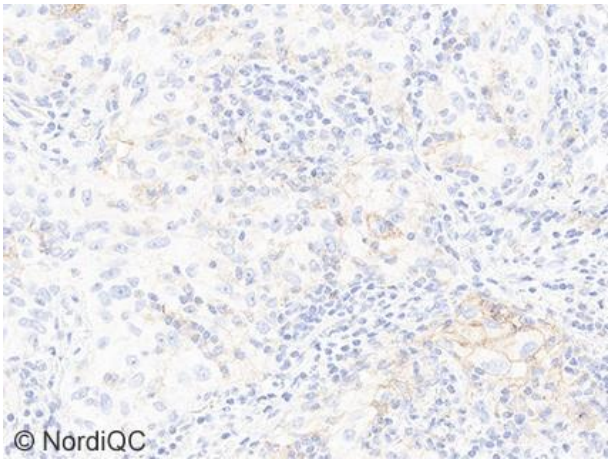


Fig. 3b
Insufficient staining result of the NSCLC, tissue core no. 6, using the same protocol as in Figs. 1b and 2b.
Only dispersed tumour cells show a membranous staining reaction and the PD-L1 category being changed from the expected TPS High to TPS Low and the tumour not being eligible for first line immune therapy.
Compare to the expected result as shown in Fig. 3a.

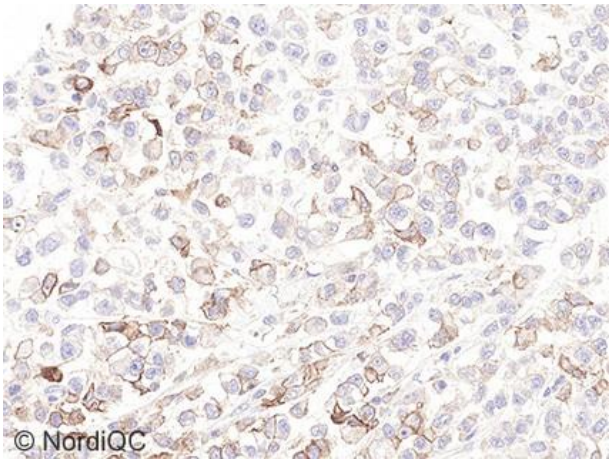


Fig. 4a
Optimal staining result of the TNBC, tissue core no. 10, using the same protocol as in Figs. 1a - 3a.
A weak to strong and distinct staining reaction is seen in most tumour cells.
The tumour was categorized as CPS≥10 and thus eligible for immune therapy with KEYTRUDA®.

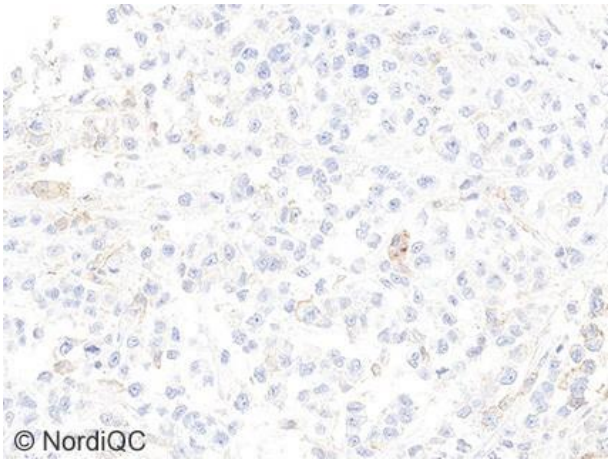


Fig. 4b
Insufficient staining result of the TNBC, tissue core no. 10, using the same protocol as in Figs. 1b – 3b.
Only dispersed tumour cells are demonstrated and significantly reduced compared to the level expected and overall the result categorized as CPS<10. The PD-L1 category consequently being changed from the expected positive level to negative and the tumour not being eligible for immune therapy.
Compare to the expected result as shown in Fig. 4a.

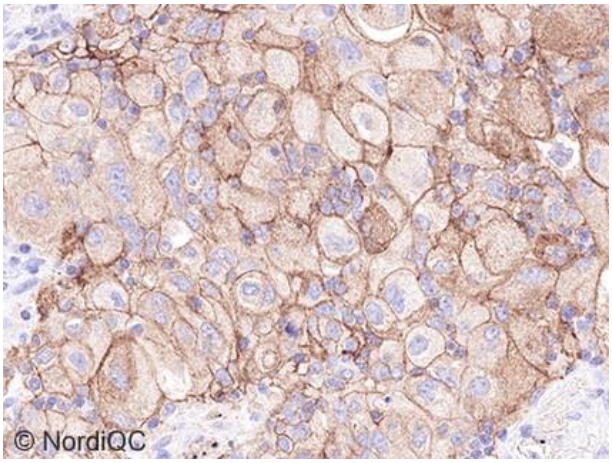


Fig. 5a
Optimal staining result of the NSCLC, tissue core no. 6, using the mAb clone 22C3 as a concentrate using EnVision FLEX+ as detection system and performed on Dako Omnis.
A moderate, distinct membranous staining reaction is seen in the majority of tumour cells.
The tumour was categorized as TPS High (≥50%) and thus eligible for first line immune therapy with KEYTRUDA® (different regional cut-offs occur).

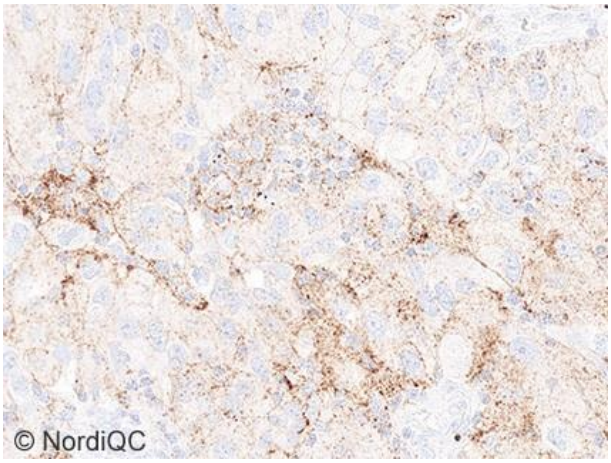


Fig. 5b
Insufficient staining result of the NSCLC, tissue core no. 6, using the mAb clone 22C3 as a concentrate using OptiView plus amplification kit as detection system and performed on BenchMark Ultra.
A diffuse granular and impreceise staining reaction is seen compromising the read-out and scoring of the specific membrane reaction for PD-L1. Both the participant and the NordiQC assessor group scored the tumour as TPS Low and thus incorrect category compared to the result expected (TPS High).
Compare to the expected result as shown in Fig. 5a.

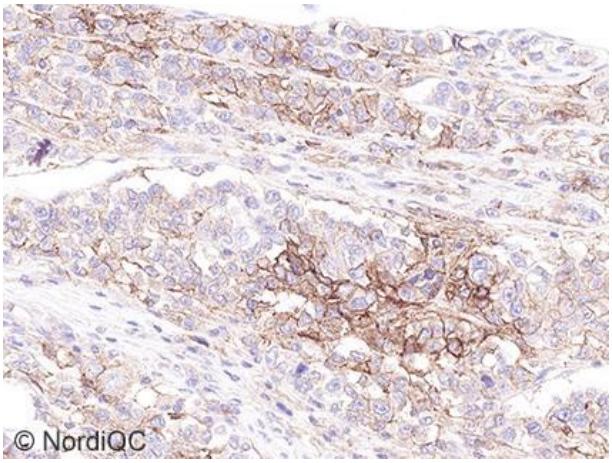


Fig. 6a
Optimal staining result of the TNBC, tissue core no. 10, using the the 22C3 pharmDx, Agilent/ Dako following the vendor recommended protocol settings.
A weak to strong and distinct membranous staining reaction is seen in most tumour cells.
The tumour was categorized as CPS≥10 and thus eligible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

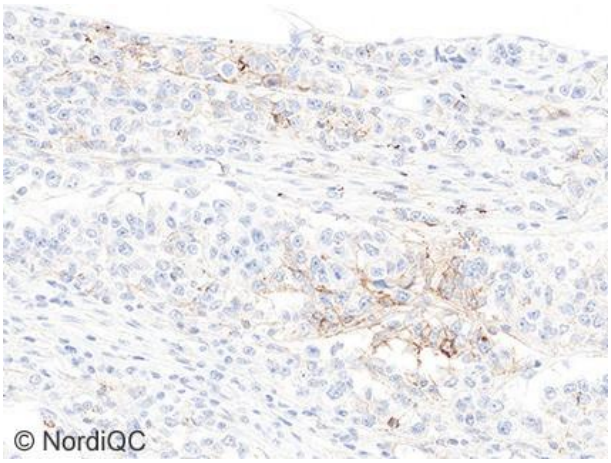


Fig. 6b
Staining result assessed as "Good" of the TNBC, tissue core no. 10, using the SP263 PD-L1 IHC assay, Ventana/Roche following the vendor recommended protocol settings.
A weak to moderate and distinct membranous staining reaction is seen in many tumour cells.
The tumour was categorized as CPS≥10 and thus eligible for immune therapy with KEYTRUDA®.
However, both the intensity and proportion of cells demonstrated was reduced compared to the level expected, but without consequence regarding treatment with KEYTRUDA®. This pattern was typically observed for SP263 compared to both other companion diagnostic assays as 22C3 Dako/Agilent and optimized laboratory developed tests based on concentratred formats or RTU without predictive claims.

JH/LE/SN 06.07.2023