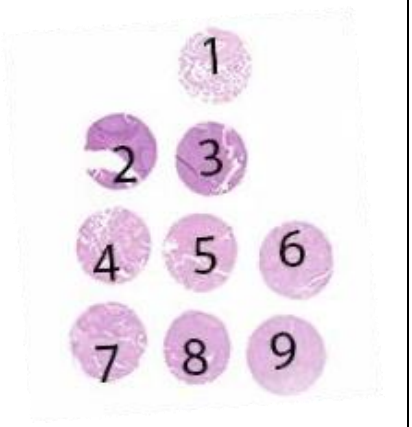
Assessment Run C16 2024
PD-L1 TPS/CPS

Purpose
This was the sixteenth 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.

Material
Table 1. **Content of the TMA used for the NordiQC PD-L1 TPS/CPS (KEYTRUDA®) C16 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; 95%
7. TNBC	CPS: <10 IC#
8. TNBC	CPS: ≥10; 30-70 IC#
9. TNBC	CPS: ≥10; 70-100 IC+TC#



* 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 seven 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.

KEY POINTS FOR PD-L1 TPS/CPS IMMUNOASSAYS

- The **CDx** IHC assays with one or more predictive claims provided an overall pass rate of 93% compared to 82% for LD assays.
- The **22C3 CDx** assay GE006, Dako/Agilent was most successful with a pass rate of 100%, 95% optimal.
- Insufficient results were mainly caused by extensive cytoplasmic staining reaction, poor signal-to-noise ratio compromising the read-out and reduced proportion of expected expression level.

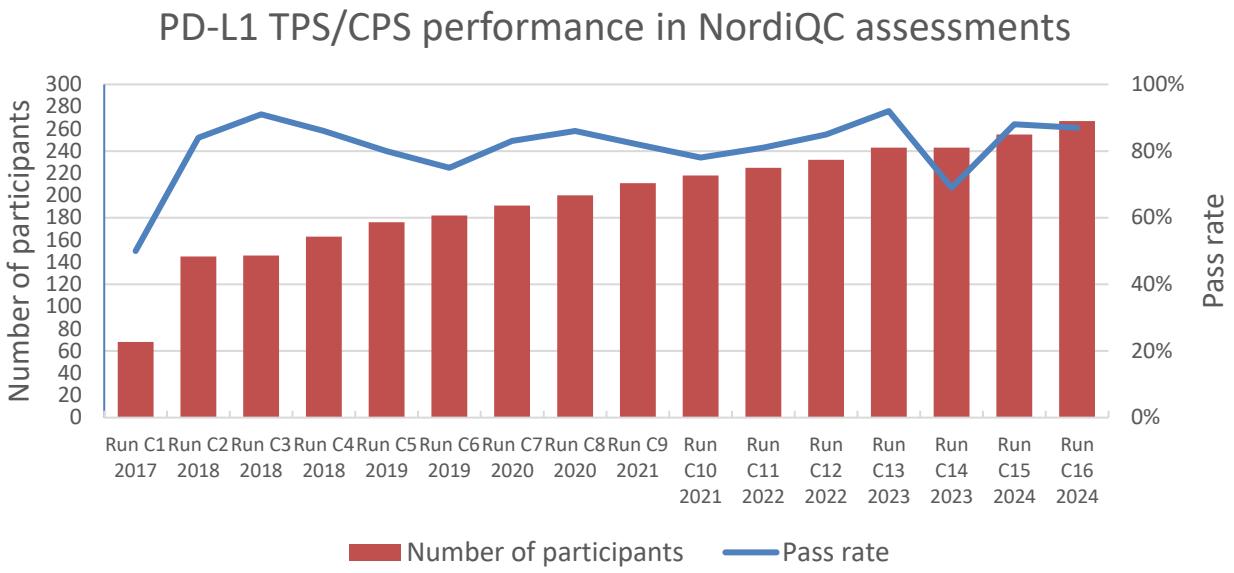
Participation

Number of laboratories registered for PD-L1 KEYTRUDA IHC C16	279
Number of laboratories returning PD-L1 KEYTRUDA IHC slides	267 (96%)
Number of laboratories returning PD-L1 scoring sheet	230

Results
267 laboratories participated in this assessment and returned slides. 87% of the participants achieved a sufficient mark. Assessment marks for IHC PD-L1 assays and PD-L1 antibodies are summarized in Table 2a-2d (see page 3-5). 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
A relatively consistent pass rate has been observed (with the exception of C14) with an upward trend seen overall, but in the latest runs it has remained steady with only 1% difference between C15 and C16 as shown in Graph 1 (see page 3). The number of new participants has recently been consistently increasing by about 3-5% in each run but had remained the same for runs C13 and C14 but increased again in the latest runs C15 and C16.

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



Conclusion

This was the sixteenth NordiQC assessment of PD-L1 for TPS/CPS status with focus on NSCLCs and TNBCs. 267 laboratories participated and a pass rate of 87% was observed.

The PD-L1 IHC pharmDx assay, 22C3 GE006, Dako/Agilent applied in concordance to the vendor recommended guidelines, was the most successful companion diagnostic assay providing a pass rate of 100%, with an optimal rate of 95%, being superior to the other companion diagnostic assays. The widely applied Ventana/Roche PD-L1 IHC assays 741-4905 and 740-4907 for BenchMark (Ultra/XT/GX) based on the rmAb clone SP263 provided an overall pass rate of 87% being significantly superior to the level seen in run C14 and virtually on par to the mean level of 86% obtained in run C1-C16.

In this assessment run the majority of insufficient results were related to technical issues e.g. related to extensive cytoplasmic staining reaction, poor signal-to-noise ratio, etc., observed in one or more of the NSCLCs and TNBC. This observation was also seen in run C15 and is again in contrast to the results obtained and described in previous NordiQC PD-L1 TPS/CPS assessments where false negative staining results were observed.

Table 2a. **Overall results for PD-L1 TPS/CPS, run C16**

	n	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
CE-IVD / FDA approved PD-L1 assays*	146	77	57	9	3	92%	53%
Laboratory developed PD-L1 assays based on concentrated antibodies	65	27	23	14	1	77%	42%
PD-L1 assays based on Ready-To-Use antibodies without predictive claims	56	13	34	7	2	84%	23%
Total	267	117	114	30	6		
Proportion		44%	43%	11%	2%	87%	

1) Proportion of sufficient stains (optimal or good) (≥5 assessed protocols).
2) Proportion of optimal results (≥5 assessed protocols).
* Including all protocol settings - both performed as per recommneded guidelines or modified settings.

Table 2b. **Assessment marks for CE-IVD / FDA approved PD-L1 assays for PD-L1 TPS/CPS, run C16**

CE-IVD / FDA approved PD-L1 assays	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
rmAb clone SP263, **741-4905 (VRPS)[3]**	39	Ventana/Roche	6	28	4	1	87%	15%
rmAb clone SP263, **741-4905 (LPMS)[4]**	5	Ventana/Roche	-	4	-	1	80%	-
rmAb clone SP263, **740-4907 (VRPS)[3]**	8	Ventana/Roche	-	7	-	1	86%	-
mAb clone 22C3 pharmDX, **SK006 (VRPS)[3]**	23	Dako/Agilent	14	8	1	-	96%	61%
mAb clone 22C3 pharmDX, **SK006 (LMPS)[4]**	14	Dako/Agilent	5	7	2	-	86%	36%
mAb clone 22C3 pharmDX, **GE006 (VRPS)[3]**	41	Dako/Agilent	39	2	-	-	100%	95%
mAb clone 22C3 pharmDX, **GE006 (LMPS)[4]**	12	Dako/Agilent	12	-	-	-	100%	100%
rmAb clone 28-8 pharmDX, **SK005 (VRPS)[3]**	2	Dako/Agilent	-	1	1	-	-	-
rmAb clone 28-8 pharmDX, **SK005 (LPMS)[4]**	2	Dako/Agilent	1	-	1	-	-	-
Total	146		77	57	9	3		
Proportion			53%	39%	6%	2%	92%	

1) Proportion of sufficient stains (optimal or good) (≥5 assessed protocols).
2) Proportion of optimal results (≥5 assessed protocols).
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.

Table 2c. **Assessment marks for concentrated antibodies for PD-L1 TPS/CPS, run C16**

Antibodies[5] for laboratory developed PD-L1 assays, concentrated antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
mAb clone **22C3**	56	Dako/Agilent	26	17	13	-	77%	46%
rmAb **CAL10**	1 1	Zytomed Systems Biocare Medical	-	2	-	-	-	-
rmAb clone **E1L3N**	2	Cell Signaling	-	1	1	-	-	-
rmAb clone **QR1**	4	Quartett	1	2	-	1	-	-
rmAb clone **28-8**	1	Dako/Agilent	-	1	-	-	-	-
Total	65		27	23	14	1		
Proportion			42%	35%	22%	1%	77%	

1) Proportion of sufficient stains (optimal or good) (≥5 assessed protocols).
2) Proportion of optimal results (≥5 assessed protocols).
5) mAb: mouse monoclonal antibody, rmAb: rabbit monoclonal antibody.

Table 2d. **Assessment marks for Ready-To-Use antibodies[6] for PD-L1 TPS/CPS, run C16**

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	1	10	4	1	67%	6%
rmAb clone SP263, **790-4905[6] (LMPS)[4]**	27	Ventana/Roche	7	18	2	-	93%	26%
rmAb clone 73-10 **PA0832**	5	Leica Biosystems	3	2	-	-	100%	60%
rmAb MX070C **MAB-0854**	1	Fuzhou Maixin	-	1	-	-	-	-
rmAb clone AC37 **PA168**	1	Abcarta	1	-	-	-	-	-
rmAb clone BP6099 **I12052E**	1	Biolynx	-	-	1	-	-	-
rmAb clone RM320 **8263-C010**	1	Sakura Finetek	-	1	-	-		
rmAb **CAL10**	1 1	Zytomed Systems Biocare Medical	-	1	-	1		
rmAb clone QR1 **2-PR292-13**	1	Biocyc	-	1	-	-		
rmAb clone 5D3 **CAA-B001**	1	Arco Biosystems	1	-	-	-		
Total	56		13	34	7	2		
Proportion			23%	61%	12%	4%	84%	

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

Detailed Analysis

CE IVD / FDA approved assays

SP263 (741-4905, Ventana/Roche): In total, 6 of 39 (15%) 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, 34 of 39 (87%) laboratories produced a sufficient staining result (optimal or good).
SP263, 741-4905 was also applied on other non-intended platforms as Leica Biosystems Bond Prime and BenchMark Ultra Plus with an overall performance as shown in Table 2b (LMPS).

PD-L1 IHC 22C3 pharmDx (SK006, Dako/Agilent): In total, 14 of 23 (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, 22 of 23 (96%) laboratories produced a sufficient staining result.
SK006 was also used with 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 2b.

PD-L1 IHC 22C3 pharmDx (GE006, Dako/Agilent): In total, 39 of 41 (95%) 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, 41 of 41 (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**	34/39 (87%)	6/39 (15%)	4/5	0/5
Ventana BenchMark Ultra rmAb SP263, **740-4907**	7/8 (87%)	0/8 (0%)	-	-
Dako Autostainer Link 48+ mAb 22C3 pharmDX, **SK006**	22/23 (96%)	14/23 (61%)	12/14	5/14
Dako Omnis mAb 22C3 pharmDX, **GE006**	41/41 (100%)	39/41 (95%)	12/12 (100%)	12/12 (100%)
Dako Autostainer Link 48+ rmAb 28-8 pharmDX, **SK005**	1/2	0/2	1/2	1/2

*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 clone **22C3**: In total, 26 of 56 (46%) protocols were assessed as optimal of which 32 were stained on the BenchMark Ultra platform (Ventana/Roche), 4 on the BenchMark Ultra Plus platform (Ventana/Roche), 2 on BenchMark XT platform (Ventana/Roche), 10 on the Omnis platform (Dako/Agilent), 2 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, 12 of 32 (36%) laboratories produced optimal staining results, and 27 of 32 (84%) 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 30-50 min. and EnVision™ FLEX+ as detection system. Using these protocol settings, 7 of 10 (70%) laboratories produced optimal results and 8 of 10 (80%) laboratories produced a sufficient staining result.

rmAb clone **QR1**: 1 of 4 protocols was assessed as optimal.
The optimal protocol for this clone was based on a titre of 1:100 of the primary Ab, incubation time of 60 min., HIER in BERS2 (Leica Biosystems) pH 9 at 100°C for 30 min. and Bond™ Refine as the detection system and performed on Bond MAX platform (Leica Biosystems).

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

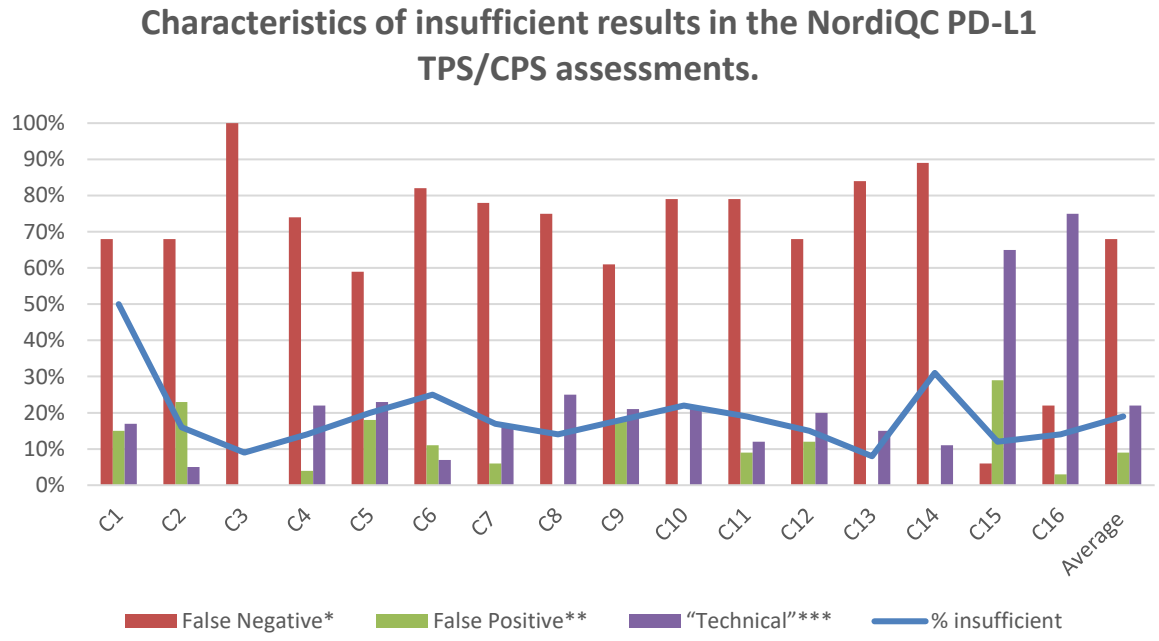
Concentrated antibodies	**Ventana/Roche BenchMark[1]**		**Dako/Agilent Autostainer[2]**		**Dako/Agilent Omnis**		**Leica Biosystems Bond III/ MAX**	
	CC1 pH 8.5	CC2 pH 6.0	TRS pH 9.0	TRS pH 6.1	TRS High pH	TRS Low pH	BERS2 pH 9.0	BERS1 pH 6.0
mAb clone **22C3**	15/38** (39%)	-	-	2/2**	-	7/10**	0/5**	1/5**

*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.
1) BenchMark, XT, Ultra, Ultra Plus
2) Autostainer Link 48

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 GE006 (Dako/Agilent).
In total, nine identical TMA blocks were constructed and seven of these were 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, 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 sixteenth NordiQC assessment for PD-L1 TPS/CPS (KEYTRUDA®), the prevalent feature of an insufficient staining result was technical issues such 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, being observed in 75% of the insufficient results. 22% of the insufficient results were caused by a false negative staining result, and 3% by a false positive staining result. As shown in Graph. 2, a false negative staining result has been the most common reason for insufficient staining results up until C15 of the NordiQC PD-L1 TPS/CPS (KEYTRUDA®) assessments.

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 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 by the NordiQC reference standard method, was the most challenging to obtain an optimal result.

43% (n=114) of the slides submitted were marked as "Good". In 70% of these (80 of 114), 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 114) 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 25% (29 of 114) 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 18% (47 of 267) of the participants and in total provided an overall pass rate of 87% (41 of 47), with 13% (6 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. The proportion of optimal results seen are still inferior to the level seen for the 22C3 IHC pharmDx assays, Dako/Agilent. Both in this assessment run and the runs from C10, a relatively high number of SP263 results have been characterized by a reduced analytical sensitivity providing a lower TPS level compared to the level seen for the 22C3 pharmDx assays. At present, no explanation for this discrepancy has been identified.

The Dako/Agilent **22C3** pharmDx assay GE006 for Dako Omnis was used by 15% (41 of 267) of the participants providing a pass rate of 100% (95% 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 19% (23 of 267) of the participants and provided a pass rate of 96% (61% optimal) when applied by protocol settings in compliance with vendor recommendations (see Table 3). The 22C3 SK006 assay was also applied off-label (n=14), 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 2b in this run with similar performance when applied as per recommendations or by modified off-label settings.

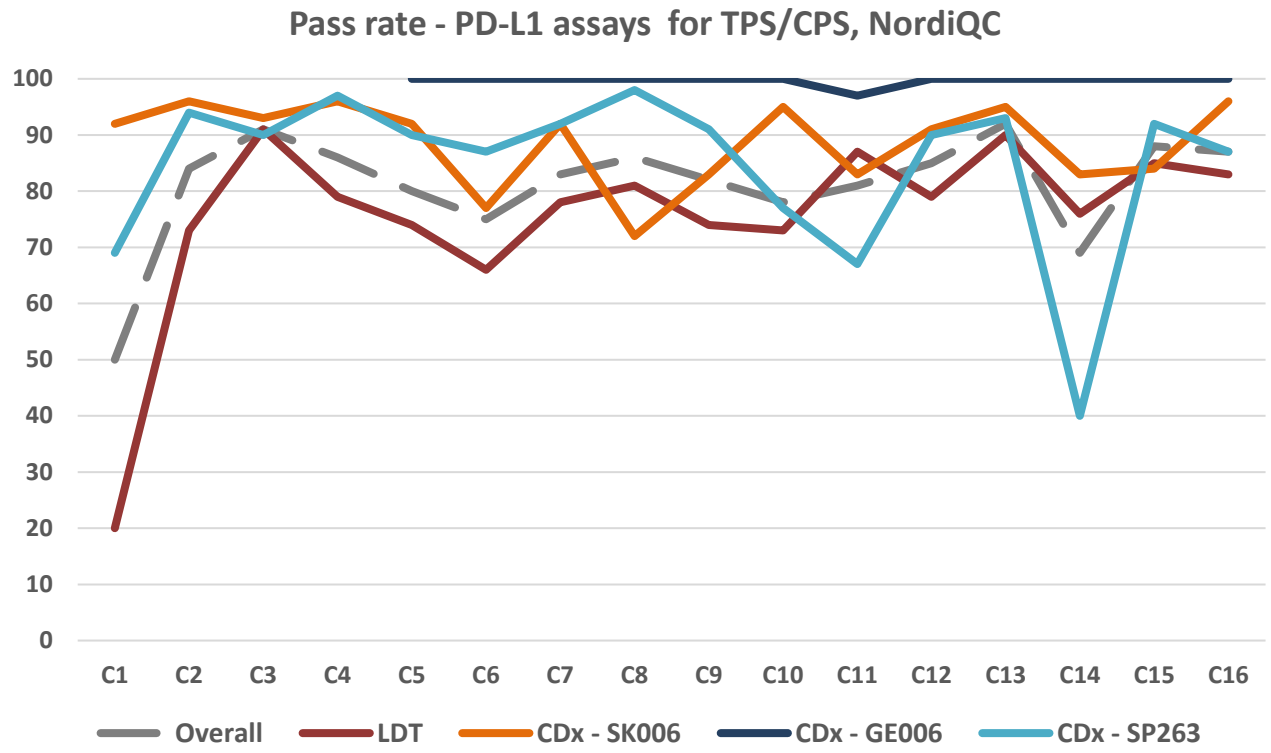
The Dako/Agilent pharmDx SK005 rmAb **28-8** for Autostainer Link 48 was used by 4 laboratories, 2 using the recommended protocol settings with 1 being assessed as good, and 1 as borderline. The other 2 were used as LDT's with 1 being assessed as optimal, and 1 as borderline.

Overall, 113 participants used one of the PD-L1 IHC CDx assays with one or more predictive claims for immune-oncology (22C3 SK006/GE006, SP263 741-4905/740-4907 and 28-8, SK005) by VRPS and a pass rate of 93% (105/113), was obtained.

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 58% (154 of 267) of the participants. For this group a pass rate of 82% was observed which is comparable to the level of 85% seen in the last assessment run – C15. Focusing on the performance of PD-L1 LD assays from C2-C16, 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), 89% for 22C3 SK006 pharmDx (Dako/Agilent) and 86% 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=57) providing a pass rate of 77% and an optimal rate of 46%, which is decreased compared to C15 (81%, 63% respectively), and a comparable pass rate to that of C14 (77%, 25% respectively) but with an increased optimal rate.

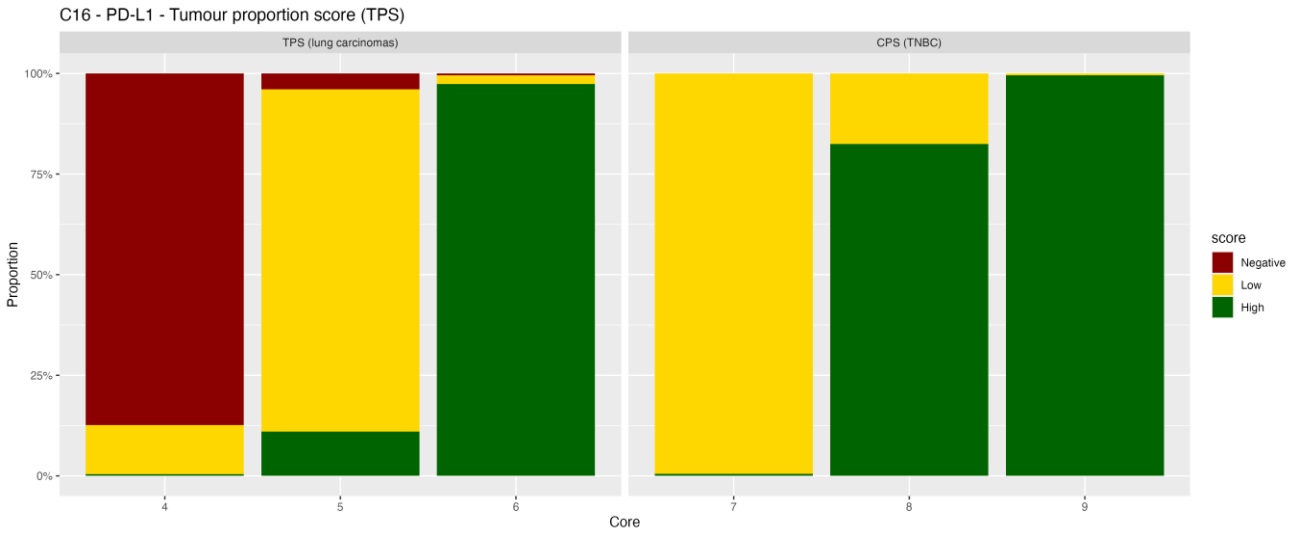
As described above for optimal protocol settings for mAb clone 22C3 as concentrated format, successful and interlaboratory reproducible settings have been identified for BenchMark (Ventana/Roche) and Omnis (Dako/Agilent) and these seem to be widely consolidated within the laboratories providing a pass rate largely comparable to most 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. This was not repeated in Run C14 or C15 as none of the participants achieved an optimal result. However, in C16 it was repeated with 1 participant achieving an optimal result. Cumulated data for runs C8 - C16 focusing on the performance of mAb clone 22C3 on the Bond platforms have shown a pass rate of 37% (13 of 35), with only 2 optimal results achieved to date. There is still only a small number of data observations generated so far and so conclusions are to be taken with caution.

The Leica Biosystems PD-L1 IHC RTU assay based on rmAb clone **73-10** (PA0832) with intended use on Bond III, was used by 6 participants in run C13, 3 participants in C14, 4 participants in C15 and 5 participants in C16 (with 2 participants achieving an optimal result). Overall, a pass rate of 100% was obtained when used by vendor recommended protocol settings.

The commonly used Ventana/Roche IHC RTU assay 790-4905, **SP263** without predictive claim showed an inferior performance compared to the corresponding locked assay 741-4905 giving a pass rate of 67%, 6% optimal when applied by vendor recommended protocol settings. When used by laboratory modified protocol settings a significantly improved pass rate of 93% and 26% optimal scores was observed. In the data analysis no single protocol changes explaining the improvement could be identified.

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 C16: 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. 6, 7 and, 9, whereas the consensus rate were lower in tissue cores no. 4, 5 and 8. In tissue core no. 4 intermingling macrophages within the tumor component of the NSCLC most likely were scored as PD-L1 positive changing the expected status from TPS negative to TPS low. As mentioned, tissue core 5 showed heterogenicity in some TMA's with TPS above 50% in some areas. This could most likely explain some of the disagreement in core 5. Concerning tissue core no. 8, 18% of the participants scored this a low, while the NordiQC assessor team only scored this as low in 2% of slides (all with insufficient score). We have no good explanation for this discrepancy.

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 (trophoblasts) 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.

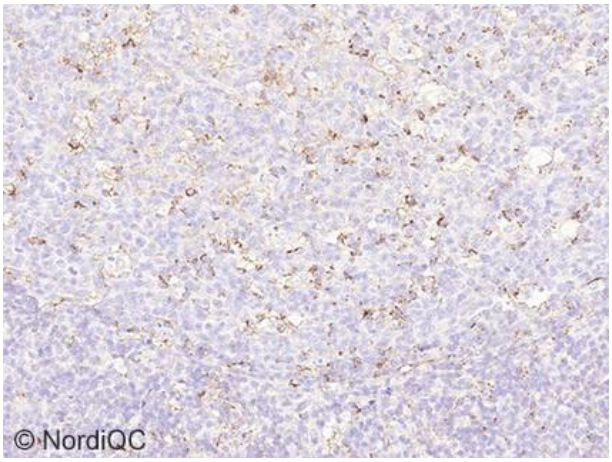


Fig. 1a
Optimal staining result of tonsil using the PD-L1 IHC 22C3 GE006 pharmDx kit, Dako/Agilent following the vendor recommended protocol settings.
A weak to moderate, but distinct punctuated membranous staining reaction of germinal centre macrophages and dispersed lymphocytes is seen.
No staining reaction is seen in the vast majority of lymphocytes.
Also compare with Figs. 2a – 4a, same protocol.

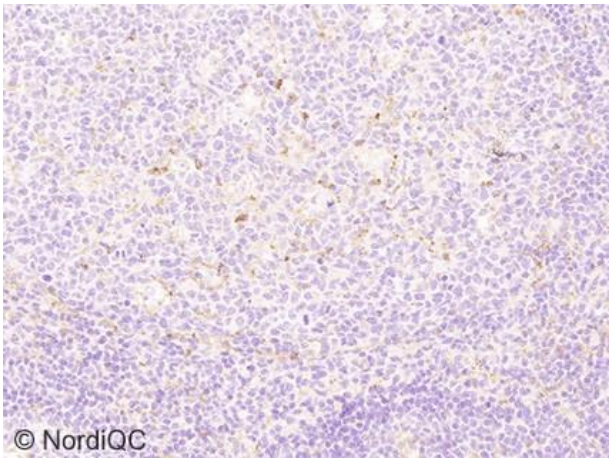


Fig. 1b
Staining result of tonsil, using the mAb clone 22C3 as concentrate within a laboratory developed assay performed on Bond III, Leica Biosystems.
Mainly dispersed T-cells are demonstrated showing a weak granular punctuated membranous staining reaction.
Virtually no staining reaction is seen the germinal centre macrophages indicating a too low level of analytical sensitivity.
Also compare with Figs. 2b - 4b, same protocol.

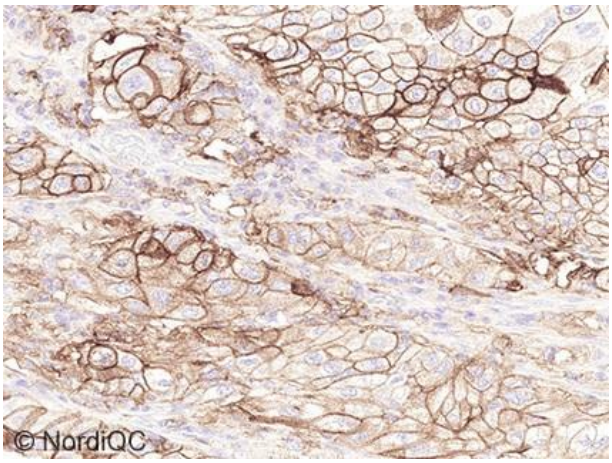


Fig. 2a
Optimal staining result of the NSCLC, tissue core no. 6, using the same protocol as in Fig. 1a.
A weak to strong, distinct membranous staining reaction is seen in virtually all tumour cells.
The tumour was categorized as TPS High (≥50%) and thus eligible for first line immune therapy with KEYTRUDA®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

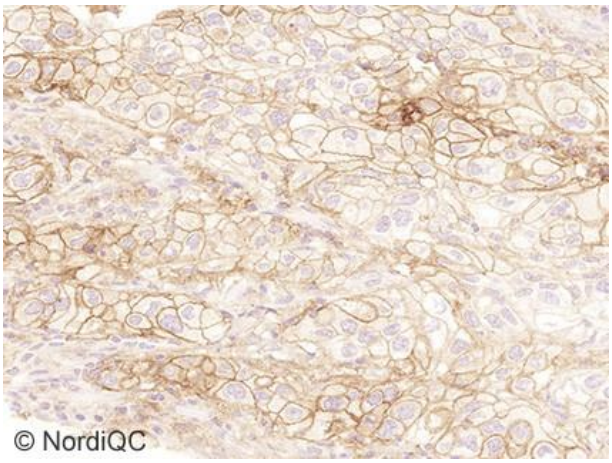


Fig. 2b
Staining result of the NSCLC, tissue core no. 6, using the same protocol as in Fig. 1b.
Overall a reduced intensity of the positive tumour cells is observed but the tumour still categorized as TPS High (≥50%) and thus eligible for first line immune therapy with KEYTRUDA® (different regional cut-offs occur).
However, also compare with Fig. 3b and 4b, same protocol.

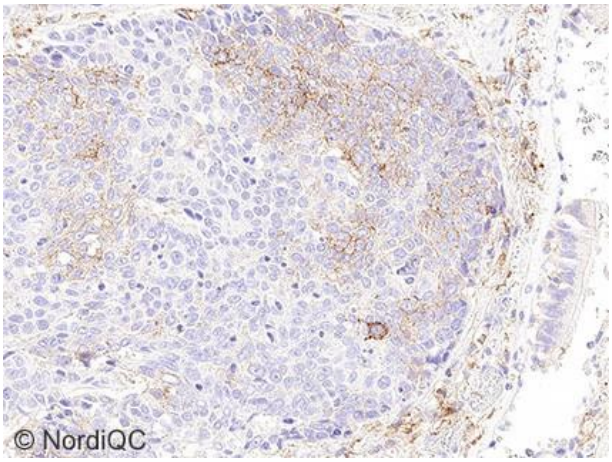


Fig. 3a
Optimal staining result of the NSCLC, tissue core no. 5, using the same protocol as in Figs. 1a and 2a.
In this area of the tumour a weak to moderate membranous staining reaction is seen in 20-25% of the neoplastic cells.
Overall, the tumour was categorized as TPS Low (≥1-49%) and thus eligible for second line immune therapy with KEYTRUDA® (different regional cut-offs occur).

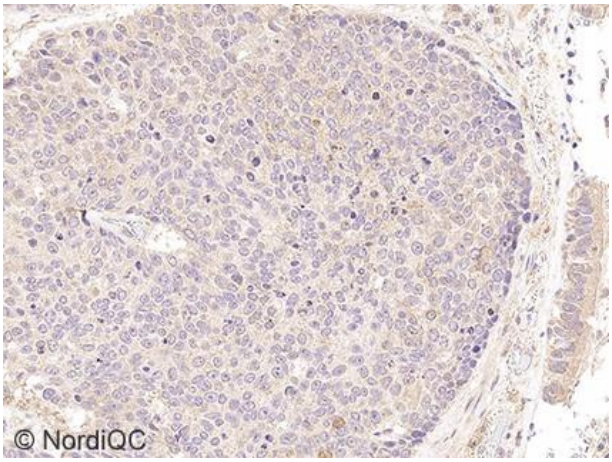


Fig. 3b
Insufficient staining result of the NSCLC, tissue core no. 5, using the same protocol as in Figs. 1b and 2b.
Less than 1% of the tumour cells show a membranous staining reaction and the PD-L1 category being changed from the expected TPS Low to TPS Negative and the tumour not being eligible for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

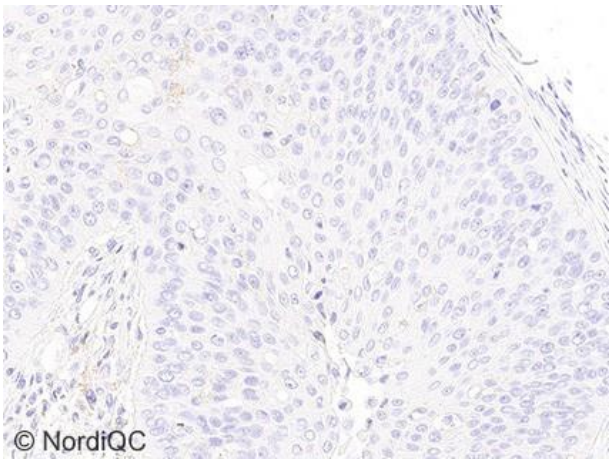


Fig. 4a
Optimal staining result of the NSCLC, tissue core no. 4, using the same protocol as in Figs. 1a - 3a.
Only intermingling immune cells in the NSCLC are demonstrated (verified by IHC for CD45).
The tumour was categorized as TPS Negative (<1%) and thus not eligible for immune therapy with KEYTRUDA®.

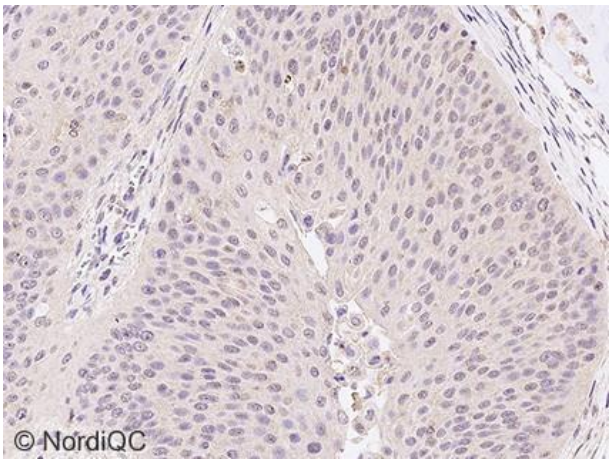


Fig. 4b
Insufficient staining result of the NSCLC, tissue core no. 4, using using same protocol as in Figs. 1b – 3b providing a poor signal-to-noise ratio hampering the read-out for PD-L1 status.
A diffuse cytoplasmic staining reaction is observed and the PD-L1 status cannot be settled with confidence.
Compare to the expected result as shown in Fig. 4a.

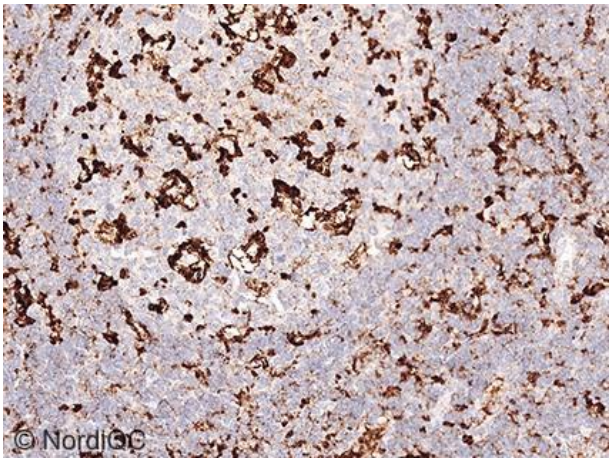


Fig. 5a
Insufficient staining result of tonsil using the Ventana/Roche SP263 assay, 741-4905 with OptiView + Amplification kit (tyramide amplification).
An excessive number of immune cells are demonstrated compared top the level expected (see Fig. 1a) and it is not possible to identify the specific membranous staining result due to an extensive cytoplasmic staining reaction. In addition, a diffuse granular staining reaction is seen comprising the evaluation of the quality of the result.

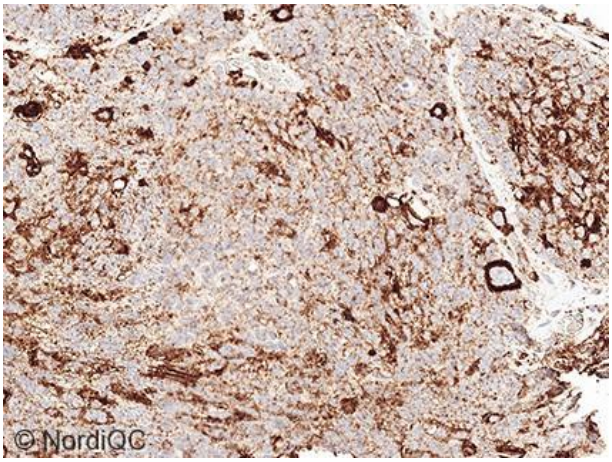


Fig. 5b
Insufficient staining result of the NSCLC, tissue core no. 5, using same insuffcient protocol as in Fig. 5a providing an extensive and granular staining reaction hampering the read-out for PD-L1 status.
A diffuse granular, membranous and cytoplasmic staining reaction is observed and PD-L1 status cannot be settled with confidence. The result was scored as TPS High (≥50%) by the participant. In this TMA block nearest NordiQC reference slide indicated a PD-L1 status as TPS Low (≥1-49%) with a TPS score level at 30%.

JH/LE/SN 29.11.2024