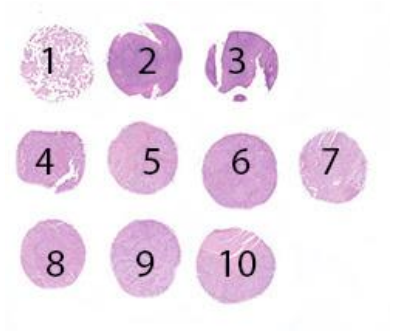
Assessment Run C10 2021
PD-L1 TPS/CPS

Purpose
This was the tenth assessment for PD-L1 in the NordiQC Companion module. This and the previous assessments 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 urothelial carcinomas to be treated with KEYTRUDA® as immune therapy. PD-L1 22C3 pharmDx, SK006 and GE006 (Dako/Agilent) and SP263 741-4905 (Ventana/Roche) were used as reference standard methods, and accuracy was evaluated in carcinomas with the dynamic and critical relevant expression levels of PD-L1 characterized by TPS and CPS. The obtained score in NordiQC is indicative of the performance of the IHC tests but due to the limited number and composition of samples additional internal validation and extended quality control, e.g. regularly measuring the PD-L1 results, is needed.

Material
Table 1. **Content of the TMA used for the NordiQC PD-L1 TPS/CPS (KEYTRUDA®) C10 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; 10-60%**
6. NSCLC	TPS: High; 90-100%
7. NSCLC	TPS: High; 90-100%
8. Urothelial carcinoma	CPS: <10
9. Urothelial carcinoma	CPS: ≥10; 30-40 IC***
10. Urothelial carcinoma	CPS: ≥10; 100 IC+TC***



* Tumour proportion score (TPS) and combined positive score (CPS) determined by PD-L1 IHC 22C3, SK006, GE006 (Dako/Agilent) and SP263 741-4905 (Ventana/Roche) performed in NordiQC reference lab.
** The tumour showed heterogeneity in the different levels within and in between the TMA's used and focally a TPS ≥50% was observed.
*** IC, Immune cells - TC; Tumour cells

All tissues were fixed in 10% neutral buffered formalin.

The participating laboratories were asked to perform the PD-L1 IHC assay for treatment with KEYTRUDA®, evaluate the PD-L1 expression level using the TPS and CPS scoring system and submit the stained slides and scores to NordiQC. This allowed assessment of the technical performance (analytical accuracy) of the PD-L1 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 throughout the blocks were characterized by the CE IVD / FDA approved 22C3 pharmDx kits SK006 (Dako/Agilent) for Autostainer Link 48, CE IVD approved 22C3 pharmDx kit GE006 (Dako/agilent) for Omnis, and also by the CE IVD approved assay (NSCLC, KEYTRUDA®) SP263 741-4905 (Ventana/Roche) for BenchMark in a NordiQC reference laboratory. During the assessment, TPS and CPS categories for each tissue core on the submitted slides were compared to the level in the nearest reference slides.

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 <u>Good</u> 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.
<u>TPS/CPS is still concordant to the NordiQC reference data obtained in all carcinomas</u>.

Criteria for assessing a staining as <u>Borderline</u> 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.
<u>TPS/CPS is **not** concordant to the NordiQC reference data in one of the carcinomas</u>

Criteria for assessing a staining as <u>Poor</u> 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.
<u>TPS/CPS is **not** concordant to the NordiQC reference data in two or more of the carcinomas.</u>

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 C10	240
Number of laboratories returning PD-L1 KEYTRUDA IHC slides	218 (91%)
Number of laboratories returning PD-L1 scoring sheet	209

Results: 218 laboratories participated in this assessment and returned slides. 78% 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 3). All slides returned after the assessment were assessed and laboratories received advice if the result was insufficient, but the data were not included in this report.

Performance history
This was the tenth NordiQC assessment of PD-L1 TPS/CPS (KEYTRUDA®). A slightly reduced pass rate has been observed in the latest two runs as shown in Graph. 1 below. The number of new participants seems to be consistently increasing about 3-5% in each run.

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

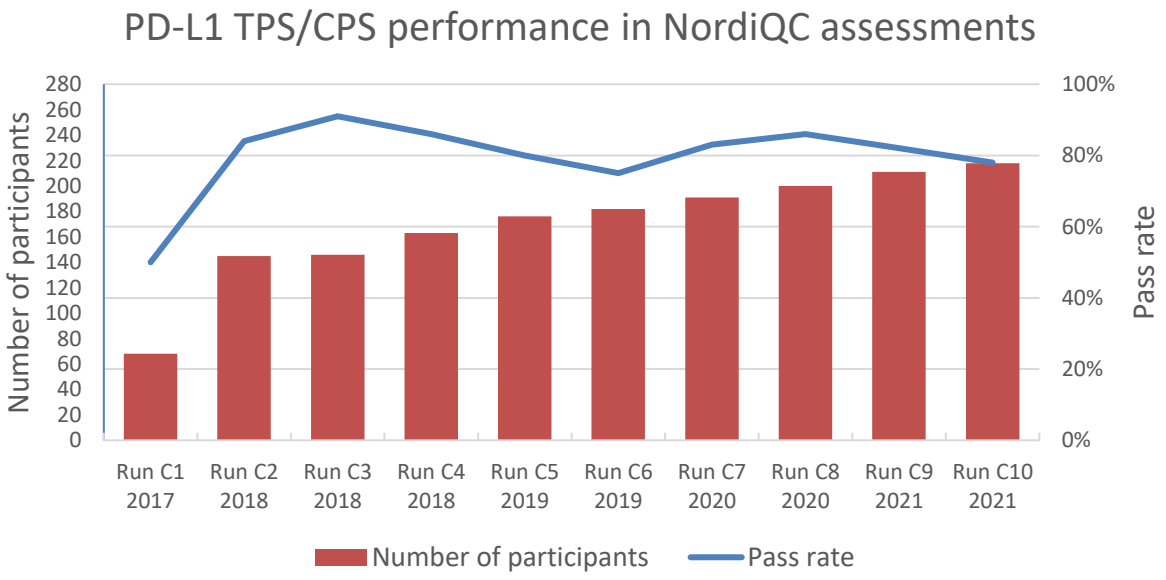


Table 2. **Assessment marks for IHC assays and antibodies run C10, PD-L1 TPS/CPS (KEYTRUDA®)**

CE-IVD / FDA approved PD-L1 assays	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
rmAb clone SP263, **741-4905 (VRPS)[3]**	44	Ventana/Roche	7	24	12	1	71%	16%
rmAb clone SP263, **741-4905 (LPMS)4**	1	Ventana/Roche	1	-	-	-	-	-
rmAb clone SP263, **740-4907 (VRPS)[3]**	12	Ventana/Roche	1	11	-	-	100%	8%
rmAb clone SP142, **741-4860 (VRPS)[3]**	2	Ventana/Roche	-	-	-	2	-	-
mAb clone 22C3 pharmDX, **SK006 (VRPS)[3]**	22	Dako/Agilent	19	2	1	-	95%	86%
mAb clone 22C3 pharmDX, **SK006 (LMPS)[4]**	9	Dako/Agilent	6	2	-	1	89%	67%
mAb clone 22C3 pharmDX, **GE006 (VRPS)[3]**	23	Dako/Agilent	23	-	-	-	100%	100%
mAb clone 22C3 pharmDX, **GE006 (LMPS)[4]**	12	Dako/Agilent	9	2	1	-	92%	75%
rmAb clone 28-8 pharmDX, **SK005 (VRPS)[3]**	2	Dako/Agilent	2	-	-	-	-	-
Antibodies[5] for laboratory developed PD-L1 assays, concentrated antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
mAb clone **22C3**	47	Dako/Agilent	21	13	7	6	72%	45%
mAb clone **E1L3N**	3	Cell Signaling	1	1	-	1	-	-
mAb clone **405-9A11**	1	Diagnostic Biosystems	-	-	1	-	-	-
rmAb clone **BP6099**	1	Biolynx	1	-	-	-	-	-
rmAb **CAL10**	3 1	Biocare Medical Zytomed Systems	2	1	-	1	-	-
rmAb clone **QR1**	2	Quartett	1	1	-	-	-	-
rmAb clone **SP142**	1	Abcam	-	-	1	-	-	-
rmAb clone **ZR3**	1	Gene Tech	-	1	-	-		
Ready-To-Use antibodies[6]	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
rmAb clone SP263, **790-4905[6] (VRPS)[3]**	18	Ventana/Roche	8	4	6	-	75%	44%
rmAb clone SP263, **790-4905[6] (LMPS)[4]**	7	Ventana/Roche	4	1	2	-	71%	57%
mAb **405-9A11 PDM572**	1	Diagnostic Biosystems	-	-	1	-	-	-
mAb **IHC441/431 IHC441-7**	1	GenomeMe	-	-	1	-	-	-
rmAb clone 73-10, **PA0832 (VRPS)[3]**	1	Leica Biosystems	-	1	-	-	-	-
rmAb clone **QR1 2-PR292-13**	2	Diagomics	-	-	1	1	-	-
rmAb clone **RM320**, **8263-C010**	1	Sakura Finetek	1	-	-	-	-	-
Total	218		108	64	34	13		
Proportion			49%	29%	16%	6%	78%	

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, 7 of 44 (16%) 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 95-100°C for 64 min., 16 min. incubation of primary Ab and OptiView as detection system. Using these protocols settings and applied on BenchMark platform, 31 of 44 (71%) laboratories produced a sufficient staining result (optimal or good).

SP263 (740-4907, Ventana/Roche): In total, 1 of 12 (8%) 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 95-100°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, 25 of 31 (81%) 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, 21 of 22 (95%) laboratories produced a sufficient staining result.
SK006 was frequently used by modified protocol settings e.g. mitigation to other platform as Ventana BenchMark or performed manually with overall slightly inferior performance as shown in Table 2.

PD-L1 IHC 22C3 pharmDx (GE006, Dako/Agilent): In total, 32 of 35 (91%) 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, 23 of 23 (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 IHC stainer device 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**	31/44 (71%)	7/44 (16%)	-	-
Ventana BenchMark Ultra rmAb SP263, **740-4907**	12/12 (100%)	8/13 (8%)	-	-
Dako Autostainer Link 48+ mAb 22C3 pharmDX, **SK006**	21/22 (95%)	19/22 (86%)	1/1	1/1
Dako Omnis mAb 22C3 pharmDX, **GE006**	23/23 (100%)	23/23 (100%)	4/4	4/4
Dako Autostainer Link 48+ rmAb 28-8 pharmDX, **SK005**	2/2	2/2	-	-

*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**: 21 of 47 (45%) protocols were assessed as optimal of which 14 were stained on the BenchMark Ultra platform (Ventana/Roche), five on the Omnis platform (Dako/Agilent) and two on Autostainer.
On BenchMark Ultra, the protocols providing optimal results were based on a titre of 1:20-50 of the primary Ab, incubation time of 40-120 min., HIER in CC1 (efficient heating time 48-64 min.) and OptiView as detection system. Using these protocol settings, 16 of 19 (84%) laboratories produced a sufficient staining result.
On Omnis, the protocols providing optimal results were based on a titre of 1:15-30 of the primary Ab, incubation time of 30-40 min., HIER in TRS low pH 6.1 at 97°C (efficient heating time 30-50 min.) and EnVision FLEX+ as detection system. Using these protocol settings, 7 of 7 (100%) laboratories produced a sufficient staining result.

On Autostainer (Dako/Agilent & Thermo) the protocols providing an optimal result were based on a titre of 1:50-55 of the primary Ab, incubation time of 30 min., HIER in TRS low pH 6.1 (Dako/Agilent) at 97°C (efficient heating time 20-30 min.) and EnVision FLEX+ (Dako/Agilent) as detection system. Using these protocol settings, 3 of 3 (100%) laboratories produced a sufficient staining result.

rmAb **CAL10**: 2 of 4 protocols (50%) were assessed as optimal.
The two protocols were based on HIER using an alkaline buffer Bond Epitope Retrieval Solution 2 (BERS2, Leica Biosystems) at 99°C for 20-30 min. The rmAb clone CAL10 was diluted 1:40-50, incubated for 15-20 min. at room temp. and visualized by Leica 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**	14/30** (47%)	-	0/1	1/3	-	7/7 (100%)	0/3	-

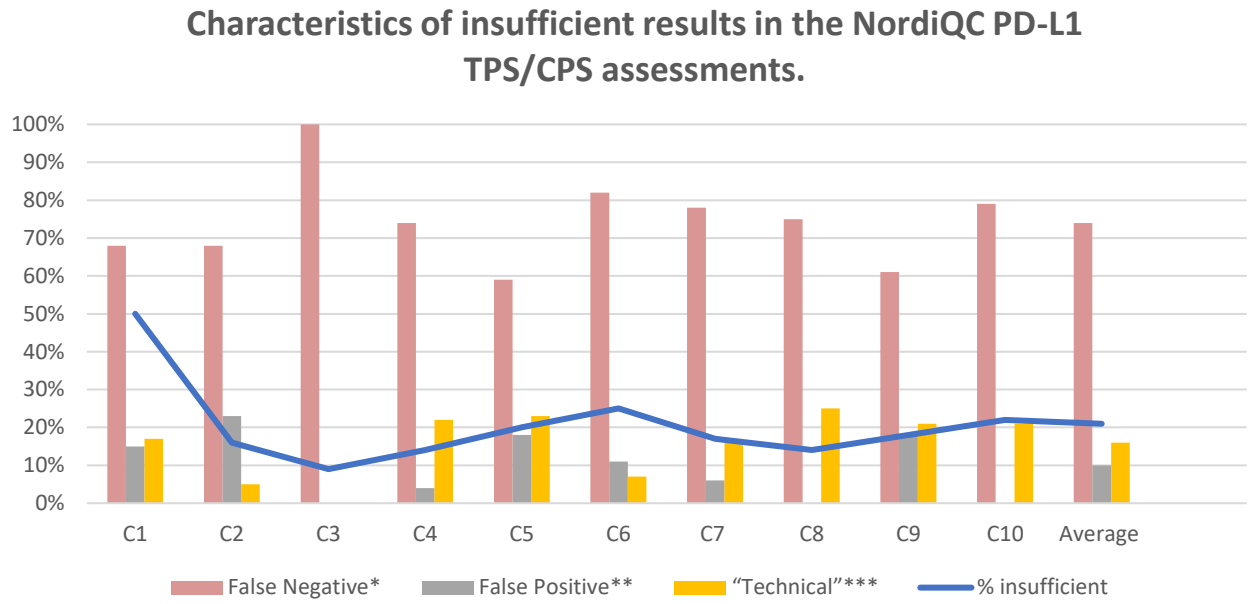
*Antibody concentration applied as listed above, HIER buffers and detection kits used as provided by the vendors of the respective platforms.
**number of optimal results/number of laboratories using this buffer.

Ready-To-Use antibodies for laboratory developed (LD) assays
rmAb **SP263** (790-4905, Ventana/Roche): In total, 12 of 25 (48%) protocols provided an optimal result. Protocols with optimal results were typically based on HIER in CC1 at 95-100°C, efficient heating time 52-64 min., 12-20 min. incubation of the primary Ab, OptiView as detection system and performed on BenchMark Ultra or XT. Using these protocols settings, 15 of 20 (75%) 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 4 NSCLCs, 3 urothelial carcinomas, 2 tonsils and 1 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 urothelial carcinomas were selected to comprise 1 carcinoma with CPS<10 and 2 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 assays 22C3 pharmDX SK006 (Dako/Agilent), 22C3 pharmDx GE006 (Dako/Agilent) and SP263 741-4905 (Ventana/Roche). 22C3 pharmDX SK006 (Dako/Agilent) was used to characterize PD-L1 for both TPS and CPS levels, whereas 22C3 pharmDx GE006 and SP263 for were mainly used to characterize TPS (reflecting the EU/FDA approved predictive claims for KEYTRUDA® at the assessment). In total, eight identical TMA blocks were constructed and six of these used for this assessment.
Reviewing the reference slides from the blocks, slight heterogenic expression of PD-L1 was seen in one of the tumor cores. In the NSCLC, tissue core no. 5, predominantly scored as TPS low (≥1-49%), focal areas with TPS high ≥50% were identified.
During the assessment, TPS and CPS categories for each tissue core on the submitted slides were compared to the level in the nearest reference slides.
Heterogeneity in PD-L1 expression is well known in NSCLCs and the assessment in this sense emulated clinical settings.

Comments
In this tenth 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 79% (37 of 47) 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%. In this run, the remaining 21% (10 of 47) 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. In contrast to the previous assessment, false positive results were not observed.

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 urothelial carcinomas only were affected in a few cases. PD-L1 IHC demonstration in the NordiQC assessments with combined tumour material has thus been more successful in urothelial carcinomas versus NSCLCs. No plausible reasons for this difference have been identified. The expression levels in the combined tumour materials used for the assessments in combination with different cut-off values and scoring methods might have favoured consistent PD-L1 demonstration in urothelial carcinomas compared to NSCLCs. 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 7 characterized as TPS Low and TPS 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 contrast, virtually all protocols provided the expected PD-L1 status in both the NSCLC, tissue core no. 6, characterized by NordiQC to show a strong membranous staining reaction in all tumour cells and the urothelial carcinomas, tissue cores no. 9 and 10 with CPS≥10 expressed in both immune cells and tumour cells. The NSCLC, tissue core no. 4, and the urothelial carcinoma, tissue core no. 8, were consistently negative by all protocols submitted.

29% (n=64) of the results submitted were marked as "Good". In 73% of these (47 of 64), 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 5% (3 of 64) 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 22% (14 of 64) 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 and was used by 8% of the participants in total.

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 26% of the participants and in total provided an overall pass rate of 77%, 14% 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 detection system. Despite

the locked protocol conditions for the two assays, some laboratories submitted protocols with reported modified settings typically indicating a change for HIER and/or incubation time of primary Ab. The different protocol settings submitted were disregarded for the two assays product no. 741-4905 and 740-4907 in this report and all protocols thus compiled as used by vendor recommended protocol settings as shown in Tables 2 and 3.
In contrast to previous assessments, the two Ventana/Roche PD-L1 assays based on the rmAb clone SP263 provided an inferior performance. This was unexpected, as the circulated proficiency material and reference slides were verified concerning the level for PD-L1 with both the Ventana/Roche 741-4905 SP263 assay and the two Dako/Agilent PD-L1 IHC assays based on 22C3, SK006 and GE006 for Autostainer and Omnis, respectively. The Ventana/Roche 741-4905 assay did in the reference slides show a slightly reduced PD-L1 expression in the two NSCLCs tissue cores no. 5 and 7 compared to the two 22C3 based assays, but still with a robust margin in the expected TPS categories and thus no indications to be as challenging as observed for the slides returned. No plausible reason for the general reduced analytical sensitivity and accuracy for the two SP263 IHC assays could be identified. In addition, it was also observed that the assay 740-4907 provided a pass rate of 100% compared to 71% for the assay 741-4905 – both with a low proportion of optimal results. Laboratories obtaining an insufficient score are recommended to continue to use the two PD-L1 assays with vendor guided protocol settings, as they historically in the NordiQC assessmens have generated high qualitative results, but also to perform in-house metrics of the PD-L1 results obtained to monitor and document these.

The Dako/Agilent 22C3 pharmDx assay GE006 for Omnis was used by 16% of the participants and in this assessment by far the most successful assay providing a pass rate of 100% and 100% 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 5 successive runs has shown a pass rate of 100% (84 of 84) for laboratories using GE006 by vendor recommended protocol settings. In comparison a pass rate of 84% (81 of 97) for laboratories using SK006 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 five successive runs performed.
In this context it has to be emphasized that the 22C3 GE006 assay for 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 for any indication with CPS as scoring system including urothelial carcinoma.

The Dako/Agilent 22C3 pharmDx assay SK006 for Autostainer Link 48 was used by 14% of the participants and provided a pass rate of 95% and 86% optimal when applied by protocol settings in compliance with vendor recommendations (see Table 3). The overall results for SK006 in this run significantly improved to the level seen in the previous runs and especially compared to run C8.

The Dako/Agilent pharmDx SK005 28-8 for Autostainer Link 48 was used by two laboratories. Both used the recommended protocol settings and both results being assessed as optimal.

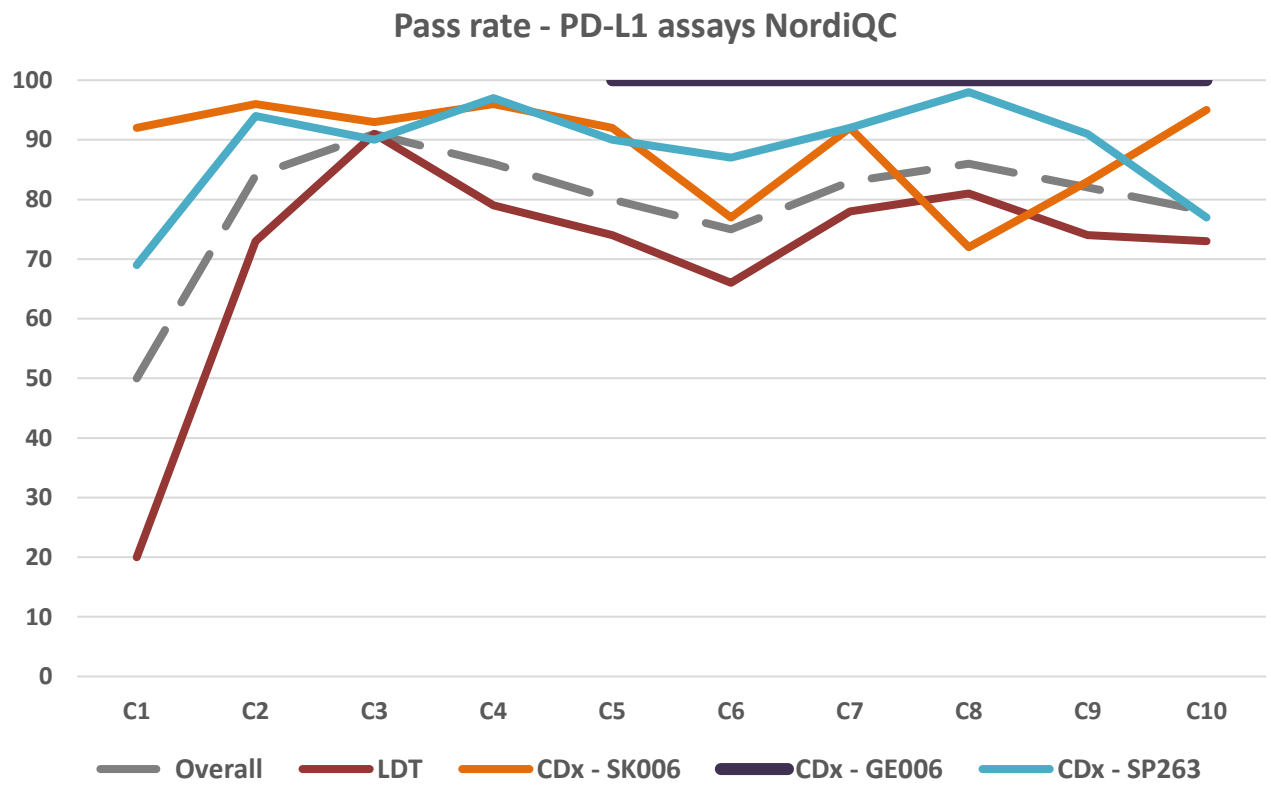
Grouped together, and using vendor recommended protocol settings, the five above mentioned CE IVD approved PD-L1 IHC assays with predictive claims - irrespective of indication and drug associated - provided a pass rate of 86% (89 of 103). However as described above, the Dako/Agilent 22C3 based PD-L1 IHC assays were more successful compared to the Ventana/Roche SP263 assays. Nevertheless, data observed in all previous runs indicate a possibility of interchangeability between the 22C3, 28-8 and SP263 based assays for PD-L1 status for KEYTRUDA® using the present cut-off values and scoring methods for TPS/CPS in the two indications addressed in this module. This must be validated by end-user according to local regulations.

The Ventana CDx assay based on SP142 was used by two participants and the result submitted assessed as insufficient and false negative changing the TPS status in the NSCLC tissue cores no. 5 and 7. This being concordant to observations and data for the SP142 CDx assay in previous NordiQC runs for PD-L1 TPS/CPS. Several publications inclusive Blueprint studies 1 and 2 (Hirsch, Tsao et al) have indicated poor analytical concordance for SP142 compared to the other CDx assays for TPS and hereby lack of interchangeability for SP142.

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 50% (113 of 216) of the participants. For this group a pass rate of 73% was observed.

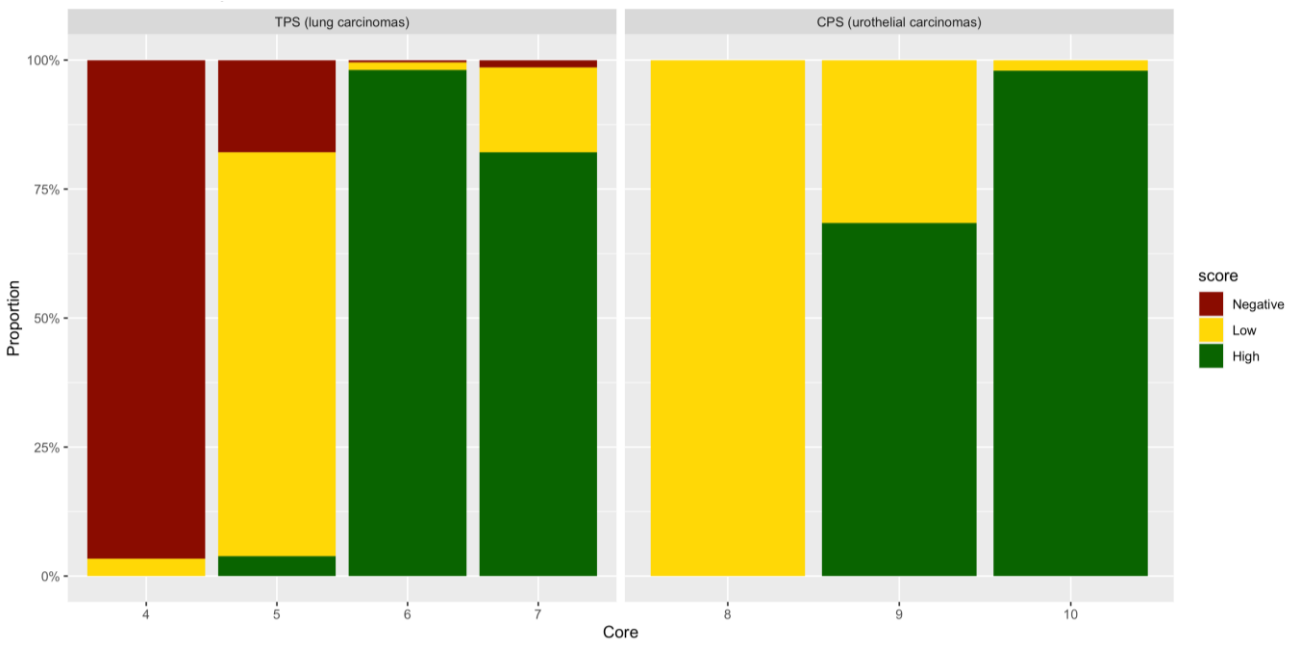
The performance of most commonly used IHC CDx and LD assays for PD-L1 is summarized and shown in Graph 3 below.

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



The mAb clone 22C3 was the most widely used concentrated Ab within a LD assay (n=47) providing a pass rate of 72%, 45% optimal.
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), whereas the performance on BOND III / BOND MAX (Leica Biosystems) has shown to be inferior. Cumulated data for run C8, C9 and C10 focusing on the performance of mAb clone 22C3 on the BOND platforms have shown a pass rate of 17% (2 of 12), 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 few data observations generated and conclusions to be taken with caution, but as mentioned same trend have now been observed in 3 successive runs.
For the BOND platform, the rmAb CAL10 as concentrated format by protocol settings as described above and the Ready-To-Use system from Leica Biosystems PA0832, based on rmAb 73-10, seem to be preferable.

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 urothelial carcinomas.



Graph 1. **NordiQC PD-L1 run C10: Tumour Proportion scores (TPS) in NSCLCs (core no. 4-7) and Combined Positive Score (CPS) in urothelial carcinomas (core no. 8-10).**

As seen in Graph 1, a relatively high consensus rates were observed for the tissue core 4, 6, 8 and 10, whereas the consensus rates were significantly lower in tissue core 9.

Controls
Tonsil and placenta were used as positive and negative tissue controls. In this and previous assessments, tonsil was found to be superior to placenta, as tonsil displayed a dynamic and clinical relevant range of PD-L1 expression levels, whereas placenta virtually only contained cells (throphoblasts) with high level PD-L1 expression.
In tonsil, protocols with optimal results typically provided the following reaction pattern:
A moderate to strong predominantly membranous staining reaction in dispersed crypt epithelial cells, a weak to moderate, typically punctuated membranous staining reaction of the majority of germinal centre macrophages and scattered interfollicular lymphocytes and macrophages. No staining reaction in the vast majority of lymphocytes and normal stratified squamous epithelial cells.
It was observed that rmAb SP263 (741-4905, 790-4905/4907, Ventana/Roche) typically provided 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. It was also observed that the two clones could provide a result in tonsil fully comparable to the two 22C3 PD-L1 IHC assays, but at the same time false negative results in carcinomas – see Figs. 1b–4b. 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.

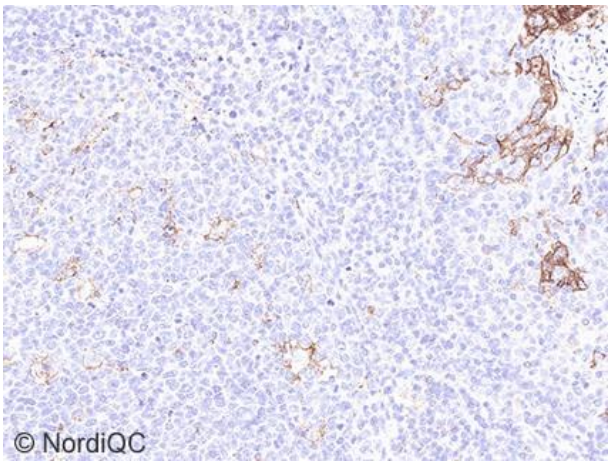


Fig. 1a
Optimal staining result of tonsil using the PD-L1 IHC 22C3 pharmDx kit, GE006, Dako/Agilent on Omnis following the vendor recommended protocol settings.
A weak to moderate, but distinct punctuated membranous staining reaction of germinal centre macrophages and dispersed lymphocytes is seen.
Crypt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

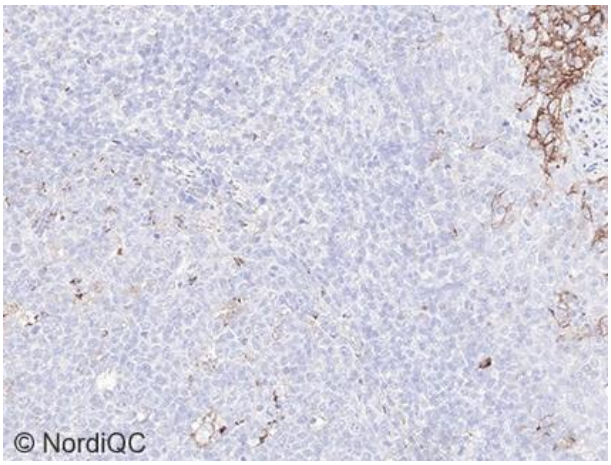


Fig. 1b
Staining result of tonsil, using the mAb clone E1L3N within a laboratory developed test for PD-L1.
The result fully comparable to the expected level as seen in Fig. 1a and obtained by the 22C3 pharmDx kit.
However still an insufficient result was seen in two of the three NSCLCs, as shown in Figs. 3b and 4.
The protocol as such provided an overall insufficient result characterized by a too low level of analytical sensitivity and indicate that a generic described reaction pattern for PD-L1 cannot be established for quality assurance and the expected pattern must be related for the clone / system used.

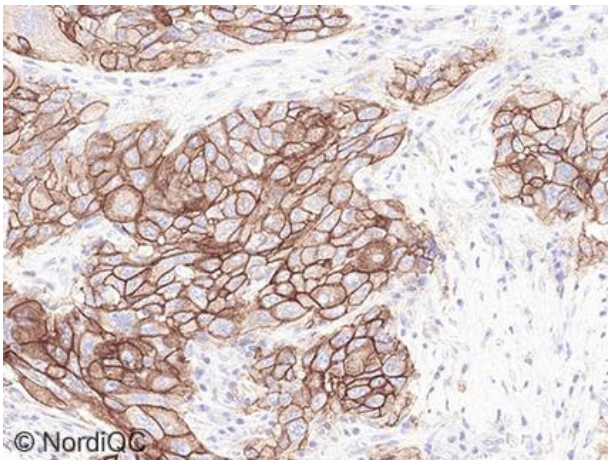


Fig. 2a
Optimal staining result of the NSCLC, tissue core no. 6, using the same protocol as in Fig. 1a.
Virtually all tumour cells show a moderate to strong membranous staining reaction.
The tumour was categorized as TPS high (≥50%) and thus eligible for first line immune therapy with KEYTRUDA® (different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

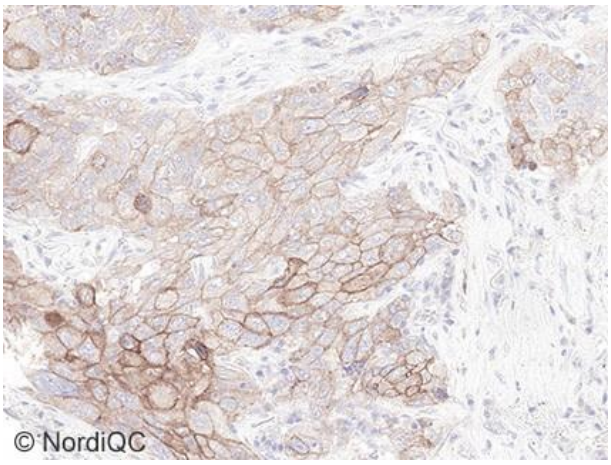


Fig. 2b
Staining result of the NSCLC, tissue core no. 6, using the same protocol as in Fig. 1b
The vast majority of tumour cells show a weak to moderate membranous staining reaction.
The tumour was despite a reduced intensity of the tumour cells still categorized as TPS high (≥50%) and thus eligible for first line immune therapy with KEYTRUDA® (different regional cut-offs occur).
However, also compare with Figs. 3b and 4b, same protocol.

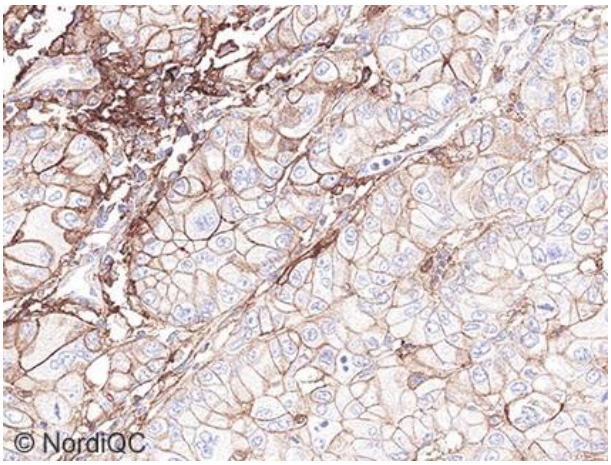


Fig. 3a
Optimal staining result of the NSCLC, tissue core no. 7, using the same protocol as in Figs. 1a and 2a.
Virtually all the tumour cells show a weak to moderate membranous staining reaction.
The tumour was categorized as TPS high (≥50%) and thus eligible for first line immune therapy with KEYTRUDA®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

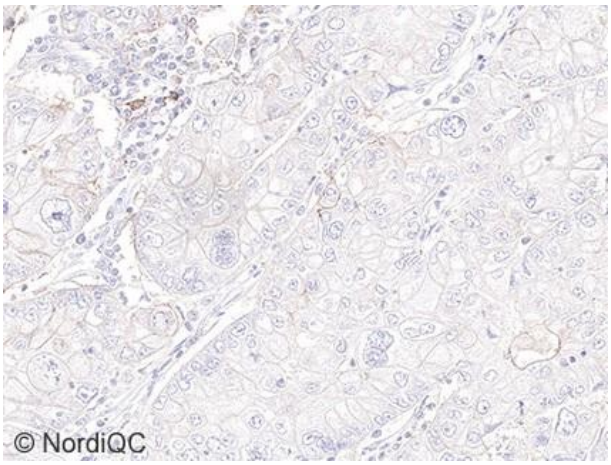


Fig. 3b
Insufficient staining result of the NSCLC, tissue core no. 7, using the same protocol as in Figs. 1b and 2b.
Only scattered tumour cells show a weak membranous staining reaction changing the TPS category from the expected high to low – same field as Fig. 3a.
The intensity of the membranes in this tumour was reduced compared to the NSCLC tissue core no. 6 (see Fig. 2a) and thus more "technically" challenging, but of diagnostic importance.

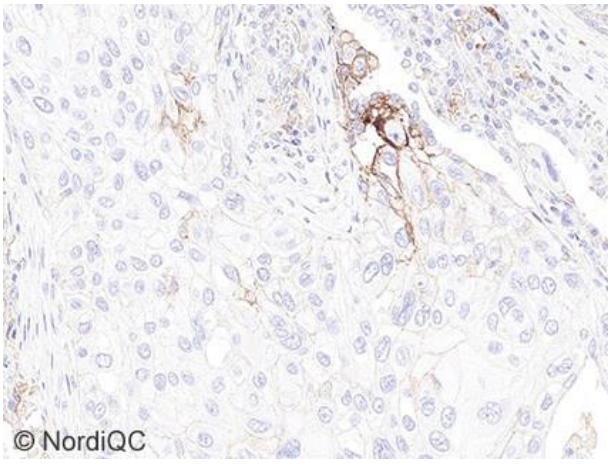


Fig. 4a
Optimal staining result of the NSCLC, tissue core no. 5, using the same protocol as in Figs. 1a - 3a.
A weak to moderate, but distinct staining reaction is seen in dispersed tumour cells.
The tumour was categorized as TPS low (≥1-49%) and thus eligible for second line immune therapy with KEYTRUDA® (different regional cut-offs occur).

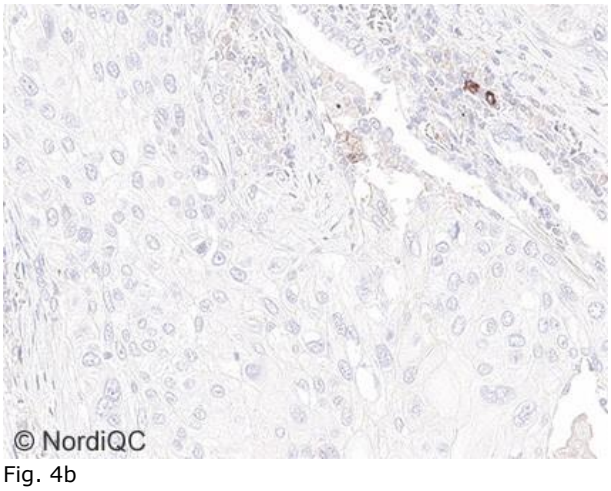


Fig. 4b
Insufficient staining result of the NSCLC, tissue core no. 5, using the same protocol as in Figs. 1b – 3b.
Virtually no staining reaction in the tumour cells is seen, whereas few immune cells are clearly demonstrated.
The PD-L1 category changed from the expected TPS low to negative and not being eligible for immune therapy.

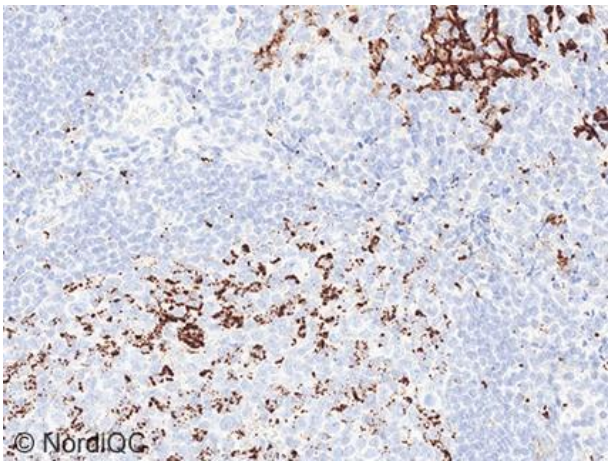


Fig. 5a
Staining result of the tonsil, using a protocol based on the mAb clone 22C3 as concentrated format by protocol settings based on OptiView + Amplification kit as detection system.
The crypt epithelial cells, germinal centre macrophages and most likely T-cells show a strong granular staining reaction.
The protocol provided an insufficient and false negative result in two of three NSCLCs as shown in Figs. 5b and 6b.

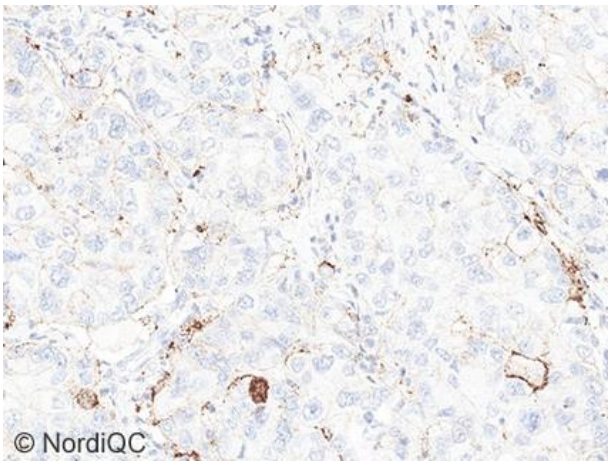


Fig. 5b
Insufficient staining result of the NSCLC, tissue core no. 7, using the same protocol as in Fig. 5a
A significant reduced proportion of neoplastic cells are demonstrated and only scattered tumour cells show a distinct membranous staining reaction changing the TPS category from the expected high to low –
Compare to the expected result as shown in Fig. 3a.

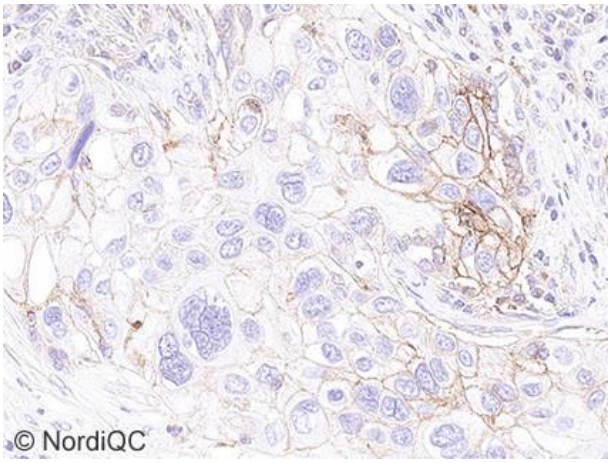


Fig. 6a
Optimal staining result in the NSCLC, tissue core no. 5, using the CDx assay SP263, 741-4905 Ventana/Roche on BenchMark Ultra following the vendor recommended protocol settings.
Approximately 20-30% of the tumour cells show a weak to moderate membranous staining reaction.
The tumour was categorized as TPS low (≥1-49%) and thus eligible for second line immune therapy with KEYTRUDA® (different regional cut-offs occur).

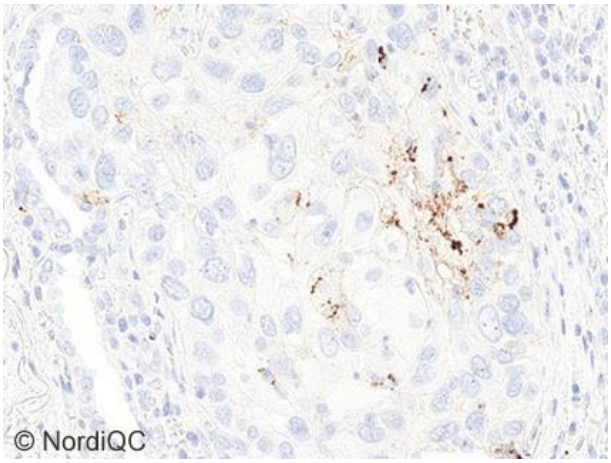


Fig. 6b
Insufficient staining result of the NSCLC, tissue core no. 5, using the same protocol as in Figs. 5a and 5b.
A granular and indistinct staining reaction for PD-L1 is seen and it is not possible with certainty to identify origin of positive cells being either tumour cells to be encountered or immune cells to be discarded.

SN/LE/HLK 08.12.2021