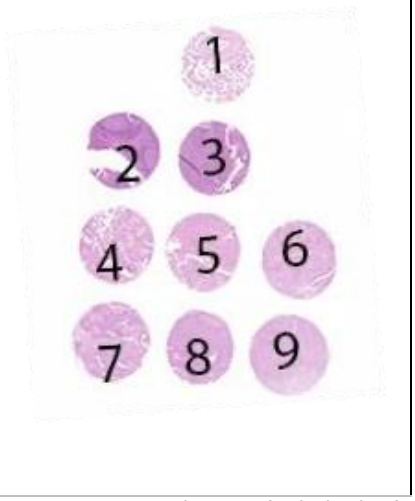
Assessment Run C15 2024
PD-L1 TPS/CPS

Purpose
This was the fifteenth 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 fifth 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®) C15 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; 20-45%**
6. NSCLC	TPS: High; 90-100%
7. TNBC	CPS: <10 IC# †
8. TNBC	CPS: ≥10; 20-30 IC#
9. TNBC	CPS: ≥10; 60-100 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 five of the seven TMA's used for the assessment, areas with TPS 55-80% were observed.
# IC, Immune cells - TC; Tumour cells.
† The CPS score varied in different levels within and in between the TMA's used. In two of the seven TMA's used for the assessment, areas with CPS 15-20% were observed.

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 91% compared to 77% for LD assays.
- The **22C3 CDx** assay GE006, Dako/Agilent was most successful with a pass rate of 100%, 97% optimal.
- Insufficient results were mainly caused by extensive cytoplasmic staining reaction, poor signal-to-noise ratio compromising the read-out.

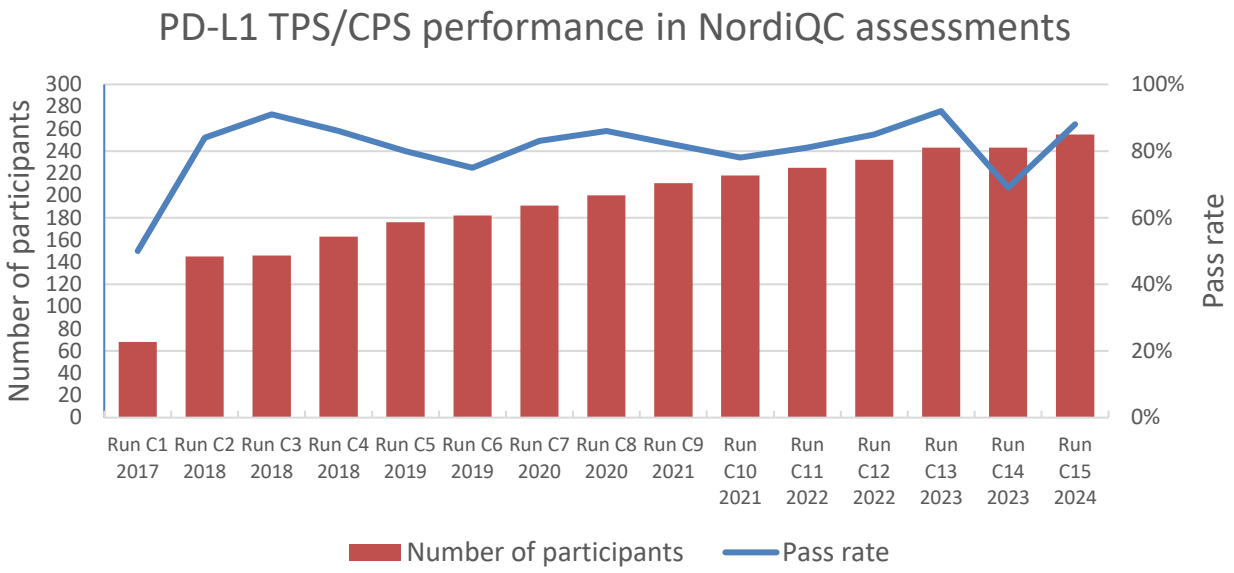
Participation

Number of laboratories registered for PD-L1 KEYTRUDA IHC C15	271
Number of laboratories returning PD-L1 KEYTRUDA IHC slides	255 (94%)
Number of laboratories returning PD-L1 scoring sheet	227

Results
255 laboratories participated in this assessment and returned slides. 88% 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
This was the fifteenth NordiQC assessment of PD-L1 TPS/CPS (KEYTRUDA®). Until C14 a relatively consistent pass rate had been observed with an upward trend seen. However, at C15 an increase in pass rate was observed similar to runs prior to C14. The latest runs results are included and shown in Graph 1 below. 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 run C15.

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



Conclusion

This was the fifteenth NordiQC assessment of PD-L1 for TPS/CPS status with focus on NSCLCs and TNBCs. 255 laboratories participated and a pass rate of 88% 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 97%, being superior to the other companion diagnostic assays and LD assays based on concentrated Abs. The 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 90% being much improved to the level of 40% obtained in C14.

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 is in contrast to the results obtained and described in previous NordiQC PD-L1 TPS/CPS assessments with the combination of NSCLCs, TNBC's and urothelial carcinomas where false negative staining results were most frequently characterized by a reduced proportion of PD-L1 positive cells compared to the level expected and defined by the NordiQC reference standard method resulting in false negative results.

Table 2a. **Overall results for PD-L1 TPS/CPS, run C15**

	n	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
CE-IVD / FDA approved PD-L1 assays*	143	97	33	9	4	91%	68%
Laboratory developed PD-L1 assays based on concentrated antibodies	66	37	14	12	3	77%	56%
PD-L1 assays based on Ready-To-Use antibodies without predictive claims	46	32	11	3	0	93%	69%
Total	255	166	58	24	7		
Proportion		65%	23%	9%	3%	88%	

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 C15**

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	16	19	1	3	90%	41%
rmAb clone SP263, **741-4905 (LPMS)[4]**	5	Ventana/Roche	3	1	1	-	80%	60%
rmAb clone SP263, **740-4907 (VRPS)[3]**	13	Ventana/Roche	9	4	-	-	100%	69%
mAb clone 22C3 pharmDX, **SK006 (VRPS)[3]**	25	Dako/Agilent	17	4	3	1	84%	68%
mAb clone 22C3 pharmDX, **SK006 (LMPS)[4]**	8	Dako/Agilent	6	1	1	-	88%	75%
mAb clone 22C3 pharmDX, **GE006 (VRPS)[3]**	39	Dako/Agilent	38	1	-	-	100%	97%
mAb clone 22C3 pharmDX, **GE006 (LMPS)[4]**	11	Dako/Agilent	8	2	1	-	91%	73%
rmAb clone 28-8 pharmDX, **SK005 (VRPS)[3]**	3	Dako/Agilent	-	1	2	-	-	-
Total	143		97	33	9	4		
Proportion			68%	23%	6%	3%	91%	

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 C15**

Antibodies[5] for laboratory developed PD-L1 assays, concentrated antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
mAb clone **22C3**	57	Dako/Agilent	36	10	8	3	81%	63%
rmAb **CAL10**	1 2	Zytomed Systems Biocare Medical	1	1	1	-	-	-
rmAb clone **E1L3N**	2	Cell Signaling	-	-	2	-	-	-
rmAb clone **QR1**	2	Quartett	-	2	-	-	-	-
rmAb clone **28-8**	1	Dako/Agilent	-	-	1	-	-	-
rmAb clone **711R**	1	MSVA	-	1	-	-	-	-
Total	66		37	14	12	3		
Proportion			56%	21%	18%	5%	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 C15**

Ready-To-Use antibodies[6]	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
rmAb clone SP263, **790-4905[6] (VRPS)[3]**	14	Ventana/Roche	12	2	-	-	100%	86%
rmAb clone SP263, **790-4905[6] (LMPS)[4]**	21	Ventana/Roche	15	5	1	-	95%	71%
rmAb clone SP142, **790-4860 (LMPS)[4]**	1	Ventana/Roche	1	-	-	-	-	-
rmAb clone 73-10 **PA0832**	4	Leica Biosystems	1	3	-	-	-	-
rmAb MX070C **MAB-0854**	1	Fuzhou Maixin	1	-	-	-	-	-
rmAb clone AC37 **PA168**	1	Abcarta	1	-	-	-	-	-
rmAb clone BP6099 **I12052E**	1	Biolynx	-	1	-	-	-	-
rmAb clone E1L3N **P05B01**	1	MEDx Translational Medicine	-	-	1	-	-	-
rmAb GR110 **GT256202**	1	Gene Tech	-	-	1	-		
rmAb clone RM320 **8263-C010**	1	Sakura Finetek	1	-	-	-		
Total	46		32	11	3	0		
Proportion			69%	24%	7%	0%	93%	

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, 16 of 39 (41%) 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, 35 of 39 (90%) laboratories produced a sufficient staining result (optimal or good).
SP263, 741-4905 was also applied on other non-intended platforms as Leica Biosystems Bond and BenchMark Ultra plus with an overall performance as shown in Table 2b (LMPS).

SP263 (740-4907, Ventana/Roche): In total, 9 of 13 (69%) protocols were assessed as optimal. This product has a locked protocol on the 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, 13 of 13 (100%) laboratories produced a sufficient staining result.

PD-L1 IHC 22C3 pharmDx (SK006, Dako/Agilent): In total, 17 of 25 (68%) 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 25 (84%) 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, 38 of 39 (97%) 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, 39 of 39 (100%) laboratories produced a sufficient staining result.

Table 3 summarizes the proportion of sufficient and optimal marks for the most commonly used CE IVD / FDA approved assays. The performance was evaluated both as "true" plug-and-play systems performed strictly accordingly to the vendor recommendations and by laboratory modified systems changing basal protocol settings. Only protocols performed on the specific automated IHC platform are included.

Table 3. **Comparison of pass rates for vendor recommended and laboratory modified protocols**

CDx assay*	**Vendor recommended protocol settings***		**Laboratory modified protocol settings****	
	Sufficient	Optimal	Sufficient	Optimal
Ventana BenchMark XT, GX, Ultra rmAb SP263, **741-4905**	35/39 (90%)	16/39 (41%)	-	-
Ventana BenchMark Ultra rmAb SP263, **740-4907**	13/13 (100%)	9/13 (69%)	-	-
Dako Autostainer Link 48+ mAb 22C3 pharmDX, **SK006**	21/25 (84%)	17/25 (68%)	7/8	6/8
Dako Omnis mAb 22C3 pharmDX, **GE006**	39/39 (100%)	38/39 (97%)	10/11 (91%)	8/11 (73%)
Dako Autostainer Link 48+ rmAb 28-8 pharmDX, **SK005**	1/3	0/3	-	-

*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, 36 of 57 (63%) protocols were assessed as optimal of which 36 were stained on the BenchMark Ultra platform (Ventana/Roche), 1 on the BenchMark Ultra Plus platform (Ventana/Roche), 1 on BenchMark XT platform (Ventana/Roche), 10 on the Omnis platform (Dako/Agilent), 3 on Autostainer Link 48 (Dako/Agilent), 5 on Bond III platform (Leica Biosystems), 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 32-120 min., HIER in CC1 for 32-95 min. and OptiView as the detection system. Using these protocol settings, 26 of 36 (72%) laboratories produced optimal staining results, and 31 of 36 (86%) 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 9 of 10 (90%) laboratories produced a sufficient staining result.

rmAb clone **CAL10**: Overall, 1 of 3 protocols were assessed as optimal.
The optimal protocol for this clone was based on HIER using EnVision™ FLEX Target Retrieval Solution (TRS) low pH 6.1 at 95-99°C for 20 min. in PT Link, a titre of 1:50 of the primary Ab, 20 min. incubation of the primary Ab, EnVision™ FLEX+ as the detection system and performed on Autostainer Link 48.

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**	
	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**	26/36** (72%)	-	-	2/3**	-	7/10**	0/5**	0/5**

*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

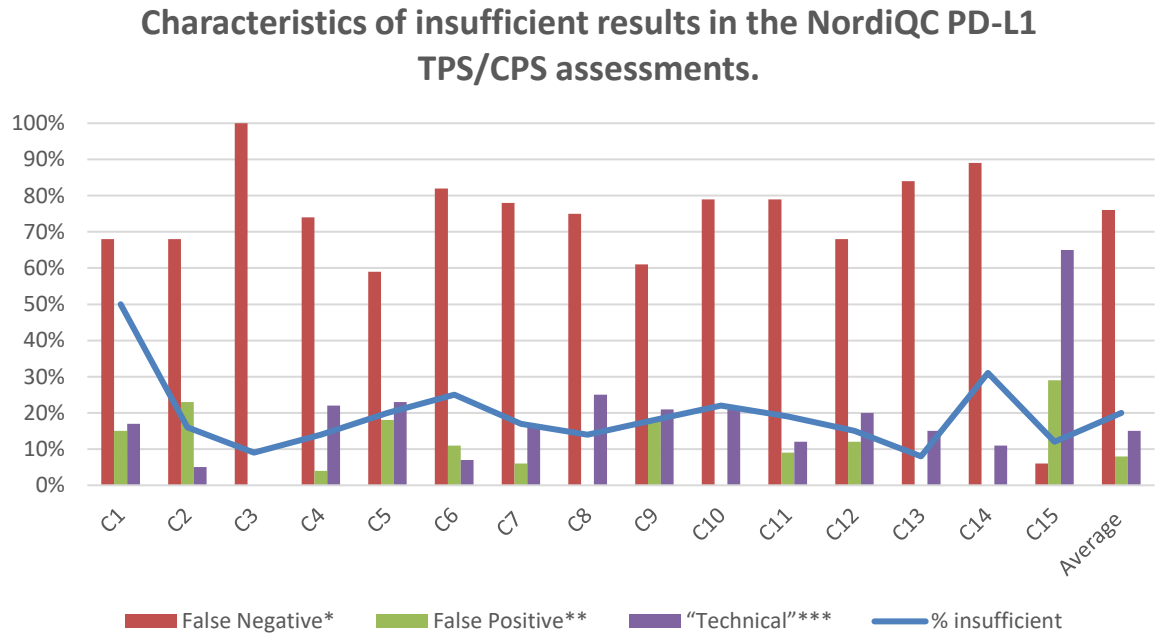
The tissue micro array (TMA) blocks constructed for this PD-L1 run consisted of three NSCLCs, three TNBCs, two tonsils and one placenta. The NSCLCs were selected to comprise PD-L1 expression levels for each TPS category: TPS negative (<1% PD-L1 positive tumour cells), TPS low (≥1-49%) and TPS high (≥50%). 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 (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 55-80% (TPS High) were observed and as such increased to the main level of 20-45% (TPS Low). Of particular importance for the TNBC, tissue core no. 7, areas with IC 15-20% (≥10) were observed and as such increased to the main level of <10. 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 fifteenth 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 65% of the insufficient results. 29% of the insufficient results were caused by a false positive staining result (which is also increased compared to previous runs) and 6% by a false negative 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.

23% (n=58) of the results submitted were marked as "Good". In 44% of these (26 of 58), 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 14% (8 of 58) 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 41% (24 of 58) 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 20% (52 of 255) of the participants and in total provided an overall pass rate of 92% (48 of 52), with 48% (25 of 52) 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 sufficient and optimal results for the SP263 IHC assay has increased in this assessment run compared to results especially seen in the previous run – C14. The increased pass rate observed in this run brings these assays back to similar levels as seen in earlier assessment runs C2-C9, C12 and C13 (see Graph 3). However, the proportion of optimal results still being 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% (39 of 255) of the participants providing a pass rate of 100% (97% 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 10% (25 of 255) of the participants and provided a pass rate of 84% (68% optimal) when applied by protocol settings in compliance with vendor recommendations (see Table 3). The 22C3 SK006 assay was also applied off-label (n=8), 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.

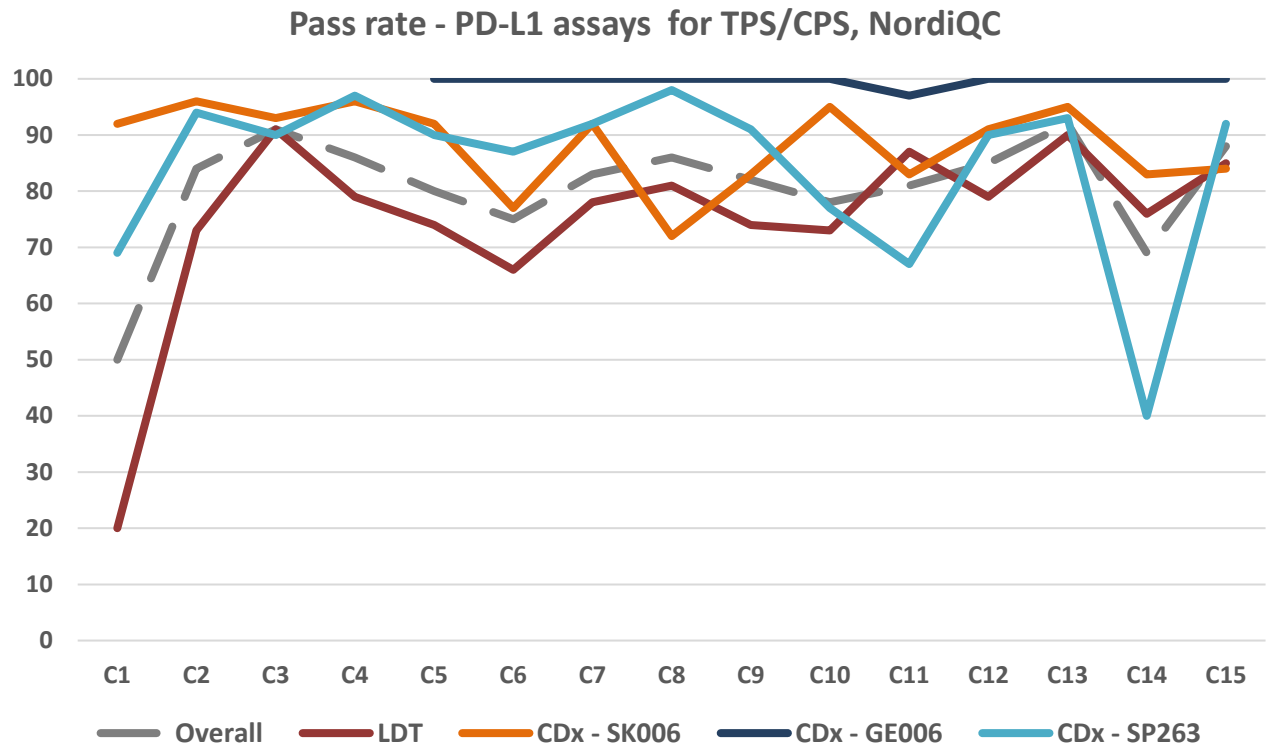
The Dako/Agilent pharmDx SK005 rmAb **28-8** for Autostainer Link 48 was used by 3 laboratories. All used the recommended protocol settings with 1 being assessed as sufficient, and 2 as borderline.

Overall 119 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) and when used by VRPS a pass rate of 92% (109/119), 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 53% (136 of 255) of the participants, which is comparable to 54% in C14 and 52% in runs C13 and C12. For this group a pass rate of 85% was observed which is an increase to the level of 76% seen in the last assessment run – C14. Focusing on the performance of PD-L1 LD assays from C2-C15, excluding the initial run C1 and start-up phase to identify "best practice LD assays", the mean pass rate for LD assays has been 79% (range 66%-91%) compared to e.g., 100% for the 22C3 GE006 pharmDx (Dako/Agilent), 88% for 22C3 SK006 pharmDx (Dako/Agilent) and 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 81% and an optimal rate of 63%, which is increased compared to C14 (77%, 25% respectively), and comparable to that of C13 (84%, 41%, respectively).

As described above for optimal protocol settings for mAb clone 22C3 as concentrated format, successful and interlaboratory reproducible settings have been identified for BenchMark (Ventana/Roche) and Omnis (Dako/Agilent) and these 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.

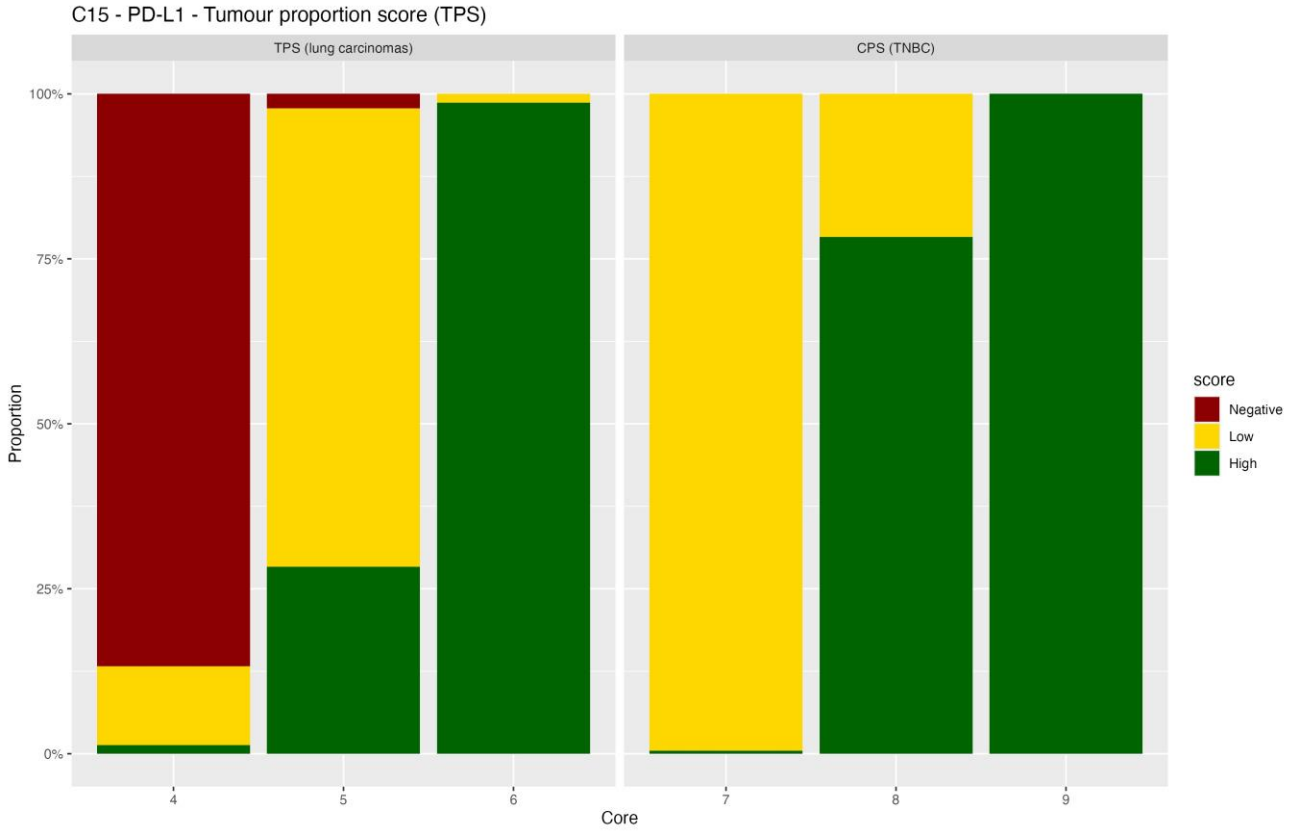
As mentioned in previous reports the performance of mAb clone **22C3** on Bond III / Bond MAX (Leica Biosystems) has shown to be inferior, however, in run C13 there was a 100% sufficient pass rate, with 1 participant achieving an optimal result. This was not repeated in Run C14 or C15 as non of the participants achieved an optimal result. Cumulated data for runs C8 - C15 focusing on the performance of mAb clone 22C3 on the Bond platforms have shown a pass rate of 40% (12 of 30), with only 1 optimal result achieved. There is still only a small number of data observations generated so far and so conclusions are to be taken with caution.

For unexplained reasons, the rmAb clone **E1L3N** (3/3), applied by protocol settings giving sufficient results in previous assessment runs provided an extensive false positive result with a TPS level at 80-90% in the NSCLC tissue core no 4, being TPS negative by all CDx assays applied by NordiQC and the participants. At present no explanation on the deviant result has been identified, but as many publications have stressed, despite overall a certain degree of interchangeability for PD-L1 IHC assays have been documented, some tumours will change PD-L1 status depending on IHC assay applied. In this assessment NordiQC used the results generated with the 22C3 CDx IHC assay GE006 Dako/Agilent as "ground truth" for the PD-L1 status, but also the CDx assays 22C3 SK006 Dako/Agilent and SP263, Ventana/Roche all characterized the NSCLC tissue core no. 4 as TPS negative and thus the positive result obtained by clone E1L3N assessed as insufficient – see Figs. 7a and 7b.

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 and 4 participants in C15 (with 1 participant 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 a slightly superior performance compared to the corresponding locked assay 741-4905 giving a pass rate of 100%, 86% optimal when applied by vendor recommended protocol settings. No data available to explain the discrepancy compared to the corresponding CDx assay 741-4905 (pass rate 90%, 41% optimal) as platforms, protocols and reagents applied being similar for the two assays.

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 C15: 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 being reduced 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. The spread of scores in the NSCLC tissue core no. 4 being reflected to the heterogeneity of this tumour ranging from TPS low in the majority of slides circulated but also with TPS high expression in a minor fraction of slides.

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.

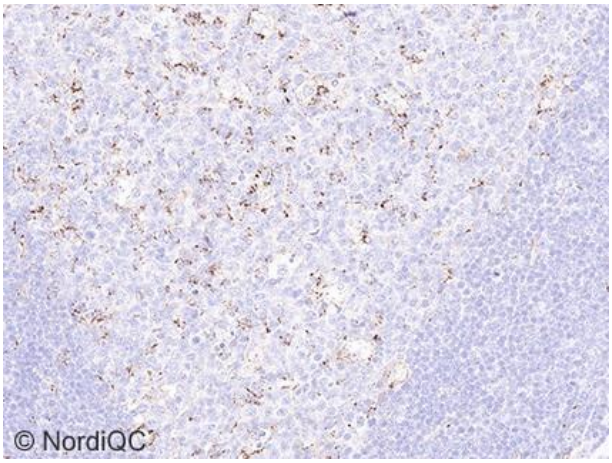


Fig. 1a
Optimal staining result of tonsil using the PD-L1 IHC 22C3 SK006 pharmDx kit, Dako/Agilent following the vendor recommended protocol settings.
A weak to moderate, but distinct punctuated membranous staining reaction of germinal centre macrophages and dispersed lymphocytes is seen.
No staining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

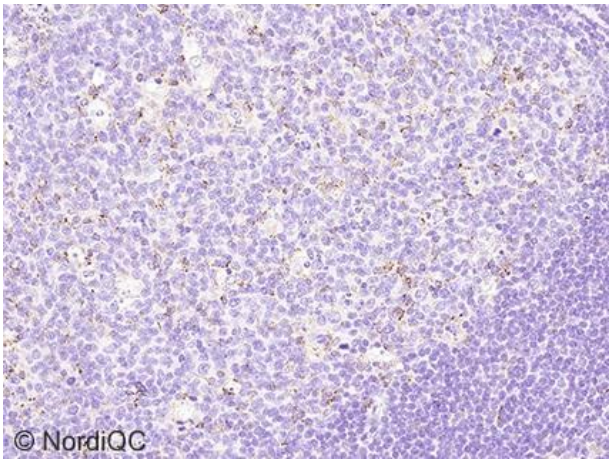


Fig. 1b
Staining result of tonsil, using the mAb clone 22C3 as concentrate within a laboratory developed assay performed on Bond III, Leica Biosystems.
Mainly dispersed T-cells are demonstrated showing a weak granular punctuated membranous staining reaction.
Virtually no staining reaction is seen the germinal centre macrophages indicating a too low level of analytical sensitivity – see also Figs. 2b and 3b.

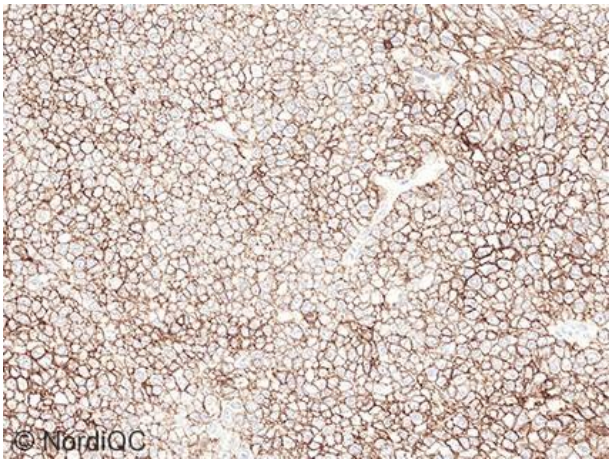


Fig. 2a
Optimal staining result of the NSCLC, tissue core no. 6, using the same protocol as in Fig. 1a.
A moderate to strong, distinct membranous staining reaction is seen in virtually all tumour cells.
The tumour was categorized as TPS High (≥50%) and thus eligible for first line immune therapy with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

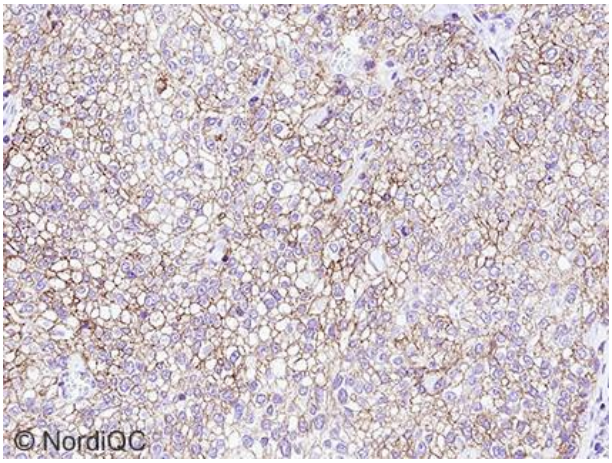


Fig. 2b
Staining result of the NSCLC, tissue core no. 6, using the same protocol as in Fig. 1b.
Overall a reduced intensity and proportion of positive tumour cells is observed but the tumour still categorized as TPS High (≥50%) and thus eligible for first line immune therapy with KEYTRUDA® (different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

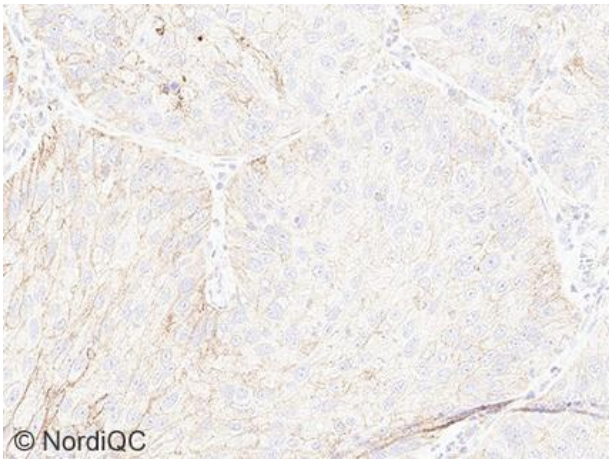


Fig. 3a
Optimal staining result of the NSCLC, tissue core no. 5, using the same protocol as in Figs. 1a and 2a.
In this area of the tumour a weak to moderate, but distinct staining reaction is seen in most tumour cells.
Overall, the tumour was categorized as TPS Low (≥1-49%) and thus eligible for second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

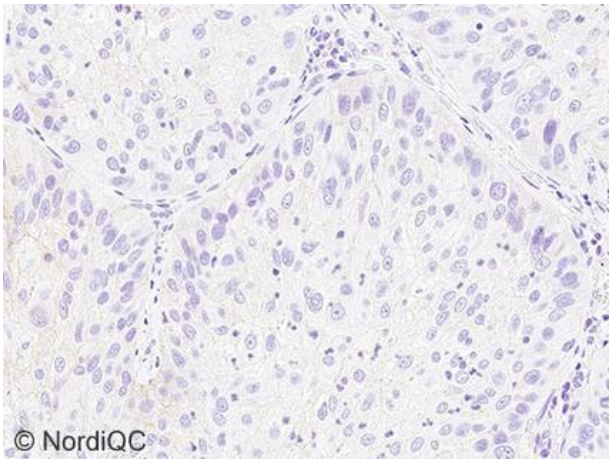


Fig. 3b
Insufficient staining result of the NSCLC, tissue core no. 5, using the same protocol as in Figs. 1b and 2b.
Less than 1% of the tumour cells show a membranous staining reaction and the PD-L1 category being changed from the expected TPS Low to TPS Negative and the tumour not being eligible for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

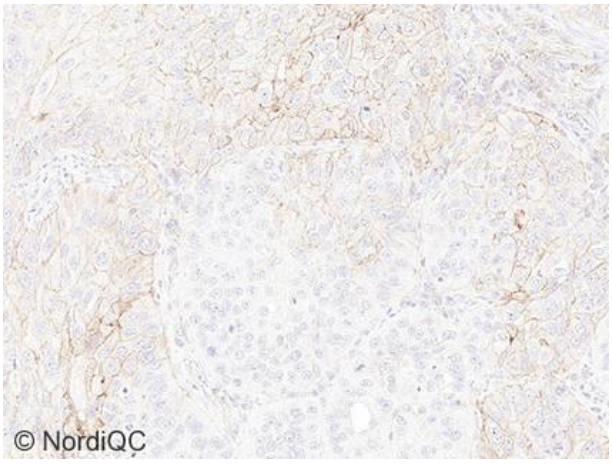


Fig. 4a
Optimal staining result of the NSCLC, tissue core no. 5, using the same protocol as in Figs. 1a - 3a.
A weak to moderate, but distinct staining reaction is seen in most tumour cells.
Overall, the tumour was categorized as TPS Low (≥1-49%) and thus eligible for second line immune therapy with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

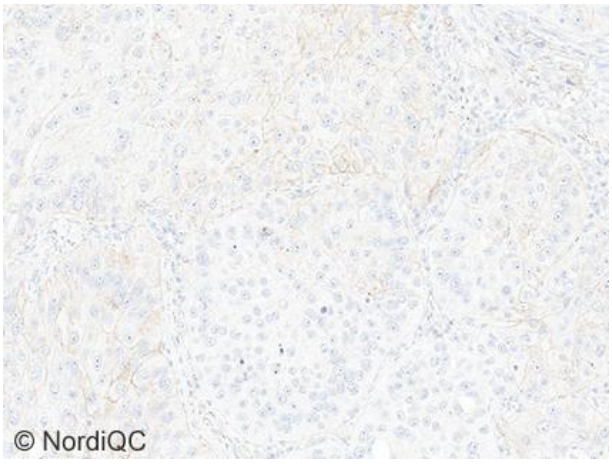


Fig. 4b
Staining result of the NSCLC, tissue core no. 5, assessed as good with a protocol providing a lowered analytical sensitivity compared to the level expected.
Overall a reduced intensity and proportion of tumour cells is observed but the tumour still to be categorized as TPS Low (≥1-49%) and thus eligible for second line immune therapy with KEYTRUDA® (different regional cut-offs occur).
Compare to the expected result as shown in Fig. 4a.

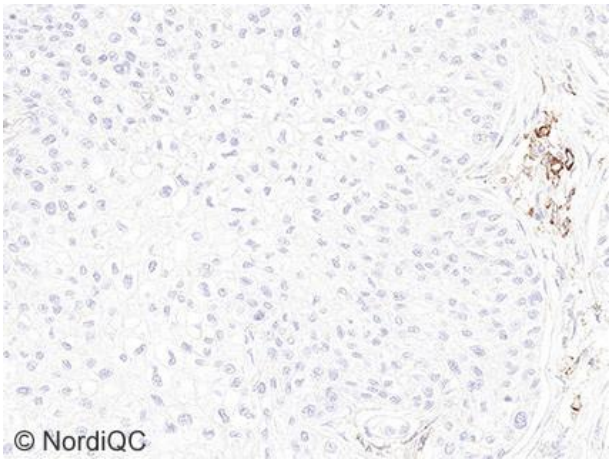


Fig. 5a
Optimal staining result of the NSCLC, tissue core no. 4, using the same protocol as in Figs. 1a – 4a.
A moderate, distinct membranous staining reaction is seen in immune cells while the neoplastic cells are negative.
The tumour was categorized as TPS Negative (<1%) and thus not eligible for immune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

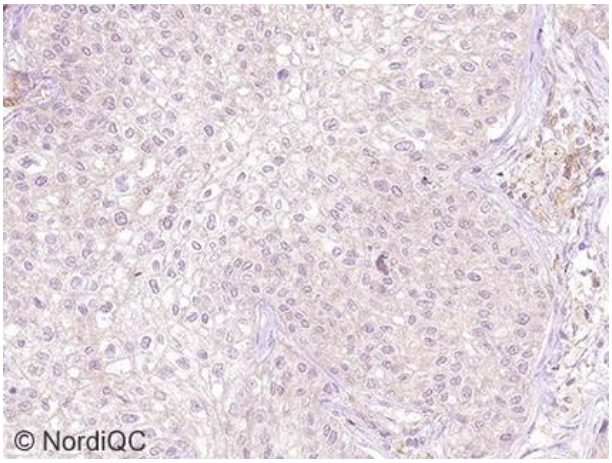


Fig. 5b
Insufficient staining result of the NSCLC, tissue core no. 4, using an inadeqaute calibrated protocol providing a poor signal-to-noise ratio hampering the read-out for PD-L1 status.
A diffuse cytoplasmic staining reaction with focal enhancement at the membranes is observed and PD-L1 status cannot be settled with confidence.

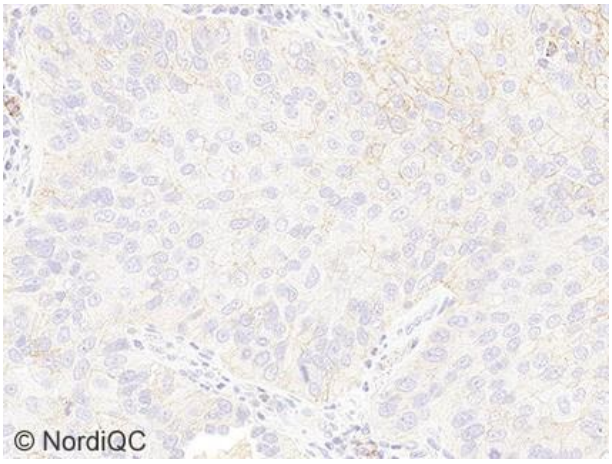


Fig. 6a
Optimal staining result of the NSCLC, tissue core no. 5, using the the 22C3 pharmDx, Dako/Agilent GE006 following the vendor recommended protocol settings.
A weak but distinct membranous staining reaction is seen focally in the neoplastic cells .
the tumour was categorized as TPS Low (≥1-49%) and thus eligible for second line immune therapy with KEYTRUDA® (different regional cut-offs occur).

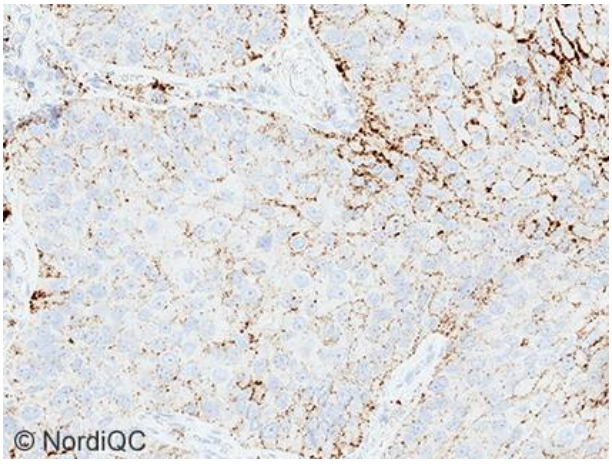


Fig. 6b
Insufficient staining result (borderline) of the NSCLC, tissue core no. 5.
Overall a significantly increased proportion of tumour cells is demonstrated changing the status from TPS Low to TPS High. In addition the the read-out is slightly compromised as the staining reaction being primarily coarsely granulated with difficulties to evaluate the PD-L1 TPS status.

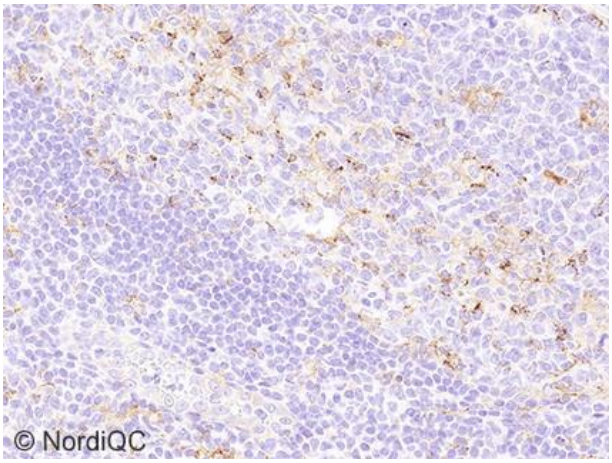


Fig. 7b
Staining result of tonsil, using the rmAb clone E1L3N as concentrate within a laboratory developed assay performed on Bond III, Leica Biosystems.
Overall, the expected result is obtained as dispersed T-cells are demonstrated showing a moderate granular punctuated membranous staining reaction. Also, a weak but distinct membranous staining reaction is seen in the germinal centre macrophages indicating an appropriate level of analytical sensitivity. Finally, the analytical specificity also seems appropriate as no staining reaction is seen in the background and most lymphocytes.
However, despite tonsil showed the expected result for PD-L1, the rmAb clone E1L3N (n=3 participants) gave an aberrant false positive staining result in the NSCLC, tissue core no. 4, characterized as TPS negative by all other CDx assays and LD assays based on "non-E1L3N" antibodies. See Fig. 7b.

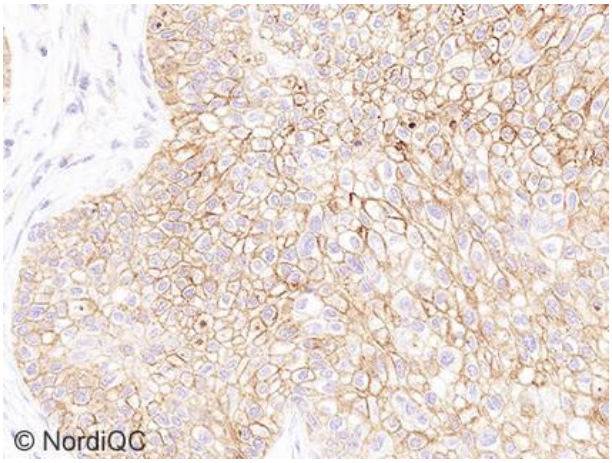


Fig. 7b
Insufficient false positive staining result of the NSCLC, tissue core no. 4, using the same protocol based on rmAb clone E1L3N as in Fig. 7a
A distinct membranous staining reaction is seen in virtually all neoplastic cells changing the expected status from TPS negative (not eligible for immune therapy) to TPS positive and eligible for first line immune therapy.
In addition to the extensive membrane reaction an focally an aberrant dot like accentuation is seen.
This staining pattern was observed for all three protocols based on rmAb clone E1L3N in this NSCLC. In all other tissue cores the overall results were as expected.
At present no explanation on the deviant result has been identified, but as many publications have stressed, despite overall a certain degree interchangeability of PD-L1 IHC assays have been documented, some tumours will change PD-L1 status depending on IHC assay applied.

JH/LE/HLK/SN 05.07.2024