Assessment Run C4 2018
PD-L1

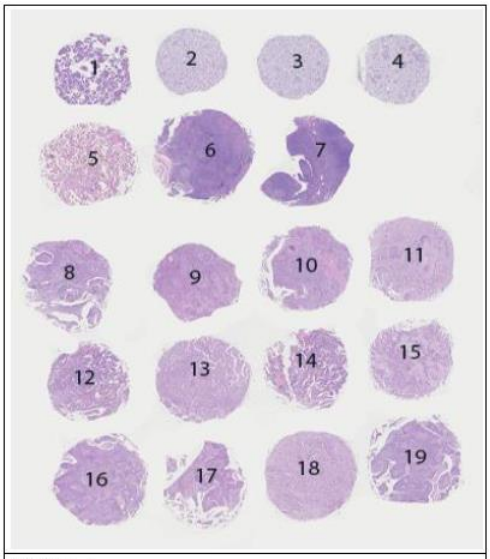
The fourth assessment in the NordiQC Companion module (C4) focused on the accuracy of the PD-L1 IHC assays performed by the participating laboratories to identify patients with non-small cell lung carcinoma (NSCLC) benefitting from immune therapy with Pembrolizumab [Keytruda®], based on the tumour proportion score (TPS) either as first line treatment (TPS ≥50%) or second line treatment (TPS ≥1%). The PD-L1 expression levels in the circulated material used for the assessment were characterized by the FDA approved companion IHC assays, 22C3 pharmDx, SK006 Dako/Agilent, and the complementary CE IVD approved assays 28-8 pharmDx, SK005 Dako/Agilent, and Ventana PD-L1 (SP263) assay, 790-4905, considered to be equivalent with the 22C3 assay[1]. Evaluation of the individual tissue cores and the associated cut-off values were used according to the interpretation guideline for the 22C3 pharmDx, indicating cut-off levels at 50% and 1%, respectively.

1) Tsao MS et al. PD-L1 Immunohistochemistry Comparability Study in Real-Life Clinical Samples: Results of Blueprint Phase 2 Project. J Thorac Oncol. 2018 Sep;13(9):1302-1311.

Material

Table 1. **Content of the TMA used for the NordiQC PD-L1 C4 assessment**

	PD-L1 IHC TPS score*
Cell line controls**	
1-4. Cell lines	
Tissue controls	
5. Placenta	
6-7. Tonsil	
NSCLC	
8. NSCLC	No <1%
9. NSCLC	No <1% / Low 1-49%***
10. NSCLC	No <1%
11. NSCLC	No <1% / Low 1-49%***
12. NSCLC	Low 1-49%
13. NSCLC	Low 1-49%
14. NSCLC	No <1% / Low 1-49%***
15. NSCLC	Low 1-49%
16. NSCLC	Low 1-49% / High ≥50%***
17. NSCLC	High ≥50%
18. NSCLC	High ≥50%
19. NSCLC	High ≥50%



PD-L1
1-4: Cell lines
5: Placenta
6-7: Tonsil
8-19: Non small lung cell carcinomas

* Tumour proportion score (TPS) determined by PD-L1 IHC 28-8, SK005 and 22C3, SK006 Dako performed in NordiQC reference lab.

** Cell lines, HistoCyte (1-4). The series included a cell line with a negative TPS, very low TPS, intermediate/low TPS and high TPS. The cell lines were not included in the assessment but will later be used for digital image analysis.

*** These tumor cores displayed tumor heterogeneity. Depending on the blocks from which sectioned slides were circulated, interpretation of the PD-L1 should be categorized into one of the percentage ranges (TPS) highlighted in the Table 1 (for further elaboration - see description below).

All tissues were fixed in 10% neutral buffered formalin.

The participating laboratories were asked to perform the PD-L1 IHC assay, interpret the PD-L1 expression level using the TPS scoring system and submit these scores to NordiQC. This allowed assessment of the technical performance (analytical accuracy) of the PD-L1 IHC assays and provided information on the reproducibility and concordance of the PD-L1 interpretation results among the laboratories.

PD-L1 IHC, Technical assessment
In order to account for heterogeneity of PD-L1 expression in the individual tumour cores included in the tissue TMA blocks, reference slides were made for every twentieth slide sent to the participants. These slides were stained for PD-L1 using the FDA approved 22C3 pharmDx kit (Dako) in a NordiQC reference laboratory. During the assessment, TPS categories for each tissue core on the submitted slides were compared to the nearest reference slide.

Criteria for assessing a staining as Optimal included:
The staining is considered perfect or close to perfect in all of the included tissues.
TPS (as estimated by NordiQC assessors based on reference staining) is concordant with the nearest NordiQC reference slide obtained in all 12 NSCLC cores.

Criteria for assessing a staining as Good included:
The staining is considered acceptable in all of the included tissues. However, the protocol may be optimized to ensure the best staining intensity, counter staining, morphology and signal-to-noise ratio.
TPS is still concordant with the nearest NordiQC reference slide in all 12 NSCLC cores.

Criteria for assessing a staining as Borderline included:
The staining is considered insufficient, e.g., because of a generally too weak staining, a false negative staining or a false positive staining reaction of one of the included tissues. The protocol should be optimized.
TPS is **not** found concordant with the nearest NordiQC reference slide in all 12 NSCLC cores.

Criteria for assessing a staining as Poor included:
The staining is considered insufficient due to a false negative or a false positive staining reaction staining of more than one of the included tissues. An optimization of the protocol is urgently needed.
TPS is **not** found concordant with the nearest NordiQC reference slide in all 12 NSCLC cores.

Participation

Number of laboratories registered for PD-L1 IHC C4	174
Number of laboratories returning PD-L1 IHC	163 (94%)
Number of laboratories returning PD-L1 scoring sheet	155 (89%)

Results: 164 laboratories participated in this assessment and 85% achieved a sufficient mark. Assessment marks for IHC PD-L1 assays and PD-L1 antibodies are summarized in Table 3 (see next page).

Table 2. **Proportion of sufficient results for PD-L1 in the four NordiQC runs performed**

	C1 2017	C2 2018	C3 2018	C4 2018
Participants, n=	68	145	146	163
Sufficient results	50%	84%	91%	86%

Performance history
This was the fourth NordiQC assessment of PD-L1. A reduced pass rate was obtained in C4 (see Table 2) compared to C3. However, run C3 was challenged by less than optimal circulated material, which may have provided an overall higher pass rate since no participants were downgraded based on the quality of the circulated slides. Despite this, the pass rate in C4 only decreased 5% to 86%.

Conclusion
In this fourth NordiQC run for PD-L1 in the companion module, C4, an overall pass rate of 85% was achieved. Insufficient PD-L1 IHC staining results were most frequently characterized by a reduced proportion of PD-L1 positive cells compared to the level expected as defined by the two PD-L1 IHC pharmDx assays, SK005 and SK006 (Dako/Agilent). This resulted in a too low TPS in one or more of NSCLC cores.
Several clones could be used to provide an optimal result: SP263, 22C3, 28-8, E1L3N, CAL10, ZR3, BSR90 and SP142. The companion diagnostic PD-L1 IHC assays from Dako/Agilent and Ventana/Roche provided a high proportion of sufficient results. Within LD-assays and no matter which Ab clone is used, meticulous calibration and validation of the assay is required.

Table 3. **Assessment marks for IHC assays and antibodies run C4, PD-L1 IHC**

CE-IVD / FDA approved PD-L1 assays	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
rmAb clone SP263, **790-4905**	50	Ventana/Roche	39	9	2	-	96%	95%
rmAb clone SP263, **790-4905**[3]	1	Ventana/Roche	1	-	-	-	-	-
rmAb clone SP263, **741-4905**	4	Ventana/Roche	4	-	-	-	-	-
rmAb clone SP263, **740-4907**	10	Ventana/Roche	9	1	-	-	100%	100%
mAb clone 22C3 pharmDX, **SK006**	28	Dako/Agilent	15	12	1	-	96%	96%
mAb clone 22C3 pharmDX, **SK006**[4]	9	Dako/Agilent	2	4	3	-	67%	-
rmAb clone 28-8 pharmDX, **SK005**	5	Dako/Agilent	2	2	-	1	80%	80%
rmAb clone 28-8 pharmDX, **SK005**[4]	1	Dako/Agilent	-	-	1	-	-	-
Antibodies[5] for laboratory developed PD-L1 assays, concentrated antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
mAb clone **22C3**	34	Dako/Agilent	11	16	6	1	79%	92%
mAb clone **E1L3N**	5	Cell Signaling	2	2	1	-	80%	100%
rmAb **CAL10**	1 3	Biocare Zytomed Systems	-	-	1	3	-	-
rmAb clone **28-8**	3	Abcam	3	-	-	-	-	-
rmAb clone **ZR3**	1 1 1	Cell Marque Zeta Corporation Nordic Biosite	2	1	-	-	-	-
rmAb clone **QR1**	1	Quartett	-	-	1	-	-	-
rmAb **BSR90**	1	Nordic Biosite	1	-	-	-	-	-
rmAb clone **SP142**	1	Spring Biosystems	1	-	-	-	-	-
Ready-To-Use antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
rmAb CAL10, **API3171**	1	Biocare	1	-	-	-	-	-
rmAb **QR1, 2-PR292-13**	1	Biocyc	-	-	-	1	-	-
rmAb clone **MXR003, RMA-0732**	1	Maixin	-	-	-	1	-	-
Total	163		93	47	16	7	-	-
Proportion			57%	29%	10%	4%	86%	-

1) Proportion of sufficient stains (optimal or good).
2) Proportion of sufficient stains with optimal protocol settings only, see below.
3) RTU system developed for the Ventana/Roche's automated systems (BenchMark) but used by laboratories on a different platform (Leica Bond).
4) RTU system developed for the Agilent/Dako's semi-automated systems (Autostainer Link48) but used by laboratories on different platforms (Ventana BenchMark and Dako Omnis).
5) mAb: mouse monoclonal antibody, rmAb: rabbit monoclonal antibody.

Detailed Analysis
CE IVD / FDA approved assays

SP263 (790-4905, Ventana): 39 of 50 (78%) assays were assessed as optimal. Protocols with optimal results were typically based on heat induced epitope retrieval (HIER) in Cell Conditioning 1 (CC1), efficient heating time 32-72 min. on BenchMark Ultra, 16-32 min. incubation of the primary Ab and OptiView as detection system. Using these protocol settings, 40 of 42 (95%) laboratories produced a sufficient staining result (optimal or good).

8 laboratories applied the same protocol settings as described above but also used an additional Amplification step. 7 of 8 (88%) produced a sufficient staining result.

SP263 (740-4907, Ventana): 9 of 10 (90%) assays were assessed as optimal. Protocols with optimal results were based on HIER in CC1, efficient heating time 32-64 min. on BenchMark Ultra, 16-32 min. incubation of the primary Ab and OptiView as detection system. Using these protocols settings, 9 of 9 (100%) laboratories produced a sufficient staining result. One laboratory used an additional Amplification step providing an optimal result.

PD-L1 IHC 22C3 pharmDx (SK006, Dako): 15 of 28 (54%) protocols were assessed as optimal. Protocols with optimal results were based on HIER using EnVision™ Flex Target Retrieval Solution (TRS) low pH 6.1 (SK006) at 95-99°C for 20 min. (PT Link), 30 min. incubation of the primary Ab and EnVision Flex+ as the detection system on Autostainer Link 48. Using these protocol settings, 27 of 28 (96%) laboratories produced a sufficient staining result.

PD-L1 IHC 28-8 pharmDx (SK005, Dako): 2 of 5 (40%) protocols were assessed as optimal. Protocols with optimal results were based on HIER using EnVision™ Flex TRS low pH 6.1 at 97°C for 20 min. (PT Link), 30 min. incubation of the primary Ab and EnVison Flex+ as the detection system on Autostainer Link 48. Using these protocol settings, 4 of 5 (80%) laboratories produced a sufficient staining result.

Table 4 summarizes the proportion of sufficient and optimal marks for the most commonly used RTU systems. 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 4. **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, **790-4905**	22/22 (100%)	17/22 (77%)	26/28 (93%)	22/28 (79%)
Ventana BenchMark XT, GX, Ultra rmAb SP263, **740-4907**	4/4	4/4	6/6 (100%)	5/6 (83%)
Dako Autostainer Link 48+ mAb 22C3 pharmDX, **SK006**	27/28 (96%)	15/28 (54%)	-	-
Dako Autostainer Link 48+ rmAb 28-8 pharmDX, **SK005**	4/5 (80%)	2/5 (40%)	-	-

*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**: 11 of 34 (32%) protocols were assessed as optimal of which six and five were stained on the Benchmark (Ventana) and Omnis (Dako) platforms, respectively.
On BenchMark GX/XT/Ultra, Ventana, the protocols providing optimal results were typically based on a titre of 1:30-40, primary Ab incubation time of 32-64 min., HIER in CC1 (efficient heating time 48-60 min.) and OptiView as detection system. Using these protocol settings, 6 of 7 (86%) laboratories produced a sufficient staining result. 7 laboratories used OptiView with Amplification and one of these obtained an optimal staining result.
On Omnis, Dako, the protocols providing optimal results were typically based on a titre of 1:20 of the primary Ab, incubation time of 30-45 min., HIER in TRS low pH 6.1 (Dako) at 97°C (efficient heating time 40 min.) and EnVision FLEX+ (10 min. in linker and 20-40 min. in polymer) as detection system. Using these protocol settings, 5 of 5 (100%) laboratories produced a sufficient staining result.

mAb **E1L3N**: Two protocols provided an optimal result. Both were based on HIER using an alkaline-buffer at 95-100°C for 20-30 min. The mAb clone E1L3N was diluted 1:100, incubated for 30 min. at room temp. and a 3-layer technique was used as detection system. Using these protocol settings, 2 of 2 (100%) laboratories produced a sufficient staining result.

Table 5. **Optimal results for PD-L1 for the most commonly used antibodies as concentrates on the 3 main IHC systems***

Concentrated antibodies	**Ventana/Roche BenchMark GX/XT/Ultra**		**Dako/Agilent Autostainer**		**Dako/Agilent Omnis**		**Leica 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**	5/16** (31%)	1/1	-	0/5	-	5/9	0/1	0/1
mAb clone **E1L3N**	0/1	-	-	0/1	-	-	2/2	0/1

*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

Block construction and assessment challenges
The tissue micro array (TMA) blocks constructed for this PD-L1 run consisted of 4 cell lines, 12 NSCLCs, 2 tonsils and 1 placenta. The NSCLCs were selected so the slides cut from the blocks would contain 4 NSCLCs in each of the TPS groups: TPS negative (<1% PD-L1 positive tumour cells), TPS low (≥1-49%) and TPS high (≥50%). Reference slides throughout the block were stained using the approved pharmDx kits (SK006 and SK005, Dako/Agilent) and SP263 (790-4905, Ventana) for the assessment. Additionally, reference slides for each 20 sent to the participants were stained using the pharmDx kit (SK006, Dako). In total 4 blocks was constructed but only slides from the first three blocks were sent to the participants.

Reviewing the reference slides from the blocks, heterogenic expression of PD-L1 were seen in four of the tumor cores. For tissue core no. 9 and 11 interpreted as TPS negative (<1% PD-L1 positive tumour cells in the primary clinical material), areas with more than 1% PD-L1 positive tumour cells were identified. In tissue core no. 14, initially scored as TPS low (1-49%) large negative areas was identified (mainly seen in slides cut from block number 2 and 3). Finally, in tissue core no. 16 (initially scored as High / ≥50%), a TPS of 1-49% could be displayed.
Heterogeneity in PD-L1 expression is well known in NSCLCs and the assessment emulated clinical settings in this way. However, the inconsistent expression of PD-L1 in one or more of the included tissue cores was challenging for the assessment settings, defining protocols providing a sufficient result from protocols giving an insufficient result.

Comments
In this fourth NordiQC assessment for PD-L1, the prevalent feature of an insufficient staining result was a too weak or false negative staining result, which were seen in 74% of the insufficient staining results (17 of 23). Only one (4%) of the insufficient results was caused by a false positive staining result. Poor-signal-to-noise ratio or technical issues were observed in the remaining 22% of the insufficient results (5 of 23). 29% of the participants obtained Good as score mostly due to a weak staining result (83% (39 of 47)), but with no change in the TPS-category. No obvious reason for this observation could be identified.

The Ventana PD-L1 IHC assay 790-4905, based on the SP263 clone, was the most widely used assay for demonstration of PD-L1 and provided a pass rate of 96%. Applying protocol settings in compliance with the vendor recommendations the pass rate was 100% (22 of 22). In comparison, protocols based on laboratory modified protocol settings obtained a pass rate of 93% (26 of 28).
One laboratory used the assay off label on a Leica Bond platform and obtained a sufficient result.

The Dako Agilent 22C3 pharmDx assay SK006 provided an overall pass rate of 96% (27 of 28). All laboratories applied protocol settings in compliance with the vender recommendations. Although the pass rate was high, the proportion of optimal results was significantly lower compared to the performance obtained in the previous assessments. In this run C4, only 54% (15 of 28) gave an optimal score whereas 76% (47 of 62) were assessed as optimal in run C1-C3 (data grouped together). As mentioned above, no technical parameters could be identified explaining for the overall decrease in optimal performance. Nine laboratories applied the SK006 RTU product on another stainer platform than the Dako Autostainer. One laboratory used the Dako Omnis and obtained a sufficient result, while 8 laboratories used the Ventana BenchMark Ultra, providing a pass rate of 63% (5 of 8).

The Dako/Agilent 28-8 pharmDx assay SK005 applying protocol settings in compliance with the vendor recommendations had an overall pass rate of 80% (4 of 5). One laboratory used the SK005 assay on the Ventana BenchMark Ultra, obtaining an insufficient result. It must be emphasized that off-label use of approved assays cannot be recommended as it requires an extended and often challenging internal validation.

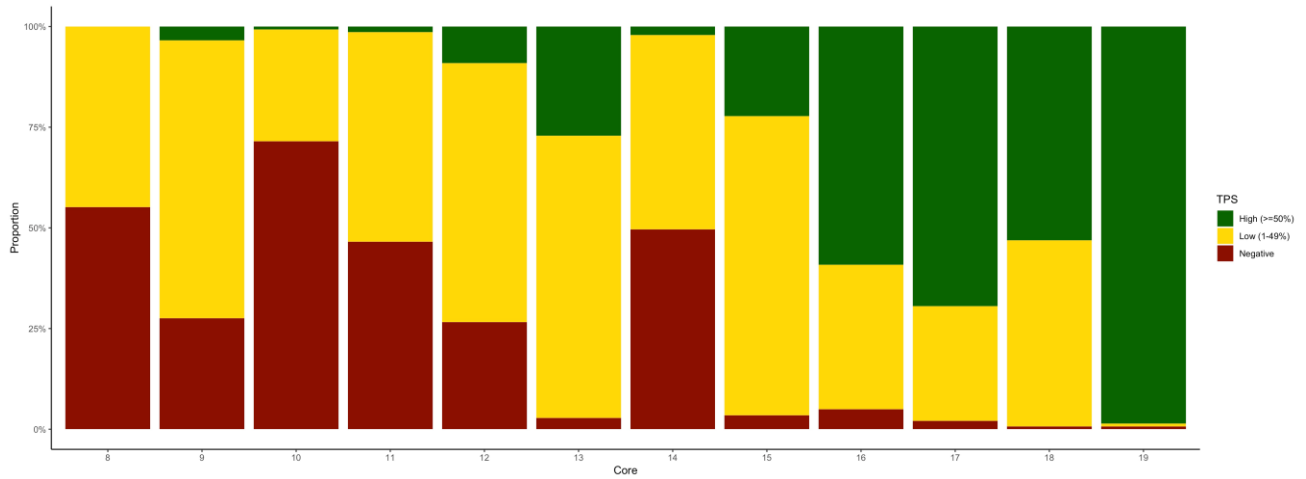
Grouped together, and using vendor recommended protocol settings, the approved PD-L1 IHC assays gave a pass rate of 95% (59 of 62). Note, that SP263 products 790-4905, 740-4907 and 741-4905 from Ventana is CE marked but not FDA approved in relation to NSCLC.

Laboratory developed (LD) assays either based on a concentrated Ab, an approved RTU product not used strictly accordingly to vendor recommendations or a non-approved RTU product, were used by 62% (101 of 164) of the participants. For this group a pass rate of 79% (81 of 102) was observed. The overall pass rate for LD assays is significantly lower than the approved assays used as recommended, which underlines the difficulties by validating LD assays.

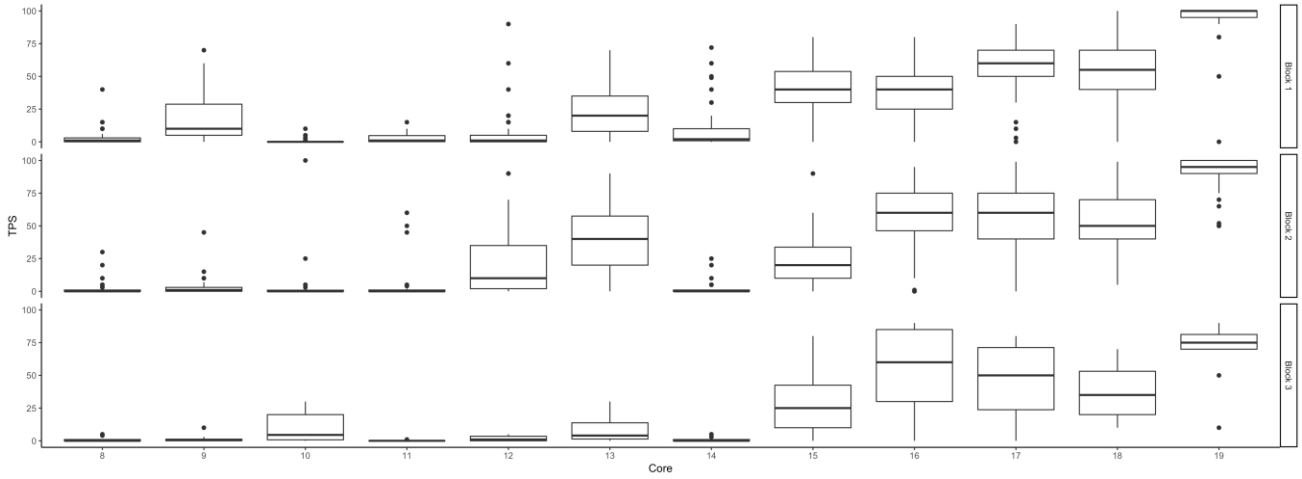
The mAb clone 22C3 was the most widely used Ab within a LD assay (n=34) and the pass rate was 79% (27 of 34). This is a significant decrease compared to the C3 run, where 97% of the LD assays based on this clone were sufficient. However, the circulated material in run C3 were less optimal. The results for run C4 with mAb clone 22C3 are comparable with the results for run C2, where 76% of the LD assays based on this clone were sufficient.
17 laboratories used the mAb clone 22C3 on the Ventana BenchMark stainer platform obtaining a pass rate of 82% (14 of 17). 7 laboratories applied OptiView with Amplification providing a pass rate of 71% (5 of 7), but only 1 laboratory obtained an optimal staining result. The remaining 6 laboratories displayed an aberrant granular staining pattern. This pattern was accepted as it did not significantly interfere with the interpretation. However, since future scoring systems, as the CPS (combined positive score), also takes positive immune cells in account, this aberrant staining pattern will prevent correct scoring and should be avoided. It is well known from previous assessments in NordiQC, both for other epitopes (general module) and PD-L1, that assays based on tyramide amplification can be challenging as low-level expressing tissue structures may be negative and if not carefully calibrated, can cause false positive staining result. In general, tyramide amplification will enhance high level expressing cellular structures and may add a fine granular staining of structures expected to be negative.
The Dako Omnis stainer platform was used by 10 laboratories applying mAb clone 22C3 as concentrate. 5 laboratories used the NordiQC validated protocol (Røge R, Vyberg M, Nielsen S. Accurate PD-L1 Protocols for Non-Small Cell Lung Cancer can be Developed for Automated Staining Platforms With Clone 22C3. Appl Immunohistochem Mol Morphol. 2017 Jul;25(6):381-385.) and all obtained an optimal result. The remaining 5 laboratories used the mAb clone 22C3 too diluted compared to the recommended dilution from the article. The article also include recommendable protocols for BenchMark Ultra (Ventana) and Bond III (Leica).

PD-L1 interpretation and scoring consensus:

Participants were asked to evaluate the percentage of PD-L1 positive tumour cells in each of the 12 NSCLCs included in the assessment. The overall interpretation of PD-L1 expression is shown in Graph 1.



Graph 1. **NordiQC PD-L1 run C4: participant interpretation of PD-L1 TPS – impact on treatment**

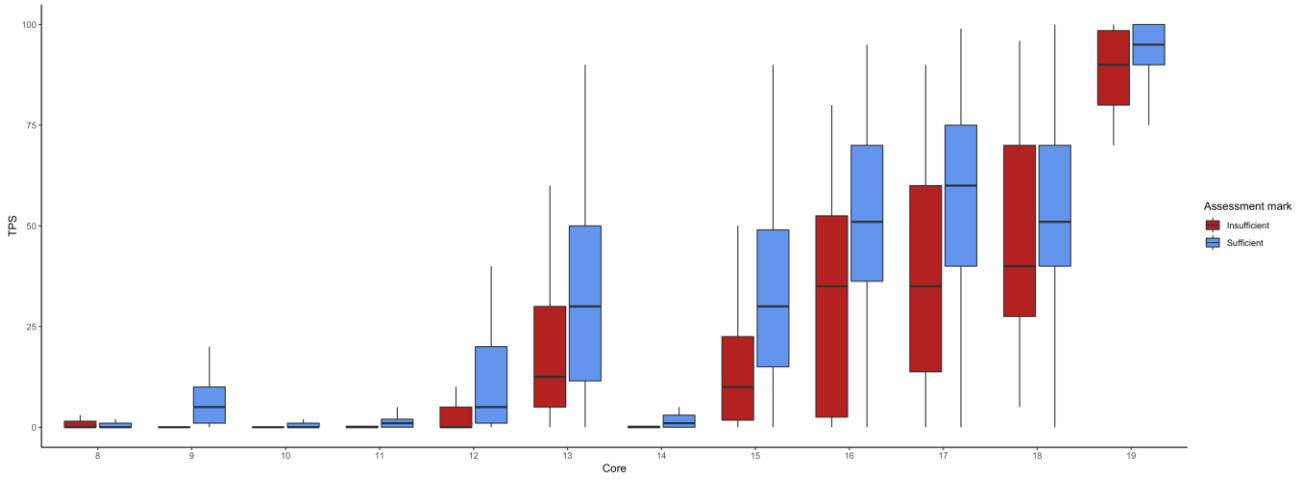


Graph 2. **NordiQC PD-L1 run C4: Tumour proportion scores across tissue core split up by different TMA blocks**

As seen in Graph 1, agreement in PD-L1 scoring by the participants was low and varied between the different cores. As shown in Graph 2, relative high consensus rates were observed for the tissue cores 8, 10, 12-13, 15 and 19 (in all blocks), whereas the consensus rates were significantly lower in the cores 9, 11, 14, and 16-18.

For the tissue core 9 (and to lesser extent core 11), a large group of participants scored this as TPS low (1-49%). This could be explained by PD-L1 expression heterogeneity seen in core no 9 across the tissue blocks. The majority of participants that received slides from Block 1 scored the slide as TPS low (1-49%), while slides from block 2 and 3 were scored as TPS negative (<1%), see Graph 2. Core 14 was for the majority of labs receiving slides from block 1 scored as TPS low (1-49%), while the remaining labs primarily scored this block TPS negative (<1%). Finally, core 16 was for labs receiving slides cut from block 1 primarily scored as TPS low (1-49%), while the remaining labs scored this core as TPS high (>50%). These expression level patterns were also seen in the neighboring reference slides. The cores no 17 and 18 were challenging, and no significant parameters (including specific TMA blocks/cores) could be identified unravelling the discrepancy in TPS scoring rates between participants own assessments and the results obtained in neighboring NordiQC reference slides. Overall, when taking heterogeneity of PD-L1 in specific cores of the different blocks into account, there was a moderate to good inter laboratory agreement in PD-L1 scoring in line with the predetermined reference scores.

When stratifying for the assessment marks, analysis revealed that scores among participants that had received an insufficient mark (borderline or poor) reported lower TPS scores than laboratories that had received a sufficient mark, see Graph 3. However, this difference was not statistically significant and a marked overlap in scores between the two groups was seen. Additionally, laboratories that received an insufficient mark tended to have more diverse TPS scores.



Graph 3. **NordiQC PD-L1 run C4: interpretation concordance for labs with sufficient vs. insufficient results**

Controls

Tonsil and placenta were used as positive and negative tissue controls. In this assessment, tonsil was found to be superior to placenta, as tonsil displayed a range of PD-L1 expression levels. Using PD-L1 IHC 28-8 (SK005, Dako/Agilent), 22C3 (SK006, Dako/Agilent) or SP263 (790-4905, Ventana/Roche) and obtaining an optimal staining result, tonsil displayed the following reaction pattern: No staining reaction in the vast majority of lymphocytes including mantle zone and germinal centre B-cells, no staining reaction in superficial epithelial cells, a weak to moderate, typically punctuated membranous staining reaction of the majority of germinal centre macrophages and finally a moderate to strong staining reaction of the majority of epithelial crypt cells. In addition, SP263 (790-4905, Ventana/Roche) provided higher proportion of positive immune cells compared to the two FDA approved kits from Dako/Agilent (SK005 and SK006).
However, it was observed that a fully acceptable staining pattern in tonsil could be obtained together with insufficient and false negative result in the NSCLCs. This underlines the need to identify more reliable positive tissue controls for PD-L1 and/or improve the interpretation criteria for a sufficient staining reaction in tonsil e.g. more accurately specify number and intensity of cells expected to be demonstrated.

Cell lines from HistoCyte (Newcastle, UK) were included in this assessment, primarily to evaluate if this material, in combination with digital image analysis, can be used to evaluate staining quality for PD-L1 and potentially be used as standard reference material for the validation of the precision of PD-L1 IHC assays. Subsequent analysis will be performed by NordiQC and published at a later stage.

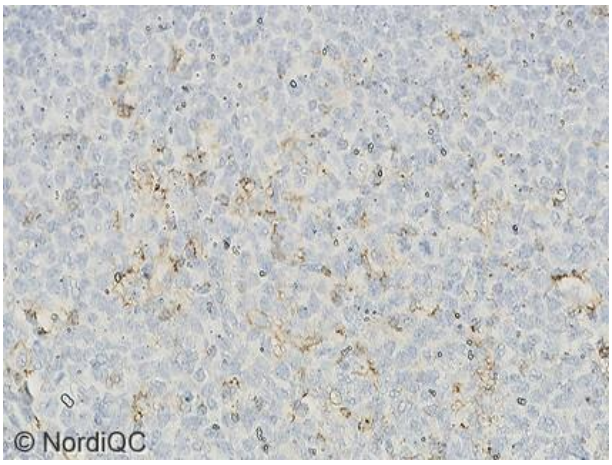


Fig. 1a. Optimal staining result of tonsil (germinal centre) using the pharmDX IHC PD-L1 assay, SK006, Dako/Agilent based on the mAb clone 22C3. Same protocol used in Figs. 2a–4a. The majority of germinal centre macrophages show an at least weak but distinct membranous staining reaction. The vast majority of lymphoid cells are negative.

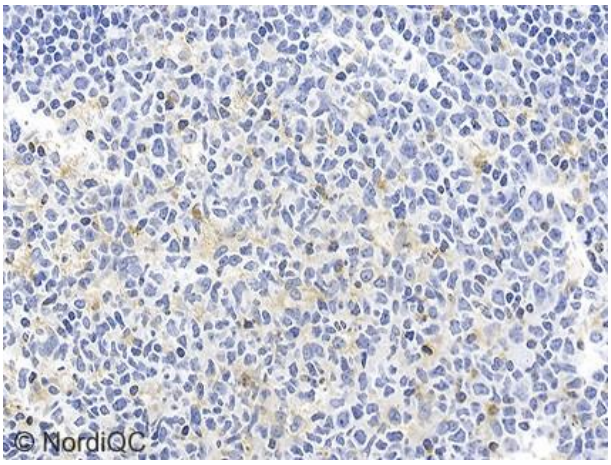
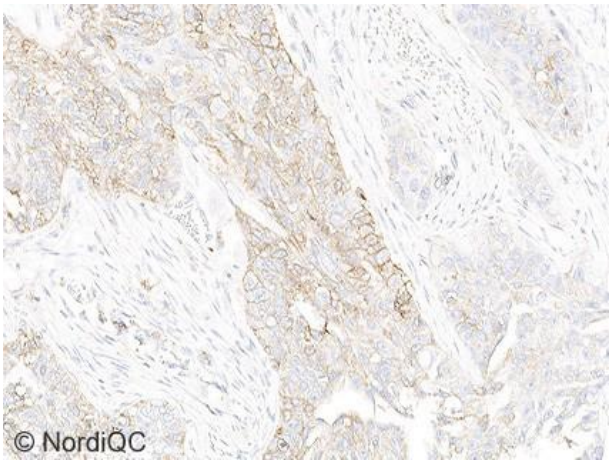


Fig. 1b. Insufficient staining result of tonsil (germinal centre) using the rmAb clone CAL10 within a laboratory developed assay. The protocol was based on HIER in an alkaline buffer and a 3-layer detection system. Same protocol used in Fig. 2b. The germinal centre macrophages display an aberrant cytoplasmic staining and membranous staining reaction is difficult to identify. Compare with optimal staining in Fig.1a.



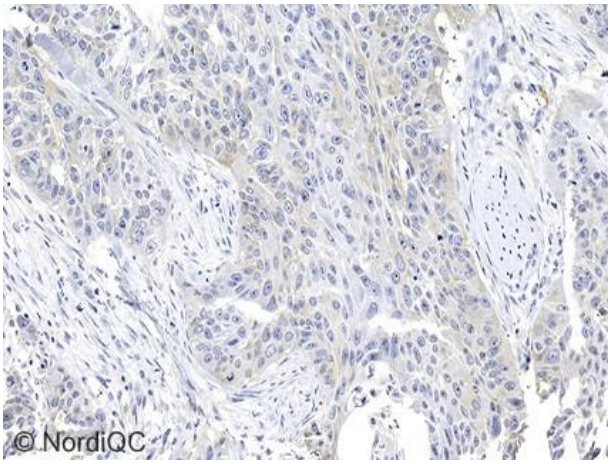


Fig. 2a. Optimal staining result of the tissue core no. 15 using same protocol as in Fig. 1a. Approximately, 45% of the neoplastic cells in the whole core show a weak but distinct membranous staining reaction. The tumour was categorized as TPS low (1-49%) and thus eligible for second line immune therapy. The same staining pattern was seen in reference slides.

Fig. 2b. Insufficient staining result of the tumour core no. 15 using same protocol as in Fig. 1b. Virtually all tumour cells are negative providing a TPS below 1%. Only an unspecific background staining of some tumour cells are seen. Compare with optimal staining in Fig. 2a.

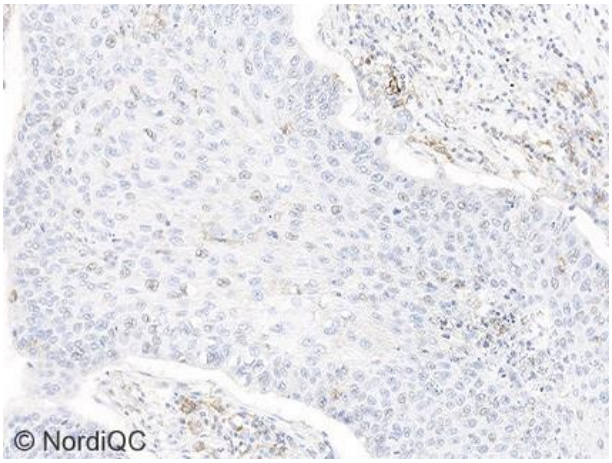


Fig. 3a. Optimal staining result of the tissue core no. 8 using same protocol as in Figs. 1a-2a.
The neoplastic cells are as expected negative for PD-L1. Only macrophages and dispersed lymphocytes show a distinct weak membranous staining reaction. The tumour was categorized as TPS negative (No <1%). The same staining pattern was observed in reference slides.

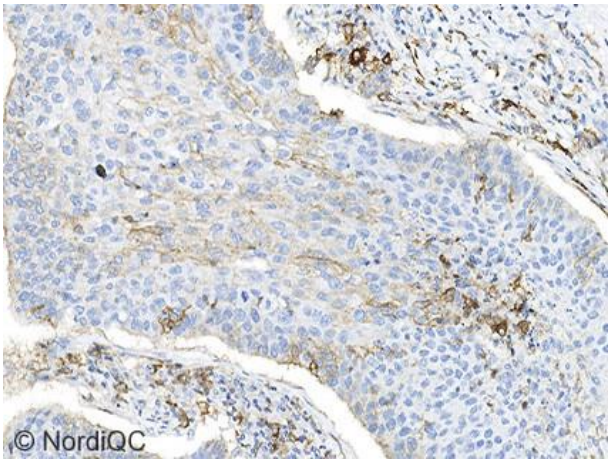


Fig. 3b. Insufficient staining result of the tissue core no. 8 using a laboratory developed assay based on rmAb clone MXR003 as RTU using HIER in an alkaline buffer and a 2-layer detection system.
Tumour cells expected to be negative display a membranous and false positive staining reaction. Increased proportions of PD-L1 positive cells were also observed in other tumour cores.
Compare with optimal result in Fig. 3a.

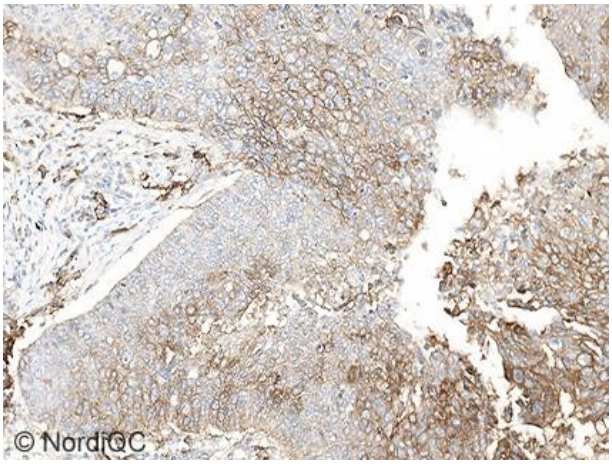


Fig. 4a. Optimal staining result of the tissue core no. 17 using same protocol as in Figs. 1a-3a.
The majority of the neoplastic cells are as expected positive for PD-L1. The tumour was categorized as TPS high (≥ 50%). The same staining pattern was observed in reference slides.

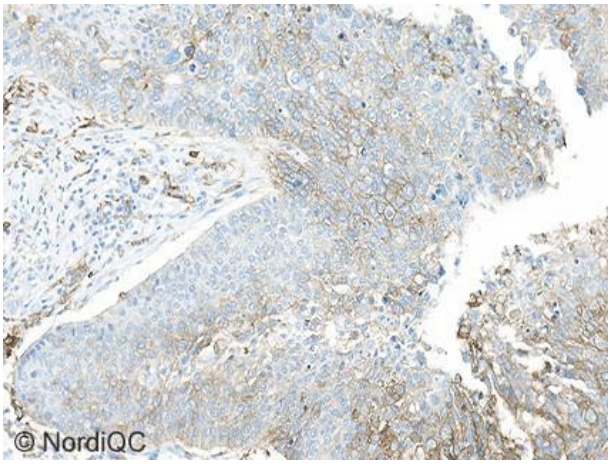


Fig. 4b. Good staining result of the tissue core no. 17 using same Ab and protocol settings as in Fig. 4a. The neoplastic cells shows a much weaker staining result compared with optimal result in Fig. 4a. Despite the weaker staining result in general, all tumour cores were assigned in the correct TPS category.

HLK/MB/ON/LE/MV/RR 12.12.2018