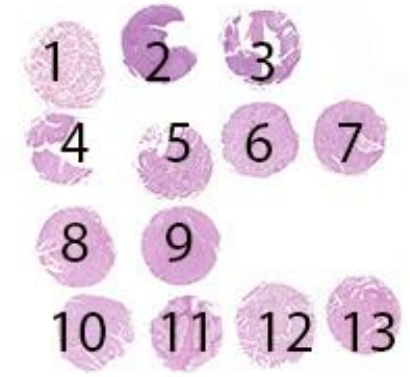
Assessment Run C7 2020
PD-L1 KEYTRUDA®

Purpose
This was the seventh assessment for PD-L1 in the NordiQC Companion module. This and the previous assessments for PD-L1 KEYTRUDA® primarily focused on the evaluation of the analytical accuracy of the IHC assays performed by the NordiQC participants to identify patients with NSCLCs and urothelial carcinomas to be treated with KEYTRUDA® as immune therapy. PD-L1 22C3 pharmDx, SK006 (Dako/Agilent) was used as reference standard method, and accuracy was evaluated in carcinomas with the dynamic and critical relevant expression levels of PD-L1 characterized and evaluated 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 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 KEYTRUDA® C7 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: No; <1%
6. NSCLC**	TPS: Low; 5-40% / High; 60-80%
7. NSCLC	TPS: Low; ≥1-40%
8. NSCLC	TPS: High; 60-80%
9. NSCLC	TPS: High; 90-100%
10. Urothelial carcinoma	CPS: <10
11. Urothelial carcinoma	CPS: ≥10; 20-40 IC***
12. Urothelial carcinoma	CPS: ≥10; 50-100 TC***
13. Urothelial carcinoma	CPS: ≥10; 20-40 IC+TC



* Tumour proportion score (TPS) and combined positive score (CPS) determined by PD-L1 IHC 22C3, SK006 (Dako/Agilent) performed in NordiQC reference lab.
** The tumour showed high level of heterogeneity in the different levels within and in between the TMA's used
*** 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®, interpret 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 KEYTRUDA® assays and provided information on the reproducibility and concordance of the PD-L1 interpretation results among the laboratories.

PD-L1 KEYTRUDA®, 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. Every twentieth slide throughout the blocks were stained for PD-L1 using the CE IVD / FDA approved 22C3 pharmDx kit SK006 (Dako/Agilent), and also by the CE IVD approved assay (NSCLC, KEYTRUDA®) SP263 741-4905 (Ventana/Roche) 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 slide of the 22C3 pharmDX SK006 (Dako/Agilent).

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 10 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 10 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 of one of the included carcinomas and/or the interpretation is compromised by e.g. a poor signal-to-noise ratio, impaired morphology etc.
The protocol should be optimized.
<u>TPS/CPS is **not** concordant to the NordiQC reference data for the carcinomas and/or interpretation is hampered.</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 and/or the interpretation being compromised by e.g. a poor signal-to-noise ratio, impaired morphology etc.
Optimization of the protocol is urgently needed.
<u>TPS/CPS is **not** concordant to the NordiQC reference data for the carcinomas and/or interpretation is hampered.</u>

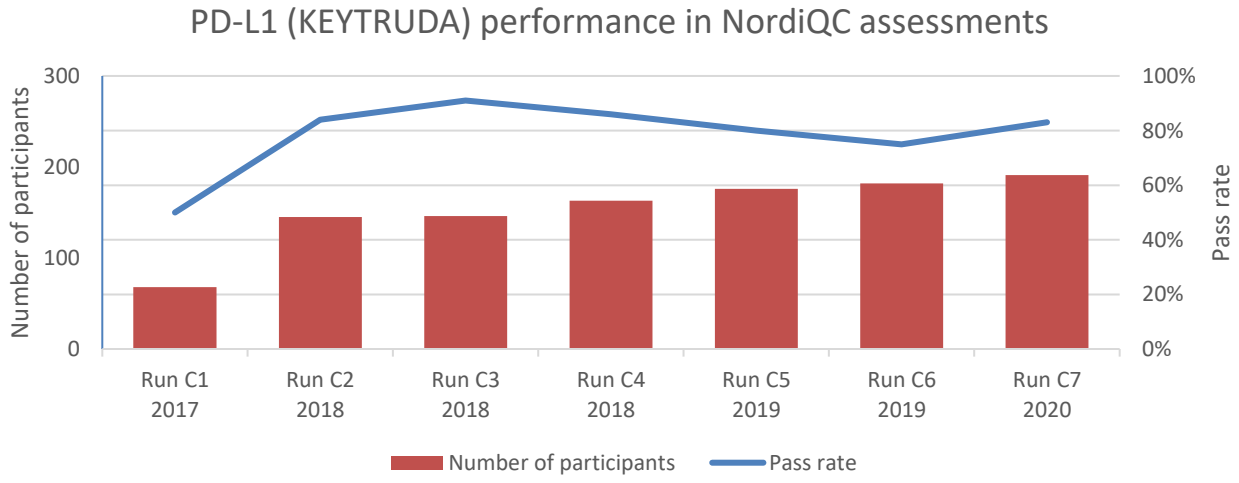
Participation

Number of laboratories registered for PD-L1 KEYTRUDA IHC C7	211
Number of laboratories returning PD-L1 KEYTRUDA IHC	191 (91%)
Number of laboratories returning PD-L1 scoring sheet	162 (89%)

The number of laboratories returning slides decreased in this run C7 compared to previous assessments, due to the COVID-19 pandemic. All slides returned after the assessment will be assessed, and receive advice if the result is insufficient, but will not be included in this report.

Results: 191 laboratories participated in this assessment and 83% achieved a sufficient mark. Assessment marks for IHC PD-L1 assays and PD-L1 antibodies are summarized in Table 3 (see next page).

Fig. 1. **Proportion of sufficient results for PD-L1 (lung)/KEYTRUDA in the seven NordiQC runs performed**



Performance history
This was the seventh NordiQC assessment of PD-L1 (lung)/KEYTRUDA®. A slightly increased pass rate was obtained in C7 (see Fig. 1) compared to the previous assessment. The number of new participants seems to be consistently increasing.

Conclusion
In this run, and similar to the observations seen in the previous runs in the PD-L1 companion module, the insufficient PD-L1 IHC staining results were most frequently characterized by a reduced proportion of PD-L1 positive cells compared to the level expected and defined by the PD-L1 IHC pharmDx assay, SK006 (Dako/Agilent). This resulted in a too low TPS/CPS changing the PD-L1 status in one or more of the carcinomas included in the assessment material.
Several companion diagnostic assays and laboratory developed (LD) assays based on concentrated antibodies as mAb clone 22C3 or Ready-To-Use formats of the clones 73-10 and MAB-0854 could all provide optimal results. The companion diagnostic PD-L1 IHC assays, 22C3 SK006 / GE006 from Dako and SP263 741-4905 / 740-4907 from Ventana/Roche performed in concordance to the product guidelines, provided a high proportion of sufficient results and was superior to LD assays. Within LD assays, irrespectively of Ab clone being used, meticulous calibration and validation of the assay is required.

Table 3. **Assessment marks for IHC assays and antibodies run C7, PD-L1 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]**	31	Ventana/Roche	25	4	1	1	94%	81%
rmAb clone SP263, **741-4905 (LMPS)[4]**	12	Ventana/Roche	8	2	1	1	83%	75%
rmAb clone SP263, **740-4907 (VRPS)[3]**	4	Ventana/Roche	2	2	-	-	-	-
rmAb clone SP263, **740-4907 (LMPS)[4]**	3	Ventana/Roche	2	1	-	-	-	-
mAb clone 22C3 pharmDX, **SK006 (VRPS)[3]**	12	Dako/Agilent	5	6	-	1	92%	42%
mAb clone 22C3 pharmDX, **SK006 (LMPS)[4]**	20	Dako/Agilent	11	7	1	1	88%	55%
mAb clone 22C3 pharmDX, **GE006 (VRPS)[3]**	12	Dako/Agilent	10	2	-	-	100%	83%
mAb clone 22C3 pharmDX, **GE006 (LMPS)[4]**	3	Dako/Agilent	1	2	-	-	-	-
rmAb clone 28-8 pharmDX, **SK005 (VRPS)[3]**	3	Dako/Agilent	1	1	-	1	-	-
Antibodies[5] for laboratory developed PD-L1 assays, concentrated antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
mAb clone **22C3**	44	Dako/Agilent	13	17	7	7	68%	30%
mAb clone **E1L3N**	3	Cell Signaling	1	1	-	1	-	-
rmAb clone **28-8**	1	Abcam	1	-	-	-	-	-
rmAb **CAL10**	3 3	Biocare Zytomed Systems	-	3	1	2	50%	-
rmAb clone **QR1**	2	Diagomics	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]**	14	Ventana/Roche	11	1	-	2	86%	79%
rmAb clone SP263, **790-4905[6] (LMPS)[4]**	14	Ventana/Roche	8	3	2	1	79%	57%
rmAb clone 73-10, **PA0832 (VRPS)[3]**	2	Leica Biosystems	2	-	-	-	-	-
mAb clone MX070C, **MAB-0854**	3	Maixin	2	1	-	-	-	-
Total	191		106	53	14	18		
Proportion			55%	28%	7%	9%	83%	

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): In total, 33 of 43 (77%) protocols 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 32-64 min., 12-16 min. incubation of the primary Ab, OptiView as detection system and performed on BenchMark Ultra, GX or XT. Using these protocols settings, 31 of 34 (91%) laboratories produced a sufficient staining result (optimal or good).

SP263 (740-4907, Ventana): In total, 4 of 7 (57%) protocols were assessed as optimal. The protocols were typically based on HIER in CC1, efficient heating time 52-64 min., 16 min. incubation of the primary Ab, OptiView with or without Amplification as detection system and performed on BenchMark Ultra or GX. Using these protocol settings, 6 of 6 (100%) laboratories produced a sufficient staining result.

PD-L1 IHC 22C3 pharmDx (SK006, Dako): In total, 16 of 32 (50%) protocols were assessed as optimal. Protocols with optimal results were typically based on HIER using EnVision™ Flex Target Retrieval Solution (TRS) low pH 6.1 (SK006) at 95-99°C for 20-40 min. (PT Link or On-board, Omnis), 30-40 min. incubation of the primary Ab, EnVision Flex+ as the detection system and performed on Autostainer Link 48 or Dako Omnis. Using these protocol settings, 23 of 25 (92%) laboratories produced a sufficient staining result.

PD-L1 IHC 22C3 pharmDx (GE006, Dako): In total, 11 of 15 (73%) protocols were assessed as optimal. Protocols with optimal results were based on HIER using EnVision™ Flex TRS low pH 6.1 (GV805) at 95-99°C for 30-40 min, 40 min. incubation of the primary Ab, EnVision Flex+ as the detection system and performed on Omnis. Using these protocol settings, 13 of 13 (100%) laboratories produced a sufficient staining result.

PD-L1 IHC 28-8 pharmDx (SK005, Dako): In total, 1 of 3 protocols was assessed as optimal. The protocol with optimal result was 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, EnVison Flex+ as the detection system and performed on Autostainer Link 48. Using these protocol settings, 2 of 3 laboratories produced a sufficient staining result.

Table 4 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 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, **741-4905**	29/31 (94%)	25/31 (81%)	10/12 (83%)	8/12 (75%)
Ventana BenchMark Ultra rmAb SP263, **740-4907**	4/4	2/4	3/3	2/3
Dako Autostainer Link 48+ mAb 22C3 pharmDX, **SK006**	11/12 (92%)	5/12 (42%)	7/8 (88%)	4/8 (50%)
Dako Omnis mAb 22C3 pharmDX, **GE006**	12/12 (100%)	10/12 (83%)	1/1	1/1
Dako Autostainer Link 48+ rmAb 28-8 pharmDX, **SK005**	2/3	1/3	-	-

*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 **22C3**: 13 of 44 (30%) protocols were assessed as optimal of which eight were stained on the BenchMark (Ventana) and three on the Omnis (Dako) platform.
On BenchMark XT/Ultra (Ventana), the protocols providing optimal results were typically based on a titre of 1:30-50, primary Ab incubation time of 32-80 min., HIER in CC1 (efficient heating time 48-80 min.) and OptiView as detection system. Using these protocol settings, 16 of 18 (89%) laboratories produced a sufficient staining result. Four laboratories applied OptiView Amplification to the protocol settings listed above. 4 of 5 produced a sufficient staining result, none optimal.
On Omnis (Dako), the protocols providing optimal results were typically based on a titre of 1:20-30 of the primary Ab, incubation time of 30-40 min., HIER in TRS low pH 6.1 (Dako) at 97°C (efficient heating time

30-40 min.) and EnVision FLEX+ as detection system. Using these protocol settings, 4 of 5 (80%) laboratories produced a sufficient staining result.

mAb **E1L3N**: One protocol provided an optimal result. The protocol was based on HIER using an alkaline-buffer (Bond Epitope Retrieval Solution 2, BERS2) at 100°C for 30 min. The mAb clone E1L3N was diluted 1:100, incubated for 30 min. at room temp. and visualized using a 3-layer detection system (Leica Refine) and performed on a Leica Bond III platform. Using these protocol settings, 2 of 2 laboratories obtained a sufficient staining result.

mAb **28-8**: One protocol provided an optimal result. The protocol was based on HIER using a TRS low pH 6.1 buffer (Dako) at 97°C for 30 min. The rmAb clone 28-8 was diluted 1:100, incubated for 30 min. at room temp. and visualized using a 3-layer detection system (Dako Flex+) and performed on a Dako Omnis platform.

rmAb **QR1**: One protocol was assessed as optimal. The protocol was based on HIER using an alkaline-buffer (BERS2) at 99°C for 20 min. The rmAb clone QR1 was diluted 1:50, incubated for 15 min. at room temp. and visualized using a 3-layer detection system (Leica Refine) and performed on a Leica Bond III platform.

rmAb **ZR3**: One protocol was assessed as optimal. The protocol was based on HIER using Tris-EDTA/EGTA pH 9 at 95°C for 15 min. in a waterbath. The rmAb clone ZR3 was diluted 1:100, incubated for 50 min. at room temp. and visualized using a 2-layer detection system (GT Vision, GeneTech) and performed on a GeneTech Genestainer.

rmAb **SP142**: One protocol provided an optimal result. The protocol was based on HIER using a TRS low pH 6.1 buffer (Dako) at 100°C for 24 min. in PT link. The rmAb clone SP142 was diluted 1:200, incubated for 30 min. at room temp. and visualized using a 3-layer detection system (Thermo, Quanto) and performed on a Dako Autostainer platform.

Table 5. **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 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**	8/25 (32%)	-	1/1	0/4	0/2	3/6 (50%)	0/4	-

*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): In total, 19 of 28 (68%) protocols provided an optimal result. Protocols with optimal results were typically based on Heat Induced Epitope Retrieval (HIER) in Cell Conditioning 1 (CC1), efficient heating time 48-64 min., 16-24 min. incubation of the primary Ab, OptiView as detection system and performed on BenchMark Ultra or XT. Using these protocols settings, 18 of 21 (86%) laboratories produced a sufficient staining result (optimal or good).

rmAb **73-10** (PA0832, Leica): 2 out of 2 protocols provided an optimal result. The protocols were based on HIER in Bond Epitope retrieval Solution 1 for 20 min, 15 min. incubation of the primary Ab, Refine as detection system and performed on Bond III.

mAb **MX070C** (MAB-0854, Maixin): 2 out of 3 protocols provided an optimal result (67%). The protocols were based on HIER in Tris-EDTA/EGTA pH 9 for 20-30 min., 30 min. incubation of the primay Ab, DAB Titan Super as detection system and performed on Lumatas Autostaining System. Using these protocol settings 3 of 3 (100%) produced a sufficient staining result.

Block construction and assessment reference standards
The tissue micro array (TMA) block constructed for this PD-L1 run consisted of 6 NSCLCs, 4 urothelial carcinomas, 2 tonsils and 1 placenta. The NSCLCs were selected to comprise 2 NSCLCs 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 3 with CPS≥10, one with PD-L1 expression primarily in immune cells, one with PD-L1 expression primarily in tumour cells and one with PD-L1 expression in both tumour cells and immune cells, respectively. Reference slides throughout the individual TMA blocks (interval at each twentieth slide) were stained using the companion diagnostic assays 22C3 pharmDX SK006 (Dako) and SP263 741-4905 (Ventana). 22C3 pharmDX SK006 (Dako) was used to characterize both TPS and CPS and SP263 for TPS (reflecting the EU/FDA approved predictive claims for KEYTRUDA® at the assessment). In total, seven identical blocks were constructed and used for this assessment.
Reviewing the reference slides from the blocks, heterogenic expression of PD-L1 was seen in two of the tumor cores. In the NSCLC, tissue core no. 6, initially scored as TPS low (≥1-49%), slides with TPS high ≥50% was identified. The NSCLC, tissue core no. 7, also showed PD-L1 heterogeneity and different proportions of TPS observed ranging from 1 to 40%, but still being TPS low in all reference slides.
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 of 22C3 pharmDX SK006 (Dako).
Heterogeneity in PD-L1 expression is well known in NSCLCs and the assessment in this sense emulated clinical settings.

Comments
In this seventh NordiQC assessment for PD-L1 for (lung)/KEYTRUDA®, the prevalent feature of an insufficient staining result was a false negative staining result, being observed in 78% (25 of 32) of the insufficient staining results. As shown in Table 6, a false negative staining result has been the most common reason for insufficient staining results in all NordiQC PD-L1 (lung)/KEYTRUDA® assessments with an average occurrence of 76%. In this run, only 6% (2 of 32) of the insufficient results were caused by a false positive staining result. Poor-signal-to-noise ratio, excessive cytoplasmic staining reaction interfering the interpretation or technical issues were observed in the remaining 16% of the insufficient results (5 of 32).

Table 6. **Characteristics of insufficient results in the seven NordiQC PD-L1 (lung)/KEYTRUDA® runs.**

	False Negative*	False Positive**	"Technical"***
C1 (50% insufficient)	68%	15%	17%
C2 (16% insufficient)	68%	23%	5%
C3 (9% insufficient)	100%	-	-
C4 (14% insufficient)	74%	4%	22%
C5 (20% insufficient)	59%	18%	23%
C6 (25% insufficient)	82%	11%	7%
C7 (17% insufficient)	78%	6%	16%
Average (22% insufficient)	**76%**	**11%**	**13%**

* TPS change from high to low or low to negative. And/or CPS change from ≥10 to <10.
** TPS change from negative to low or low to high. And/or CPS change from <10 to ≥10.
*** Interpretation compromised e.g. by poor-signal-to noise ratio, poor morphology excessive cytoplasmic staining reaction etc.

In run C7 28% (53 of 191) of the participants obtained a score as Good. In 57% of these (30 of 53), this was due to a general weak staining result or a reduced TPS/CPS, but with no change in the TPS/CPS-category and thus still accurate PD-L1 status. In 23% (12 of 53), poor signal-to-noise ratio, impaired morphology or excessive counterstaining was observed. And finally in 11% (6 of 53) a coarse granular staining reaction compromising the interpretation of the membranous staining reaction was observed. This pattern was only seen for protocols based on OptiView with amplification kit.

In this assessment and in concordance to the previous run C6, virtually all insufficient results were related to incorrect TPS categories in one or more of the NSCLCs, whereas the CPS scores only were affected in a few cases of the urothelial carcinomas. PD-L1 IHC demonstration has in the two latest assessments with combined assessment material been more successful in urothelial carcinomas versus NSCLCs. No plausible reasons for this difference can be identified. Expression levels in the 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. 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.

The NSCLC, tissue cores no. 6, 7 and 8 were most challenging to obtain an optimal result and required an accurate and carefully calibrated protocol. The majority of false negative results were especially seen in the NSCLC, tissue core no. 6, changing the TPS category from low to negative, but also frequently observed in the NSCLC, tissue core no. 8, changing TPS category from high to low, which would change the status for first line immune therapy using the present guidelines for NSCLCs as e.g. in EU. The two NSCLCs were both characterized by the NordiQC reference standard methods to display a weak to moderate membranous staining reaction. It was observed that protocols providing a reduced analytical sensitivity were less successful in these two tissue cores. In contrast, virtually all protocols provided the expected PD-L1 status in the NSCLC, tissue core no 9, characterized by NordiQC to show a strong membranous staining reaction in all tumour cells.

The Ventana/Roche PD-L1 IHC assays 741-4905 and 740-4907 for BenchMark with predictive claims, based on the SP263 clone, were the most widely used assays for demonstration of PD-L1 and provided an overall pass rate of 92% (46 of 50). Applying protocol settings in compliance with the vendor recommendations, the pass rate was 94% (33 of 35) and 77% optimal (27 of 35). 15 participants used laboratory modified protocol settings and for this group a pass rate of 87% was seen, 67% optimal (see Table 4). The protocol modifications reported were typically related to change in HIER time, incubation time in primary Ab and/or use of OptiView with amplification kit as detection system.

The Dako/Agilent 22C3 pharmDx assay SK006 for Autostainer Link 48 provided an overall pass rate of 91% (29 of 32). Applying the recommended protocol settings from Dako, the pass rate was 92% (11 of 12) and 42% optimal (5 of 12) (see Table 4). The proportion of optimal results was reduced compared to both the corresponding Dako 22C3 pharmDx assay for Omnis and also the Ventana PD-L1 assays based on SP263. The results assessed as "Good" were all characterized either by a generally weak staining reaction and/or a reduced number of cells being demonstrated compared to the level expected. No plausible reason for the reduced analytical sensitivity could be identified. Laboratory modified protocol settings provided a comparable proportion of sufficient and optimal results as shown in Tables 1 and 4.

The Dako/Agilent 22C3 pharmDx assay GE006 for Omnis provided an overall pass rate of 100% (15 of 15). 12 protocols were applied in compliance with the recommended protocol settings from Dako and for this group 83% being optimal (see Table 4).
Similar to data generated in run C6, it was observed that the recently launched PD-L1 22C3 GE006 assay for Omnis was more successful compared to 22C3 pharmDx SK006. As mentioned, no data concerning protocol settings submitted to NordiQC could explain the difference observed. It has to be emphasized that the 22C3 GE006 assay for Omnis is only validated by the vendor 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 pharmDx SK005 28-8 for Autostainer Link 48 was used by three laboratories. All used the Dako recommended protocol settings and two of three with a sufficient result. One result was assessed as optimal.

Grouped together, and using vendor recommended protocol settings, the CE IVD approved PD-L1 IHC assays with predictive claims provided a pass rate of 94% (58 of 62) and 69% being optimal (43 of 62).

Laboratory developed (LD) assays either based on a concentrated Ab, an RTU Ab without any predictive claim or a companion diagnostic assay not used strictly accordingly to vendor recommendations were applied by 68% (129 of 191) of the participants. For this group a pass rate of 78% (101 of 129) was observed, 49% optimal (63 of 129).
LD assays based on concentrated formats provided a relatively low pass rate of 67% (38 of 57), only 30% being optimal (17 of 57) (see Table 1).
Protocols based on RTU Abs without predictive claims as rmAb clone SP263, 790-4905, Ventana and 73-10, PA0835, Leica were more successful. For LD based assays on RTU Abs, the overall pass rate and proportion of sufficient and optimal results were 85% and 70%, respectively.

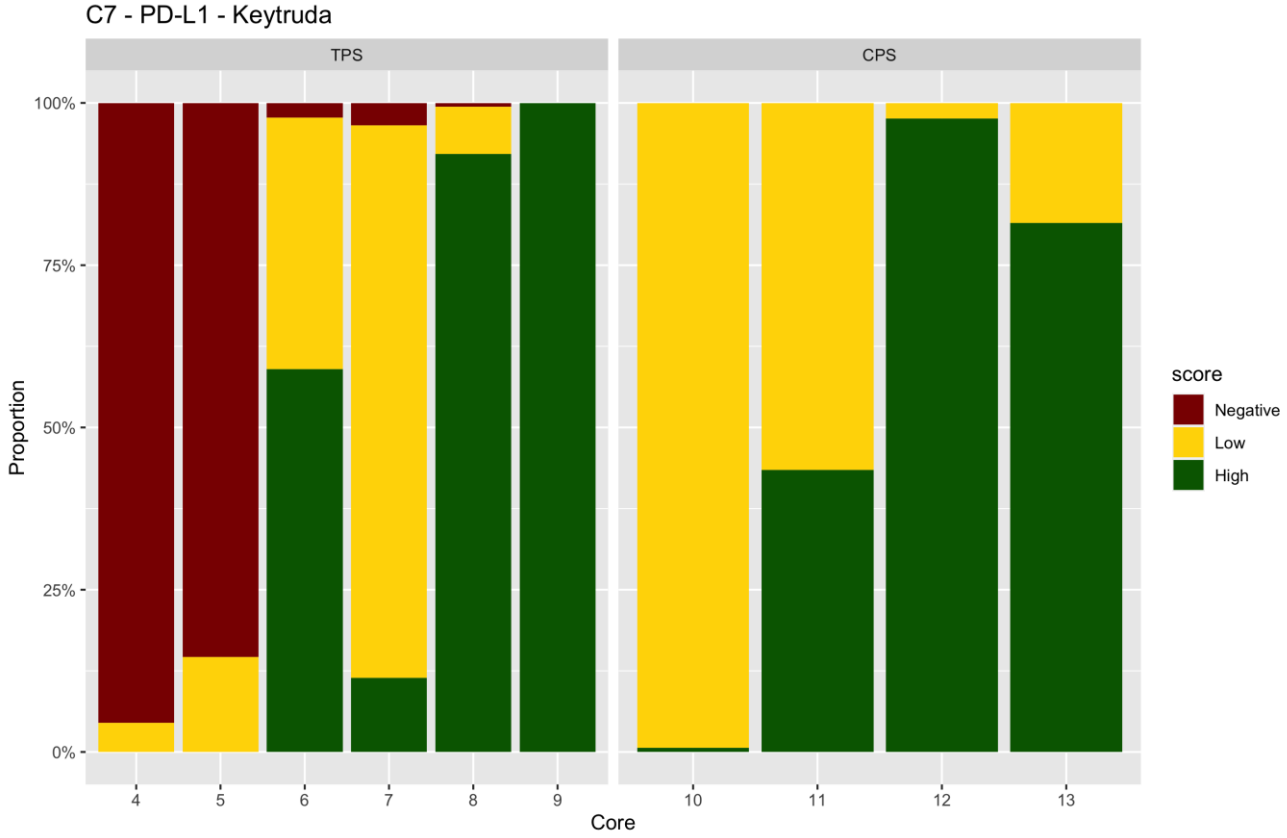
The mAb clone 22C3 was the most widely used concentrated Ab within a LD assay (n=44) providing a pass rate of 68% (30 of 44), 30% optimal – similar levels as in run C6.
25 laboratories used the mAb clone 22C3 on the Ventana BenchMark stainer platform and obtained a pass rate of 84% (21 of 25). All 8 optimal protocols were based on OptiView as detection system. 5 laboratories included amplification kit for OptiView, but similar to previous observations for PD-L1 KEYTRUDA®, with inferior performance resulting in an excessive granular staining reaction compromising interpretation of the membranous PD-L1 expression for especially TPS. Overall the use of mAb clone 22C as concentrate with OptiView + amplification kit gave a pass rate of 80%, 0% optimal.

Other concentrated Abs as e.g. clones E1L3N, QR1 and ZR3 could also be used to generate sufficient and optimal results with LD assays.

RTU Abs without predictive claims were in this assessment categorized separately in Table 1 and included the Ventana rmAb clone SP263, 790-4905 previously being listed as CE IVD /FDA approved PD-L1 assays. The RTU format is similar to the two other Ventana SP263 products but positioned as an open "analytical product" for PD-L1 and option to validate the protocol by the end-user. Laboratories using the RTU format as recommended by Ventana obtained a pass rate and proportion of optimal results comparable to the "locked PD-L1 assays" 741-4905 and 740-4907 (see Table 1).

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 C7: Tumour Proportion scores (TPS) in NSCLCs (core no. 4-9) and Combined Positive Score (CPS) in urothelial carcinomas (core no. 10-13).**

As seen in Graph 1, a relatively high consensus rates were observed for the tissue core 4,5,7-10,12-13, whereas the consensus rates were significantly lower in tissue core 6 and 11.

For the tissue core no. 6, tissue heterogeneity was observed throughout the blocks and the TPS category changed between TPS low and TPS high. This also seems to be reflected in the scoring results from the participants. Concerning tissue core no 11, which only had positive immune cells (and not any positive tumour cells), a relative high number of laboratories scored this as CPS low. One possible explanation could be that some laboratories scored the core using the TPS scoring system and not the CPS scoring system, which also includes positive immune cells in the CPS score.

When stratifying for assessment marks, analysis indicated that participants that had received an insufficient mark (borderline or poor) for the technical assessment of their PD-L1 result also had a higher tendency to perform an incorrect read-out of TPS and/or CPS in the submitted slides.

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 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 squamous epithelial cells.

It was observed that rmAb SP263 (741-4905, 790-4905/4907, Ventana) 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).

However, it was observed that occasionally a fully acceptable staining pattern in tonsil could be obtained together with an insufficient result in the carcinomas. This underlines the need to identify more reliable positive and negative (tissue) controls for PD-L1 and/or improve the interpretation criteria for a sufficient staining reaction in tonsils e.g. more accurately specific number and intensity of cells expected to be demonstrated and correlated to the PD-L1 IHC test/clone used. Some clones, e.g. mAb clone CAL10 typically gave a stronger staining reaction in more germinal centre macrophages compared to mAb clone 22C3.

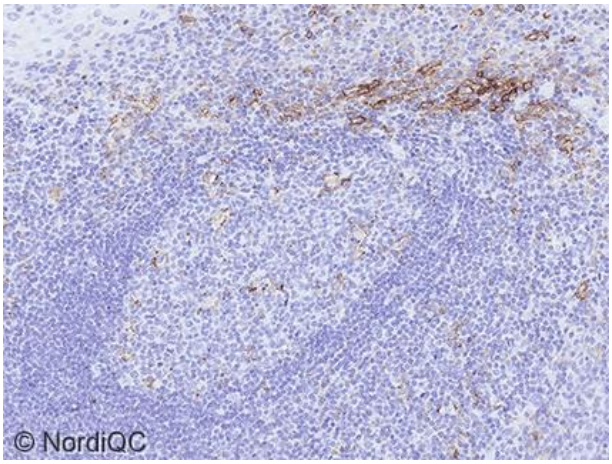


Fig. 1a
Optimal staining result of tonsil using the PD-L1 IHC 22C3 pharmDx kit, GE006, Dako/Agilent on Omnis following the recommended protocol settings.
Dispersed epithelial crypt cells show a moderate to strong membranous staining reaction, while the majority of germinal centre macrophages show a weak to moderate and distinct membranous staining reaction.
A few interfollicular T-cells and macrophages are weakly positive, whereas the vast majority of lymhocytes are negative.

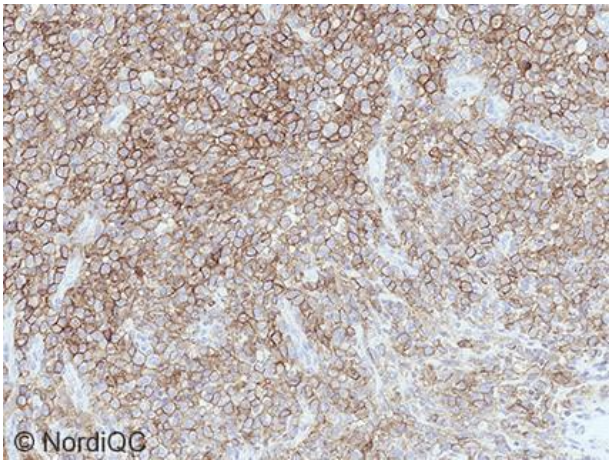


Fig. 1b
Optimal staining result of the NSCLC, tissue core no. 9, using the PD-L1 IHC 22C3 pharmDx kit, GE006, Dako/Agilent on Omnis following the recommended protocol settings.
Virtually all tumour cells show a moderate to strong membranous staining reaction.
The tumour was categorized as TPS high (≥50%) and thus eligible for first line immune therapy with KEYTRUDA® (different regional cut-offs occur).

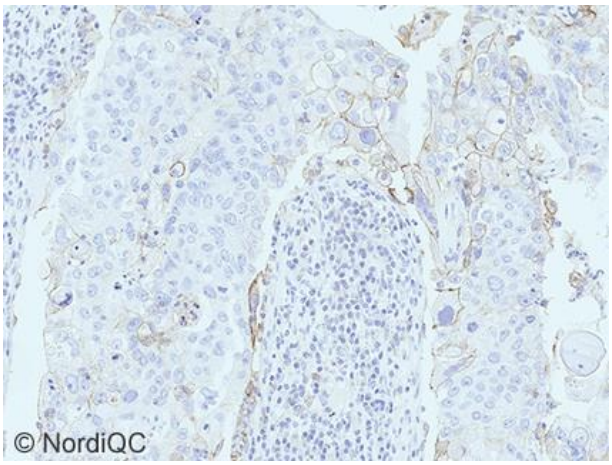


Fig. 2a
Optimal staining result of the NSCLC, tissue core no. 7, using the PD-L1 IHC 22C3 pharmDx kit, GE006, Dako/Agilent on Omnis following the recommended protocol settings.
Approximately 5-10% of tumour cells show a weak to moderate membranous staining reaction.
The tumour was categorized as TPS low (≥1-49%) and thus eligible for second line immune therapy with KEYTRUDA® (different regional cut-offs occur).

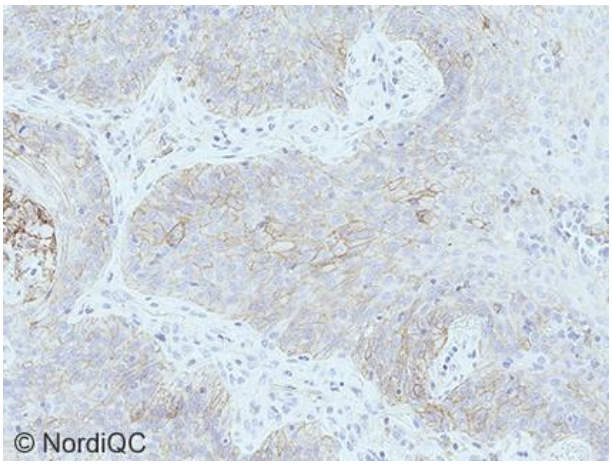


Fig. 2b
Optimal staining result of the NSCLC, tissue core no. 8, using the PD-L1 IHC 22C3 pharmDx kit, GE006, Dako/Agilent on Omnis following the recommended protocol settings.
Approximately 60-70% of tumour cells show a weak to moderate membranous staining reaction.
The tumour was categorized as TPS high (≥50%) and thus eligible for first line immune therapy with KEYTRUDA® (different regional cut-offs occur).

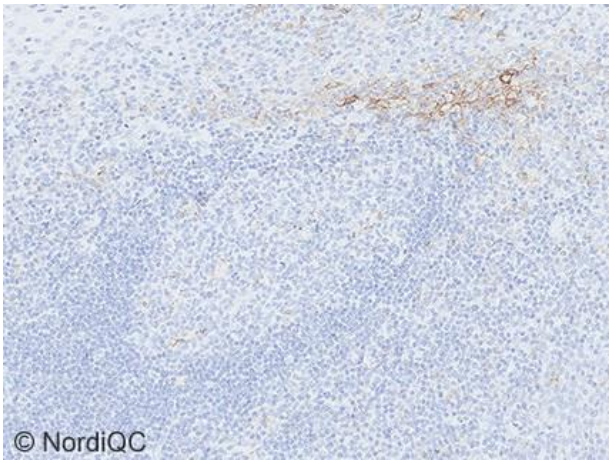


Fig. 3a
Insufficient staining result of tonsil using the PD-L1 IHC 22C3 pharmDx kit, SK006, Dako/Agilent "off-label" on a BenchMark Ultra with laboratory modified protocol settings.
Only few epithelial crypt cells show a moderate membranous staining reaction, while virtually all germinal centre macrophages, interfollicular T-cells and macrophages are negative.
Also see Figs. 3b, 4a and 4b, same protocol.

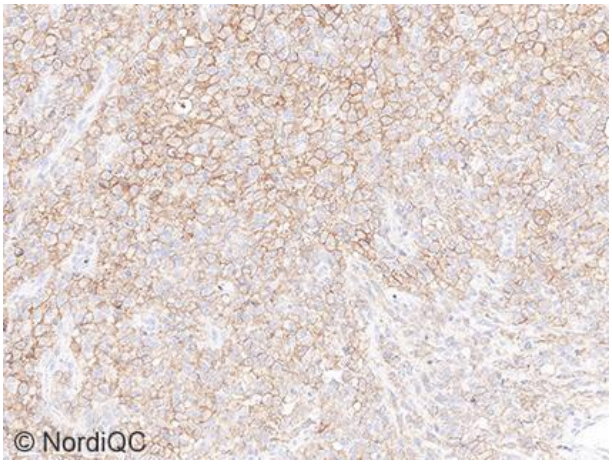


Fig. 3b
Staining result of the NSCLC, tissue core no. 9, using the PD-L1 IHC 22C3 pharmDx kit, SK006, Dako/Agilent "off-label" on a BenchMark Ultra with laboratory modified protocol settings.
Virtually all tumour cells show a weak to moderate membranous staining reaction.
The tumour was categorized as TPS high (≥50%) and thus eligible for first line immune therapy with KEYTRUDA® (different regional cut-offs occur).
However compare with Figs. 4a and 4b, same protocol.

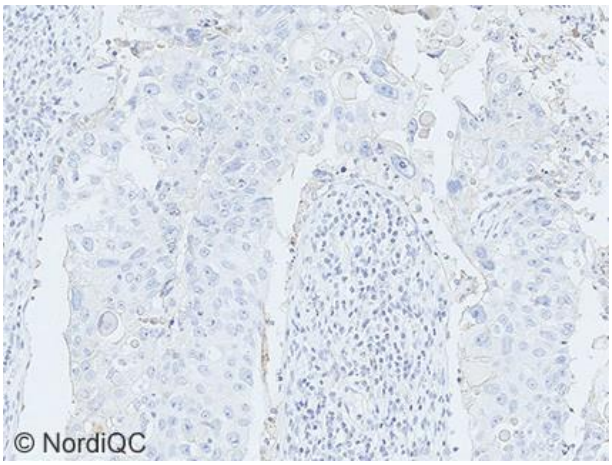


Fig. 4a
Insufficient staining result of the NSCLC, tissue core no. 7, using the PD-L1 IHC 22C3 pharmDx kit, SK006, Dako/Agilent "off-label" on a BenchMark Ultra with laboratory modified protocol settings.
<1% of tumour cells are demonstrated and categorized as "PD-L1 negative" and not eligible for immune therapy with KEYTRUDA®
Also compare with Fig. 2a, same field.

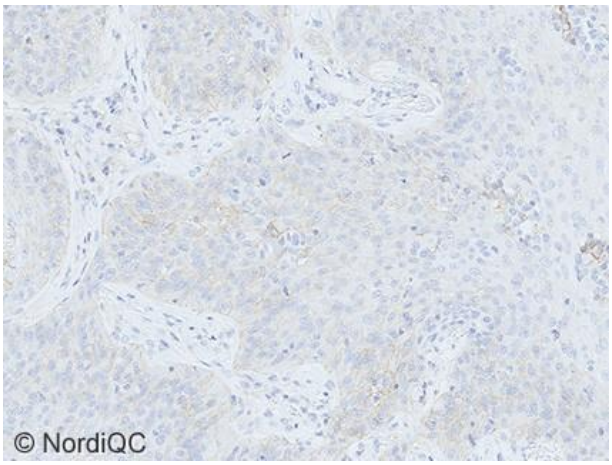


Fig. 4b
Insufficient staining result of the NSCLC, tissue core no. 8, using the PD-L1 IHC 22C3 pharmDx kit, SK006, Dako/Agilent "off-label" on a BenchMark Ultra with laboratory modified protocol settings.
A significant reduced proportion of tumour cells are demonstrated changing the TPS category from the expected high to low.
Also compare with Fig. 2b, same field.

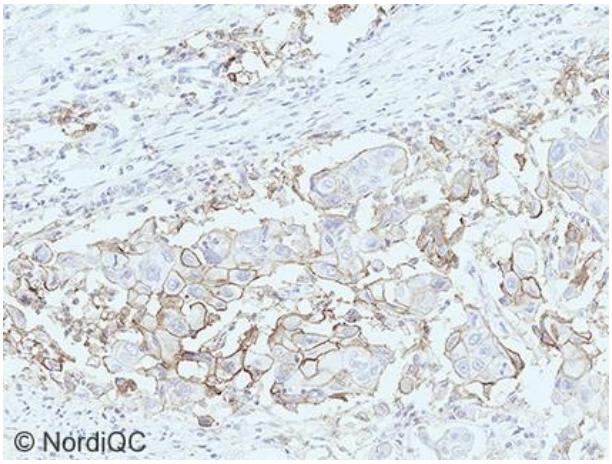


Fig. 5a
Optimal staining result of the urothelial carcinoma, tissue core no. 12, using the 22C3 pharmDx IHC PD-L1, SK006, Dako/Agilent on Autostainer Link 48 following the recommended protocol settings.
The majority of tumour cells and scattered immune cells in the adjacent stroma show a weak to moderate membranous staining reaction.
The tumour was categorized as CPS ≥10 and thus eligible for immune therapy with KEYTRUDA®.

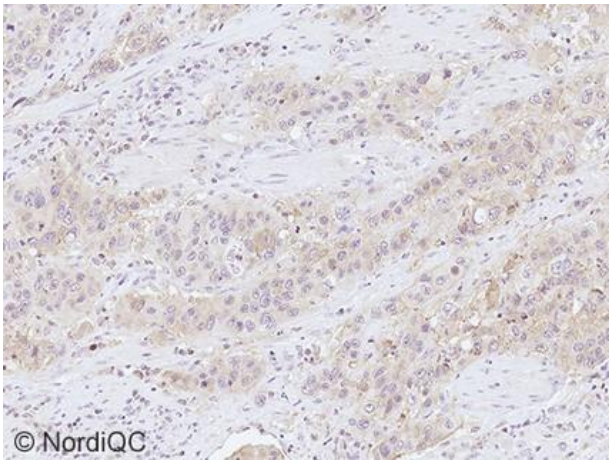


Fig. 5b
Insufficient staining result of the urothelial carcinoma, tissue core no. 12, using the mAb CAL10 by inappropriate protocol settings providing a poor-signal-to-noise ratio compromising the interpretation.
The PD-L1 status cannot reliable be determined due to an excessive cytoplasmic and aberrant nuclear staining reaction.

SN/LE/HLK/RR 07.07.2020