Assessment Run C6 2019

PD-L1 TECENTRIQ®

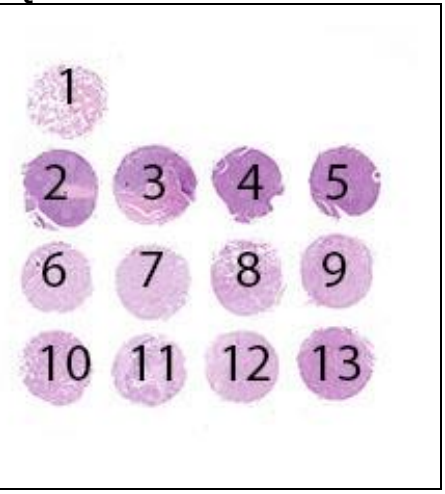
Purpose

This first assessment in the NordiQC Companion module of PD-L1 TECENTRIQ® primarily focused on evaluation of the analytical accuracy of the PD-L1 IHC assays performed by the participating laboratories to identify patients with urothelial carcinomas or triple negative breast carcinomas (TNBC) to be treated with TECENTRIQ® as immune therapy. The PD-L1 SP142 IHC assay (741-4860, Ventana) was used as reference standard method. Accuracy was evaluated in eight carcinomas with the dynamic and critical relevant expression levels of PD-L1 characterized by tumour-infiltrating immune cell score (IC). The assessment mark obtained 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, are needed.

Material

Table 1. **Content of the multi-block used for the NordiQC PD-L1 TECENTRIQ® C6 assessment**

Tissue controls	**PD-L1 IHC reaction pattern**
1. Placenta	See control section
2-5. Tonsil	See control section
Carcinomas	**IC score***
6. Urothelial carcinoma	<5%
7. Urothelial carcinoma	<5%
8. Urothelial carcinoma	≥5% (IC 5-10%)
9. Urothelial carcinoma	≥5% (IC 10-15%)
10. TNBC**	<1%/≥1%***
11. TNBC	<1%
12. TNBC	≥1% (IC 10-20%)
13. TNBC	≥1% (IC 5-15%)



* tumour-infiltrating immune cell score (IC) determined by PD-L1 SP142 IHC (741-4860, Ventana) performed in NordiQC reference lab.
** Triple negative breast carcinoma
*** Due to heterogeneity the IC category changed throughout one block.

All tissues were fixed in 10% neutral buffered formalin.

The participating laboratories were asked to perform the PD-L1 IHC assay accordingly to the protocol used in the laboratory and also interpret the PD-L1 expression level using IC as scoring method and submit these scores to NordiQC.
This allowed both an assessment of the technical performance / analytical accuracy of the PD-L1 IHC assays but also information on the reproducibility and concordance of the interpretation of PD-L1 expression among the laboratories.

PD-L1 TECENTRIQ® 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 through out the blocks. Reference slides were stained for PD-L1 using the CE IVD / FDA approved PD-L1 SP142 IHC assay (741-4860, Ventana). During the assessment, IC categories for each tissue core on the submitted slides were compared to the level in the nearest reference slide of PD-L1 (SP142) IHC assay (741-4860 Ventana).

Criteria for assessing a staining as Optimal include:
The staining is considered perfect or close to perfect in all of the included tissues.
IC score is concordant to the NordiQC reference data in all 8 carcinomas.

Criteria for assessing a staining as Good include:
The staining is considered acceptable in all of the included tissues. However, the protocol may be optimized to provide improved analytical accuracy, e.g. counter staining, morphology or signal-to-noise ratio.
IC score is still concordant to the NordiQC reference data in all 8 carcinomas.

Criteria for assessing a staining as Borderline include:
The staining is considered insufficient, e.g. because of a generally too weak staining, a false negative staining or false positive staining reaction of one of the included tissues. The protocol should be optimized.
IC score is **not** concordant to the NordiQC reference data in 1 of the 8 carcinomas.

Criteria for assessing a staining as Poor include:
The staining is considered very insufficient, e.g. because of a false negative or false positive staining reaction staining of more than one of the included tissues. Optimization of the protocol is urgently needed.
IC score is **not** found concordant to the NordiQC reference data in all 8 carcinomas.

A staining can also be assessed as **borderline/poor** in case the interpretation and scoring is significantly hampered by impaired morphology or excessive background reaction etc.

PD-L1 IHC, Interpretation
All participating laboratories were asked to submit a scoring sheet with their interpretation of the tumour-infiltrating immune cell score (IC) in the eight carcinomas. Results were compared to NordiQC data from the reference laboratory to analyze scoring consensus.

Participation

Number of laboratories registered for PD-L1 TECENTRIQ® IHC C6	95
Number of laboratories returning PD-L1 TECENTRIQ® IHC	85 (89%)*
Number of laboratories returning PD-L1 TECENTRIQ® scoring sheet	73 (86%)

*One laboratory was excluded from the assessment due to contamination with a CD45-like antibody, which compromised the assessment of the staining.

Results: 84 laboratories participated in this assessment and 76% achieved a sufficient mark. Assessment marks for IHC PD-L1 assays and PD-L1 antibodies are summarized in Table 2 (see page 3).

Conclusion
This was the first NordiQC assessment of PD-L1 for **TECENTRIQ®** in urothelial carcinoma and TNBC in the companion module. 84 laboratories participated and a pass rate of 76% was observed.
The PD-L1 SP142 companion diagnostic (CDx) IHC assays from Ventana were most successful assays for evaluation of PD-L1 status in urothelial carcinomas and TNBCs to guide treatment with TECENTRIQ® as immune therapy providing a pass rate of 94%.
Other CDx assays based on SP263, Ventana, 22C3, Dako and laboratory developed (LD) tests provided significantly lower pass rates of 67% and 65%, respectively. Insufficient results were typically characterized by a too strong staining reaction in tumour cells compromising the interpretation of immune cells. – most likely because these protocols have been developed and calibrated to primarily demonstrate PD-L1 expression in tumour cells and to imitate the performance of the Dako SK006 PharmDx 22C3 assay for treatment with KEYTRUDA®.
Tonsil is at present the preferred choice as positive and negative tissue control for IHC methods with the purpose to demonstrate PD-L1 for TECENTRIQ® using IC as scoring method. Tonsillar epithelial crypt cells must show a strong staining reaction, while germinal centre macrophages and lymphocytes and scattered interfollicular lymphocytes must show a moderate membranous staining reaction. No staining must be seen in superficial squamous epithelial.

Table 2. **Assessment marks for IHC assays and antibodies run C6, PD-L1 TECENTRIQ® for IC scoring**

CE-IVD / FDA approved PD-L1 assays	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
rmAb clone SP142, **740-4859**[4]	12	Ventana/Roche	11	-	-	1	92%	100%
rmAb clone SP142, **741-4860**[4]	17	Ventana/Roche	15	1	1	-	94%	94%
rmAb clone SP142, **790-4860**[5]	5	Ventana/Roche	5	-	-	-	100%	100%
rmAb clone SP263, **740-4907**[6]	2	Ventana/Roche	-	2	-	-	-	-
rmAb clone SP263, **741-4905**[7]	9	Ventana/Roche	1	6	2	-	78%	83%
rmAb clone SP263, **790-4905**[5]	9	Ventana/Roche	-	5	4	-	56%	-
mAb clone 22C3 pharmDX, **SK006**[8]	1	Dako/Agilent	-	-	1	-	-	-
mAb clone 22C3 pharmDX, **SK006**[3]	2	Dako/Agilent	-	1	-	1	-	-
mAb clone 22C3 pharmDX, **GE006**[9]	2	Dako/Agilent	1	-	1	-	-	-
mAb clone 22C3 pharmDX, **GE006**[3]	1	Dako/Agilent	-	1	-	-	-	-
rmAb clone 28-8 pharmDX, **SK005**[10]	1	Dako/Agilent	-	1	-	-	-	-
Antibodies[11] for laboratory developed PD-L1 assays, concentrated antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
mAb clone **22C3**	8	Dako/Agilent	-	6	2	-	75%	-
mAb clone **E1L3N**	4	Cell Signaling	-	3	1	-	-	-
rmAb **CAL10**	1 3	Biocare Zytomed Systems	1	1	2	-	-	-
rmAb clone **ZR3**	1 1 1	Zeta Corporation Nordic Biosite Gene Tech	-	1	2	-	-	-
rmAb **BSR90**	1	Nordic Biosite	-	1	-	-	-	-
Ready-To-Use antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
rmAb CAL10, **API3171**	1	Biocare	-	-	1	-	-	-
rmAb 73-10, **PA0832**	1	Leica Biosystems	-	1	-	-	-	-
rmAb clone MXR003**, RMA-0732**	1	Maixin	-	-	1	-	-	-
Total	84		34	30	18	2		
Proportion			40%	36%	22%	2%	76%	

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 Agilent/Dako's semi-automated system (Autostainer Link48) but used on different platforms.
4) Approved for TECENTRIQ in UCs and TNBCs in EU/US.
5) Analytical claim.
6) Approved for IMFINZI in UCs in US.
7) Approved for KEYTRUDA, IMFINZI and OPDIVO in NSCLCs and UCs in EU.
8) Approved for KEYTRUDA in NSCLCs, GEJs, ESCCs, Cervix carcinomas, UCs and HNSCCs in EU/US.
9) Approved for KEYTRUDA in NSCLCs in EU.
10) Approved for OPDIVO in NSCLCs, HNSCCs and UCs in EU/US.
11) mAb: mouse monoclonal antibody, rmAb: rabbit monoclonal antibody

Detailed Analysis
CE IVD / FDA approved assays

SP142 (740-4859/741-4860, Ventana): 26 of 29 (90%) 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 24-90 min.), 16-60 min. incubation of the primary Ab and OptiView with OptiView Amplification as detection system and performed on the BenchMark GX/Ultra. Using these protocol settings, 26 of 27 (96%) laboratories produced a sufficient staining result (optimal or good).

SP142 (790-4860, Ventana): 5 of 5 protocols were assessed as optimal. All protocols were based on HIER in CC1 typically for 32-64 min., 16 min. incubation of primary Ab and OptiView with OptiView Amplification as detection system.

SP263 (741-4905/740-4907, Ventana): One protocol was assessed as optimal. The protocol was based on HIER in CC1 (efficient heating time 64 min.), 16 min. incubation of the primary Ab, OptiView as detection

system and performed on the BenchMark Ultra. Using these protocol settings, 6 of 8 laboratories produced a sufficient staining result.

PD-L1 IHC 22C3 pharmDx (GE006, Dako): One protocol was assessed as optimal. The protocol was based on HIER in EnVision™ Flex Target Retrieval Solution low pH 6.1 (GV805) at 97°C for 40 min., 40 min. incubation of the primary Ab, linker and polymer and performed on the Omnis. Using these protocol settings, 1 of 2 laboratories produced a sufficient staining result.

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

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

CDx assays	**Vendor recommended protocol settings[1]**		**Laboratory modified protocol settings[2]**	
	Sufficient	Optimal	Sufficient	Optimal
Ventana BenchMark GX, Ultra rmAb SP142, **740-4859**	5/5	5/5	6/7	6/7
Ventana BenchMark Ultra rmAb SP142, **741-4860**	9/10	8/10	7/7	7/7
Ventana BenchMark Ultra rmAb SP142, **790-4860**	-	-	5/5	5/5
Ventana BenchMark Ultra rmAb SP263, **740-4907**	1/1	0/1	1/1	0/1
Ventana BenchMark XT, GX, Ultra rmAb SP263, **741-4905**	5/7	1/7	2/3	0/3
Ventana BenchMark XT, GX, Ultra rmAb SP263, **790-4905**	2/6	0/6	2/2	0/2
Dako Autostainer Link 48+ mAb 22C3 pharmDX, **SK006**	0/1	0/1	-	-
Dako Omnis mAb 22C3 pharmDX, **GE006**	1/2	1/2	-	-
Dako Autostainer Link 48+ mAb 28-8 pharmDX, **SK005**	0/1	0/1	-	-

1) Protocol settings recommended by vendor – Retrieval method and duration, Ab incubation times, detection kit, IHC stainer/equipment.
2) 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

rmAb **CAL10**: One protocol was assessed as optimal. The staining was performed on a BenchMark Ultra (Ventana), based on HIER in CC1 for 64 min., Ab titre of 1:50, primary Ab incubation time of 32 min. and OptiView with OptiView Amplification as detection system. Only this laboratory used these protocol settings.

Table 4. **Optimal results for PD-L1 for the most commonly used antibody as concentrate on the 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
rmAb clone **CAL10**	1/1**	-	-	-	-	-	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

Comments – accuracy of PD-L1 IHC using IC scoring to guide treatment with TECENTRIQ®
In this first NordiQC run for PD-L1 TECENTRIQ® in the companion module C6, a pass rate of 76% was observed for the participants performing PD-L1 IHC assays to identify patients with urothelial carcinomas and triple negative breast carcinomas (TNBC) to be treated with TECENTRIQ® as immune therapy using the tumour-infiltrating immune cell score (IC) as scoring method.

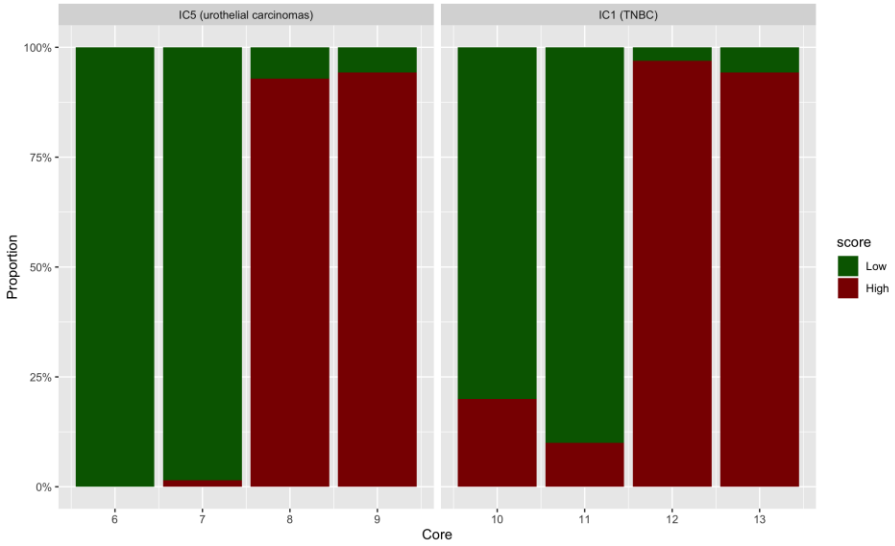
Insufficient results were most frequently characterized by a reduced proportion and too weak specific staining of immune cells combined with an excessive staining of tumour cells compromising the interpretation. This was observed in 75% (15 of 20) of the insufficient staining results. The remaining insufficient results were caused by either a too weak or false negative staining (4 of 20, 20%), or a poor-signal-noise ratio (1 of 20, 5%)

The Ventana PD-L1 SP142 assays 740-4859/741-4860 provided a pass rate of 93% (14 of 15) when applying protocol settings in compliance with the vendor recommendations. When modifying the protocol, a pass rate of 93% (13 of 14) was obtained (see Table 3). The most frequent modification was reduced HIER time. Despite the high pass rate, it must be emphasized that modifications of approved companion diagnostic assays can provide inaccurate test results and require an extended and often challenging internal validation.

Non-SP142 companion diagnostic assays as SP263 (Ventana), 22C3 pharmDx (Dako) and 28-8 pharmDx (Dako) gave an overall inferior performance and reduced pass rate compared to the SP142 assays from Ventana. This was observed both using these assays according to the vendor recommended protocol settings and by modified protocols as shown in Table 3. If the assays were used strictly to the recommended protocol settings, a pass rate of 50% (9 of 18) was observed and only 12% optimal (2 of 18). For SP263, Ventana a higher pass rate of 100% (5 of 5) was obtained by laboratory modified protocols, but none was assessed as optimal. Although caution must be taken due to the low number of observations, these data indicate a challenge for interchangeability of SP142 with other companion diagnostic assays. Most likely, the detection system applied for SP142 based on OptiView with amplification (tyramide based) and the calibration of the antibody SP142 in the Ventana assay provides a performance that intensifies demonstration of immune cells and reduces staining of tumour cells.

Overall, laboratory developed (LD) assays based on concentrated Abs, "non-companion diagnostic approved" RTU formats or a companion diagnostic assay not used strictly accordingly to the recommended protocol settings provided by the vendor were used by 64% (54 of 84) of the participants, and for this group a pass rate of 74% (40 of 54) was observed.
For the LD assays based on concentrates, only one clone obtained an optimal staining result. The rmAb CAL10 carefully calibrated using OptiView with amplification as detection system (same as the SP142 assay) provided a similar staining pattern as the NordiQC reference stains using rmAb SP142, 741-4860, as recommended by Ventana.
However, it has to be emphasized that both off-label use of companion diagnostic assays e.g. using a companion diagnostic assay for other intended use as approved claims, modifying the protocol settings for an approved companion diagnostic assay or using a LD assay for PD-L1 status to guide treatment stratification must be meticulously validated by the laboratory.

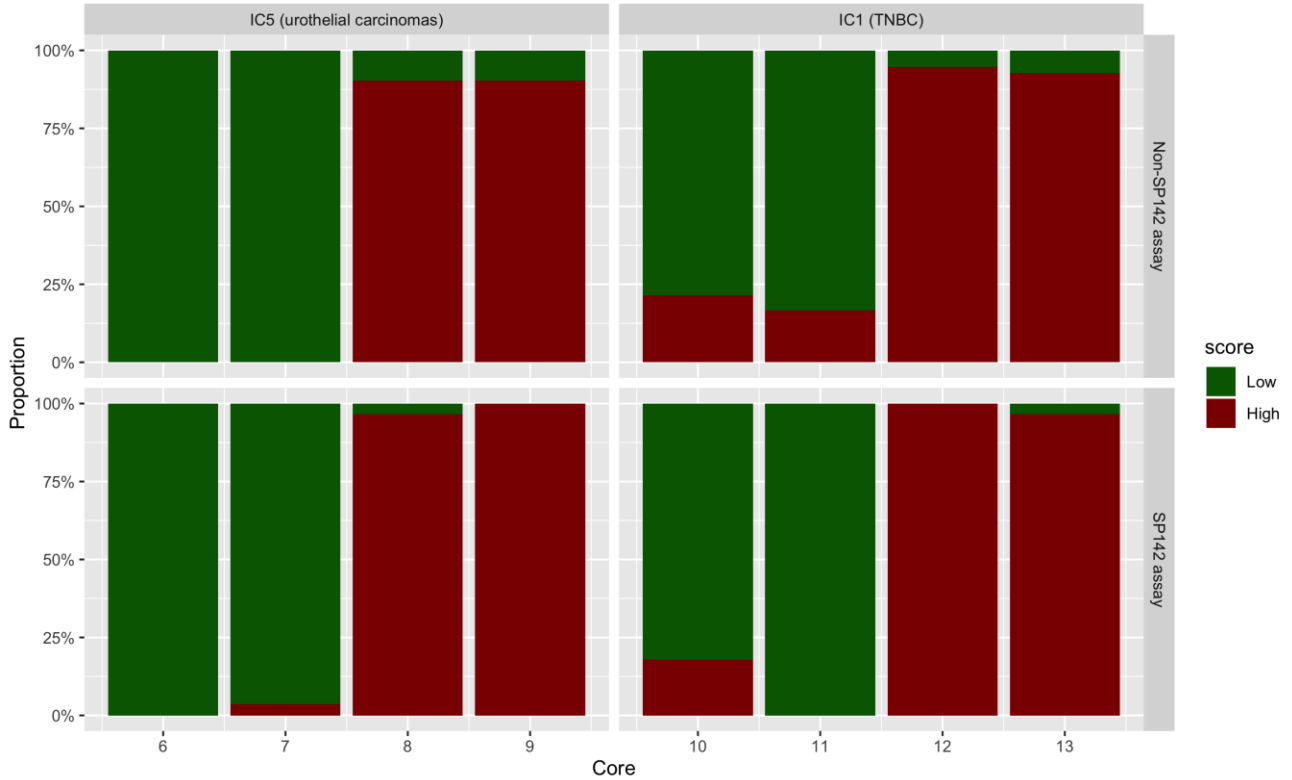
PD-L1 scoring
Participants were asked to evaluate the IC score in each of the four urothelial carcinomas (IC with 5% cut-off) and four TNBC (IC with 1% cut-off) included in the assessment material. The overall interpretation of the PD-L1 expression among the participants is shown in Graph 1.



Graph 1. NordiQC PD-L1 run C6: Interpretation of IC in four urothelial carcinomas and four TNBC.

As seen in Graph 1, tissue core no. 10 (TNBC) had the highest level of disagreement in IC scoring. This was not surprising, as this core due to heterogeneity in PD-L1 expression changed IC-category through the block. It is important to underline that submitted slides during the assessment was compared to the nearest reference slide and no participants was downgraded based on this heterogeneity. Tissue cores 11, 8 and 9 also challenged the laboratories but to a lesser extent than tissue core no. 10.

The agreement rate and concordance to the NordiQC reference scores were somewhat higher among the laboratories that stained the slides using a protocol based on the SP142 Ab compared to "non-SP142" assays, see Graph 2. This was most markedly seen in tissue core no. 11 (TNBC), where all participants using a protocol based on SP142 correctly scored the core as IC <1%, whereas 16% of participants using non-SP142 based protocols marked this as IC ≥1%.



Graph 2. NordiQC PD-L1 run C6: Interpretation of IC in four urothelial carcinomas and four TNBC separated on assays based on either SP142 or other Abs

Controls

Tonsil and placenta were used as positive and negative tissue controls. In this assessment and in concordance with the official scoring guidelines from Ventana, tonsil was found to be superior to placenta. Crypt epithelium in tonsil should display a strong staining reaction, while a moderate to strong staining reaction should be seen in lymphocytes and macrophages in germinal centres and scattered lymphocytes in the interfollicular regions. No staining reaction should be seen in superficial squamous epithelium. In this assessment, it was observed that a moderate staining reaction in scattered lymphocytes in the interfollicular region were most challenging for the participants and could only be detected with an optimal protocol. Placenta did not contain the same dynamic PD-L1 expression range. Only throphoblast cells with a high-level PD-L1 expression were stained.

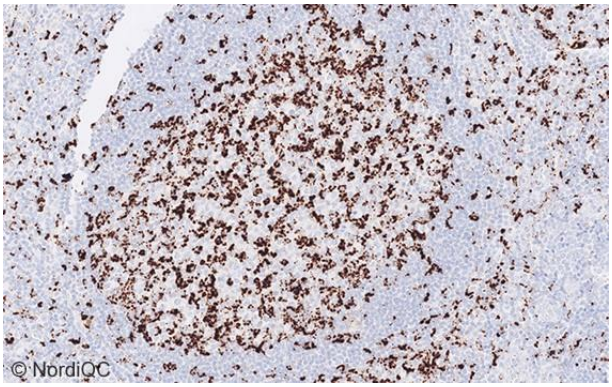


Fig. 1a
Optimal staining result of tonsil (germinal centre) using the PD-L1 assay 741-4860, Ventana, based on the rmAb clone SP142 following the recommended protocol settings. Same protocol used in Figs. 2a-6a. Germinal centre macrophages and lymphocytes show a moderate to strong staining reaction.

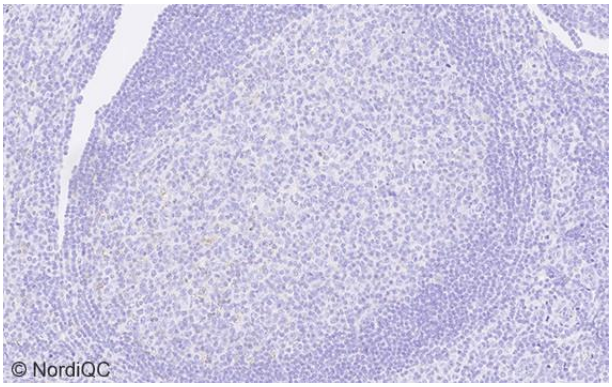


Fig. 1b
Insufficient staining result of tonsil (germinal centre) using the pharmDx IHC PD-L1 assay, SK006, Dako/Agilent based on the mAb clone 22C3 approved for treatment with KEYTRUDA®. Same protocol in Figs. 2b-6b. Virtually all germinal centre lymphocytes are negative and only a faint staining reaction is seen in scattered macrophages. Compare with optimal result shown in Fig. 1a – same area.

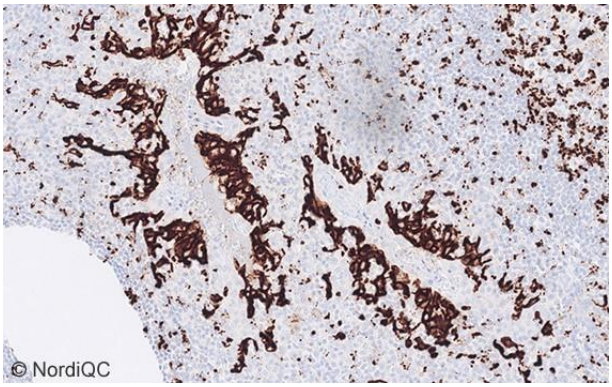


Fig. 2a
Optimal staining result of tonsil (crypt epithelium) using same protocol as in Fig. 1a. A strong staining reaction is seen in the majority of epithelial crypt cells.

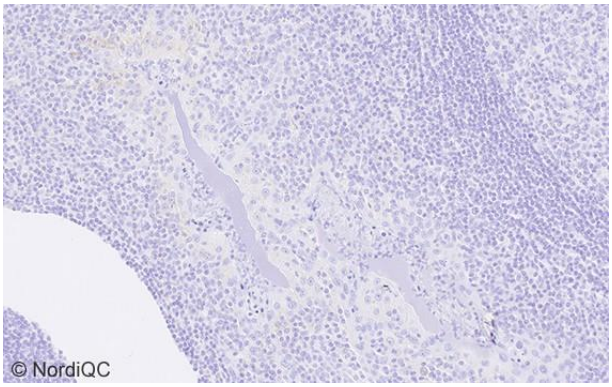


Fig. 2b
Insufficient staining result of tonsil (crypt epithelium) using same protocol as in Fig. 1b. Virtually all epithelial crypt cells are negative. Compare with optimal result shown in Fig. 2a – same area.

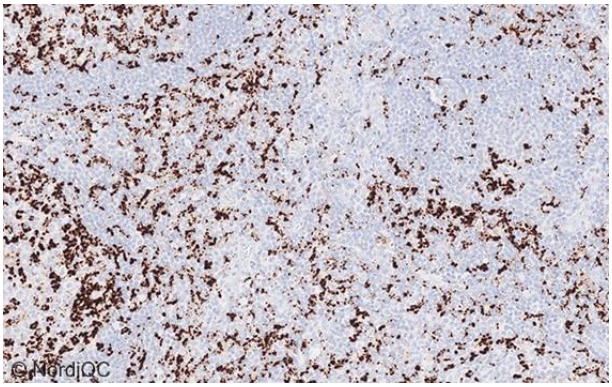


Fig. 3a
Optimal staining result of tonsil (interfollicular lymphocytes) using same protocol as in Figs. 1a-2a. A moderate staining reaction is seen in scattered lymphocytes.

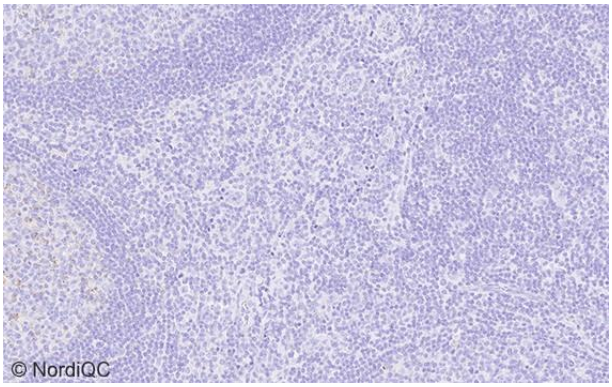


Fig. 3b.
Insufficient staining result of tonsil (interfollicular lymphocytes) using same protocol as in Figs. 1b-2b. Virtually all lymphocytes are negative. Compare with optimal result shown in Fig. 3a – same area.

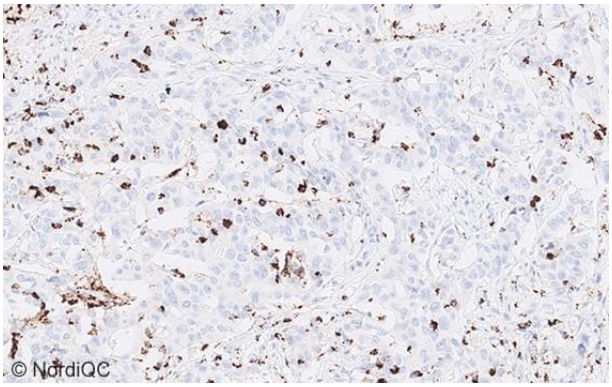


Fig. 4a
Optimal staining result of urothelial carcinoma tissue core no. 9, using same protocol as in Figs. 1a-3a. Virtually all tumour cells are negative and immune cells show a moderate to strong staining reaction giving an IC score ≥5%.

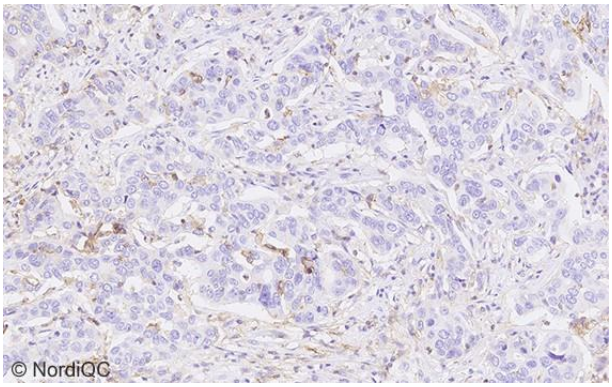


Fig. 4b
Staining result of urothelial carcinoma tissue core no. 9, using same protocol as in Figs. 1b-3b. A faint staining reaction is observed in scattered tumour cells. Immune cells show a predominantly moderate to staining reaction, however the proportion of positive lymphocytes is reduced when comparing to optimal result shown in Fig. 4a.

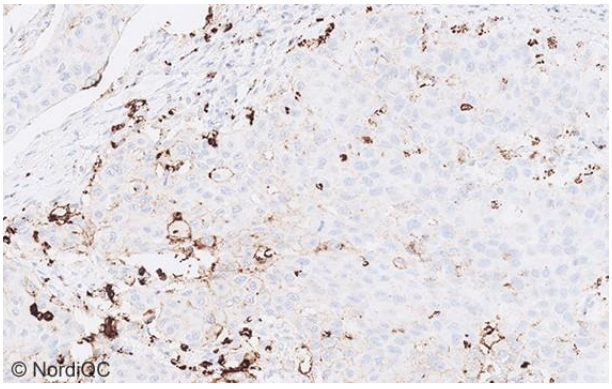


Fig. 5a
Optimal staining result of urothelial carcinoma tissue core no. 8, using same protocol as in Figs. 1a-4a. A faint staining reaction is seen in scattered tumour cells. Immune cells display a moderate to strong staining reaction giving an IC score ≥5%.

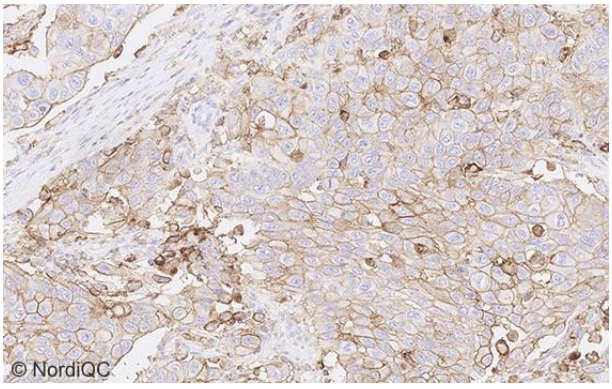


Fig. 5b
Insufficient staining result of urothelial carcinoma tissue core no. 8, using same protocol as in Figs. 1b-4b. A too strong staining reaction is obtained in tumour cells interfering the interpretation of immune cells. Compare with optimal result shown in Fig. 5a – same area.

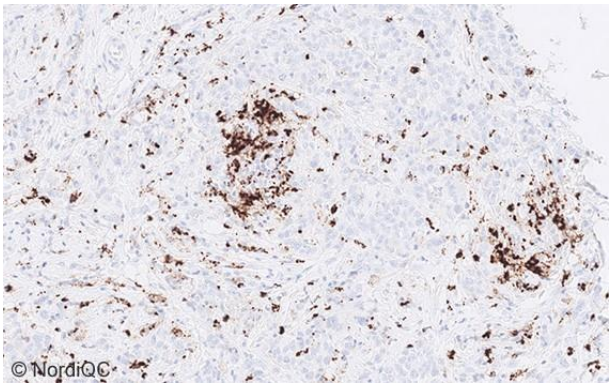


Fig. 6a
Optimal staining result of TNBC tissue core no. 12, using same protocol as in Figs. 1a-5a. A strong staining reaction is seen in immune cells giving an IC score ≥1%.

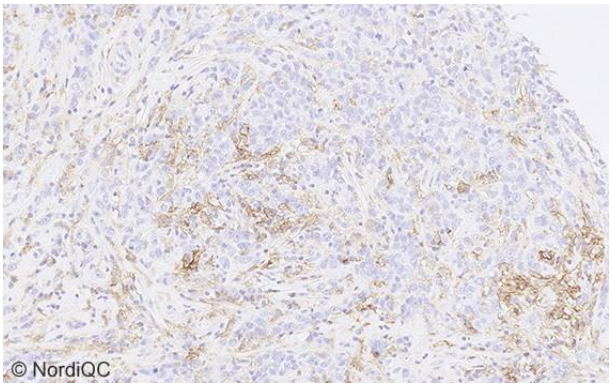


Fig. 6b
Staining result of TNBC tissue core no. 12, using same protocol as in Figs. 1b-5b. A weaker and less distinct staining reaction is observed compared to optimal staining result in Fig. 6a.

HLK/SN/LE/RR 10-12-19