

Assessment Run 67 2023 p53

Purpose

Evaluation of the technical performance and level of analytical sensitivity and specificity of p53 IHC tests among NordiQC participants for the demonstration of corresponding TP53 mutations in endometrial carcinomas. Relevant clinical tissues, both normal and neoplastic, were selected to display a broad spectrum of antigen densities for p53 (see below).

Material

The slide to be stained for p53 comprised:

- 1. Appendix, 2. Tonsil, 3. Endometrial serous carcinoma with p53 overexpression,
- 4. Low grade endometrial carcinoma p53 wild type, 5. Endometrial serous carcinoma with absence of p53.



All tissues were fixed in 10% neutral buffered formalin.

Criteria* for assessing a p53 staining as optimal included:

- A weak to moderate nuclear staining reaction in ≥50% of the germinal centre B-cells of the tonsil.
- A weak to moderate nuclear staining reaction in dispersed epithelial cells in the basal crypts of the appendix.
- A moderate to strong, distinct nuclear staining reaction in virtually all the neoplastic cells of the ovarian serous carcinoma with p53 overexpression (tissue core no. 3).
- No staining reaction in the neoplastic cells in the endometrial serous carcinoma with absence of p53 expression (tissue core no. 5). Dispersed stromal cells, lymphocytes and endothelial cells must show an at least weak nuclear staining reaction.
- A weak to moderate, distinct nuclear staining reaction in the majority of neoplastic cells in the low grade endometrial carcinoma (tissue core no. 4). Dispersed stromal cells, lymphocytes and endothelial cells should show an at least weak nuclear staining reaction.
- No staining of the luminal epithelial cells in the appendix and <50% of the mantle zone B-cells showing maximum a weak to moderate nuclear staining reaction.

Participation

Number of laboratories registered for p53, run 67	417
Number of laboratories returning slides	372 (89%)

Results

At the date of assessment, 89% of the participants had returned the circulated NordiQC slides. All slides returned after the assessment were assessed and laboratories received advice if the result was insufficient, but the data were not included in this report.

372 laboratories participated in this assessment. 65% achieved a sufficient mark (optimal or good). Table 1 summarizes the antibodies (Abs) used and assessment marks (see page 3).

The most frequent causes of insufficient staining reactions were:

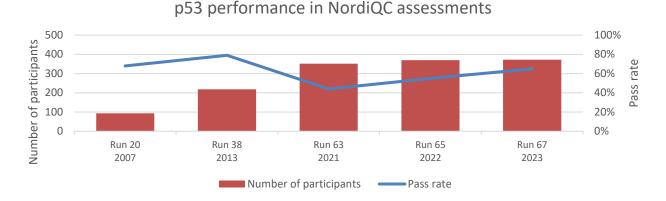
- Use of a less sensitive detection system.
- Poor-signal-to-noise ratio (mainly due to problems with the Envision Flex detection system).
- Too low concentration of the primary antibody.
- Use of less successful vendor recommended protocol settings for Ready-To-Use systems.

^{*} The criteria and expected staining patterns were based on the previous NordiQC assessments and the publication by Köbel et al; Interpretation of P53 Immunohistochemistry in Endometrial Carcinomas: Toward Increased Reproducibility. Int J Gynecol Pathol Vol. 38, No. 1 Supplement 1, January 2019, S123-S131

Performance history

This was the fifth NordiQC assessment of p53. An increased pass rate was seen compared to run 63 and 65, but still reduced compared to the previous assessments (see Graph 1). From run 63 both the purpose, scoring criteria of the included neoplasias and composition of the assessment material has changed and thus being more challenging than previously.

Graph 1. Proportion of sufficient results for p53 in the five NordiQC runs performed



Conclusion

The mAb clones **BP53-12** and **DO-7** could both be used to obtain optimal staining result for p53. The most widely used antibody, the mAb clone DO-7 gave optimal staining results on all the main IHC systems from Dako/Agilent, Ventana/Roche and Leica Biosystems. For all the clones efficient HIER, careful calibration of the primary antibody titer and in particular usage of a 3-layer detection system were mandatory for optimal performance. 75% (310 of 372) of the participants used a 3-layer detection system, with an overall pass rate at 74% (230 of 310), 34% optimal (n=106) compared to a pass rate of 19% (12 of 62), 3% optimal (n=2) if using a 2-layer detection system.

Table 1. Antibodies and assessment marks for p53, Run 67

Table 1. Antibodies and	asse	ssment marks for p53	, Run 67					
Concentrated antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff. ¹	OR ²
mAb clone DO-7	Diagnostic Biosystems I Immunologic Zeta Corporation Epredia		35	30	9	9	78%	42%
mAb clone DO-7+BP53-12	3	Thermo Scientific	-	2	1	-	-	-
mAb clone BP53-12		PathnSitu Zytomed Systems GMD	-	1	2	-	-	-
Ab clone BPM6168	1	Biolynx Biotechnology	-	-	-	1	-	-
Conc total			35	33	12	10	76%	39%
Ready-To-Use antibodies							Suff. ¹	OR. ²
mAb clone BP53-11 760-2542 (VRPS) ³	5	Ventana/Roche	-	3	2	-	60%	-
mAb clone BP53-11 760-2542 (LMPS) ⁴	45	Ventana/Roche	7	16	16	6	51%	16%
mAb clone DO-7 800-2912 (VRPS) ³	7	Ventana/Roche	1	3	2	1	57%	14%
mAb clone DO-7 800-2912 (LMPS) ⁴	85	Ventana/Roche	40	23	16	6	74%	47%
mAb clone DO-7 IS/IR616 (VRPS) ³	4	Dako/Agilent	1	1	-	2	-	-
mAb clone DO-7 IS/IR616 (LMPS) ⁴	28	Dako/Agilent	11	7	4	6	64%	39%
mAb clone DO-7 GA616 (VRPS) ³	10	Dako/Agilent	-	1	-	9	10%	-
mAb clone DO-7 GA616 (LMPS) ⁴			9	39	9	5	77%	15%
mAb clone DO-7 PA0057 (VRPS) ³		Leica Biosystems	1	5	6	1	46%	8%
mAb clone DO-7 PA0057 (LMPS) ⁴		Leica Biosystems	2	3	5	2	42%	17%
mAb clone DO-7 453M-9 x	1	Cell Marque	-	-	-	1	-	-
mAb clone DO-7 PDM013	1	Biosystems	-	-	1	-	-	-
mAb clone BP53-12 BMS064	1	Zytomed Systems	-	-	1	-	-	-
mAb clone C2H10 CPM-0142	1	Celnovte	1	-	-	-	-	-
mAb clone DO-7 AM239-5M	1	BioGenex	-	-	-	1	-	-
rmAb clone SP5 MAD-000309QD	2	Master Diagnostica	-	-	1	1	-	-
mAb clone 882F5H1 PA172	1	Abcarta	-	-	-	1	-	-
mAb clone DO-7 BFM-0002	1	Bioin Biotechnology	-	-	1	-	-	-
Other	2		-	-	2	-	-	-
RTU total	282		73	101	66	42	62%	26%
Total	372		108	134	78	52		
Proportion	L		29%	36%	21%	14%	65%	

¹⁾ Proportion of sufficient results (optimal or good). (≥5 asessed protocols).

²⁾ Proportion of Optimal Results (OR).

³⁾ Vendor Recommended Protocol Settings (VRPS) to a specific RTU product applied on the vendor recommended platform(s) (≥5

asessed protocols).

4) Laboratory Modified Protocol Settings (LMPS) to a specific RTU product applied either on the vendor recommended platform(s), non-validated semi/fully automatic systems or used manually (≥5 asessed protocols).

Detailed analysis of p53, Run 67

The following protocol parameters were central to obtain optimal staining:

Concentrated antibodies

mAb clone **DO-7**: Protocols with optimal results were based on Heat Induced Epitope Retrieval (HIER) using Cell Conditioning 1 (CC1, Ventana/Roche) (28/41)*, Bond Epitope Retrieval 2 (BERS2, Leica Biosystems) (5/18), Target Retrieval Solution (TRS) pH 9 (Dako/Agilent) (2/16) as retrieval buffer. The mAb was diluted in the range of 1:50-1:2.000 depending on the total sensitivity of the protocol employed. Using these protocol settings 57 of 71 (80%) laboratories produced a sufficient staining (optimal or good). * (number of optimal results/number of laboratories using this buffer)

Table 2. Proportion of optimal results for p53 for the most commonly used antibody as concentrate on the four main IHC systems*

Concentrated antibody	Dako/Agilent Autostainer		Dako/Agilent Omnis		Ventana/Roche BenchMark XT / Ultra		Leica Biosystems Bond III / Max	
	TRS pH	TRS pH	TRS pH	TRS pH	CC1 pH	CC2 pH	ER2 pH	ER1 pH
	9.0	6.1	9.0	6.1	8.5	6.0	9.0	6.0
mAb clone	0/5**	0/1	2/11	_	28/41	_	5/18	0/6
DO-7	(0%)	0/1	(18%)	_	(68%)	-	(28%)	(0%)

^{*} Antibody concentration applied as listed above, HIER buffers and detection kits used as provided by the vendors of the respective systems.

** (number of optimal results/number of laboratories using this buffer).

Ready-To-Use antibodies and corresponding systems

mAb clone BP53-11, product no. 760-2542, Ventana/Roche, BenchMark XT/Ultra:

Protocols with optimal results were based on HIER using CC1, efficient heating time 40-48 min. and 20-36 min. incubation of the primary Ab. OptiView (760-700) was used as detection system. Using these protocol settings 12 of 14 (86%) laboratories produced a sufficient staining (optimal or good).

mAb clone DO-7 product no. 800-2912, Ventana/Roche, BenchMark XT/Ultra/GX:

Protocols with optimal results were typically based on HIER using CC1, efficient heating time 32-64 min., 12-48 min. incubation of the primary Ab and OptiView (760-700) or UltraView (760-500) with amplification (760-080) as detection system. Using these protocol settings 58 of 67 (87%) laboratories produced a sufficient staining.

mAb clone **DO-7**, product no. **IS/IR616**, Dako/Agilent, Dako Autostainer:

Protocols with optimal results were based on HIER in PT-Link using TRS pH 9 (3-in-1), efficient heating time 20 min. at 97°C, 20 min. incubation of the primary Ab and EnVision FLEX+ (K8002) as detection system. Using these protocol settings 7 of 9 (78%) laboratories produced a sufficient staining.

mAb clone **DO-7**, product no. **GA616**, Dako/Agilent, Dako Omnis:

Protocols with optimal results were typically based on HIER in PT-Link using TRS High pH, efficient heating time 30 min., and 20-28 min. incubation of the primary Ab and EnVision FLEX+ (GV800/GV823) as detection system. Using these protocol settings 35 of 40 (88%) laboratories produced a sufficient staining.

mAb clone **DO-7** product no. **PA0057**, Leica Biosystems, Leica Bond III:

Protocols with an optimal result were based on HIER using BERS2, efficient heating time 20-30 min., 15-30 min. incubation of the primary Ab and Bond Polymer Refine Detection (DS9800) as detection system. Using these protocol settings 11 of 19 (58%) laboratories produced a sufficient staining.

Table 3 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 according to the vendor recommendations and by laboratory modified systems changing basal protocol settings. Only protocols performed on the intended IHC stainer device are included.

Table 3. Proportion of sufficient and optimal results for p53 for the most commonly used RTU IHC systems

RTU systems	Recommended protocol settings*		Laboratory modified protocol settings**		
	Sufficient			Optimal	
Ventana Benchmark mAb clone BP53-11 , 760-2542	60% (3/5)	0% (0/5)	51% (23/45)	16% (7/45)	
Ventana Benchmark mAb clone DO-7 , 800-2912	57% (4/7)	14% (1/7)	74% (63/85)	47% (40/85)	
Dako Autostainer mAb clone DO-7, IS/IR616	50% (2/4)	25% (1/4)	60% (12/20)	40% (8/20)	
Dako Omnis mAb clone DO-7, GA616	10% (1/10)	0% (0/10)	79% (44/56)	13% (7/56)	
Leica Bond mAb clone DO-7 , PA0057	46% (6/13)	8% (1/13)	46% (5/11)	18% (2/11)	

^{*} Protocol settings recommended by vendor – Retrieval method and duration, Ab incubation times, detection kit, IHC stainer/equipment.

** Significant modifications: retrieval method, retrieval duration and Ab incubation time altered, detection kit – only protocols performed on the specified vendor IHC stainer integrated.

Comments

In this assessment and in concordance with the previous assessment of p53, the prevalent feature of an insufficient result was a too weak or false negative staining reaction of the cells and structures expected to be demonstrated. Too weak or a complete false negative staining reaction was seen in 92% of the insufficient results (120/130). The remaining 8% (10/130) insufficient results were characterized by either a false positive staining reaction (n=5) or poor signal-to-noise ratio/excessive background (n=5). Virtually all laboratories were able to demonstrate p53 in the endometrial serous carcinoma with p53 overexpression (tissue core no. 3) and high-level antigen expression. On the contrary, the demonstration of p53 in low-level structures as stromal cells, lymphocytes and endothelial cells in the endometrial serous carcinoma with loss of p53 (tissue core no. 5), the neoplastic cells in the p53 wild-type low grade endometrial carcinoma (tissue core no. 4) and germinal centre and mantle zone B-cells in tonsil were much more challenging and required a carefully calibrated protocol.

24% (90/372) of the laboratories used an Ab as concentrated format within a laboratory developed (LD) assay for p53. 92% (83/90) of the LD assays were based on the mAb clone DO-7 with a pass rate of 78% and optimal results could be obtained on the three main IHC systems (see Table 2). The main prerequisites for optimal and sufficient staining results were HIER in an alkaline buffer, careful calibration of the titer of the primary Ab and use of a sensitive detection system. If using a 3-layer polymer/multimer based system as EnVision FLEX+ (Dako/Agilent), OptiView (Ventana/Roche) or Refine (Leica Biosystems), a pass rate of 74% (64/86) was observed, 41% optimal (n=35) compared to a pass rate of 25% (1/4), none optimal when using a 2-step detection system as EnVision FLEX (Dako/Agilent) or UltraView (Ventana/Roche) (also see Table 4). Especially OptiView (Ventana/Roche) with a carefully calibrated titer of the primary Ab performed very well with a pass rate of 88% (36/41), 68% optimal.

Ready-To-Use (RTU) antibodies were used by 76% (282 of 372) of the laboratories. Overall, it was observed that the pass rates and proportion of optimal results were low for the RTU systems from the three main IHC providers, Dako/Agilent, Ventana/Roche and Leica Biosystems, when these were applied by vendor recommended protocol settings (VRPS) - see Tables 1 and 3. If the RTU systems from these three vendors were used by VRPS an overall pass rate of 38% was seen and only 8% optimal.

The Dako/Agilent RTU system for Omnis, based on mAb clone DO-7 was most successful if modifying the protocol settings, giving an overall pass rate of 79% (44 of 56). However, if following the vendor recommended protocol settings, a pass rate of only 10% was seen (1 of 10) (see Table 3). The vendor recommended protocol was based on HIER in TRS High pH for 30 min., 20 min. incubation of the primary Ab and EnVision FLEX as detection system. The most successful modification was adding a mouse linker to the detection system and thus "upgrading to EnVision FLEX+. If using EnVision FLEX+, a pass rate at 85% (44 of 52) was obtained.

The same pattern was seen for the Dako Autostainer RTU system, also based on mAb clone DO-7. If using the 3-layer EnVision FLEX+, a pass rate 75% (12 of 16) was observed, compared to 25% (2 of 8) if using EnVision FLEX as detection system (see Table 4).

For protocols performed on Dako Omnis a poor-signal-to-noise ratio or excessive background frequently was observed influencing the staining performance and in total 31 of the participants had issues where the extent and intensity of the background/cytoplasmic staining reaction impacted the interpretation of p53 status (see Figs. 6a+b). At present NordiQC and the assessors have been informed by Dako/Agilent that certain lot numbers of EnVision Flex detection systems can cause an increased background reaction and Dako/Agilent is working on a solution. Participants observing these problems should contact Dako/Agilent to solve the issue.

Data from Table 4 underlines the importance of using a 3-layer detection system applying the different p53 clones both as concentrates and as RTU formats.

Table 4. Summarization of the proportion of sufficient and optimal marks using either 2- or 3-layer detection systems**.

		2-layer dete	ection system	3-layer detection system			
Antibodies n		Sufficient	Optimal	Sufficient	Optimal		
mAb conc DO-7	90	25% (1/4)	0% (0/4)	74% (64/86)	41% (35/86)		
mAb RTU BP53-11 760-2542* Ventana/Roche	50	15% (2/13)	0% (0/13)	65% (24/37)	19% (7/37)		
mAb clone RTU DO-7 800-2912* Ventana/Roche	92	8% (1/12)	8% (1/12)	83% (66/80)	50% (40/80)		
mAb clone RTU DO-7 IS/IR616* Dako/Agilent	32	30% (3/10)	10% (1/10)	77% (17/22)	50% (11/22)		
mAb clone RTU DO-7 GA616* Dako/Agilent	72	16% (3/17)	0% (0/17)	84% (46/55)	16% (9/55)		

^{*}Only protocols performed on the intended IHC stainer device are included.

The Ventana/Roche RTU system based on mAb clone DO-7 was the most widely used RTU system. The VRPS were based on either UltraView or OptiView as detection system. Using UltraView, the protocol was based on HIER in CC1 for 64 min. and primary Ab incubation time of 24 or 28 min. for BenchMark Ultra(plus) or XT, respectively. Using OptiView, the protocol was based on HIER in CC1 for 32 min. and primary Ab incubation time of 16 min. The majority of laboratories modified the protocol settings as shown in Table 3. The most common modifications were prolonging HIER and incubation time of primary Ab. If using UltraView, only one laboratory was able to produce a sufficient staining result. If using OptiView or UltraView with Amplification, a pass rate at 83% (66 of 80) was obtained.

The RTU system based on mAb clone BP53-11 from Ventana/Roche displayed very similar results as the DO-7 clone although the pass-rate was slightly decreased compared to the more popular DO-7 product, providing a pass rate of 52% (26 of 50). The VRPS were very similar also based on either UltraView or OptiView as detection system. Five laboratories used the vendor recommended protocol settings with a pass rate of 60%, none optimal. 45 laboratories used a modified protocol typically prolonging incubation time of both HIER and primary Ab with a pass rate of 51%, 16% optimal. Overall the performance of the Ventana/Roche RTU system mAb BP53-11 was inferior to the Ventana/Roche RTU system based on mAb clone DO-7 (see Tables 1 and 3).

The Leica Biosystems RTU system based on mAb clone DO-7 provided almost identical performance when applying the VRPS (IHC protocol F) compared to laboratory modified settings. A total of 13 laboratories applied VRPS with a pass-rate of 46% but only one optimal. The remaining 12 laboratories applying laboratory modified protocol settings obtained a pass rate of 42%, 18% optimal.

In this assessment it was observed that protocols based on a concentrated format provided a slightly higher pass rate (76%) than the corresponding RTU systems (62%). The inferior performance of RTU systems was mainly caused by the application of VRPS giving a too low analytical sensitivity and not being calibrated for the updated and extended diagnostic use to identify both p53 overexpression and p53 protein loss. A small increase in the overall pass-rate was observed from the last NordiQC assessment run 65 in 2022 (see Graph 1). However, the pass-rate is still low compared to run 38 with a pass-rate of 79%. The clear discrepancy is most likely influenced by the altered focus for usage of IHC for p53 in endometrial

^{**}Regardless of the protocol settings applied e.g., HIER time and/or incubation time in the primary Ab (≥10 protocols assessed).

carcinomas and awareness of more TP53 mutations being present with different p53 expression patterns. Previously the intended use of IHC for p53 mainly focused on the demonstration of p53 overexpression caused by TP53 mutations, but at present also the TP53 mutations with loss of p53 expression must be identified. In the latter an increased demand for the p53 IHC test also to consistently demonstrate p53 expression in internal cells is induced and a recalibration of the IHC test must typically be performed.

Controls

Tonsil and appendix are the most recommendable external positive and negative tissue controls. As a guideline for an accurate p53 IHC test more than 50% of germinal centre B-cells must show a weak to moderate nuclear staining reaction, while less than 50% of the mantle zone B-cells should be demonstrated in tonsil. In appendix, dispersed epithelial cells in the basal parts of the crypts must show a weak to moderate nuclear staining reaction, while the luminal epithelial cells must be negative. In addition, it has to be emphasized, that stromal cells, lymphocytes and endothelial cells in the clinical samples are essential as internal positive tissue controls especially for carcinomas with TP53 mutations causing absence and loss of p53 expression in the tumour cells (see Figs. 5a and 5b).

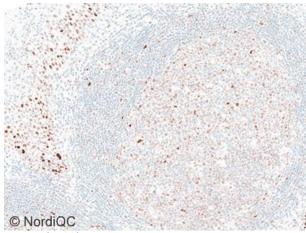


Fig. 1a (x100)

Optimal p53 staining of the tonsil using the mAb clone DO-7 as concentrate performed on Benchmark Ultra, Ventana/Roche, using the Ab in a dilution of 1:1000 with OptiView as detection system.

A weak to moderate nuclear staining reaction is seen in most of the germinal centre B-cells, whereas <50% of the mantle zone B-cells are demonstrated. Also compare with Figs. 2a-5a, same protocol.

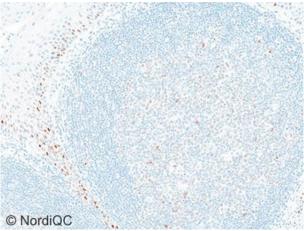


Fig. 1b (x100)

Insufficient p53 staining of the tonsil using the mAb clone DO-7 as RTU for Benchmark Ultra, Ventana/Roche, using the recommended protocol settings with HIER for 32 min. in an alkaline buffer, a short Ab incubation and OptiView as detection system.

Less than 50% nuclear staining reaction for p53 is seen in the germinal centre B-cells compared to the optimal result in Fig. 1a – same area. Also compare with Figs. 2b-5b, same protocol.

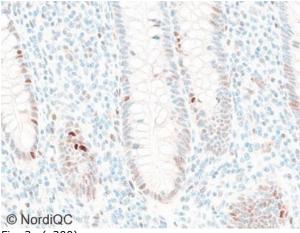


Fig. 2a (x200)

Optimal p53 staining of the appendix using same protocol as in Fig. 1a.

Dispersed epithelial cells of the basal parts of the crypts show a weak to moderate nuclear staining reaction.

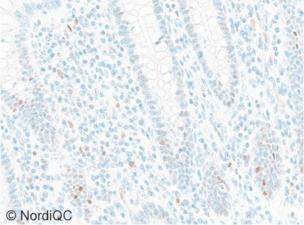


Fig. 2b (x200)

Insufficient p53 staining of the appendix using same protocol as in Fig. 1b.

Only a very faint nuclear staining reaction for p53 is seen in the crypt epithelial cells compared to the optimal result in Fig. 2a – same area. Also compare with Figs. 3b-5b, same protocol.

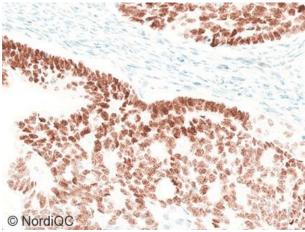


Fig. 3a (x200) Optimal p53 staining of the endometrial serous carcinoma with p53 overexpression, tissue core no. 3, using same protocol as in Figs. 1a – 2a. Virtually all neoplastic cells show a strong, nuclear staining reaction.

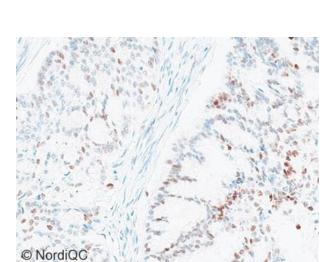


Fig. 4a (x200) Optimal p53 staining of the low grade endometrial carcinoma (p53 wild-type), tissue core no. 4, using same protocol as in Figs. 1a – 3a.

Virtually all neoplastic cells show a weak to moderate,

nuclear staining reaction.

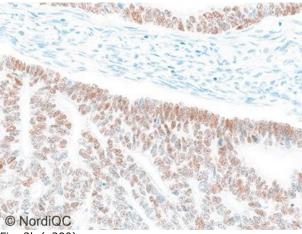


Fig. 3b (x200) p53 staining of the endometrial serous carcinoma with p53 overexpression, tissue core no. 3, using same protocol as in Figs. 1b - 2b.

Virtually all neoplastic cells show a moderate nuclear staining reaction - same area as Fig 3a. Overall all cells are demonstrated as these have high-level p53 expression. Also compare with Figs. 4b and 5b, same protocol.

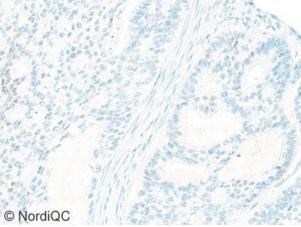


Fig. 4b (x200)

Insufficient p53 staining of the low grade endometrial carcinoma (p53 wild-type), tissue core no. 4, using same protocol as in Figs. 1b – 3b.

A significantly reduced number of neoplastic and stromal cells are stained for p53 compared to the optimal result in Fig. 4a – same area – and result can be interpreted as TP53 mutation with loss of p53.

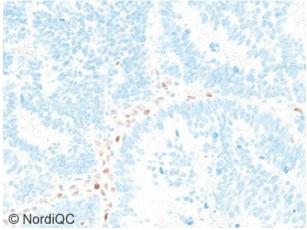


Fig. 5a (x200)
Optimal p53 staining of the endometrial serous carcinoma with absence of p53 expression, tissue core no. 5, using same protocol as in Figs. 1a – 4a.
No nuclear staining reaction is seen in the neoplastic cells. Stromal cells display a weak to moderate reaction, serving as internal positive tissue control.

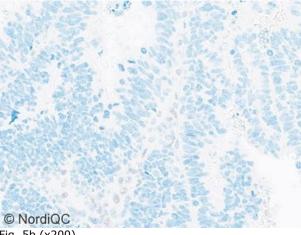
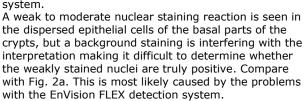
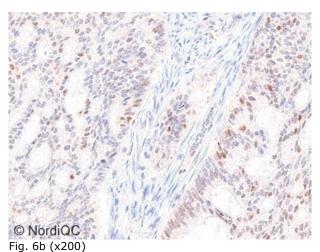


Fig. 5b (x200)
Insufficient p53 staining of the endometrial serous carcinoma with absence of p53 expression, tissue core no. 5, using same protocol as in Figs. 1b – 4b.
Virtually no nuclear staining reaction for p53 is seen in the stromal cells and p53 status cannot be determined.



Fig. 6a (x200)
Sufficient p53 staining of the appendix using the mAb clone DO-7 as concentrate for Dako Omnis,
Dako/Agilent, using the EnVision FLEX+ as detection system.





Sufficient p53 staining of the low grade endometrial carcinoma (p53 wild-type), tissue core no. 4, using same protocol as in Fig. 6a.

Virtually all neoplastic cells show weak to moderate, nuclear staining reaction but only the moderately stained nuclei can be determined as truly positive. The background staining produced by the EnVision FLEX system is suppressing the weakly stained nuclei. Compare with Fig. 4a.

TJ/LE/SN 10.04.2023