

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. Low grade endometrial carcinoma - p53 wild type,
4. Endometrial serous carcinoma with absence of p53, 5. Endometrial serous carcinoma with p53 overexpression.



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 $\geq 50\%$ 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. 5).
- No staining reaction in the neoplastic cells in the endometrial serous carcinoma with absence of p53 expression (tissue core no. 4). 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. 3). 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.

* 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*

KEY POINTS FOR p53 IMMUNUASSAYS

- The mAb clones **DO-7** and **BP53-12** are recommendable Abs.
- RTUs gave in general a lower pass rate compared to concentrates and requires optimization.
- On BenchMark, HIER in CC1 should be performed for at least 48 min., and a 3-step detection system should be used.
- On Omnis, HIER in TRS high pH is required for at least 30 min., combined with EnVision Flex+.
- On Autostainer, HIER in TRS high pH is required for at least 20 min., combined with EnVision Flex+.
- The Bond platform gave an inferior performance in this assessment, however, HIER in BERS2 for at least 20 min., is recommendable.

Participation

Number of laboratories registered for p53, run 70	417
Number of laboratories returning slides	397 (95%)

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.

397 laboratories returned slides. One laboratory had stained on the wrong slide and thus not included in the following analysis. 396 laboratories participated in this assessment. 65% achieved a sufficient mark (optimal or good), see Table 1a. Tables 1b and 1c summarizes the antibodies (Abs) used and assessment marks (see page 3).

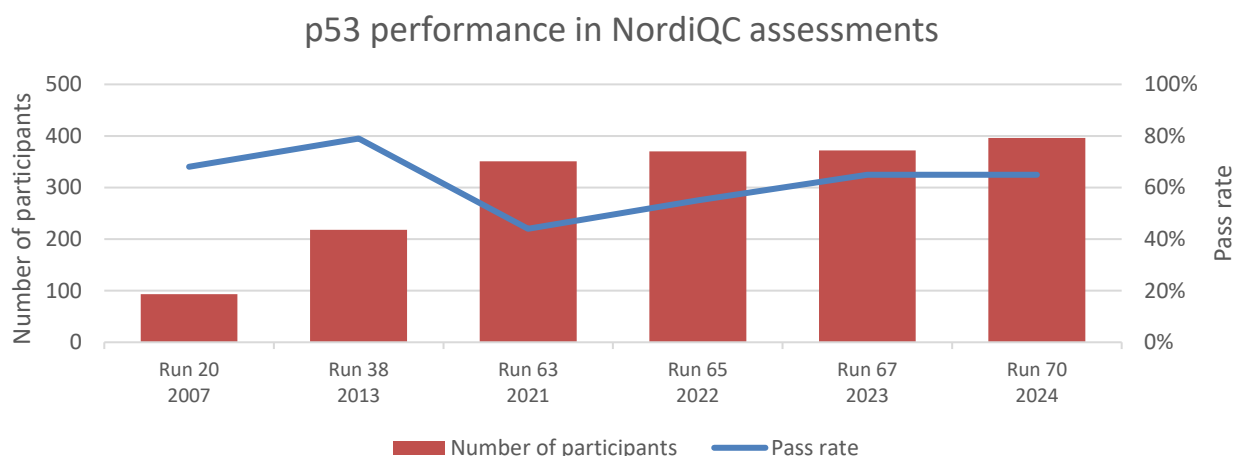
The most frequent causes of insufficient staining reactions were:

- Use of a less sensitive detection system.
- Inefficient HIER.
- Use of less successful Abs.

Performance history

This was the sixth NordiQC assessment of p53. An identical pass rate was seen compared to run 67, but still reduced compared to the first 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 six NordiQC runs 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 or loss of p53 expression in the tumour cells.

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. 84% (332 of 396) of the participants used a 3-layer detection system, with an overall pass rate at 73% (241 of 332), 18% optimal (n=60) compared to a pass rate of 23% (15 of 60), 3% optimal (n=2) if using a 2-layer detection system.

Table 1a. **Overall results for p53, run 70**

	n	Optimal	Good	Borderline	Poor	Suff. ¹	OR ²
Concentrated antibodies	90	23	41	19	7	71%	26%
Ready-To-Use antibodies	306	39	153	75	39	63%	13%
Total	396	62	194	94	46		
Proportion		16%	49%	24%	12%	65%	

1) Proportion of sufficient stains (optimal or good).

2) Proportion of Optimal Results.

Table 1b. **Concentrated antibodies and assessment marks for p53, Run 70**

Concentrated antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff. ¹	OR ²
mAb clone DO-7	4 56 1 21 1	Cell Marque Dako/Agilent Diagnostic Biosystems Leica Biosystems Thermo Scientific	23	38	15	7	73%	28%
mAb clone DO-7+BP53-12	3	Thermo Scientific	0	1	2	0	-	-
mAb clone BP53-12	1 1	ImmunoLogic PathnSitu	0	1	1	0	-	-
mAb clone IHC053	1	GenomeMe	0	1	0	0	-	-
Ab clone BPM6168	1	Biolynx Biotechnology	0	0	1	0	-	-
Total	90		23	41	19	7		
Proportion			26%	46%	21%	7%	71%	

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

2) Proportion of Optimal Results (OR).

Table 1c. **Ready-To-Use antibodies and assessment marks for p53, Run 70**

Ready-To-Use antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff. ¹	OR. ²
mAb clone BP53-11 760-2542 (VRPS)³	8	Ventana/Roche	0	3	2	3	38%	0%
mAb clone BP53-11 760-2542 (LMPS)⁴	44	Ventana/Roche	6	25	6	7	70%	14%
mAb clone DO-7 800-2912 (VRPS)³	8	Ventana/Roche	1	3	2	2	50%	13%
mAb clone DO-7 800-2912 (LMPS)⁴	104	Ventana/Roche	20	55	26	3	72%	19%
mAb clone DO-7 IS/IR616 (VRPS)³	5	Dako/Agilent	0	1	1	3	20%	0%
mAb clone DO-7 IS/IR616 (LMPS)⁴	28	Dako/Agilent	4	15	5	4	68%	14%
mAb clone DO-7 GA616 (VRPS)³	10	Dako/Agilent	0	3	0	7	30%	0%
mAb clone DO-7 GA616 (LMPS)⁴	59	Dako/Agilent	7	39	7	6	78%	12%
mAb clone DO-7 PA0057 (VRPS)³	19	Leica Biosystems	0	4	14	1	21%	0%
mAb clone DO-7 PA0057 (LMPS)⁴	12	Leica Biosystems	0	3	9	0	25%	0%
mAb clone DO-7 PDM013	1	BioSystems	0	0	1	0	-	-
mAb clone BP53-12 BMS064	1	Zytomed Systems	0	0	1	0	-	-
mAb clone MX008 MAB-0674	2	Fuzhou Maixin	1	1	0	0	-	-
mAb clone DA095 MMB1A024	1	Dartmon	0	0	1	0	-	-
rmAb clone SP5 MAD-000309QD	2	Master Diagnostica	0	0	0	2	-	-
rmAb clone MSVA-053R MAD-000307QD-R/V	1	Master Diagnostica	0	0	0	1	-	-
mAb clone 882F5H1 PA172	1	Abcarta	0	1	0	0	-	-
Total	306		39	153	75	39		
Proportion			13%	50%	24%	13%	63%	

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

Detailed analysis of p53, Run 70

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) (20/39)*, Bond Epitope Retrieval 2 (BERS2, Leica Biosystems) (2/19), Target Retrieval Solution (TRS) pH 9 (Dako/Agilent) (1/17) as retrieval buffer. The mAb was typically diluted in the range of 1:50-1:2.000 depending on the total sensitivity of the protocol employed. All optimal protocols were based on a 3-step detection system. Using these protocol settings 58 of 71 (82%) 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 ¹		Dako/Agilent Omnis		Ventana/Roche BenchMark ²		Leica Biosystems Bond ³	
	TRS pH 9.0	TRS pH 6.1	TRS pH 9.0	TRS pH 6.1	CC1 pH 8.5	CC2 pH 6.0	BERS2 pH 9.0	BERS1 pH 6.0
mAb clone DO-7	0/5** (0%)	-	1/11 (9%)	-	20/38 (53%)	-	2/18 (11%)	0/4

* 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).

1) Autostainer Classical, Link 48.

2) BenchMark GX, XT, Ultra, Ultra plus

3) Bond III, Prime

Ready-To-Use antibodies and corresponding systems

mAb clone **BP53-11**, product no. **760-2542**, Ventana/Roche, BenchMark XT/Ultra/Ultra Plus: Protocols with optimal results were based on HIER using CC1, efficient heating time 48-64 min. and 20-52 min. incubation of the primary Ab. Using these protocol settings 19 of 20 (95%) laboratories produced a sufficient staining (optimal or good).

mAb clone **DO-7** product no. **800-2912**, Ventana/Roche, BenchMark XT/Ultra/Ultra PlusGX: Protocols with optimal results were typically based on HIER using CC1, efficient heating time 32-64 min., 12-44 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 76 of 103 (74%) 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, 20-30 min. incubation of the primary Ab and EnVision FLEX+ (K8002) as detection system. Using these protocol settings 9 of 14 (64%) 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 20-30 min. incubation of the primary Ab and EnVision FLEX+ (GV800/GV823) as detection system. Using these protocol settings 40 of 46 (87%) 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	Sufficient	Optimal
Ventana Benchmark mAb clone BP53-11, 760-2542 UltraView	(0/4)	(0/4)	53% (8/15)	13% (2/15)
Ventana Benchmark mAb clone BP53-11, 760-2542 OptiView	(3/4)	(0/4)	79% (23/29)	14% (4/29)
Ventana Benchmark mAb clone DO-7, 800-2912 UltraView	(1/2)	(1/2)	50% (8/16)	6% (1/16)
Ventana Benchmark mAb clone DO-7, 800-2912 OptiView	50% (3/6)	0% (0/6)	76% (67/88)	22% (19/88)
Dako Autostainer mAb clone DO-7, IS/IR616 EnVision Flex	20% (1/5)	0% (0/5)	(0/2)	(0/2)
Dako Autostainer mAb clone DO-7, IS/IR616 EnVision Flex+	***	***	71% (12/17)	12% (2/17)
Dako Omnis mAb clone DO-7, GA616 EnVision Flex	30% (3/10)	0% (0/10)	33% (3/9)	0% (0/9)
Dako Omnis mAb clone DO-7, GA616 EnVision Flex+	***	***	86% (42/49)	14% (7/49)
Leica Bond mAb clone DO-7, PA0057	21% (4/19)	0% (0/19)	25% (3/12)	0% (0/12)

* 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.

*** Not recommended by vendor

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 84% of the insufficient results (117/140). 5% (7/140) gave a false positive staining reaction, and the remaining 11% (16/140) insufficient results were characterized by e.g. poor signal-to-noise ratio or excessive background. Virtually all laboratories were able to demonstrate p53 in the endometrial serous carcinoma with p53 overexpression (tissue core no. 5) 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. 4), the neoplastic cells in the p53 wild-type low grade endometrial carcinoma (tissue core no. 3) and germinal centre and mantle zone B-cells in tonsil were much more challenging and required a carefully calibrated protocol.

23% (90/396) 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 73% and optimal results could be obtained on the three main fully automated 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 78% (61/78) was observed compared to a pass rate of 0% (0/5) when using a 2-step detection system as EnVision FLEX (Dako/Agilent) or UltraView (Ventana/Roche). Especially OptiView (Ventana/Roche) with a carefully calibrated titer of the primary Ab performed very well with a pass rate of 85% (23/27), 52% optimal.

Ready-To-Use (RTU) antibodies were used by 77% (306 of 396) 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 1c and 3. If the RTU systems from these three vendors were used by VRPS an overall pass rate of 30% (15/50) was seen and only 2% optimal (n=1).

The Dako/Agilent **GA616** RTU system for Omnis, based on mAb clone **DO-7** was most successful if modifying the protocol settings, giving an overall pass rate of 78% (45/58). However, if following the vendor recommended protocol settings, a pass rate of only 30% was seen (3/10) (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 86% (42/49) was obtained.

Same tendency was seen for the Dako/Agilent **IR/IS616** RTU system for Autostainer, also based on mAb clone **DO-7**. Laboratories applying EnVision Flex+ as detection system had a pass rate of 71% (12/17) compared to 20% (1/5) if following the recommended protocol based on the 2-step EnVision Flex.

The Ventana/Roche **800-2912** RTU system based on mAb clone **DO-7** was the most widely used RTU system, giving an overall pass rate of 71% (79/112). 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, 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 performing HIER \geq 48 min. a pass rate of 80% (64/80) was seen, compared to HIER \leq 40 min. giving a pass rate of 47% (15/32).

The **760-2542** 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 65% (34 of 52). The VRPS were very similar also based on either UltraView or OptiView as detection system. 8 laboratories used the vendor recommended protocol settings with a pass rate of 38%, non optimal. 44 laboratories used a modified protocol typically prolonging incubation time of both HIER and primary Ab with a pass rate of 70%, 14% optimal. Overall the performance of the Ventana/Roche RTU system mAb BP53-11 was similar to the Ventana/Roche RTU system based on mAb clone DO-7 (see Tables 1c and 3).

The Leica Biosystems **PA0057** RTU system based on mAb clone **DO-7** provided an inferior pass rate compared to the latest assessment run 67. No obvious cause has been found, as similar protocol settings were applied. In this run, an almost identical performance was seen, when applying the VRPS (BERS2 20 min., IHC protocol F) compared to laboratory modified settings. A total of 19 laboratories applied VRPS with a pass-rate of 21% but no optimal. The remaining 12 laboratories applying laboratory modified protocol settings obtained a pass rate of 25%, no optimal.

In this assessment, and in concordance to last assessment run 67, it was observed that protocols based on a concentrated format provided a slightly higher pass rate (71%) than the corresponding RTU systems (63%). The overall pass-rate was the same as observed in the last NordiQC assessment run 67 in 2023 (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 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.

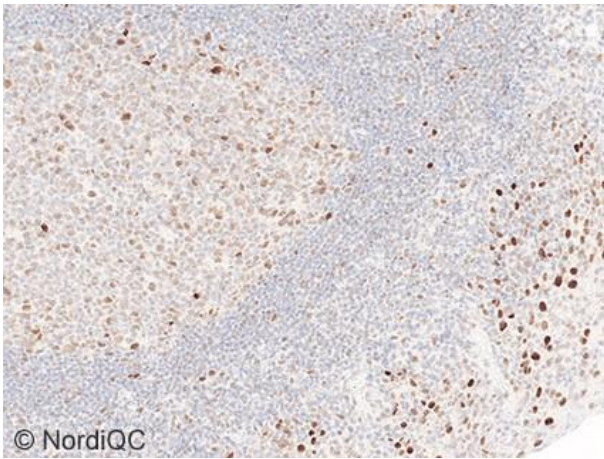


Fig. 1a
Optimal p53 staining reaction of the tonsil using the mAb clone DO-7 as RTU for Benchmark Ultra, Ventana/Roche, using modified protocol settings with HIER for 64 min. in CC1, 32 min. Ab incubation and OptiView as detection system.

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.

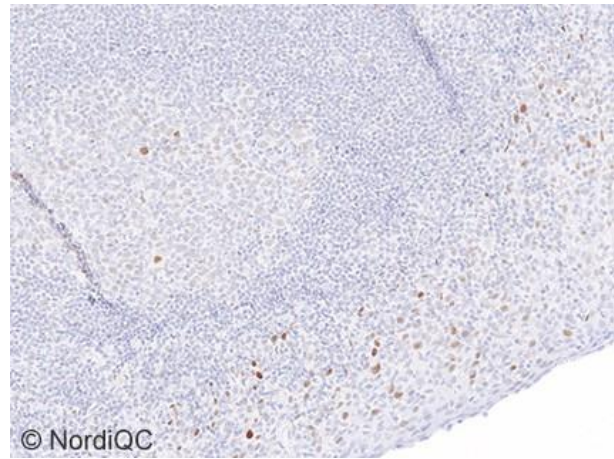


Fig. 1b
Insufficient p53 staining reaction of the tonsil using the mAb clone DO-7 as RTU for Benchmark Ultra, Ventana/Roche, using the recommended protocol settings with HIER for 32 min. in CC1, 16 min. Ab incubation and OptiView as detection system.

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.

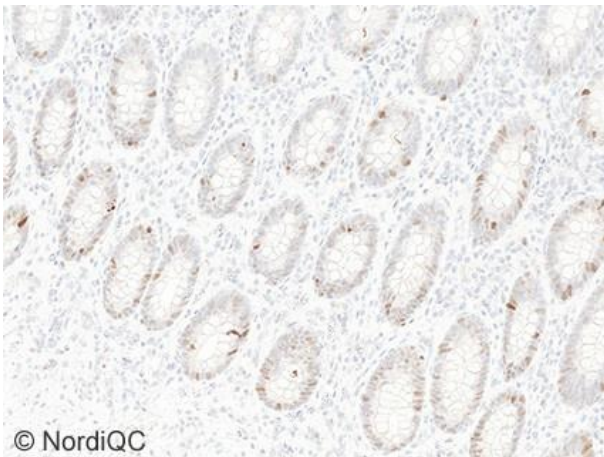


Fig. 2a
Optimal p53 staining reaction of the appendix using same protocol as in Fig. 1a. Dispersed epithelial cells of the basal parts of the crypts show a weak to moderate nuclear staining reaction.

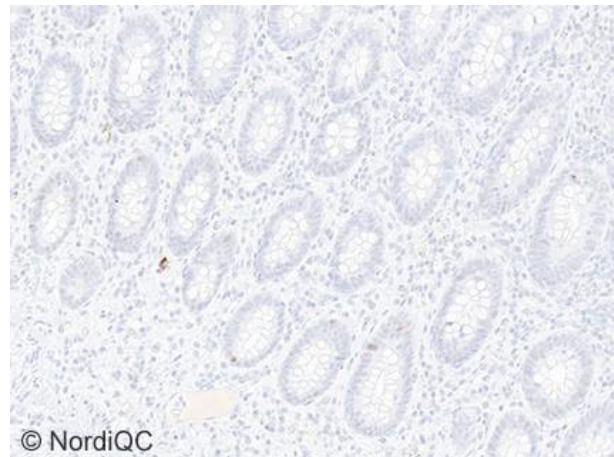
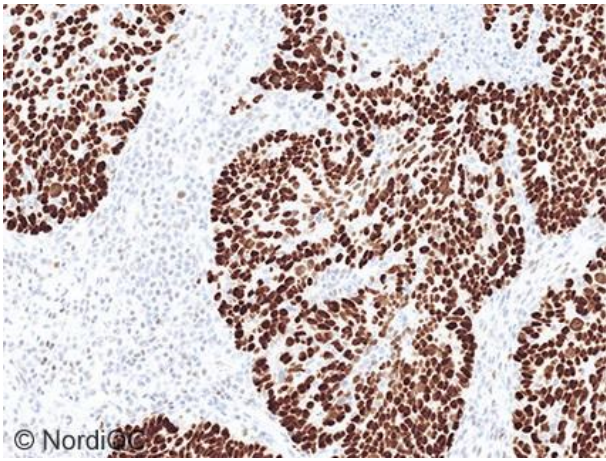
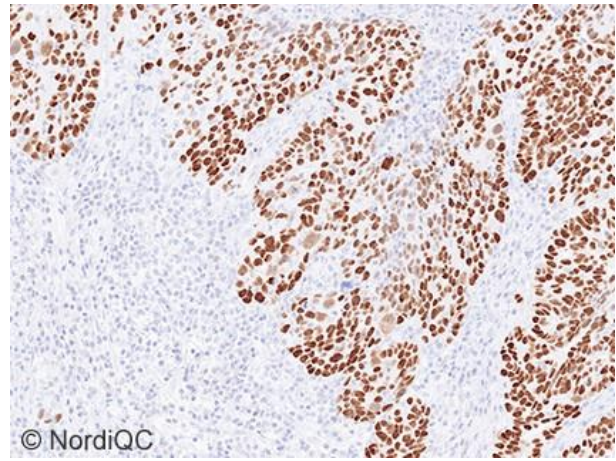


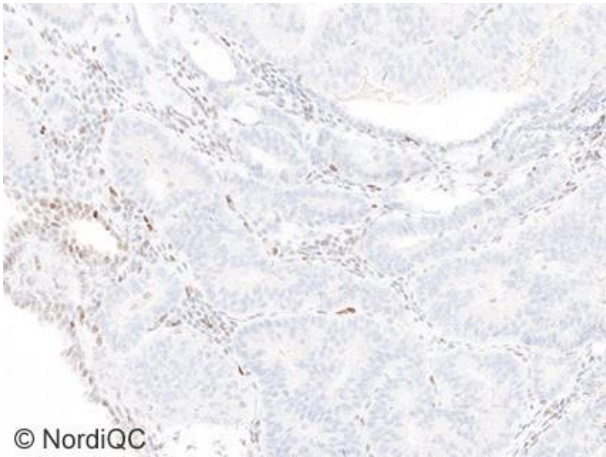
Fig. 2b
Insufficient p53 staining reaction of the appendix using same protocol as in Fig. 1b. Only a very faint nuclear staining reaction for p53 is seen in few crypt epithelial cells compared to the optimal result in Fig. 2a – same area. Also compare with Figs. 3b-5b, same protocol.



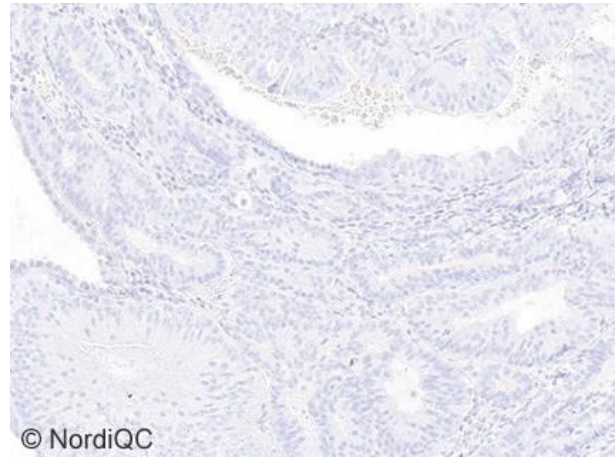
© NordiQC
 Fig. 3a
 Optimal p53 staining reaction of the endometrial serous carcinoma with p53 overexpression, tissue core no. 5, using same protocol as in Figs. 1a – 2a. Virtually all neoplastic cells show a strong, nuclear staining reaction.



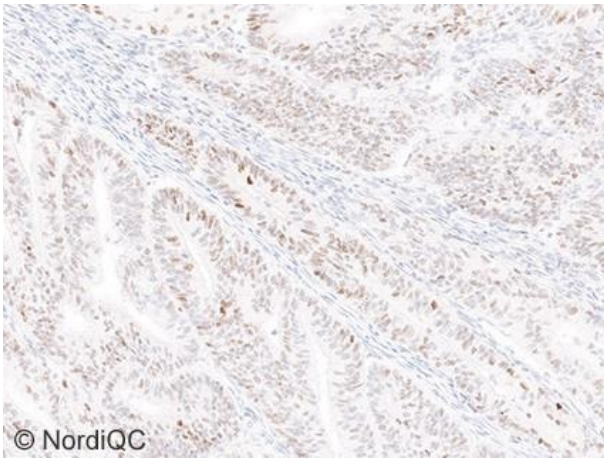
© NordiQC
 Fig. 3b
 p53 staining reaction of the endometrial serous carcinoma with p53 overexpression, tissue core no. 5, using same protocol as in Figs. 1b – 2b. Virtually all neoplastic cells show a moderate nuclear staining reaction – same area as Fig 3a. Overall all cells are demonstrated as these have high-level p53 expression. Also compare with Figs. 4b and 5b, same protocol.



© NordiQC
 Fig. 4a
 Optimal p53 staining reaction of the endometrial serous carcinoma with absence of p53 expression, tissue core no. 4, using same protocol as in Figs. 1a – 3a. No nuclear staining reaction is seen in the neoplastic cells. Stromal cells display a weak to moderate reaction, serving as internal positive tissue control.

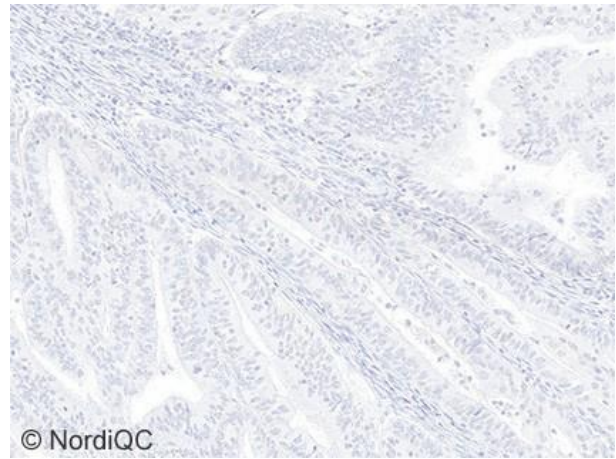


© NordiQC
 Fig. 4b
 Insufficient p53 staining reaction of the endometrial serous carcinoma with absence of p53 expression, tissue core no. 4, using same protocol as in Figs. 1b – 3b. Virtually no nuclear staining reaction for p53 is seen in the stromal cells and p53 status cannot be determined.



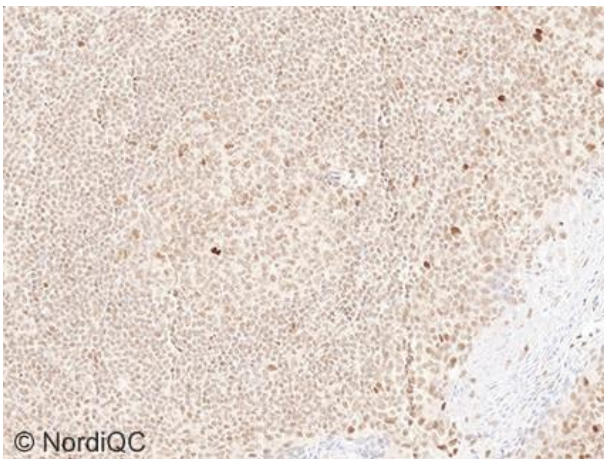
© NordiQC

Fig. 5a
Optimal p53 staining reaction of the low grade endometrial carcinoma (p53 wild-type), tissue core no. 3, using same protocol as in Figs. 1a – 4a. Virtually all neoplastic cells show a weak to moderate, nuclear staining reaction.



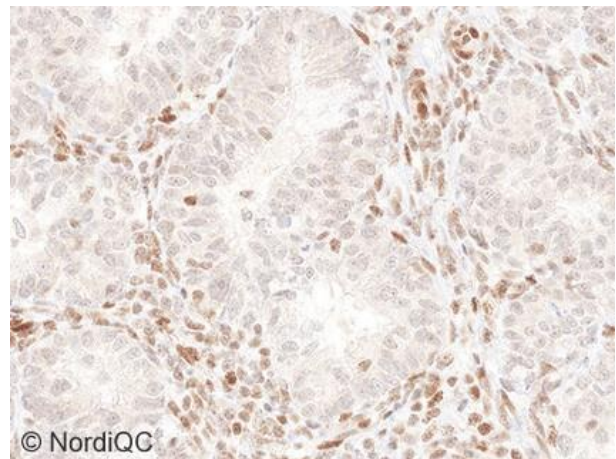
© NordiQC

Fig. 5b
Insufficient p53 staining reaction of the low grade endometrial carcinoma (p53 wild-type), tissue core no. 3, using same protocol as in Figs. 1b – 4b. A significantly reduced number of neoplastic and stromal cells are stained for p53 compared to the optimal result in Fig. 5a – same area – and result can erroneously be interpreted as TP53 mutation with loss of p53.



© NordiQC

Fig. 6a
Insufficient false positive p53 staining reaction of the tonsil using the mAb clone DO-7 as concentrate diluted 1:150, performing efficient HIER in an alkaline buffer and using a 3-step detection system. A weak to moderate nuclear staining reaction is seen in virtual all the germinal centre B-cells, but also in most mantle zone B-cells indicating a too high level of analytical sensitivity. Compare with Fig. 1a for optimal result.



© NordiQC

Fig. 6b
Insufficient false positive p53 staining reaction of the endometrial serous carcinoma with expected absence of nuclear p53 expression, tissue core no. 4, using same protocol as in Fig. 6a. Virtually all neoplastic cells show a moderate false positive nuclear staining reaction. Compare with Fig. 4a for optimal result.

HLK/LE/SN 26.03.2024