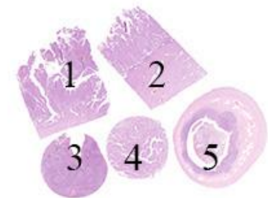
NordiQC

Assessment Run 65 2022
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. Endometrial serous carcinoma with p53 overexpression, 2. Endometrial serous carcinoma with absence of p53 expression, 3. Tonsil, 4. Low grade endometrial carcinoma – p53 wild-type, 5. Appendix.

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 ≥20% 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. 1).
- No staining reaction in the neoplastic cells in the endometrial serous carcinoma with absence of p53 expression (tissue core no. 2). 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.
- Nuclear staining of less than 10% of the mantle zone B-cells of the secondary follicles of the tonsil and no staining of the luminal epithelial cells in the appendix.

* 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*

Participation

Number of laboratories registered for p53, run 65	394
Number of laboratories returning slides	370 (94%)

Results
At the date of assessment, 94% 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.

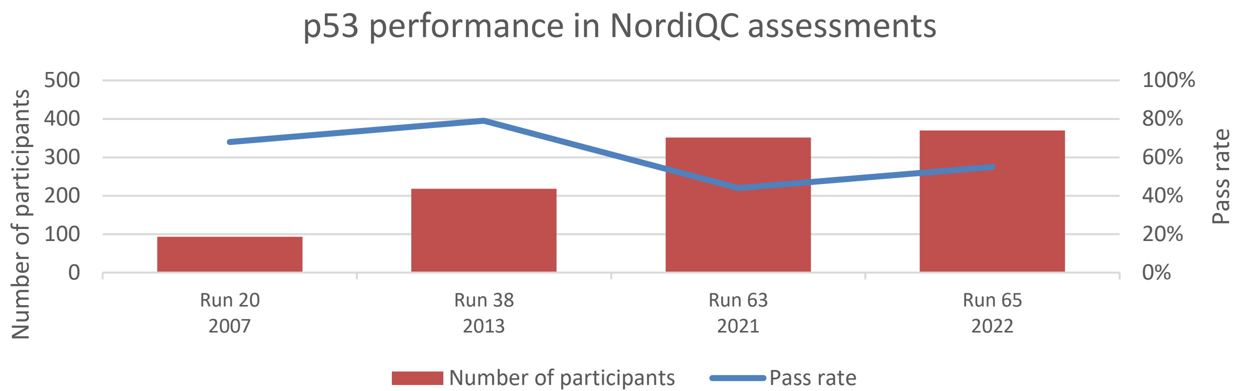
370 laboratories participated in this assessment. 55% achieved a sufficient mark (optimal or good). Table 1 summarizes the antibodies (Abs) used and assessment marks (see page 3).

The most frequent causes of insufficient staining reactions were:

- Use of a less sensitive detection system.
- Too low concentration of the primary antibody.
- Use of less successful vendor recommended protocol settings for Ready-To-Use systems.

Performance history
This was the fourth NordiQC assessment of p53. An increased pass rate was seen compared to run 63, but still reduced compared to the previous assessments (see Graph 1). In run 63 and 65 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 four 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. 77% (286 of 370) of the participants used a 3-layer detection system, with an overall pass rate at 67% (193 of 286), 28% optimal (n=81) compared to a pass rate of 11% (9 of 84), 4% optimal (n=3) if using a 2-layer detection system.

Table 1. **Antibodies and assessment marks for p53, Run 65**

Concentrated antibodies	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
mAb clone **DO-7**	58 20 6 1 1 1 1 1 1	Dako/Agilent Leica Biosystems Cell Marque DBS Diagomics Immunologic Monosan Thermo Scientific Zeta Corporation	33	23	28	6	62%	37%
mAb clone **DO-7+BP53-12**	2	Thermo Scientific	1	1	-	-	-	-
mAb clone **BP53-12**	1 1	PathnSitu Zytomed Systems	-	1	1	-	-	-
mAb clone **IHC053**	1	GenomeMe	-	-	1	-	-	-
Conc total	95		34	25	30	6	62%	36%
Ready-To-Use antibodies							Suff.[1]	OR.[2]
mAb clone **BP53-11 760-2542 (VRPS)[3]**	9	Ventana/Roche	1	2	2	4	33%	11%
mAb clone **BP53-11 760-2542 (LMPS)[4]**	39	Ventana/Roche	4	16	15	4	51%	10%
mAb clone **DO-7 800-2912 (VRPS)[3]**	10	Ventana/Roche	2	2	4	2	40%	20%
mAb clone **DO-7 800-2912 (LMPS)[4]**	84	Ventana/Roche	19	25	29	11	52%	23%
mAb clone **DO-7 IS/IR616 (VRPS)[3]**	8	Dako/Agilent	-	1	4	3	12%	-
mAb clone **DO-7 IS/IR616 (LMPS)[4]**	22	Dako/Agilent	4	10	5	3	64%	18%
mAb clone **DO-7 GA616 (VRPS)[3]**	15	Dako/Agilent	1	-	3	11	7%	7%
mAb clone **DO-7 GA616 (LMPS)[4]**	54	Dako/Agilent	16	24	10	4	74%	30%
mAb clone **DO-7 PA0057 (VRPS)[3]**	13	Leica Biosystems	1	8	4	-	69%	8%
mAb clone **DO-7 PA0057 (LMPS)[4]**	8	Leica Biosystems	-	5	3	-	63%	-
mAb clone **DO-7 453M-9x**	5	Cell Marque	1	-	3	1	20%	20%
mAb clone **DO-7 PM042**	2	Biocare Medical	1	-	-	1	-	-
mAb clone **BP53-12 BMS064**	1	Zytomed Systems	-	-	1	-	-	-
mAb clone **C2H10 CPM-0141**	1	Celnovte	-	-	1	-	-	-
mAb clone **DO-7 AM239-5M**	1	BioGenex	-	-	-	1	-	-
rmAb clone **SP5 MAD-000309QD**	2	Master Diagnostica	-	-	-	2	-	-
mAb clone **882F5H1 PA172**	1	Abcarta	-	-	-	1	-	-
RTU total	275		50	93	84	48	52%	18%
Total	370		84	118	114	54		
Proportion			23%	32%	31%	15%	55%	

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 65
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) (20/42)*, Bond Epitope Retrieval Solution 1 (BERS1, Leica Biosystems) (2/6), Bond Epitope Retrieval 2 (BERS2, Leica Biosystems) (6/19), Target Retrieval Solution (TRS) pH 9 (Dako/Agilent) (3/12) as retrieval buffer. The mAb was diluted in the range of 1:100 -1:1000 depending on the total sensitivity of the protocol employed. Using these protocol settings 45 of 60 (75%) 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 9.0	TRS pH 6.1	TRS pH 9.0	TRS pH 6.1	CC1 pH 8.5	CC2 pH 6.0	ER2 pH 9.0	ER1 pH 6.0
mAb clone **DO-7**	0/7** (0%)	-	5/12 (42%)	-	20/42 (48%)	-	6/19 (32%)	2/6 (33%)

* 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 32-48 min. and 16-28 min. incubation of the primary Ab. OptiView (760-700) was used as detection system. Using these protocol settings 13 of 15 (87%) 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., 12-40 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 44 of 61 (72%) 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 95-97°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 24-30 min., and 15-30 min. incubation of the primary Ab and EnVision FLEX+ (GV800/GV823) as detection system. Using these protocol settings 36 of 43 (84%) laboratories produced a sufficient staining.

mAb clone **DO-7** product no. **PA0057**, Leica Biosystems, Leica Bond III:
One protocol with an optimal result was based on HIER using BERS2, efficient heating time 20 min., 15 min. incubation of the primary Ab and Bond Polymer Refine Detection (DS9800) as detection system. A modified protocol was used applying the endogenous peroxidase blocking after the detection system. Using these protocol settings but without the modified protocol settings for peroxidase blocking, 9 of 13 (69%) 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**	33% (3/9)	11% (1/9)	51% (20/39)	10% (4/39)
Ventana Benchmark mAb clone **DO-7, 800-2912**	40% (4/10)	20% (2/10)	52% (44/84)	23% (19/84)
Dako Autostainer mAb clone **DO-7, IR/IS616**	12% (1/8)	0% (0/8)	59% (10/17)	18% (3/17)
Dako Omnis mAb clone **DO-7, GA616**	7% (1/15)	7% (1/15)	76% (39/51)	31% (16/51)
Leica Bond mAb clone **DO-7, PA0057**	69% (9/13)	8% (1/13)	63% (5/8)	0% (0/8)

* 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 80% of the insufficient results (135/168). The remaining 20% (33/168) insufficient results were characterized by either a false positive staining reaction (n=6) or poor signal-to-noise ratio/excessive background (n=27). Virtually all laboratories were able to demonstrate p53 in the endometrial serous carcinoma with p53 overexpression (tissue core no. 1) 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. 2), 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.

26% (95/370) of the laboratories used an Ab as concentrated format within a laboratory developed (LD) assay for p53. 95% (90/95) of the LD assays were based on the mAb clone DO-7 with a relatively low pass rate at 62%. However, 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 72% (53/74) was observed, 45% optimal (n=33) compared to a pass rate of 30% (3/10) 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 82% (23/28), 64% optimal.

Ready-To-Use (RTU) antibodies were used by 74% (275 of 370) 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 33% was seen and only 9% optimal.

The Dako/Agilent RTU system for Omnis, based on mAb clone DO-7 was most successful if modifying the protocol settings, giving a pass rate of 76% (39 of 51). However, if following the vendor recommended protocol settings, a pass rate of only 7% was seen (1 of 15) (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 79% (37 of 47) 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 77% (10 of 13) was observed, compared to 8% (1 of 12) if using EnVision FLEX as detection system (see Table 4).

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 detection system		3-layer detection system	
Antibodies	**n**	**Sufficient**	**Optimal**	**Sufficient**	**Optimal**
mAb conc **DO-7**	90	21% (3/14)	0% (-/14)	70% (53/76)	43% (33/76)
mAb RTU **BP53-11** **760-2542*** Ventana/Roche	48	0%(0/15)	0%(0/15)	70% (23/33)	15% (5/33)
mAb clone RTU **DO-7** **800-2912*** Ventana/Roche	94	5% (1/21)	0% (0/21)	64% (47/73)	29% (21/73)
mAb clone RTU **DO-7** **IS/IR616*** Dako/Agilent	25	8% (1/12)	0% (0/12)	77% (10/13)	23% (3/13)
mAb clone RTU **DO-7** **GA616*** Dako/Agilent	66	16% (3/19)	11% (2/19)	79% (37/47)	32% (15/47)

*Only protocols performed on the intended IHC stainer device are included.
** regardless of the protocol settings applied e.g., HIER time and/or incubation time in the primary Ab (≥10 protocols assessed).

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 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 64% (47 of 73) 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 48% (23 of 48). The VRPS settings were very similar also based on either UltraView or OptiView as detection system. Nine laboratories used the vendor recommended protocol settings with a pass rate of 33%. 39 laboratories used a modified protocol typically prolonging incubation time of both HIER and primary Ab with a pass rate of 51%.

The Leica Biosystems RTU system based on mAb clone DO-7 was most successful when applying vendor recommended protocol settings (protocol F). A total of 13 laboratories applied this protocol with a pass-rate of 69% but only one optimal. The optimal protocol was based on a slightly modified protocol applying the peroxide blocking step after the polymer reagent. The remaining 8 laboratories with more extensive laboratory modified protocol settings obtained a pass rate of 63%, none optimal.
In this assessment the protocols based on a concentrated format of the primary Ab provided a slightly higher pass rate (62%) than the corresponding RTU systems (52%). 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 63 in 2021 (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.

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 20% of germinal centre B-cells must show a weak to moderate nuclear staining reaction, while less than 10% 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.

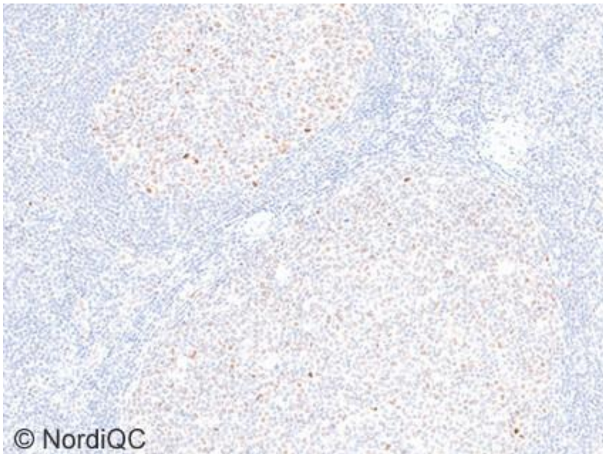


Fig. 1a (x100)
Optimal p53 staining of the tonsil using the mAb clone DO-7 as RTU for Dako Omnis, Dako/Agilent, using modified protocol settings with HIER in an alkaline buffer and EnVision FLEX+ as detection system.
A weak to moderate nuclear staining reaction is seen in >20 % of the germinal centre B-cells, whereas <10 % of the mantle zone B-cells are demonstrated. Also compare with Figs. 2a-5a, same protocol.

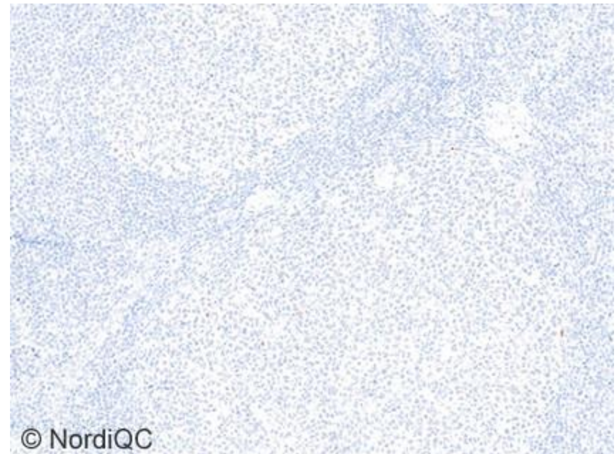


Fig. 1b (x100)
Insufficient p53 staining of the tonsil using the mAb clone DO-7 as RTU for Dako Omnis, Dako/Agilent, using the recommended protocol settings with HIER in an alkaline buffer and EnVision FLEX as detection system.
Virtually no nuclear staining reaction for p53 is seen in the germinal centre B-cells compared to the optimal result in Fig. 1a – same area. Also compare with Figs. 2b-5b, same protocol.

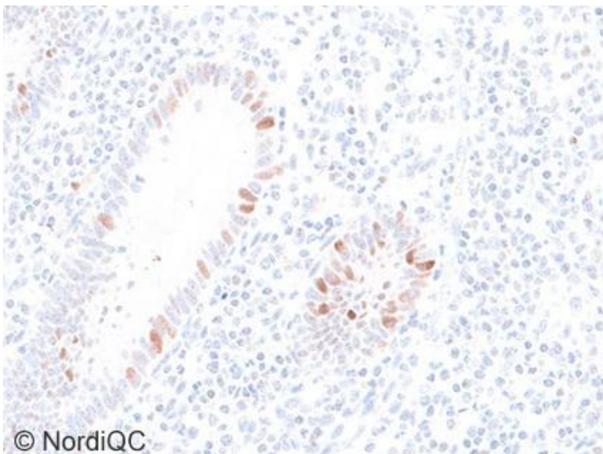


Fig. 2a (x200)
Optimal p53 staining of the appendix using same protocol as in Fig. 1a.
Dispersed epithelial cells of the basal parts of the crypts show a weak to moderate nuclear staining reaction.

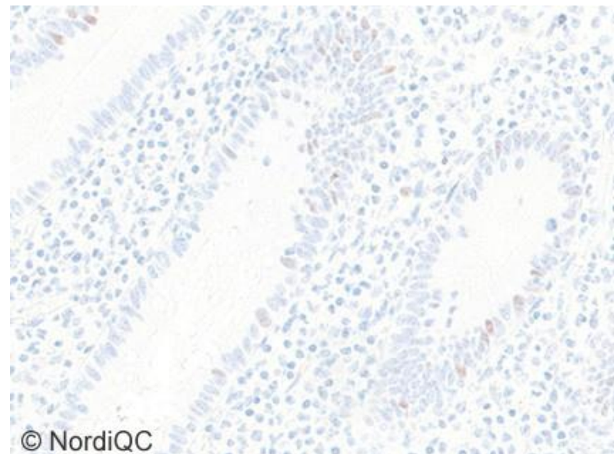


Fig. 2b (x200)
Insufficient p53 staining of the appendix using same protocol as in Fig. 1b.
Virtually no nuclear staining reaction for p53 is seen in the crypt epithelial cells compared to the optimal result in Fig. 2a – same area. Also compare with Figs. 3b-5b, same protocol.

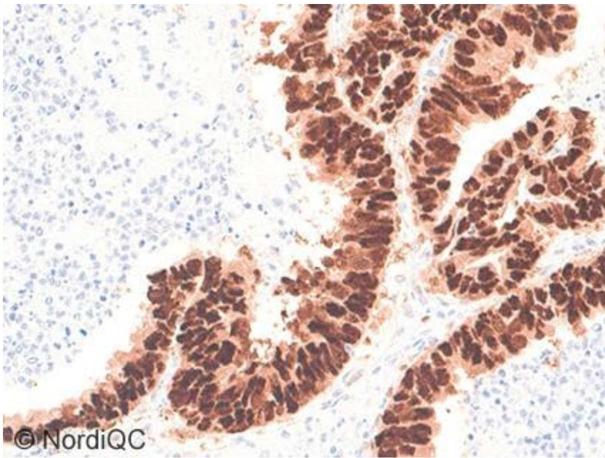


Fig. 3a (x200)
Optimal p53 staining of the endometrial serous carcinoma with p53 overexpression (tissue core no. 1), using same protocol as in Figs. 1a – 2a.
Virtually all neoplastic cells show a strong, nuclear staining reaction.

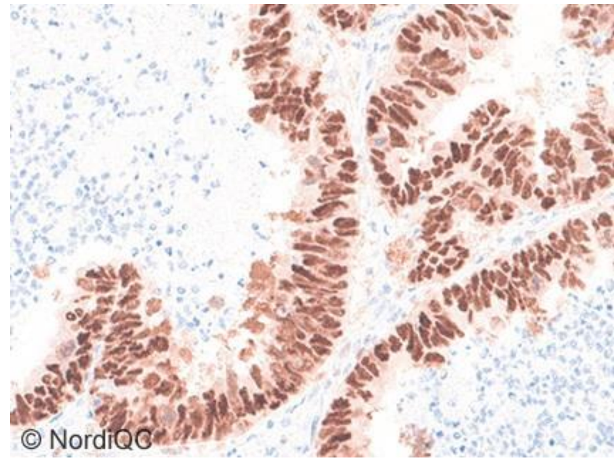


Fig. 3b (x200)
p53 staining of the endometrial serous carcinoma with p53 overexpression (tissue core no. 1), using same protocol as in Figs. 1b – 2b.
Virtually all neoplastic cells show a moderate to strong, nuclear staining reaction. The intensity is slightly reduced compared to the optimal result in Fig. 3a – same area, but overall, all cells are demonstrated as these have high-level p53 expression.

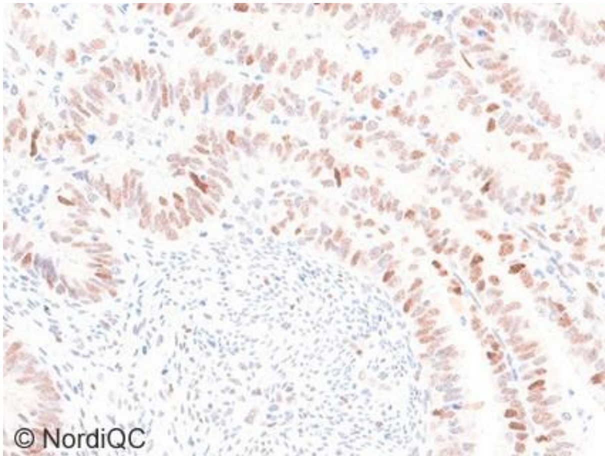


Fig. 4a (x200)
Optimal p53 staining of the low grade endometrial carcinoma, p53 wild-type (tissue core no. 4), using same protocol as in Figs. 1a – 3a.
Virtually all neoplastic cells show a weak to moderate, nuclear staining reaction.

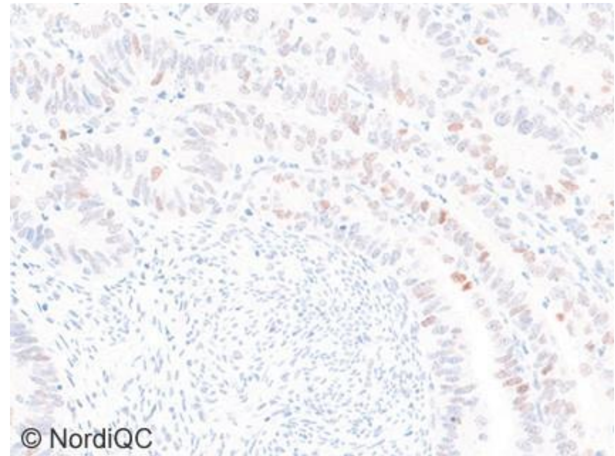


Fig. 4b (x200)
Insufficient p53 staining of the low grade endometrial carcinoma, p53 wild-type (tissue core no. 4), using same protocol as in Figs. 1b – 3b.
A significantly reduced number of neoplastic and stromal cells are stained for p53 compared to the optimal result in Fig. 4a – same area.

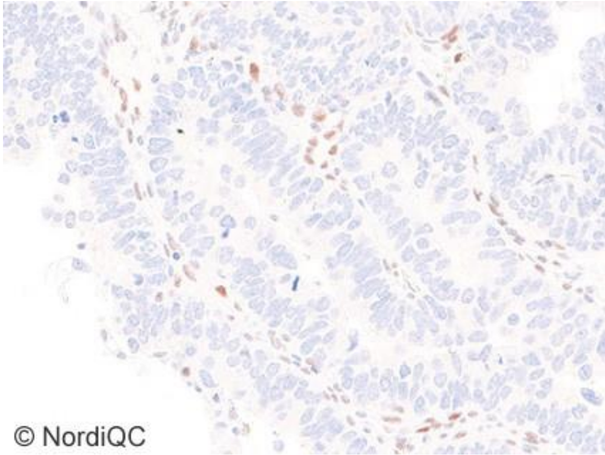


Fig. 5a (x200)
Optimal p53 staining of the endometrial serous carcinoma with absence of p53 expression (tissue core no. 2), using same protocol as in Figs. 1a – 4a.
No nuclear staining reaction is seen in the neoplastic cells. Stromal cells are weakly positive, serving as internal positive tissue control.

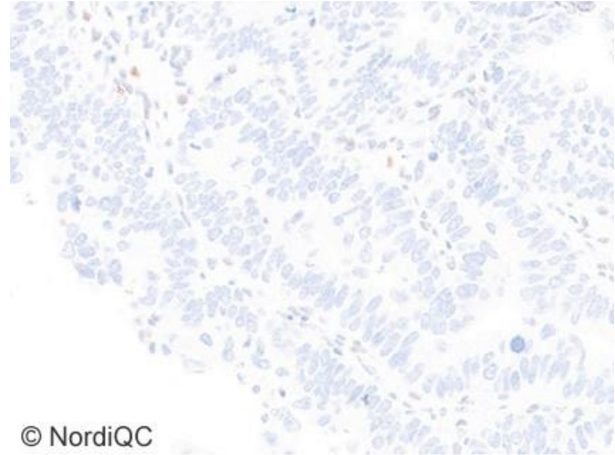


Fig. 5b (x200)
Insufficient p53 staining of the endometrial serous carcinoma with absence of p53 expression (tissue core no. 2), using same protocol as in Figs. 1b – 4b.
Virtually no nuclear staining reaction for p53 is seen in neither the neoplastic or stromal cells.

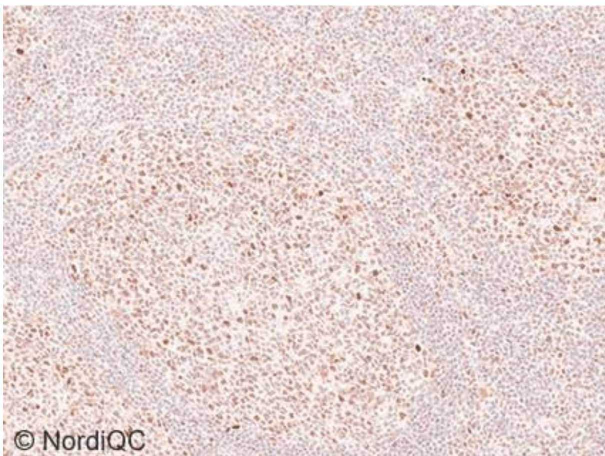


Fig. 6a (x200)
Insufficient p53 staining of the tonsil using the mAb clone DO-7 as concentrate for Dako Autostainer, Dako/Agilent, using the Ab in a dilution of 1:100 with the EnVision FLEX+ as detection system.
A weak to moderate nuclear staining reaction is seen in virtual all the germinal centre B-cells, but also in most mantle zone B-cells indicating a too high level of sensitivity. Compare with Figs. 1a.

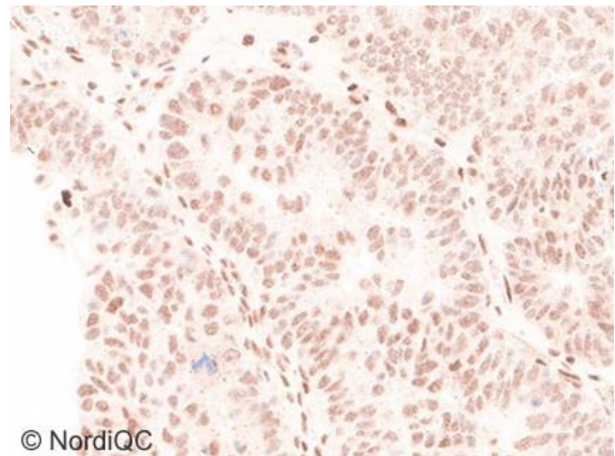


Fig. 6b (x200)
Insufficient p53 staining of the endometrial serous carcinoma with expected absence of nuclear p53 expression (tissue core no. 2), using same protocol as in Fig. 6a.
Virtually all neoplastic cells show a moderate false positive nuclear staining reaction. Compare with Fig. 5a – same field.

TJ/LE/SN 09.07.2022