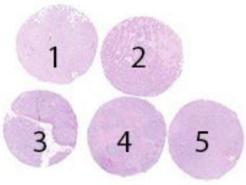
Assessment Run B25 2018

HER2 IHC

Material

The slide to be stained for HER2 comprised the following 5 materials:

	IHC: HER2 Score* (0, 1+, 2+, 3+)	FISH: HER2 gene/chr 17 ratio**
1. Breast carcinoma, no. 1	0-1+	1.2 – 1.4 (unamplified)
2. Breast carcinoma, no. 2	3+	> 6.0 (clusters) (amplified)
3. Breast carcinoma, no. 3	0-1+	1.1 – 1.4 (unamplified)
4. Breast carcinoma, no. 4	2+	5.3 – 5.8 (amplified)
5. Breast carcinoma, no. 5	2+	0.9 – 1.1 (unamplified)



* HER2 immunohistochemical score (see table below) as achieved by using the two FDA approved kits and antibodies, HercepTest™ (Dako) and PATHWAY® (Ventana), in NordiQC reference laboratories.
** HER2 gene/chromosome 17 ratios achieved using Zyto*Light* ® SPEC HER2/CEN 17 Dual Color FISH (Zytovision)

All carcinomas were fixed for 24 - 48 h in 10% neutral buffered formalin.

IHC scoring system according to the 2013 ASCO/CAP guidelines

Score 0	No staining is observed or incomplete membrane staining is observed in ≤ 10% of the tumour cells.
Score 1+	A faint perceptible and incomplete membrane staining is observed in more than 10% of the tumour cells.
Score 2+	A weak to moderate circumferential incomplete membrane staining is observed in more than 10% of the tumour cells or an intense circumferential complete membranous staining in ≤ 10% of the tumour cells.
Score 3+	An intense circumferential complete membrane staining is observed in more than 10% of the tumour cells.

Criteria for assessing a HER2 staining as **optimal** were:

- Staining corresponding to score 0 or 1+ in carcinomas no. 1 and 3.
- Staining corresponding to score 2+ in carcinoma no. 5.
- Staining corresponding to score 2+ or 3+ in carcinoma no. 4.
- Staining corresponding to score 3+ in carcinoma no. 2.
- No or only weak cytoplasmic reaction that did not interfere with the interpretation.

Staining was assessed as **good**, if (1) the HER2 gene amplified tumour no. 2 showed a 2+ reaction and the other breast carcinomas showed reaction pattern as described above (equivocal 2+ IHC staining should always be analyzed by ISH according to the ASCO/CAP guidelines) ***or*** (2) the HER2 0/1+ gene non-amplified tumour no. 1 and/or 3 showed a 2+ reaction and the other breast carcinomas showed the expected reaction pattern ***or*** (3) the HER2 2+ gene non-amplified tumor no. 5 showed a 0/1+ reaction.

Staining was assessed as **borderline,** if the signal-to-noise ratio was low, e.g., because of moderate cytoplasmic reaction, excessive counterstaining or excessive retrieval hampering the interpretation.

Staining was assessed as **poor** in case of a false negative staining (e.g., the 3+ tumour or the 2+ tumour with gene amplification showed a 0 or 1+ reaction) or a false positive staining (e.g., the 0/1+ tumors and the 2+ tumour without gene amplification showing a 3+ reaction).

Participation

Number of laboratories registered for HER2, run B25	358
Number of laboratories returning slides	342 (96%)

Results: 342 laboratories participated in this assessment and 96% achieved a sufficient mark (optimal or good). Assessment marks for IHC HER2 assays and HER2 antibodies are summarized in Table 1.

Table 1. Assessment marks for **IHC assays and antibodies run B25, HER2 IHC**

FDA approved HER2 assays	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
PATHWAY® rmAb clone **4B5, 790-2991**	195	Ventana/Roche	181	14	0	0	100%	100%
PATHWAY® rmAb clone **4B5, 790-2991[4]**	2	Ventana/Roche	2	0	0	0	-	-
CONFIRM™, rmAb clone **4B5, 790-4493**	19	Ventana/Roche	18	1	0	0	100%	100%
CONFIRM™, rmAb clone **4B5, 790-4493[4]**	1	Ventana/Roche	1	0	0	0	-	-
HercepTest™ **SK001**	33	Dako/Agilent	28	5	0	0	100%	100%
HercepTest™ **SK001**[5]	5	Dako/Agilent	3	1	1	0	80%	-
HercepTest™ **K5204**	1	Dako/Agilent	1	0	0	0	-	-
Oracle™ mAb clone **CB11, TA9145**	6	Leica	4	2	0	0	100%	100%
Antibodies[3] for laboratory developed HER2 assays, conc. antibody	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
rmAb clone **BSR44**	1	Nordic Biosite	1	0	0	0	-	-
mAb clone **CB11**	7 1	Leica/Novocastra Biogenex	0	2	4	2	25%	-
rmAB clone **EP1045Y**	2	ThermoFisher Scientific	1	1	0	0	-	-
pAb clone **A0485**	38	Dako/Agilent	25	9	0	4	89%	89%
rmAb clone **RM228**	1	RevMAB Bioscience	1	0	0	0	-	-
rmAb clone **SP3**	14 4 3 1 1	ThermoFisher Scientific Zytomed Cell Marque Immunologic Springer Bioscience	7	16	0	0	100%	100%
rmAb clone **A24-V**	1	DB Biotech	0	0	1	0	-	-
Antibodies for laboratory developed HER2 assays, RTU	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
rmAb clone **EP3, CCR-0843**	1	Celnovte	1	0	0	0	-	-
rmAb clone **EP3, RMPD049R**	1	Diagnostic Biosystems	1	0	0	0	-	-
rmAb clone **EP3, AN726**	1	Biogenex	0	0	1	0	-	-
rmAb clone **GR011, 8362-C010**	1	Sakura Finetek USA Inc	1	0	0	0	-	-
Ab clone **MXR001, RMA-0701**	1	Maixin	0	0	0	1	-	-
rmAb clone **SP3, MAD-000308QD**	1	Master Diagnostica	0	1	0	0	-	-
Total	342		276	52	7	7	-	-
Proportion			81%	15%	2%	2%	96%	-

1) Proportion of sufficient stains (optimal or good),
2) Proportion of sufficient stains with optimal protocol settings only, see below.
3) mAb: mouse monoclonal antibody, rmAb: rabbit monoclonal antibody, pAb: polyclonal antibody.
4) RTU system developed for the Roche/Ventana's fully automated systems (BenchMark) but used by laboratories on different platforms (e.g. Leica Bond)
5) RTU system developed for the Agilent/Dako`s semi-automated systems (Autostainer Link48) but used by laboratories on different platforms (Leica Bond and Dako Omnis)

Detailed Analysis
FDA/CE IVD approved assays

PATHWAY® rmAb clone **4B5** (790-2991, Ventana): 181 of 195 (93%) 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 8-76 min.) in BenchMark XT, GX or Ultra, 8-60 min. incubation of the primary Ab and Iview, UltraView or OptiView as detection kit. Using these protocol settings, 181 of 181 (100%) laboratories produced a sufficient staining result (optimal or good). 8 laboratories added UltraView/iView Amplification Kit to the above mentioned protocol settings. All 8 protocols were assessed as optimal.

CONFIRM™ rmAb clone **4B5** (790-4493, Ventana): 18 of 19 (95%) protocols were assessed as optimal. Protocols with optimal results were typically based on HIER in CC1 (efficient heating time 36-64 min.) on BenchMark XT, GX or Ultra, 18-34 min. incubation of the primary Ab and iView, UltraView or OptiView as detection kit. Using these protocol settings, 19 of 19 (100%) laboratories produced a sufficient staining result.

HercepTest™ pAb (SK001, Dako): 28 of 33 (85%) protocols were assessed as optimal. Protocols with optimal results were typically based on HIER in HercepTest™ epitope retrieval solution at 97-99°C for 20-40 min. in a water bath or PT Link and 20-40 min. incubation of the primary Ab. Using these protocol settings, 33 of 33 (100%) laboratories produced a sufficient staining result.

Oracle™ mAb clone **CB11** (TA9145, Leica): 4 of 6 (67%) protocols were assessed as optimal. Three protocols were based on HIER in Bond Epitope Retrieval Solution 1 (BERS1, Leica) for 25-30 min, and 15-30 min., incubation of the primary Ab. One protocol was based on HIER in Bond Epitope Retrieval Solution 2 (BERS2)for 30 min., and 15 min., incubation of the primary Ab. Using these protocol settings, 5 of 5 (100%) laboratories produced a sufficient staining result.

Table 2 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 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 2. **Comparison of pass rates for vendor recommended and laboratory modified protocols**

CDx assay	**Vendor recommended protocol settings***		**Laboratory modified protocol settings****	
	Sufficient	Optimal	Sufficient	Optimal
Ventana BenchMark XT, GX, Ultra PATHWAY® rmAb 4B5 **790-2991**	60/60 (100%)	54/60 (90%)	135/135 (100%)	127/135 (94%)
Ventana BenchMark XT, GX, Ultra CONFIRM™ rmAb 4B5 **790-4493**	4/4	4/4	16/16 (100%)	15/16 (94%)
Dako Autostainer Link 48+ HercepTest™ pAb **SK001**	24/24 (100%)	19/24 (79%)	9/9 (100%)	9/9 (100%)
Leica Bond MAX, III Oracle™ mAb CB11 **TA9145**	2/2	1/2	4/4	3/4

* 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 >25%, detection kit. Only protocols performed on the specified vendor IHC stainer are included.

Concentrated antibodies for laboratory developed (LD) assays
pAb **A0485**: 25 of 38 (66%) protocols were assessed as optimal. Optimal protocols were based on HIER using either TRS low pH 6.1 (Dako) (14/21), TRS pH 9 (3-in-1) (Dako) (7/11), CC1 (Ventana) (1/1) or BERS1 (Leica) (2/2). The pAb A0485 was typically diluted in the range of 1:100-1,200 depending on the total sensitivity of the protocol employed. Using these protocol settings, 31 of 35 (89%) laboratories produced a sufficient staining result.
* (number of optimal results/number of laboratories using this HIER buffer)

rmAb **SP3**: 7 of 23 (30%) protocols were assessed as optimal. Optimal protocols were based on HIER using TRS pH 9 (3-in-1) (Dako) (2/3), BERS2 (Leica) (3/12) or CC1 (Ventana) (2/5). The rmAb clone SP3

was diluted 1:40-100 depending on the total sensitivity of the protocol employed. Using this protocol setting, 16 of 16 (100%) laboratories produced a sufficient staining result.

Table 3 summarizes the overall proportion of optimal staining results when using the three most frequently used concentrated Abs on the most commonly used IHC stainer platforms.

Table 3. **Optimal results for HER2 for the most commonly used antibodies as concentrate on the main IHC systems***

Concentrated antibodies	**Dako Agilent Autostainer / Omnis**		**Ventana BenchMark XT / Ultra**		**Leica Bond III / Max**	
	TRS pH 9.0	TRS pH 6.1	CC1 pH 8.5	CC2 pH 6.0	BERS2 pH 9.0	BERS1 pH 6.0
pAb clone **A0485**	6/10 (60%)	14/20 (70%)	1/1	-	0/1	2/2
rmAb clone **SP3**	2/3	-	2/5	-	3/12 (25%)	-

* 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
In this assessment the insufficient results were typically characterized by a poor signal-to-noise ratio complicating interpretation. Poor signal-to-noise ratio were seen in 50% of the insufficient results (7 of 14).
The remaining insufficient results were based on either false positive staining results, 4 of 14 (29%), or false negative staining results, 3 of 14 (21%)
The false negative staining results were particularly and most critically observed as 0/1+ IHC reaction in the HER2 gene amplified breast carcinoma. This tumour was categorized as IHC 2+ in the NordiQC reference laboratories using two FDA/CE-IVD HER2 IHC assays: PATHWAY® (Ventana) and HercepTest™ (Dako) and showed a high level of HER2 gene amplification (ratio 5.3-5.8) by FISH.

False negative and false positive results were only seen in laboratory developed (LD) assays.
False negative results were for the LD assays typically caused by too low sensitivity of the protocol applied (e.g. too low concentration of the primary Ab and/or insufficient HIER). No single cause for the false positive staining reactions could be identified.
In this assessment all laboratory using RTU off-label obtained sufficient results. However, despite the encouraging results, off-label use must be meticulously validated by the end-users on a large cohort of breast carcinomas (n=100, ASCO/CAP 2013 guidelines).

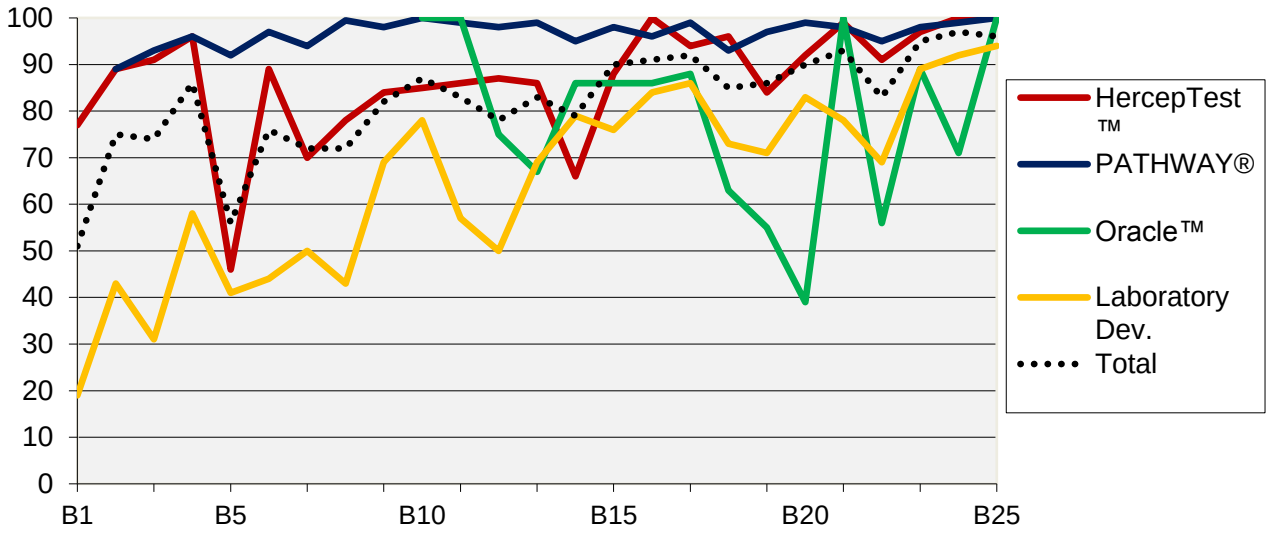
The Ventana PATHWAY® /CONFIRM™ HER2 IHC assay was also increasingly used off-label by the participants, applying OptiView as detection system and not UltraView or iView as recommended by Ventana. In this assessment, no impact on the analytical sensitivity and specificity was revealed, see Graph 2. In contrast, internal studies previously performed in the NordiQC reference laboratory indicated a less precise and robust HER2 IHC assay if UltraView was substituted by OptiView. OptiView will typically amplify the analytical sensitivity of the IHC system 3-4 times compared to UltraView. Consequently if OptiView is applied, the HER2 IHC assay must be adjusted at other parameters e.g incubation time of the primary Ab, HIER settings to provide the analytical sensitivity level validated by Ventana, which, as mentioned, can cause a less precise and robust assay.

In this assessment, the FDA-/CE-IVD approved HER2 IHC assays PATHWAY® /CONFIRM™ and HercepTest™ from Ventana and Dako respectively were most successful and provided a higher pass-rate superior to LD assays as illustrated in Graph. 1. PATHWAY®/CONFIRM™ IHC assays have provided a consistent high pass rate throughout all HER2 IHC runs in NordiQC.

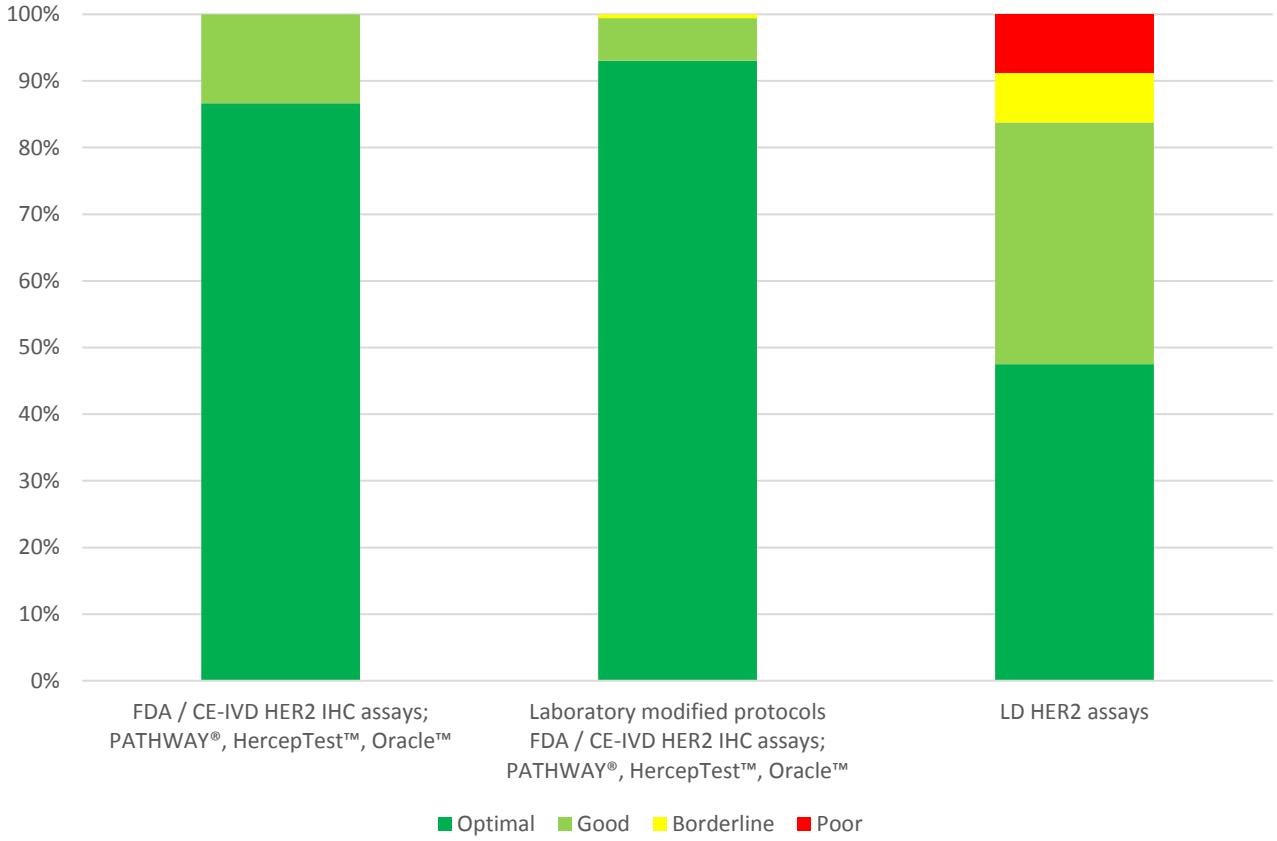
The proportion of laboratories using FDA-/CE-IVD approved HER2 IHC assays and LD assays is very consistent. In this run, 26% of the participants (n=88) used LD assays compared to 23-31% in the last 13 assessments. Despite an overall improvement of the pass rate for LD HER2 assays from run B1 to B25, the pass rate and proportion of optimal results still is inferior to the FDA/CE-IVD approved systems as PATHWAY® /CONFIRM™ and HercepTest™. In general, the three FDA-/CE-IVD approved HER2 assays provided a proportion of optimal results of 91% (232 of 254), whereas only 50% of LD HER2 assays were assessed as optimal (44 of 88). As shown in Grafp. 2, LD HER2 assays both provided a reduced proportion of sufficient results but also a shift from optimal to good, typically caused by 0/1+ staining reaction in the HER2 non-amplified tumour (tissue core no. 5) expected to show a 2+ staining reaction. The staining reaction of 0/1+ in this tumour would not directly lead to a wrong diagnosis but this indicate that the sensitivity of the protocol is low.

The overall pass rate of 96% obtained in this assessment is very satisfactory and is largely comparable to the pass rates seen in the last 5 runs indicating a relatively stable level has been reached. A significant improvement compared to the pass rate of 51% seen in run B1, 2006 has been obtained and maintained.

Graph 1. **Pass rates of 25 HER2 IHC assessments in the NordiQC breast cancer module**



Graph 2. **Proportion of assessment marks using FDA-/CD-IVD and LD assays**

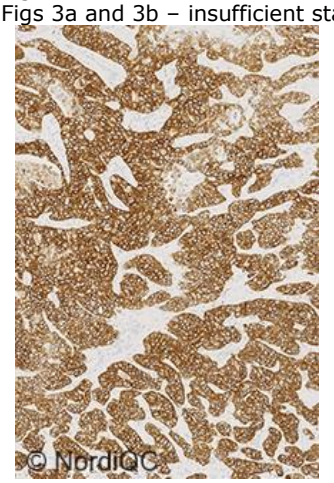


Scoring consensus B25
Laboratories were requested to submit scores (0, 1+, 2+, 3+) of their own HER2 stained slides. This was done by 84% (287 of 342) of the participants.
For 203 of the 287 (71%) responding participants, scores for all the tissues in the multi-tissue sections were in concordance with the NordiQC assessor group using the ASCO/CAP 2013 interpretation guidelines.

This was a decline compared to run B23, where 86% of the scores were in consensus with the NordiQC assessor group. This must likely due to more strict criteria in interpretation of consensus agreement in this run especially in the 2+ core without amplification. Among laboratories with sufficient staining, 74% (203 of 275) of interpretations were in agreement with the NordiQC assessors. Among participants with insufficient staining, none of the submitted scores were in consensus with the NordiQC assessor group (0 of 12). Typically, the laboratories had interpreted one (or more) of the cores in higher HER2 category than the NordiQC assessor group.

Conclusion
The FDA-/CE-IVD approved HER2 IHC assays **PATHWAY®/CONFIRM™** rmAb clone 4B5 (Ventana), **HercepTest™** (Dako) and **Oracle**™ (Leica) were in this assessment the most precise assays for the semi-quantitative IHC determination of HER2 protein expression. Laboratory developed assays produced a lower pass-rate and were less precise for the HER2 status requiring an additional ISH test for final evaluation. Inclusion of 2+ tumours with and without HER2 gene amplification in the control material for both EQA and internal quality control is essential to evaluate precision and performance stability of the IHC HER2 assays used by laboratories.

Figs 1a and 1b – optimal staining results, same protocol
Figs 2a and 2b – insufficient staining results - false negative, same protocol
Figs 3a and 3b – insufficient staining results – false positive, same protocol



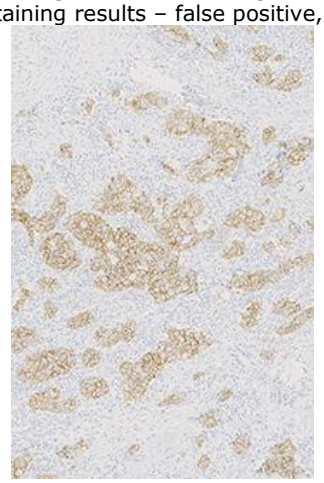


Fig 1a.
Left: Optimal staining result for HER-2 of the breast ductal carcinoma no. 2 with a ratio of HER-2 / chr17 of > 6.0.
> 10% of the neoplastic cells show an intense and complete membranous staining reaction corresponding to 3+.
Right: Optimal staining result for HER-2 of the breast ductal carcinoma no. 4 with a ratio of HER-2 / chr17 of 5.3 – 5.8.
> 10% of the neoplastic cells show a weak to moderate and complete membranous staining reaction corresponding to 2+.

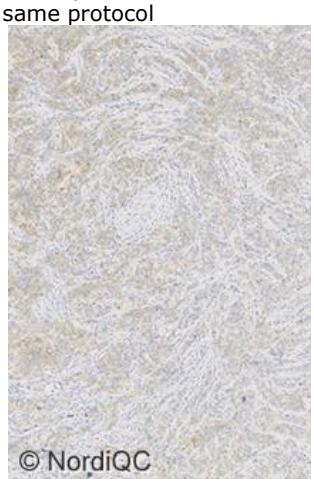


Fig 1b.
Left: Optimal staining result for HER-2 of the breast ductal carcinoma no. 5 with a ratio of HER-2 / chr17 of 0.9 – 1.1.
> 10% of the neoplastic cells show a faint membranous staining reaction corresponding to 2+.
Right: Optimal staining result for HER-2 of the breast ductal carcinoma no. 3 with a HER-2 / chr17 ratio of 1.1 – 1.4.
< 10% of the neoplastic cells show a faint membranous staining reaction corresponding to 0.

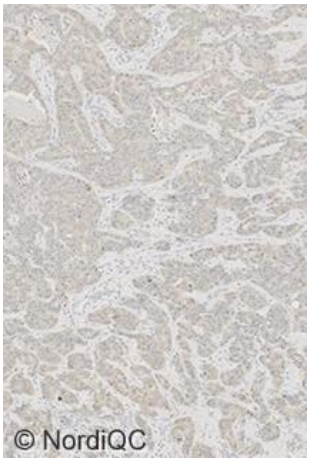


Fig 2a.
Left: Staining result for HER-2 of the breast ductal carcinoma no. 2 with a ratio of HER-2 / chr17 of > 6.0.
> 10% of the neoplastic cells show faint membranous staining and at the same time an excessive cytoplasmic staining reaction corresponding to 1+ compromising the interpretation.
Right: Staining result for HER-2 of the breast ductal carcinoma no. 4 with a ratio of HER-2 / chr17 of 5.3 – 5.8.
> 10% of the neoplastic cells show a weak to moderate and complete membranous staining reaction corresponding to 0.

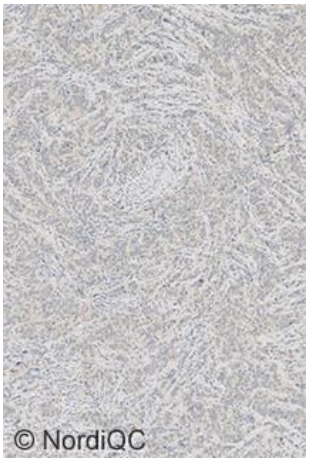
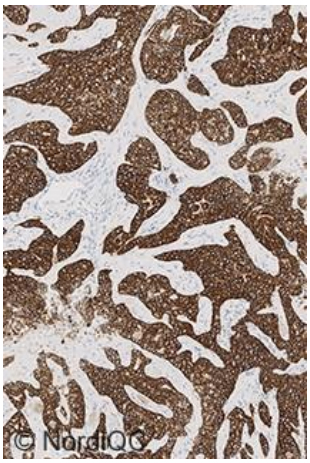


Fig 2b.
Left: Staining result for HER-2 of the breast ductal carcinoma no. 5 with a ratio of HER-2 / chr17 of 0.9 – 1.1.
< 10% of the neoplastic cells show a faint membranous staining reaction corresponding to 0.
Right: Staining result for HER-2 of the breast ductal carcinoma no. 3 with a HER-2 / chr17 ratio of 1.1 – 1.4.
< 10% of the neoplastic cells show a faint membranous staining reaction corresponding to 0.



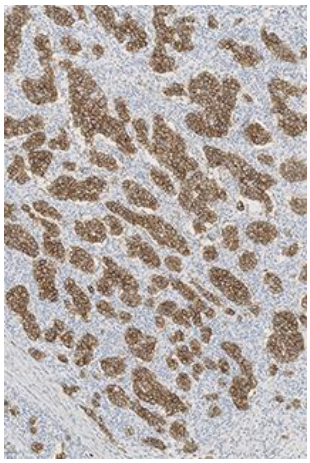


Fig 3a.
Left: Staining result for HER-2 of the breast ductal carcinoma no. 2 with a ratio of HER-2 / chr17 of > 6.0.
> 10% of the neoplastic cells show an intense and complete membranous staining reaction corresponding to 3+.
Right: Staining result for HER-2 of the breast ductal carcinoma no. 4 with a ratio of HER-2 / chr17 of 5.3 – 5.8
> 10% of the neoplastic cells show a strong and complete membranous staining reaction corresponding to 3+.

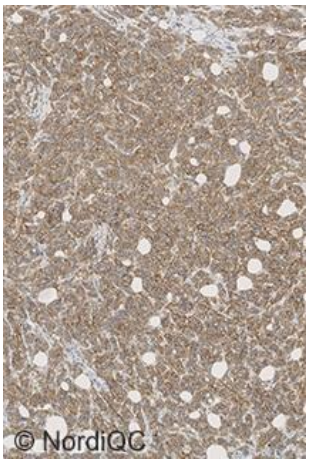


Fig 3b.
Left: Insufficient and false positive staining result for HER-2 of the breast ductal carcinoma no. 5 with a ratio of HER-2 / chr17 of 0.9 – 1.1.
> 10% of the neoplastic cells show an intense and complete membranous staining reaction corresponding to 3+.
Right: Staining result for HER-2 of the breast ductal carcinoma no. 3 with a HER-2 / chr17 ratio of 1.1 – 1.4.
> 10% of the neoplastic cells show a weak membranous staining reaction corresponding to 1+.

HLK/LE/RR 19.04.2018