Assessment Run B23 2017

HER-2 IHC

Material

The slide to be stained for HER-2 comprised the following 9 materials:

	IHC: HER-2 Score* (0, 1+, 2+, 3+)	FISH: HER-2/chr17 ratio**
1. Cell line 1, Horizon Discovery***	3+	
2. Cell line 2, Horizon Discovery***	2+	
3. Cell line 3, Horizon Discovery***	1+	
4. Cell line 4, Horizon Discovery***	0	
5. Breast carcinoma, no. 1	3+	> 6.0 (clusters) (amplified)
6. Breast carcinoma, no. 2	2+	2.3 – 2.9 (amplified)
7. Breast carcinoma, no. 3	0-1+	1.1 – 1.4 (unamplified)
8. Breast carcinoma, no. 4	1-2+	1.3 – 1.7 (unamplified)
9. Breast carcinoma, no. 5	0-1+	1.2 – 1.4 (unamplified)

* HER-2 immunohistochemical score (see table below) as achieved by using the three FDA approved kits and antibodies, HercepTest™ (Dako), Oracle™ (Leica) and PATHWAY® (Ventana), in NordiQC reference laboratories.
** HER-2/chr17 ratios achieved using Zyto*Light* ® SPEC HER2/CEN 17 Dual Color FISH (Zytovision)
*** The cell lines were not included in the assessment. Data will be analyzed subsequently by digital image analysis.

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 HER-2 staining as **optimal** were:

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

Staining was assessed as **good**, if (1) the HER-2 gene amplified tumour no. 1 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 HER-2 gene non-amplified tumour no. 3 and/or 5 showed a 2+ reaction and the other breast carcinomas showed the expected reaction pattern.

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 HER-2, run B23	380
Number of laboratories returning slides	372 (98%)

Results: 372 laboratories participated in this assessment and 95% achieved a sufficient mark. Assessment marks for IHC HER-2 assays and HER-2 antibodies are summarised in table 1.

Table 1. **Assessment marks for IHC assays and antibodies run B23, HER-2 IHC**

FDA approved HER-2 assays	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
PATHWAY® rmAb clone **4B5, 790-2991**	204	Ventana/Roche	186	13	2	3	98%	98%
CONFIRM™, rmAb clone **4B5, 790-4493**	15	Ventana/Roche	12	3	0	0	100%	100%
HercepTest™ **SK001**	33	Dako/Agilent	32	0	0	1	97%	97%
HercepTest™ **SK001**[4]	8	Dako/Agilent	8	0	0	0	100%	-
HercepTest™ **K5207**	2	Dako/Agilent	2	0	0	0	-	-
HercepTest™ **K5204**	2	Dako/Agilent	1	1	0	0	-	-
Oracle™ mAb clone **CB11, TA9145**	9	Leica	4	4	0	1	89%	88%
Antibodies[3] for laboratory developed HER-2 assays, conc. antibody	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
mAb clone **BS24**	1	Nordic Biosite	0	1	0	0	-	-
mAb clone **CB11**	7 1	Leica/Novocastra Biogenex	4	4	0	0	100%	100%
rmAb clone **EP3**	1 1 1 1 1	Biocare Cell Marque Celnovte Thermo/NeoMarkers PathnSitu	3	2	0	0	100%	100%
rmAb clone **SP3**	17 3 1 1 1	Thermo/NeoMarkers Zytomed Cell Marque Spring Bioscience Thermo/Pierce	12	8	0	3	87%	100%
pAb clone **A0485**	50	Dako/Agilent	35	11	1	3	92%	93%
Antibodies for laboratory developed HER-2 assays, RTU	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
mAb clone **CB11, NCL-L-CB11**	4	Leica/Novocastra	0	0	3	1	-	-
mAb clone **CB11, BMS014**	1	Zytomed	0	0	1	0	-	-
rmAb clone **EP3, AN726**	1	Biogenex	1	0	0	0	-	-
rmAb clone **EP3, RMPD049R**	1	Diagnostic Biosystems	1	0	0	0	-	-
rmAb clone **SP3, 237R**	2	Cell Marque	1	1	0	0	-	-
rmAb clone **SP3, MAD-000308QD**	1	Master Diagnostics	1	0	0	0	-	-
Ab clone **MXR001, RMA-0701**	2	Maixin	1	1	0	0	-	-
Total	372		304	49	7	12	-	-
Proportion			82%	13%	2%	3%	95%	-

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 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): 186 of 204 (91%) 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 8-64 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, 199 of 203 (98%) laboratories produced a sufficient staining result (optimal or good).

CONFIRM™ rmAb clone **4B5** (790-4493, Ventana): 12 of 15 (80%) protocols were assessed as optimal. Protocols with optimal results were typically based on HIER in CC1 (efficient heating time 30-64 min.) in BenchMark XT, GX or Ultra, 16 – 40 min. incubation of the primary Ab and Iview or UltraView as detection kit. Using these protocol settings, 15 of 15 (100%) laboratories produced a sufficient staining result.

HercepTest™ pAb (SK001, Dako): 40 of 41 (98%) protocols were assessed as optimal. Protocols with optimal results were typically based on HIER in HercepTest™ epitope retrieval solution at 97 - 99°C for 40 min. in a water bath or PT Link and 30 min. incubation of the primary Ab. Using these protocol settings, 28 of 29 (97%) laboratories produced a sufficient staining result.

HercepTest™ pAb (K5207, Dako): 2 of 2 protocols were assessed as optimal. Protocols with optimal result were based on HIER in HercepTest™ epitope retrieval solution at 97-98°C for 40 min. in a water bath or PT link and 30 min. incubation of the primary Ab. Using these protocol settings, 2 of 2 laboratories produced a sufficient staining result.

Oracle™ mAb clone **CB11** (TA9145, Leica): 4 of 9 (44%) protocols were assessed as optimal. Three protocols were based on HIER in Bond Epitope Retrieval Solution 1 for 25-30 min., and 30 min., incubation of the primary Ab. One protocol was based on HIER in Bond Epitope Retrieval Solution 2 for 15 min., and 15 min., incubation of the primary Ab. Using these protocol settings, 7 of 8 (88%) laboratories produced a sufficient staining result.

Concentrated antibodies for laboratory developed (LD) assays

mAb **CB11**: 4 of 8 (50%) protocols were assessed as optimal. Optimal protocols were based on HIER using Target Retrieval Solution (TRS) Low (Dako) (1/1)*, Bond Epitope Retrieval Solution 1 pH 6 (BERS1, Leica) (1/1), Tris-EDTA/EGTA pH 9 (1/1) or Citrate pH 6 (1/2). The mAb clone CB11 was diluted in a range of 1:150-400 depending on the total sensitivity of the protocol employed. Using these protocol settings, 5 of 5 (100%) laboratories produced a sufficient staining result (optimal or good).
* (number of optimal results/number of laboratories using this HIER buffer)

rmAb **EP3**: 3 of 5 (60%) protocols were assessed as optimal. Optimal protocols were based on HIER using TRS pH 9 (3-in-1) (Dako) (1/1), BERS2 (Bond, Leica) (1/1) or Tris-EDTA/EGTA pH 9 (1/2). The rmAb clone EP3 was diluted in the range of 1:70-200 depending on the total sensitivity of the protocol employed. Using these protocol settings, 4 of 4 (100%) laboratories produced a sufficient staining result.

rmAb **SP3**: 12 of 23 (52%) protocols were assessed as optimal. Optimal protocols were based on HIER using TRS pH 9 (3-in-1) (Dako) (1/3), BERS2 (Bond, Leica) (9/10), Tris-EDTA pH 9 (1/2) or Citrate pH 6 (1/2). The rmAb clone SP3 was diluted 1:50-250 depending on the total sensitivity of the protocol employed. Using this protocol setting, 17 of 17 (100%) laboratories produced a sufficient staining result.

pAb **A0485**: 35 of 50 (70%) protocols were assessed as optimal. Optimal protocols were based on HIER using either TRS low pH 6.1 (Dako) (20/29), TRS pH 9 (3-in-1) (Dako) (6/8), CC1 (BenchMark, Ventana) (2/4), BERS1 (Bond, Leica) (3/3), Citrate pH 6 (2/3) or unknown (2/2). The pAb A0485 was typically diluted in the range of 1:100-1,500 depending on the total sensitivity of the protocol employed. Using these protocol settings, 43 of 46 (93%) laboratories produced a sufficient staining result.

Comments
In this assessment and in concordance with the previous NordiQC assessments of HER-2 IHC, insufficient HER-2 staining results were characterised by too weak or false negative staining reactions. This was particularly and most critically observed as 0/1+ IHC reaction in the low-level HER-2 gene amplified breast carcinoma no. 2. This tumour was categorized as IHC 2+ in the NordiQC reference laboratories using three FDA/CE-IVD HER-2 IHC assays: PATHWAY® (Ventana), HercepTest™ (Dako) and Oracle™ (Leica) and showed a low level of HER-2 gene amplification (ratio 2.3 – 2.9) by ISH. False negative staining reaction of the breast carcinoma no. 2 was seen in 58% of the insufficient results (11 of 19).

The remaining insufficient results were typically characterised by a poor signal-to-noise ratio, impaired morphology, excessive counterstaining complicating interpretation or false positive 3+ IHC staining in the HER-2 non-amplified tumours.
False negative results were seen both in laboratory developed (LD) and FDA-/CE-IVD approved assays, while false positive results only were seen in 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, too short incubation time of the primary Ab and/or insufficient HIER). For the FDA-/CE-IVD approved systems, used according to the official package inserts for the respective systems, no single cause for insufficient and false negative staining reactions could be identified from the protocols submitted. However, compared to previous assessments an extended use of off-label protocols for the approved systems were applied by the participants. Three laboratories used the Dako HercepTest™ SK001 assay on the Bond IHC platform, while 5 laboratories performed the assay manually and as such not within the intended use and must consequently be considered as a LD assay. In this assessment, all 8 laboratories using HercepTest™ SK001 on either Bond or manually produced a result assessed as optimal. However, despite the encouraging results, off-label use must be carefully validated by the end-users on a large cohort of breast carcinomas (n=100, ASCO/CAP 2013 guidelines).

The Ventana PATHWAY® /CONFIRM™ HER-2 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. In contrast, internal studies previously performed in the NordiQC reference laboratory indicated a less precise and robust HER-2 IHC assay if UltraView was substituted by OptiView PATHWAY® /CONFIRM™. OptiView will typically amplify the analytical sensitivity of the IHC system 3-4x compared to UltraView. Consequently if OptiView is applied, the HER-2 IHC assay must be adjusted at other parameters e.g incubation time or 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 HER-2 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 Fig. 1. PATHWAY®/CONFIRM™ IHC assays have provided a consistent high pass rate throughout all HER-2 IHC runs in NordiQC.

The proportion of laboratories using FDA-/CE-IVD approved HER-2 IHC assays and LD assays is very consistent. In this run, 27% of the participants (n=99) used LD assays compared to the range of 23 - 31% in the last 11 assessments. Despite an overall improvement of the pass rate for LD HER-2 assays from run B1 to B23, 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 HER-2 assays provided a high proportion of optimal results (89%; 237 of 265), whereas only 63% of LD HER-2 assays were assessed as optimal (67 of 107). As shown in Fig. 2, LD HER-2 assays provided a reduced proportion of sufficient results but also demonstrates a shift towards lower assessment score (optimal to good), typically caused by 2+ staining reaction in one or both of the HER-2 non-amplified tumours (no. 2 and 5) expected to show a 0/1+ staining reaction. The staining reaction of 2+ in these tumours would not directly lead to a wrong diagnosis but require an additional ISH test due to the less precise IHC result.

The overall pass rate of 95% obtained in this assessment is very encouraging 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 at least for the five latest runs.

Figure 1. **Pass rates of 23 HER-2 IHC assessments in the NordiQC breast cancer module**

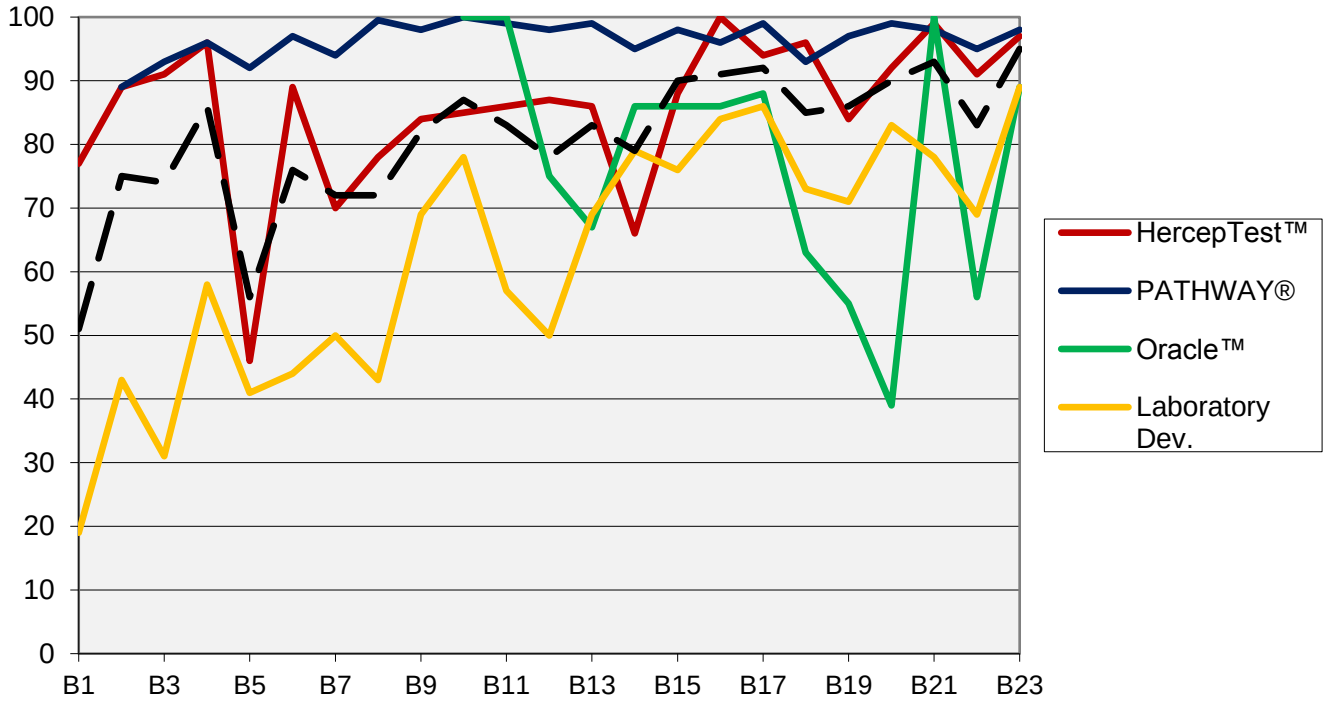
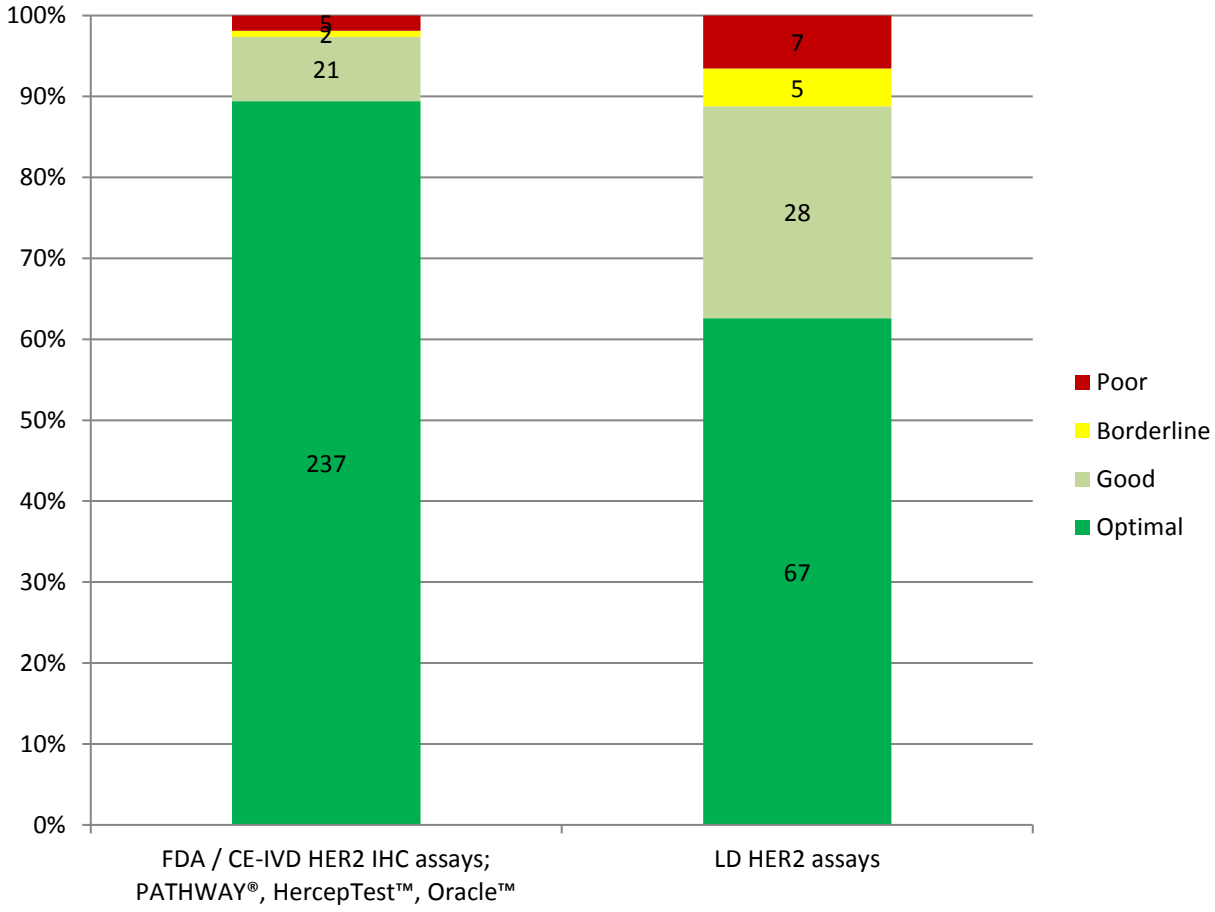


Figure 2. **Proportion of assessment marks using FDA-/CD-IVD and LD assays**



Four different cell lines (Horizon Discovery, Cambridge UK) were included in this assessment to evaluate if this material in combination with digital image analysis can be used to evaluate the accuracy and interlaboratory reproducibility of HER-2 IHC assays and potentially function as standard reference material for both EQA and internal quality control for HER-2 IHC assays. Subsequent analysis will be performed by NordiQC and published at a later stage.

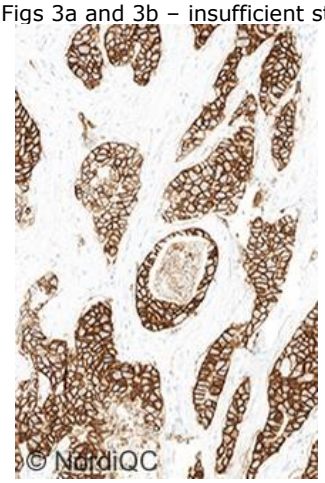
Scoring consensus B23

Laboratories were requested to submit scores (0, 1+, 2+, 3+) of their own HER-2 stained slides. This was done by 86% (320 of 372) of the participants. For 284 of the 320 (89%) 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 at the same level as in run B22, where 86% of the scores were in consensus with the NordiQC assessor group. Among laboratories with sufficient staining, 96% (274 of 284) of interpretations were in agreement with the NordiQC assessors. Interpretation in concordance with the NordiQC assessor group was seen in 78% (28 of 36) among participants with insufficient staining. Typically, breast carcinoma no. 2 was interpreted as 2+ by the laboratories, but 0-1+ by the NordiQC assessor group.

Conclusion

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



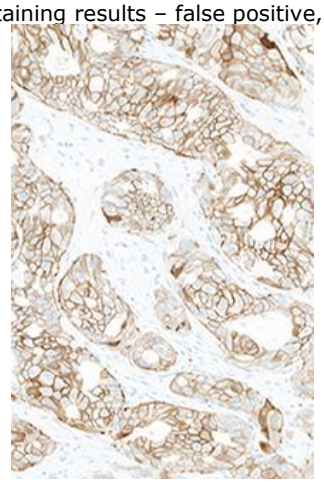
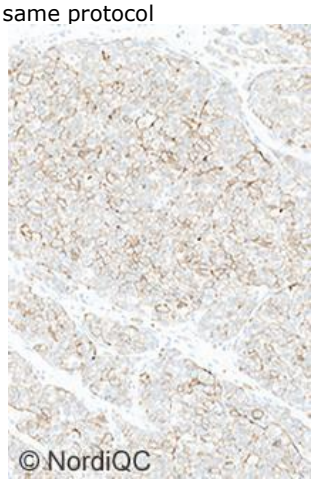


Fig 1a.
Left: Optimal staining result for HER-2 of the breast ductal carcinoma no. 1 with a ratio of HER-2 / chr17 of > 6.0.
> 10% of the neoplastic cells show an intense and complete membranous staining reaction corresponding to 3+.

Right: Optimal staining result for HER-2 of the breast ductal carcinoma no. 2 with a ratio of HER-2 / chr17 of 2.3 – 2.9.
> 10% of the neoplastic cells show a weak to moderate and complete membranous staining reaction corresponding to 2+.



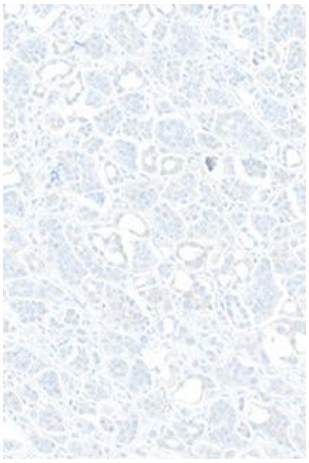
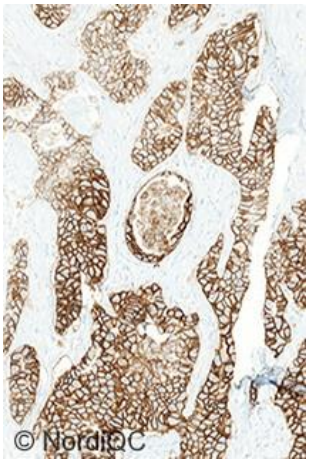


Fig 1b.
Left: Optimal staining result for HER-2 of the breast ductal carcinoma no. 4 with a ratio of HER-2 / chr17 of 1.3 – 1.7.
> 10% of the neoplastic cells show a weak to moderate membranous staining reaction corresponding to 2+.

Right: Optimal staining result for HER-2 of the breast ductal carcinoma no. 5 with a HER-2 / chr17 ratio of 1.2–1.4.
> 10% of the neoplastic cells show a faint membranous staining reaction corresponding to 1+.



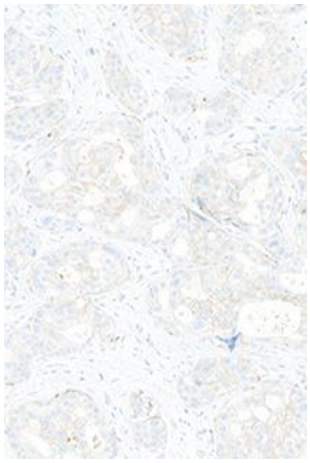
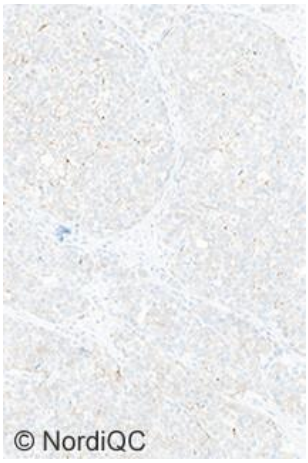


Fig 2a.
Left: Staining result for HER-2 of the breast ductal carcinoma no. 1 with a ratio of HER-2 / chr17 of > 6.0.
> 10% of the neoplastic cells show a strong and complete membranous staining reaction corresponding to 3+.

Right: Insufficient and false negative staining result for HER-2 of the breast ductal carcinoma no. 2 with a ratio of HER-2 / chr17 of 2.3 – 2.9.
> 10% of the neoplastic cells show a faint perceptible membranous staining reaction corresponding to 1+, but does not meet the criteria to be classified as 2+ and will not be referred to ISH.



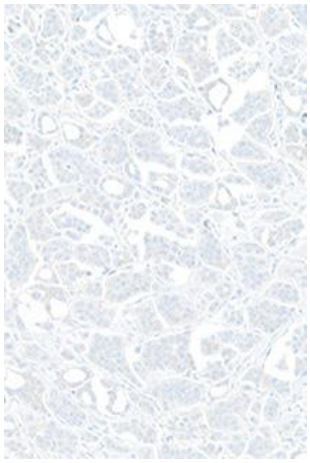
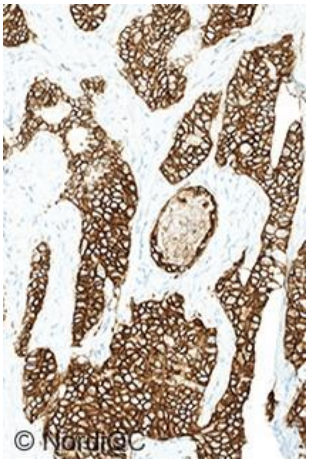


Fig 2b.
Left: Staining result for HER-2 of the breast ductal carcinoma no. 4 with a ratio of HER-2 / chr17 of 1.3 – 1.7.
> 10% of the neoplastic cells show a faint perceptible membranous staining reaction corresponding to 1+.

Right: Staining result for HER-2 of the breast ductal carcinoma no. 5 with a HER-2 / chr17 ratio of 1.2 – 1.4.
< 10% of the neoplastic cells show a faint membranous staining reaction corresponding to 0.



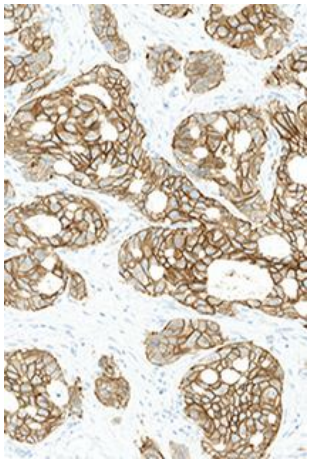
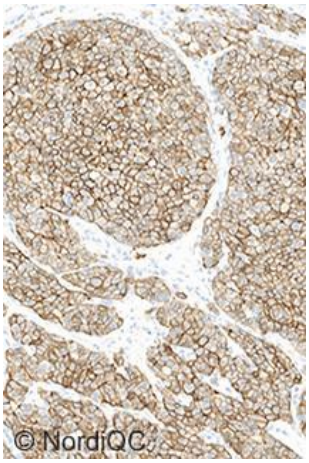


Fig 3a.
Left: Staining result for HER-2 of the breast ductal carcinoma no. 1 with a ratio of HER-2 / chr17 of > 6.0.
> 10% of the neoplastic cells show an intense and complete membranous staining reaction corresponding to 3+.

Right: Staining result for HER-2 of the breast ductal carcinoma no. 2 with a ratio of HER-2 / chr17 of 2.3 – 2.9.
> 10% of the neoplastic cells show a strong and complete membranous staining reaction corresponding to 3+.



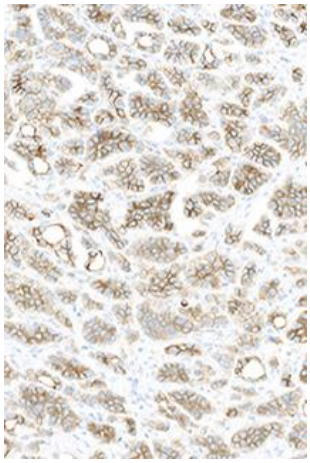


Fig 3b.
Left: Insufficient and false positive staining result for HER-2 of the breast ductal carcinoma no. 4 with a ratio of HER-2 / chr17 of 1.3 – 1.7.
> 10% of the neoplastic cells show an intense and complete membranous staining reaction corresponding to 3+.

Right: Staining result for HER-2 of the breast ductal carcinoma no. 5 with a HER-2 / chr17 ratio of 1.2 – 1.4.
> 10% of the neoplastic cells show a moderate membranous staining reaction corresponding to 2+.

SN/LE/RR 13.04.17