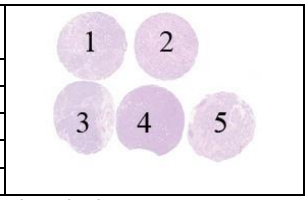
Assessment Run B24 2017
HER2 IHC

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	1-2+	1.3 – 1.5 (unamplified)***
2. Breast carcinoma, no. 2	0-1+	1.0-1.3 (unamplified)
3. Breast carcinoma, no. 3	1-2+	1.1 – 1.3 (unamplified)
4. Breast carcinoma, no. 4	0-1+	0.9-1.1 (unamplified)
5. Breast carcinoma, no. 5	3+	> 6.0 (clusters) (amplified)



* HER2 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.
** HER2 gene/chromosome 17 ratios achieved using Zyto*Light* ® SPEC HER2/CEN 17 Dual Color FISH (Zytovision)
*** By internal verification after material circulation, the tumour showed a non-amplified status and no 2+ amplified tumours were included in this assessment.

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. 2 and 4.
- Staining corresponding to score 1+ or 2+ in carcinoma no. 1 and 3.
- Staining corresponding to score 3+ in carcinoma no. 5.
- 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. 5 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 gene non-amplified tumour no. 2 and/or 4 showed a 2+ reaction and the other breast carcinomas showed the expected reaction pattern ***or*** (3) an overall weak staining reaction was observed.

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 (the 3+ tumour with gene amplification showed a 0 or 1+ reaction) or a false positive staining (e.g., the 0/1+ tumors and the 2+ tumours without gene amplification showing a 3+ reaction).

Participation

Number of laboratories registered for HER2, run B24	382
Number of laboratories returning slides	369 (97%)

Results: 369 laboratories participated in this assessment and 97% achieved a sufficient mark. Assessment marks for IHC HER2 assays and HER2 antibodies are summarized in Table 1.

Table 1. **Assessment marks for IHC assays and antibodies run B24, HER2 IHC**

FDA approved HER2 assays	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
PATHWAY® rmAb clone **4B5, 790-2991**	208	Ventana/Roche	184	22	2	0	99%	99%
PATHWAY® rmAb clone **4B5, 790-2991**[4]	1	Ventana/Roche	1	0	0	0	-	-
CONFIRM™, rmAb clone **4B5, 790-4493**	16	Ventana/Roche	13	2	1	0	94%	94%
HercepTest™ **SK001**	32	Dako/Agilent	26	6	0	0	100%	100%
HercepTest™ **SK001**[5]	7	Dako/Agilent	5	2	0	0	100%	-
HercepTest™ **K5204**	3	Dako/Agilent	2	1	0	0	-	-
Oracle™ mAb clone **CB11, TA9145**	7	Leica	1	4	2	0	71%	80%
Antibodies[3] for laboratory developed HER2 assays, conc. antibody	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
mAb clone **BS24**	1	Nordic Biosite	1	0	0	0	-	-
mAb clone **CB11**	13 1	Leica/Novocastra Biogenex	1	8	3	2	64%	-
rmAb clone **EP3**	1	Cell Marque	1	0	0	0	-	-
rmAb clone **SP3**	13 3 4 1 1 1	Thermo/NeoMarkers Zytomed Cell Marque Master Diagnostaica SL Immunologic Invitrogen	10	12	0	1	96%	100%
pAb clone **A0485**	48	Dako/Agilent	37	11	0	0	100%	100%
Ab clone **ZP1045Y, RM-2112-S**	1	Thermo Scientific	1	0	0	0	-	-
Antibodies for laboratory developed HER2 assays, RTU	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	Suff. OPS[2]
mAb clone **CB11, NCL-L-CB11**	1	Leica/Novocastra	0	0	1	0	-	-
rmAb clone **EP3, AN726**	1	Biogenex	0	1	0	0	-	-
rmAb **EP3, CCR-0843**	1	Celnovte	1	0	0	0	-	-
rmAb clone **EP3, RMPD049R**	1	Diagnostic Biosystems	1	0	0	0	-	-
rmAb **EP3, RMA-0690**	1	Maixin	1	0	0	0		
Ab clone **MXR001, RMA-0701**	1	Maixin	1	0	0	0	-	-
rmAb clone **SP3, 237R**	1	Cell Marque	0	1	0	0	-	-
Total	369		287	70	9	3	-	-
Proportion			78%	19%	2%	1%	97%	-

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 (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, Dako Omnis and manual staining)

Detailed Analysis
FDA/CE IVD approved assays

PATHWAY® rmAb clone **4B5** (790-2991, Ventana): 184 of 208 (88%) 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, 189 of 191 (99%) laboratories produced a sufficient staining result (optimal or good).

CONFIRM™ rmAb clone **4B5** (790-4493, Ventana): 13 of 16 (81%) protocols were assessed as optimal. Protocols with optimal results were typically based on HIER in CC1 (efficient heating time 30-64 min.) in BenchMark GX and Ultra, 16-32 min. incubation of the primary Ab and UltraView or OptiView as detection kit. Using these protocol settings, 15 of 16 (94%) laboratories produced a sufficient staining result.

HercepTest™ pAb (SK001, Dako): 26 of 32 (81%) 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 30 min. incubation of the primary Ab. Using these protocol settings, 30 of 30 (100%) laboratories produced a sufficient staining result.

Oracle™ mAb clone **CB11** (TA9145, Leica): 1 of 7 (14%) protocols were assessed as optimal. The optimal protocol was based on HIER in Bond Epitope Retrieval Solution 1 pH 6 (BERS1, Leica) for 25 min. and 30 min. incubation of the primary Ab. Using these protocol settings, 4 of 5 (80%) laboratories produced a sufficient staining result.

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**	87/87 (100%)	65/87 (75%)	129/131 (99%)	119/131 (91%)
Ventana BenchMark XT, GX, Ultra CONFIRM™ rmAb 4B5 **790-4493**	6/7 (86%)	5/7 (71%)	9/9 (100%)	8/9 (89%)
Dako Autostainer Link 48+ HercepTest™ pAb **SK001**	26/26 (100%)	20/26 (77%)	6/6 (100%)	6/6 (100%)
Leica Bond MAX, III Oracle™ mAb CB11 **TA9145**	4/5 (80%)	1/5 (20%)	1/2	0/2

* 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**: 37 of 48 (77%) protocols were assessed as optimal. Optimal protocols were based on HIER using either Target Retrieval Solution (TRS) low pH 6.1 (Dako) (21/27), TRS pH 9 (3-in-1) (Dako) (8/9), Cell Conditioning 1 (CC1, Ventana) (3/3), Bond Epitope Retrieval Solution 1 (BERS1, Leica) (1/2), Bond Epitope Retrieval Solution 2 (BERS2, Leica) (1/1), Tris-EDTA/EGTA pH 9 (1/1), Citrate pH 6 (1/3) and unknown (1/1). The pAb A0485 was typically diluted in the range of 1:100-1,600 depending on the total sensitivity of the protocol employed. Using these protocol settings, 45 of 45 (100%) laboratories produced a sufficient staining result (optimal or good).

rmAb **SP3**: 10 of 23 (43%) protocols were assessed as optimal. Optimal protocols were based on HIER using TRS pH 9 (3-in-1) (Dako) (1/2), BERS2 (Leica) (7/11), CC1 (Ventana) (1/4) or Citrate pH 6 (1/2). The rmAb clone SP3 was diluted 1:50-200 depending on the total sensitivity of the protocol employed. Using these protocol setting, 13 of 13 (100%) laboratories produced a sufficient staining result.

mAb **CB11**: 1 of 14 (7%) protocols was assessed as optimal. The optimal protocol was based on HIER using BERS1 (Leica) for 5 min. at 97°C. The mAb clone CB11 (Leica NCL-CB11) was diluted 1:200, incubated for 15 min. at RT, Bond Refine DS9800 used as detection system and stained on the BOND III platform.

rmAb **EP3** (Cell Marque): One protocol with an optimal result was based on HIER using BERS1 (Leica) for 30 min. at 100°C. The Ab was diluted 1:200, incubated for 15 min. at RT, Bond Refine DS9800 used as detection system and stained on the BOND III platform.

mAb clone **BS24** (Nordic Biosite): One protocol with an optimal result was based on HIER using Tris-EDTA / EGTA for 20 min. at 98°C. The Ab was diluted 1:75, incubated for 30 min. at RT and BioSite Histo Plus HRP Polymer anti-Mouse (KDB-10007) was used as detection system and stained on the Thermo Autostainer platform.

Table 3. **Optimal results for HER2 for the most commonly used antibodies as concentrate on the 3 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**	8/9 (89%)	21/27 (78%)	3/3	-	1/1	1/2
rmAb clone **SP3**	1/2	-	1/4	-	7/11 (64%)	-
mAb clone **CB11**	0/3	0/1	-	-	0/2	1/3

* 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)

RTU system
Ab clone **MXR001** (RMA-0701, Maixin): One protocol with an optimal result was based on HIER using Tris-EDTA / EGTA pH 9 for 20 min. at 100°C. The Ab was incubated for 60 min. at RT and MaxVision III Ultra DAB was used as detection system. The staining was performed manually.

Comments
In this assessment, the tumour composition of the block was different compared to the standard used by NordiQC. No HER2 2+ tumour with gene amplification was included, as the tumour tissue core no. 1 by internal validation showed to be non-amplified. This might have influenced the high overall pass rate of 97%, being the highest in the NordiQC HER2 IHC assessments performed. On the other hand, pass rates have generally increased in the last years, perhaps due to the extended and harmonized use of high quality HER2 IHC assays.

The predominant feature of an insufficient result for HER2 IHC was characterized by too weak or false negative staining reaction. This was observed as 0/1+ IHC reaction in the breast carcinoma core no. 5. This tumour was categorized as IHC 3+ in the NordiQC reference laboratory using three FDA/CE-IVD HER2 IHC assays: PATHWAY® (Ventana), HercepTest™ (Dako) and Oracle™ (Leica) and a high level of HER2 gene amplification (ratio >6) by ISH. False negative or too weak staining reaction of the breast carcinoma no. 5 was seen in 50% of the insufficient results (6 of 12).
The remaining insufficient results were typically characterized by a poor signal-to-noise ratio, impaired morphology, excessive counterstaining complicating interpretation (n=4) or false positive 3+ IHC staining in the HER2 non-amplified tumours (n=2).
An insufficient result caused by false positive staining was only seen in LD assays, while insufficient result caused by false negative staining or a poor signal-to-noise ratio were seen both in laboratory developed (LD) and FDA-/CE-IVD approved 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 use of detection system with low sensitivity). For the FDA-/CE-IVD approved systems, no single cause for insufficient staining reactions could be identified from the protocols submitted.

An extended off-label use was applied for the approved HER2 IHC systems. Three laboratories used the Dako HercepTest™ SK001 assay for manually staining, one laboratory used the Leica Bond III platform, while 3 laboratories performed the assay on Dako Omnis and, as such, not within intended use. Consequently, these were considered as LD assays. In this assessment, all 7 laboratories using HercepTest™ SK001 off-label on either Omnis, Bond or manually produced a result assessed as optimal or good. 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. 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 PATHWAY® /CONFIRM™. OptiView will typically amplify the analytical sensitivity of the IHC system 3-4x 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 or HIER settings to provide the analytical sensitivity level validated by Ventana, which as mentioned can cause a less precise and robust assay.

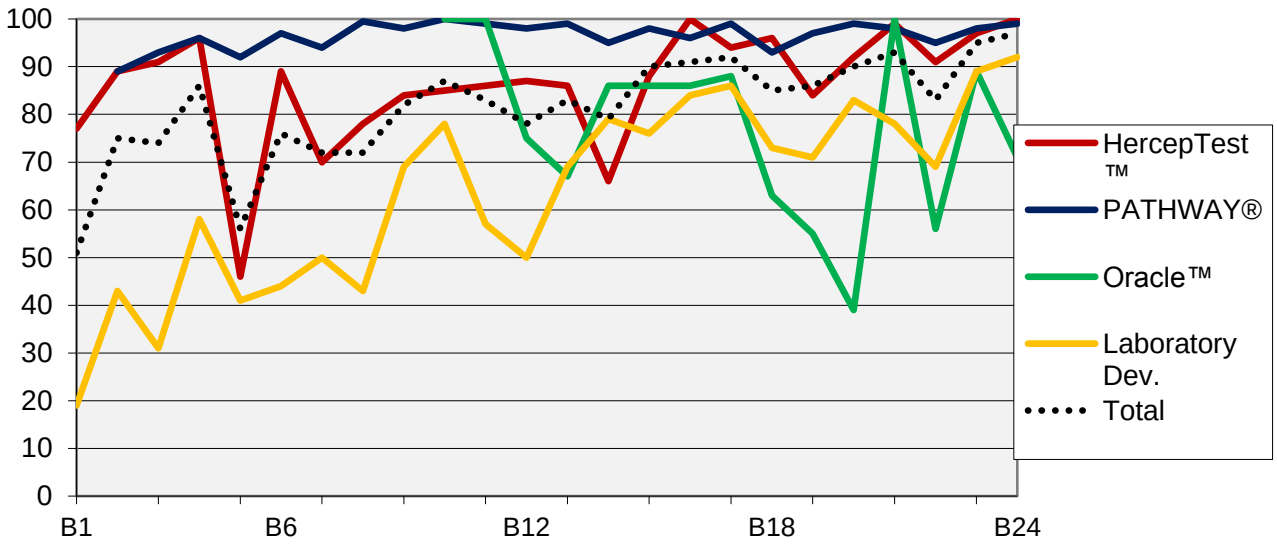
In this assessment, the FDA-/CE-IVD approved HER2 IHC assays PATHWAY® /CONFIRM™ and HercepTest™ from Ventana and Dako, respectively, were most successful and provided a higher pass-rate superior to LD assays as illustrated in Graph 1. PATHWAY®/CONFIRM™ IHC assays have provided a consistent high pass 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, 24% of the participants (n=95) used LD assays compared to 23-31% in the last 12 assessments. Despite an overall improvement of the pass rate for LD HER2 assays from run B1 to B24, the pass rate and proportion of optimal results is still 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 85% (226 of 266), whereas only 59% (61 of 103) of LD HER2 assays were assessed as optimal. As shown in Graph 2, LD HER2 assays both provided a reduced proportion of sufficient results but also a shift from optimal to good, typically caused by 2+ staining reaction in one or both of the HER2 non-amplified tumours (no. 2 and 4) 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.

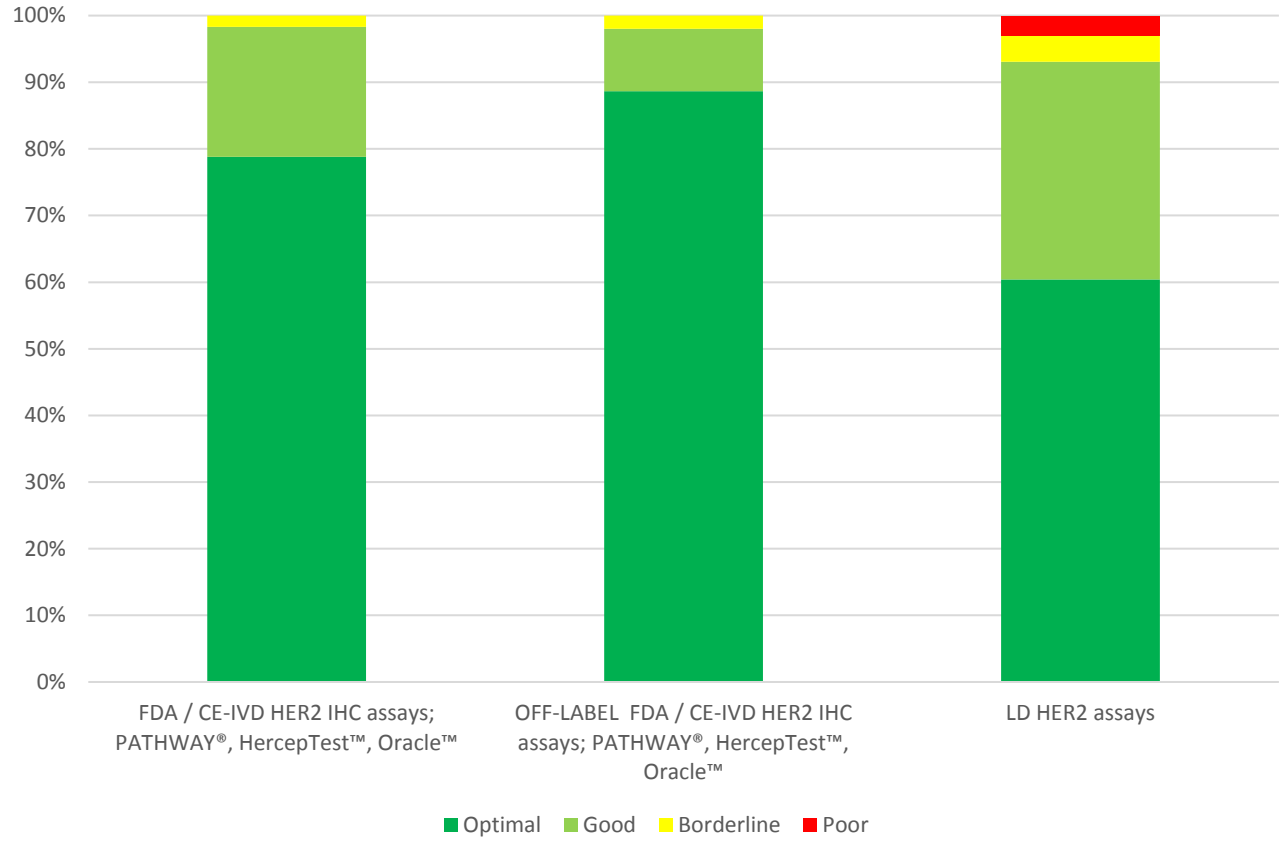
In this assessment NordiQC evaluated the performance of the FDA-/CE-IVD approved HER2 IHC assays both when used in concordance to the recommendations provided and as off-label use, as the latter in fact is considered as an LD assay. The overall result is also shown in Graph 2. Briefly, the overall pass rates were similar when the approved assays were used either as true plug-and-play systems or off-label, but in fact the off-label use provided a slightly higher proportion of optimal results. The observations must be interpreted with caution, as the composition of the material in this run was modified to the best practice composition of material for EQA of HER2 IHC. In NordiQC the inclusion of 2+ tumours with and without gene amplification have proven to be central for the assessment of the accuracy and calibration level of the HER2 IHC assays and without these, no confident conclusions for the performance of "on-label and off-label" use of approved HER2 assays can be generated.

The overall pass rate of 97% obtained in this assessment was the highest in the NordiQC breast cancer module 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. However, as mentioned above the data in this run must be interpreted with caution due to the altered composition in the material circulated.

Graph 1. **Pass rates of 24 HER2 IHC assessments in the NordiQC breast cancer module**



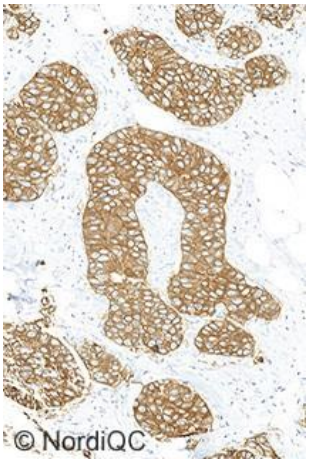
Graph 2. **Proportion of assessment marks using FDA-/CD-IVD and LD assays**



Scoring consensus B24
Laboratories were requested to submit scores (0, 1+, 2+, 3+) of their own HER2 stained slides. This was done by 85% (313 of 369) of the participants. For 283 of the 312 (91%) 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 that in run B23, where 86% of the scores were in consensus with the NordiQC assessor group. Among laboratories with sufficient staining, 98% (278 of 283) of interpretations were in agreement with the NordiQC assessors. Interpretation in concordance with the NordiQC assessor group was seen in 83% (25 of 30) among participants with insufficient staining. Typically, one (or more) of the cores without HER2 overexpression was interpreted as 3+ by the laboratory but 0-2+ by the NordiQC assessor group.

Conclusion
The FDA-/CE-IVD approved HER2 IHC assays **PATHWAY®/CONFIRM™** rmAb clone 4B5 (Ventana) and **HercepTest™** (Dako) were in this assessment the most successful assays for the semi-quantitative IHC determination of HER2 protein expression. Laboratory developed assays provided a slightly lower pass rate and were less precise for the HER2 status requiring an additional ISH test for final evaluation.

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



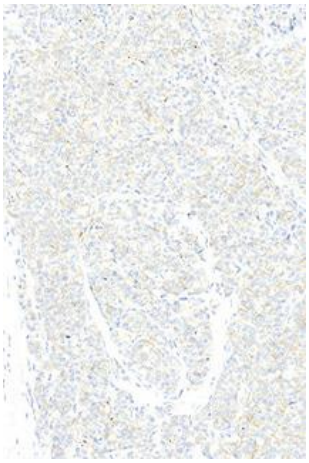
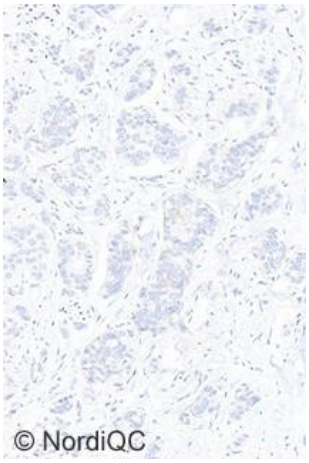


Fig 1a.
Left: Optimal staining result for HER2 of the breast ductal carcinoma no. 5 with a ratio of HER2 / chr17 of > 6.0.
> 10% of the neoplastic cells show an intense and complete membranous staining reaction corresponding to 3+.
Right: Optimal staining result for HER2 of the breast ductal carcinoma no. 3 with a ratio of HER2 / chr17 of 1.1 – 1.3.
> 10% of the neoplastic cells show a weak to moderate and complete membranous staining reaction corresponding to 2+.



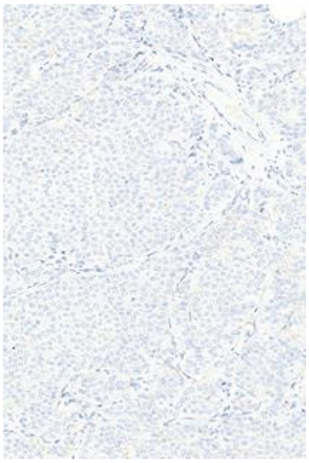
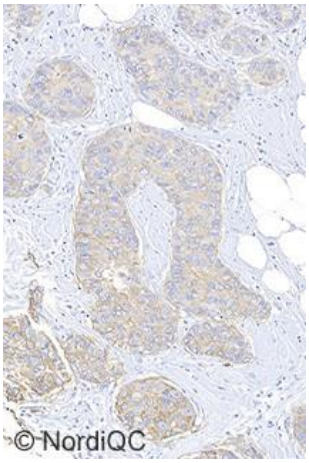


Fig 1b.
Left: Optimal staining result for HER2 of the breast ductal carcinoma no. 2 with a ratio of HER2 / chr17 of 1.0-1.3.
> 10% of the neoplastic cells show a faint membranous staining reaction corresponding to 1+.
Right: Optimal staining result for HER2 of the breast ductal carcinoma no. 4 with a HER2 / chr17 ratio of 0.9-1.1.
< 10% of the neoplastic cells show a faint membranous staining reaction corresponding to 0.



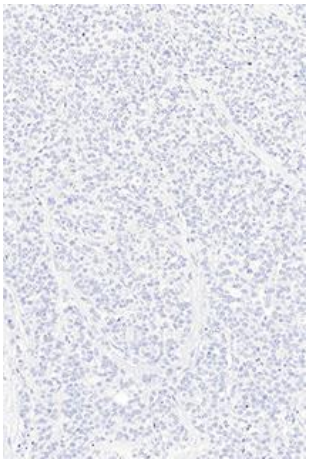
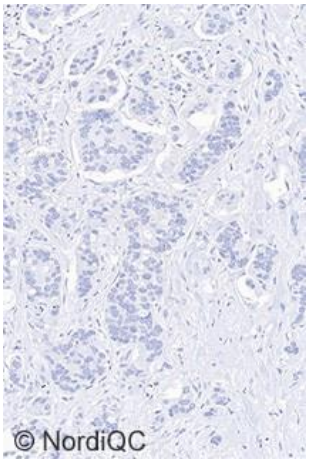


Fig 2a.
Left: Staining result for HER2 of the breast ductal carcinoma no. 5 with a ratio of HER2 / chr17 of > 6.0.
> 10% of the neoplastic cells show faint membranous staining and at the same time an excessive cytoplasmic staining reaction corresponding to 1+ compromising the interpretation.
Right: Staining result for HER2 of the breast ductal carcinoma no. 3 with a ratio of HER2 / chr17 of 1.1 – 1.3.
> 10% of the neoplastic cells show a weak to moderate and complete membranous staining reaction corresponding to 0.



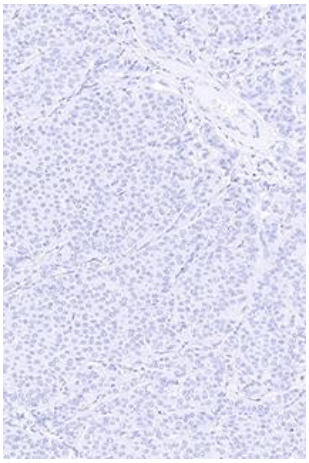
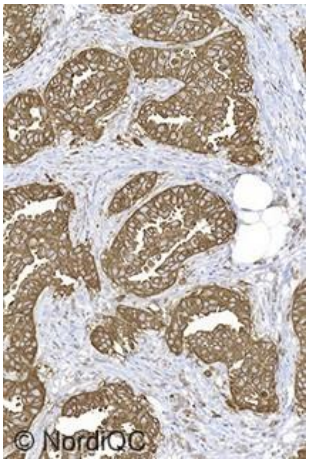


Fig 2b.
Left: Staining result for HER2 of the breast ductal carcinoma no. 2 with a ratio of HER2 / chr17 of 1.0-1.3.
< 10% of the neoplastic cells show a faint membranous staining reaction corresponding to 0.
Right: Staining result for HER2 of the breast ductal carcinoma no. 4 with a HER2 / chr17 ratio of 0.9-1.1.
< 10% of the neoplastic cells show a faint membranous staining reaction corresponding to 0.



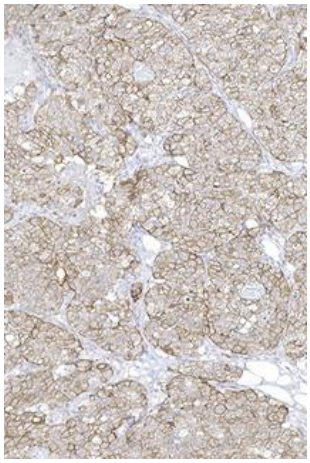
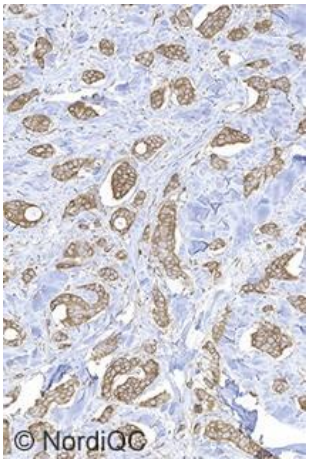


Fig 3a.
Left: Staining result for HER2 of the breast ductal carcinoma no. 5 with a ratio of HER2 / chr17 of > 6.0.
> 10% of the neoplastic cells show an intense and complete membranous staining reaction corresponding to 3+.
Right: Staining result for HER2 of the breast ductal carcinoma no. 3 with a ratio of HER2 / chr17 of 1.1 – 1.3.
> 10% of the neoplastic cells show a strong and complete membranous staining reaction corresponding to 3+.



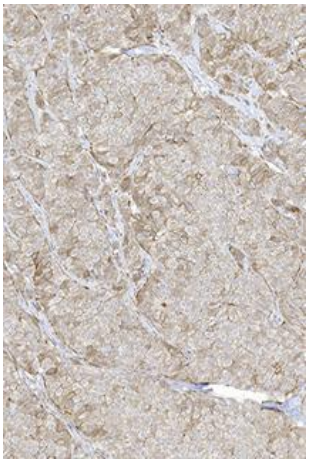


Fig 3b.
Left: Staining result for HER2 of the breast ductal carcinoma no. 2 with a ratio of HER2 / chr17 of 1.0-1.3.
> 10% of the neoplastic cells show an intense and complete membranous staining reaction corresponding to 3+.
Right: Staining result for HER2 of the breast ductal carcinoma no. 4 with a HER2 / chr17 ratio of 0.9-1.1.
> 10% of the neoplastic cells show a moderate membranous staining reaction corresponding to 2+.

HLK/LE/MV/RR 11.12.2017