

Application of Digital Image Analysis for the Assessment of HER2-low in Breast Cancer – Comparison with Visual Assessments Performed at The UK National External Quality Assessment Scheme for Immunocytochemistry and In-Situ Hybridisation (UK NEQAS ICC & ISH) HER2-low EQA Programme

Lila Zabaglo [1], Suzanne Parry [1], Dawn Wilkinson [1], Andrew Dodson [1]

[1] UK National External Quality Assessment Scheme for Immunocytochemistry and In-Situ Hybridisation, London UK. Corresponding author's email address: adodson@ukneqasiccish.org

Background and Aims

- Accurate assessment of lower ranges of HER2 immunohistochemical (IHC) expression has become important since trastuzumab deruxtecan (T-DXd) was approved for treatment of patients with HER2-low metastatic breast cancer (BC) [Ref.1].
- Findings from DESTINY-Breast06 Phase 3 trial showed that T-DXd offers a favourable outcome over standard chemotherapy in patients with HER2-low and HER2-ultra-low metastatic BC [Ref.2].
- Several studies showed low concordance among pathologists in distinguishing low levels of HER2 IHC expression.
- There is an emerging need for a standardized and reproducible HER2-low testing methodology.
- A dedicated EQA programme for the assessment of IHC stain quality of HER2-low testing was established by UK NEQAS ICC & ISH in 2023 providing support to clinical laboratories and gathering evidence about the reproducibility of HER2-low testing.
- We report results from the application of digital image analysis (DIA) for the HER2-low IHC assessment with a view of implementing DIA in the HER2-low EQA programme.

Materials and Methods

- Formalin-fixed paraffin embedded (FFPE) sections from a set of BC specimens showing a range of HER2 expression (from 0 to 2+) were prepared onto microscope slides (Fig. 1).
- Unstained sections were distributed to laboratories participating in the UK NEQAS ICC & ISH HER2-Low EQA programme.
- The participating laboratories stained the slides using their routine IHC methods for HER2 and returned for central assessment by an expert panel of assessors.
- Slides were scanned at x40 using a NanoZoomer (Hamamatsu, Japan).
- Digital image analysis (DIA) using Visiopharm HER2 APP 10185 (Visiopharm A/S, Hørsholm, Denmark) was applied to:
 - training set of 272 samples that were selected and scored by four Assessors as confident HER2 0, 1+, and 2+ samples.
 - 1554 samples from four EQA Runs (Runs 144-147 conducted at 4-monthly intervals).
- Agreement between DIA and visual scores was assessed using Spearman correlation coefficient.

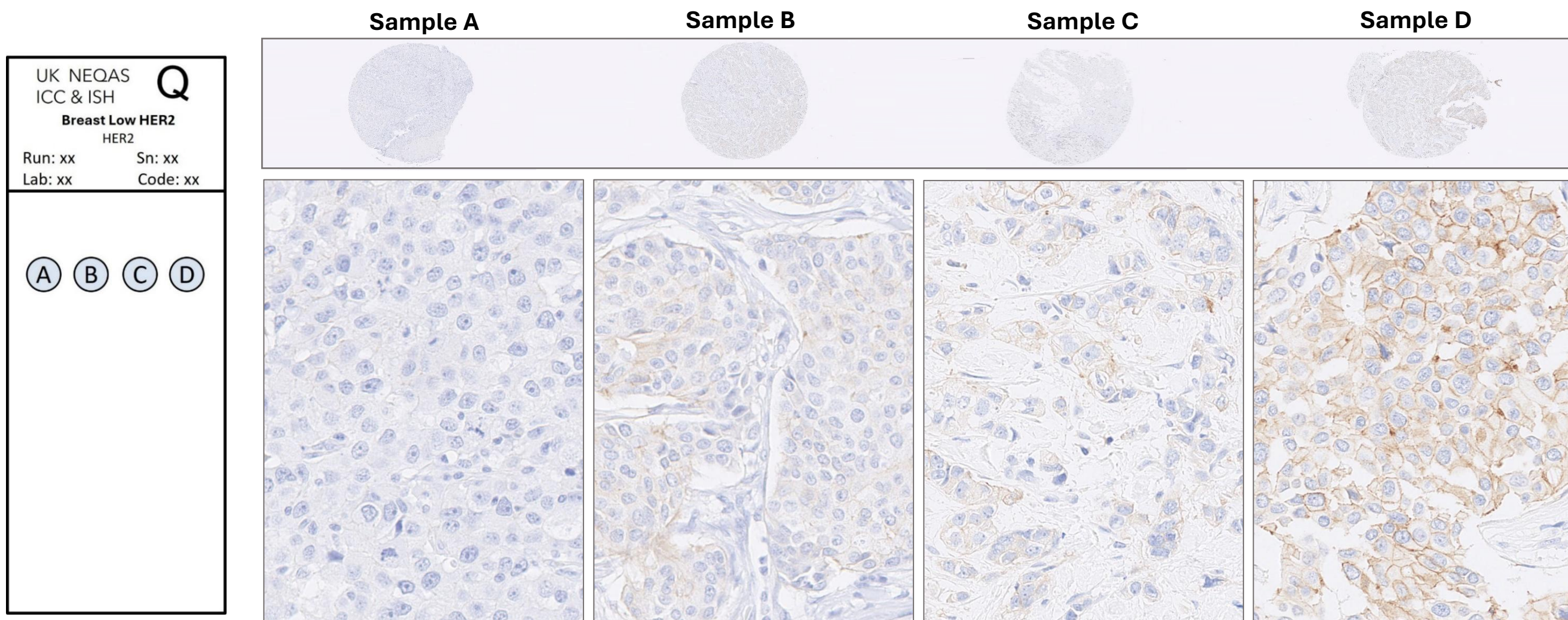
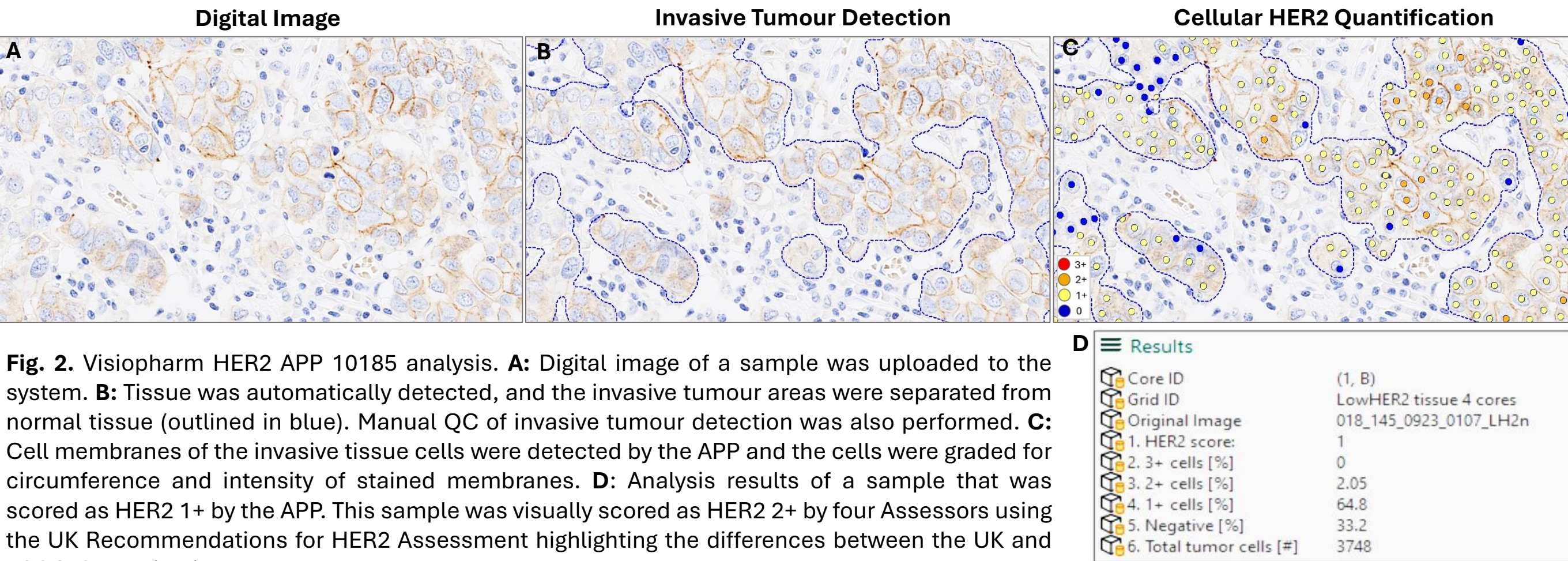


Fig. 1. UK NEQAS ICC & ISH HER2-Low EQA programme slide layout. TMA cores taken from FFPE breast cancer excision tissues (samples A, B, C, and D) showing varying levels of HER2 membrane protein expression (from HER2 negative to HER2 2+) were provided as the assessment samples.

HER2 Assessment

Visual assessment was performed by four Assessors experienced in breast HER2 IHC evaluation who scored the preparations concurrently around a multi-header microscope. Each Assessor rated the HER2 category. Scores which disagreed were resolved by consensus agreement.

DIA assessment was conducted using Visiopharm HER2 APP 10185 (Fig.2). It is an automatic single-cell approach APP for HER2 IHC quantification using an algorithm aligned with ASCO/CAP HER2 IHC guidelines for 0, 1+, 2+, and 3+ categorization.



Results

Training set: The correlation between DIA and visual scores was $r = 0.91$ ($p < 0.0001$), (Fig.3A). In the training set:

- 93% specimens assessed visually as ‘HER2 0’ and 99% specimens visually scored as ‘HER2 1+’ were DIA scored as ‘0’ and ‘1+’ respectively.
- 31% of samples visually scored as ‘2+’ were also scored as ‘2+’ by DIA, while the remaining BCs were scored as ‘1+’.

EQA Runs 144-147: Slightly lower correlation was observed in BC samples ($n = 1554$) from four EQA Runs ($r = 0.83$, $p < 0.0001$), (Fig.3B). In this set:

- 91% of samples visually classified as ‘0’ and 92% specimens visually scored as ‘1+’ were DIA scored as ‘0’ and ‘1+’ respectively (Fig.4).
- Similarly to the training set, lower percentage (34.6%) of samples that were scored visually as ‘2+’ were DIA classified as ‘2+’, while the majority of remaining BCs were DIA scored as ‘1+’. There were two cores (<1%) that were DIA scored as ‘3+’, (Fig.4).

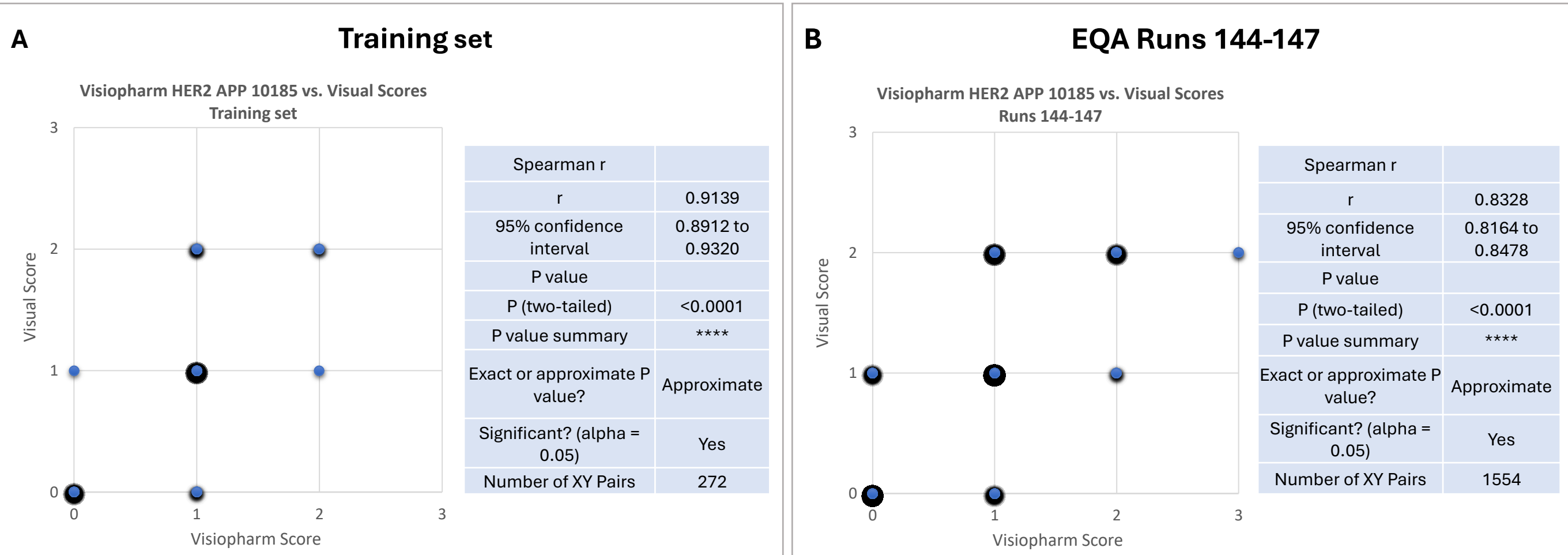


Fig. 3. Correlation between Visiopharm DIA and Assessors' visual scores in: (A) training set; (B) BC samples from four EQA runs. Agreement between DIA and visual scores was assessed using Spearman correlation coefficient.

Results

RUN 144		RUN 145		RUN 146		RUN 147		ALL RUNS 144-147	
Run 144 ALL cores		Run 145 ALL cores		Run 146 ALL cores		Run 147 ALL cores		Runs 144-147 ALL Cores	
Total	369	Total	390	Total	373	Total	422	Total	1554
Matched	270 73.17	Matched	329 84.36	Matched	319 85.52	Matched	380 90.05	Matched	1298 83.53
Mismatched:	99 26.83	Mismatched:	61 15.64	Mismatched:	54 14.48	Mismatched:	42 9.95	Mismatched:	256 16.47
DIA Upgraded from 0 to 1+	25 6.78	DIA Upgraded from 0 to 1+	5 1.28	DIA Upgraded from 0 to 1+	12 3.22	DIA Upgraded from 0 to 1+	11 2.61	DIA Upgraded from 0 to 1+	53 3.41
DIA Upgraded from 2+ to 3+	2 0.54			DIA Upgraded from 1+ to 2+	1 0.27	DIA Upgraded from 1+ to 2+	5 1.18	DIA Upgraded from 1+ to 2+	6 0.39
DIA Downgraded from 1+ to 0	6 1.62	DIA Downgraded from 1+ to 0	16 4.10	DIA Downgraded from 1+ to 0	6 1.61	DIA Downgraded from 1+ to 0	25 5.92	DIA Downgraded from 1+ to 0	53 3.40
DIA Downgraded from 2+ to 1+	66 17.89	DIA Downgraded from 2+ to 1+	40 10.26	DIA Downgraded from 2+ to 1+	35 9.38	DIA Downgraded from 2+ to 1+	1 0.24	DIA Downgraded from 2+ to 1+	142 9.14
Run 144 Cores Visually scored 0		Run 145 Cores Visually scored 0		Run 146 Cores Visually scored 0		Run 147 Cores Visually scored 0		Runs 144-147 Cores Visually scored 0	
Total	136	Total	149	Total	133	Total	151	Total	569
Matched	111 81.62	Matched	144 96.64	Matched	121 90.98	Matched	140 92.72	Matched	516 90.69
Mismatched:	25 18.38	Mismatched:	5 3.36	Mismatched:	12 9.02	Mismatched:	11 7.28	Mismatched:	53 9.31
DIA Upgraded from 0 to 1+	25 18.38	DIA Upgraded from 0 to 1+	5 3.36	DIA Upgraded from 0 to 1+	12 9.02	DIA Upgraded from 0 to 1+	11 7.28	DIA Upgraded from 0 to 1+	53 9.31
Run 144 Cores Visually scored 1+		Run 145 Cores Visually scored 1+		Run 146 Cores Visually scored 1+		Run 147 Cores Visually scored 1+		Runs 144-147 Cores Visually scored 1+	
Total	114	Total	193	Total	197	Total	261	Total	765
Matched	108 94.74	Matched	177 91.71	Matched	190 96.45	Matched	231 88.51	Matched	706 92.29
Mismatched:	6 5.26	Mismatched:	16 8.29	Mismatched:	7 3.55	Mismatched:	30 11.49	Mismatched:	59 7.71
DIA Downgraded from 1+ to 0	6 5.26	DIA Downgraded from 1+ to 0	16 8.29	DIA Downgraded from 1+ to 0	6 3.05	DIA Downgraded from 1+ to 0	25 9.58	DIA Downgraded from 1+ to 0	53 6.93
		DIA Upgraded from 1+ to 2+	1 0.50	DIA Upgraded from 1+ to 2+	1 0.50	DIA Upgraded from 1+ to 2+	5 1.91	DIA Upgraded from 1+ to 2+	6 0.78
Run 144 Cores Visually scored 2+		Run 145 Cores Visually scored 2+		Run 146 Cores Visually scored 2+		Run 147 Cores Visually scored 2+		Runs 144-147 Cores Visually scored 2+	
Total	119	Total	48	Total	43	Total	10	Total	220
Matched	51 42.86	Matched	8 16.67	Matched	8 18.60	Matched	9 90.00	Matched	76 34.55
Mismatched:	68 57.14	Mismatched:	40 83.33	Mismatched:	35 81.40	Mismatched:	1 10.00	Mismatched:	144 65.45
DIA Downgraded from 2+ to 1+	66 55.46	DIA Downgraded from 2+ to 1+	40 83.33	DIA Downgraded from 2+ to 1+	35 81.40	DIA Downgraded from 2+ to 1+	1 10.00	DIA Downgraded from 2+ to 1+	142 64.55
DIA Upgraded from 2+ to 3+	2 1.68							DIA Upgraded from 2+ to 3+	2 0.90

Fig. 4. Comparison of Assessors' visual and Visiopharm DIA scores in individual EQA Runs and in BC samples pooled from four runs. Numbers of matched and mismatched cases together with percentages are shown.

Conclusions

- Overall, strong agreement between visual and DIA scores was observed for both HER2 0 and HER2 1+ BCs in both training and EQA Runs 144-147 data sets. A significant level of disagreement was observed in the cases assigned as 2+ by visual assessment, where more than two-thirds of which were scored as 1+ by DIA.
- In samples from four EQA Runs:
 - DIA scored 53 cases (9.3%) as 1+, which visually were classified as negative. In 40% of cases DIA classified as HER2 1+, staining was close to the 10% threshold currently established by UK and ASCO/CAP guidelines.
 - The membrane staining of 53 cases (6.9%) was downgraded by DIA from HER2 1+ to 0. In approximately 50% of these cases, the number of cells with HER2 1+ staining was close to the 10% threshold.
 - Much larger variability was observed for BCs visually assessed as ‘2+’ where only 34.6% of samples were DIA classified as ‘2+’, while majority of BCs were DIA scored as ‘1+’. The results may mirror the overscoring of HER2 2+ tumours that is currently observed in the UK diagnostic community using the UK guidelines.
- DIA enhanced the HER2 scoring reproducibility and precision particularly in cases scoring close to categorical cut-off points.
- Implementing DIA in the HER2-low EQA programme will also improve workflow efficiency.

Acknowledgements

Authors gratefully acknowledge assistance of Visiopharm and contribution of Assessors participating in the EQA assessments.

References

- Ref.1 Narayan, P., et al., Clin Cancer Res, 2021. 27(16): 4478-4485
Ref.2 Bardia, A., et al., N Engl J Med, 2024. 391(22): 2110-2122