

Concordance of HER2-low scoring in breast carcinoma among expert pathologists in the United Kingdom and the republic of Ireland -on behalf of the UK national coordinating committee for breast pathology

Zaakouk, Mohamed; Quinn, Cecily; Provenzano, Elena; Boyd, Clinton; Callagy, Grace; Elsheikh, Soha; Flint, Joe; Millican-Slater, Rebecca; Gunavardhan, Anu; Mir, Yasmeen; Makhija, Purnima; Di Palma, Silvana; Pritchard, Susan; Tanchel, Bruce; Rakha, Emad; Atallah, Nehal M; Lee, Andrew H S; Pinder, Sarah; Shaaban, Abeer M

DOI:
[10.1016/j.breast.2023.06.005](https://doi.org/10.1016/j.breast.2023.06.005)

License:
Creative Commons: Attribution (CC BY)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Zaakouk, M, Quinn, C, Provenzano, E, Boyd, C, Callagy, G, Elsheikh, S, Flint, J, Millican-Slater, R, Gunavardhan, A, Mir, Y, Makhija, P, Di Palma, S, Pritchard, S, Tanchel, B, Rakha, E, Atallah, NM, Lee, AHS, Pinder, S & Shaaban, AM 2023, 'Concordance of HER2-low scoring in breast carcinoma among expert pathologists in the United Kingdom and the republic of Ireland -on behalf of the UK national coordinating committee for breast pathology', *The Breast*, vol. 70, pp. 82-91. <https://doi.org/10.1016/j.breast.2023.06.005>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

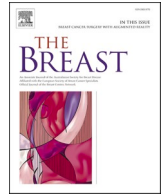
When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 26. Aug. 2023



Concordance of HER2-low scoring in breast carcinoma among expert pathologists in the United Kingdom and the republic of Ireland –on behalf of the UK national coordinating committee for breast pathology

Mohamed Zaakouk^{a,b}, Cecily Quinn^{c,d}, Elena Provenzano^{e,f}, Clinton Boyd^g, Grace Callagy^h, Soha Elsheikh^{i,j}, Joe Flint^k, Rebecca Millican-Slater^l, Anu Gunavardhan^m, Yasmeen Mirⁿ, Purnima Makhija^o, Silvana Di Palma^p, Susan Pritchard^q, Bruce Tanchel^r, Emad Rakha^s, Nehal M. Atallah^{s,t}, Andrew H.S. Lee^s, Sarah Pinder^u, Abeer M. Shaaban^{a,v,*}

^a Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, UK

^b Cancer Pathology, National Cancer Institute, Cairo University, Cairo, Egypt

^c Department of Histopathology, St Vincent's University Hospital, Elm Park, Ireland

^d UCD School of Medicine, Dublin, Ireland

^e Addenbrookes Hospital and NIHR Cambridge Biomedical Research Centre, Cambridge, UK

^f Department of Histopathology, Addenbrookes Hospital, Cambridge, UK

^g Histopathology, Belfast Health and Social Care Trust, Belfast, UK

^h Discipline of Pathology, University of Galway, School of Medicine, Lambe Institute for Translational Research, Galway, Ireland

ⁱ Department of Cellular Pathology, Royal Free Hospital, London, UK

^j Research Department of Pathology, University College London, Cancer Institute, London, UK

^k Birmingham Tissue Analytics, University of Birmingham, UK

^l Department of Cellular Pathology, St James's University Hospital, Leeds, UK

^m Department of Histopathology, Glan Clwyd Hospital Betsi Cadwaladr University Health Board, Bodelwyddan, UK

ⁿ Pathology, Liverpool University Hospitals Foundation Trust, Liverpool, UK

^o Pathology, Barts Health NHS Trust, London, UK

^p Cellular Pathology Department, Royal Surrey Hospital NHS Foundation Trust, Guildford, UK

^q Pathology, Wythenshawe Hospital Manchester Foundation Trust, Manchester, UK

^r Cellular Pathology, Heart of England NHS Foundation Trust, Birmingham, UK

^s Histopathology Department, Nottingham University Hospitals NHS Trust, City Hospital Campus, Nottingham, UK

^t Department of Pathology, Faculty of Medicine, Menoufia University, Egypt

^u School of Cancer & Pharmaceutical Sciences, Kings College London, London, UK

^v Cellular Pathology, Queen Elizabeth Hospital Birmingham, Birmingham, UK

ARTICLE INFO

Keywords:

Breast cancer
HER2-Low
HER2-Ultralow
Consistency
Concordance
Immunohistochemistry

ABSTRACT

Background: Recent clinical evidence showed that breast cancer with low HER2 expression levels responded to trastuzumab deruxtecan therapy. The HER2-low cancers comprise immunohistochemistry (IHC) score 1+ and 2+ ISH non-amplified tumours, currently classified as HER2 negative. Little data exists on the reproducibility of pathologists reporting of HER2-low cancer.

Patient and methods: Sixteen expert pathologists of the UK National Coordinating Committee for Breast Pathology scored 50 digitally scanned HER2 IHC slides. The overall level of agreement, Fleiss multiple-rater kappa statistics and Cohen's Kappa were calculated. Cases with low concordance were re-scored by the same pathologists after a washout period.

Results: Absolute agreement was achieved in 6% of cases, all of which scored 3+. Poor agreement was found in 5/50 (10%) of cases. This was due to heterogeneous HER2 expression, cytoplasmic staining and low expression spanning the 10% cut-off value. Highest concordance (86%) was achieved when scores were clustered as 0 versus others. Improvement in kappa of overall agreement was achieved when scores 1+ and 2+ were combined. Inter-observer agreement was moderate to substantial in the whole cohort but fair to moderate in the HER2-low group.

* Corresponding author. Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, UK

E-mail address: a.shaaban@bham.ac.uk (A.M. Shaaban).

<https://doi.org/10.1016/j.breast.2023.06.005>

Received 26 April 2023; Received in revised form 12 June 2023; Accepted 13 June 2023

Available online 27 June 2023

0960-9776/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Similarly, consensus-observer agreement was substantial to almost perfect in the whole cohort and moderate to substantial in the HER2-low group.

Conclusion: HER2-low breast cancer suffers from lower concordance among expert pathologists. While most cases can reproducibly be classified, a small proportion (10%) remained challenging. Refining the criteria for reporting and consensus scoring will help select appropriate patients for targeted therapy.

1. Introduction

Since the introduction of adjuvant trastuzumab anti-HER2 therapy and the subsequent development of several further anti-HER2 agents, accurate identification of patients with HER2 positive breast carcinoma has been essential to ensure appropriate personalised management of their disease [1]. This requires diagnostic tests that can correctly identify HER2 positive breast cancers. For the past 2 decades, a binary algorithm has been used that classifies breast cancer into HER2 positive (immunohistochemistry (IHC) scores 3+ and/or 2+ in situ hybridisation (ISH) amplified) and HER2 negative (IHC scores 0,1 or 2+ ISH non amplified) [2,3].

Recently, it was recognized that a group of HER2-low expressing tumours responded to trastuzumab deruxtecan (T-DXd) therapy; an antibody drug conjugate (ADC) that affects cancer cells expressing small quantities of HER2 protein and those surrounding them via a by-stander effect. DESTINY-Breast 04 was the first open label randomised controlled phase 3 trial for T-DXd for advanced and all metastatic HER2-low breast cancer including both hormonal receptor positive (n = 494, 88.7%) and negative (n = 63, 11.3%) tumours. Both progression free survival and overall survival were significantly longer in the T-DXd treated group compared with physician choice (9.9 versus 5.1 months and 23.4 versus 16.8 months, $p < 0.001$ for both) [4]. Based on the trial findings, the Food and Drug Administration (FDA) approved T-DXd for the treatment of patients with unresectable or metastatic HER2-low breast cancer. The drug has subsequently been approved by the European Medicines Agency as the first monotherapy for treating metastatic HER2-low breast cancer [5]. In the UK, T-DXd is currently licenced as monotherapy for the treatment of patients with unresectable or metastatic HER2+ breast cancer who have received one or more prior anti-HER2-based regimens [6] and its use for the same FDA approved indication is being considered by the National Institute for Health and Care Excellence (NICE).

HER2-low tumours (IHC score 1+ and/or 2+, ISH non-amplified) are presently classified as HER2 negative and include both luminal (oestrogen receptor positive) and triple negative breast cancer (TNBC). When the value of HER2-low breast cancer becomes widely applicable, clinical decision-making will require a move away from the binary paradigm for HER2 scoring. However, there are minimal data regarding the reliability of HER2-low scoring among pathologists and this category is not currently separated from the HER2 negative cancers in daily practice. The recently updated UK HER2 reporting guidelines [7] and the updated American Society of Clinical Oncology–College of American Pathologists Guideline Update [8] have recognized the importance of distinguishing the HER2-low cancers to inform therapy with trastuzumab deruxtecan. In a recent study of 105 HER2 non-amplified breast carcinomas scored by 16 pathologists, Baez-Navarro et al., showed moderate overall concordance [9]. Furthermore, Scott et al., in 2021 showed significant discordance in the reported HER2 status between local and central review laboratories with concordance rates of 70.8% and 40% for HER2 1+ and HER2 2+ groups respectively [10].

The present study aimed to assess the inter-observer concordance in the reporting of HER2-low and ultralow breast carcinoma among expert breast pathologists to assess the accuracy of identifying this group of tumours, highlight areas of reporting difficulty and provide suggestions for improving analytical consistency.

2. Materials & methods

50 invasive breast cancer core biopsies, from routine clinical practice, enriched for HER2 scores 0, 1+ and 2+ categories were selected from the archives of a single large UK institution (Queen Elizabeth Hospital Birmingham). The department is a reference centre for HER2 immunohistochemistry and FISH testing, performing approximately 4400 HER2 immunohistochemistry tests and 1200 HER2 Fluorescent in situ hybridisation (FISH) tests per annum. All slides were stained using the Ventana 4B5 assay on the Ventana Autostainer ((Ventana Bench-Mark Ultra, Roche, Indianapolis, USA) following the manufacturer's instructions. Each slide included 3 breast cancer tissue controls of HER2 scores 0, 2+ and 3+ as per the UK National External Quality Assurance Scheme (NEQAS) recommendations [11]. Paired H&E and HER2 IHC slides were reviewed prior to scanning to ensure adequate amount of tumour tissue and optimal quality of staining. Slides were then digitally scanned using a Leica Aperio AT2 slide scanner (Leica Biosystems Imaging, California, USA) at $\times 40$ magnification and uploaded to the University of Birmingham digital platform via a secure link: <https://eslidepath.bham.ac.uk>. Each participant was provided with secure access to the digital platform using a unique username and password.

The study cases were scored by 16 pathologist members of the UK National Coordinating Committee for Breast Pathology (NCCBP); a steering committee for the UK National Health Service (NHS) Breast Screening Programme and for the Royal College of Pathologists. Membership includes expert breast pathologists acting as professional clinical advisors to the NHS Breast Screening Programme and covering the geography of the UK and Republic of Ireland. In addition to the final HER2 scores of 0, 1+, 2+ or 3+, participants also scored the percentage ($< 10\%$ or $\geq 10\%$), intensity and completeness of membranous expression (complete versus incomplete) to arrive at the final scoring. This method is more detailed and is in line with the algorithm for HER2 scoring in the updated UK HER2 scoring guidelines [7]. Fig. 1 summarizes the study design and Table 1 shows the parameters assessed by each of the scoring pathologists.

Data on patients' demographics, tumour type, grade and hormonal status were collected. The original HER2 scores were obtained from the histopathology reports. The FISH amplification status, for the HER2 IHC score 2+ (equivocal) cases, was also recorded.

3. Statistical analysis

Statistical analysis was performed using the SPSS (IBMS) software version 28. For statistical analysis, cases were categorised into whole cohort as well as HER2-low group (HER2 IHC score 1+ or 2+ FISH non-amplified), HER2-ultralow (score 0 with incomplete membrane staining of less than 10%) and HER2 positive (IHC score 3+ or 2+ FISH amplified).

If all participants agreed (100%; 16/16 scorers) this was regarded as absolute agreement (AA). The HER2 preparations with agreement of 12–15 (75–94%) raters were regarded as showing high agreement, those with agreement of 9–11 raters ($> 50\%$ – $< 75\%$) were regarded as low agreement level, while those with ≤ 8 ($\leq 50\%$) were considered to represent challenging cases (Table 2). In addition to calculating the level of agreement for the 4 standard scoring categories (0, 1+, 2+, 3+), further three clustered scoring categories were used as follows: 0/1 (for IHC scores 0 and 1+), 1/2 (for IHC scores 1+ and 2+) and lastly dividing scores into 0 against all other scores.

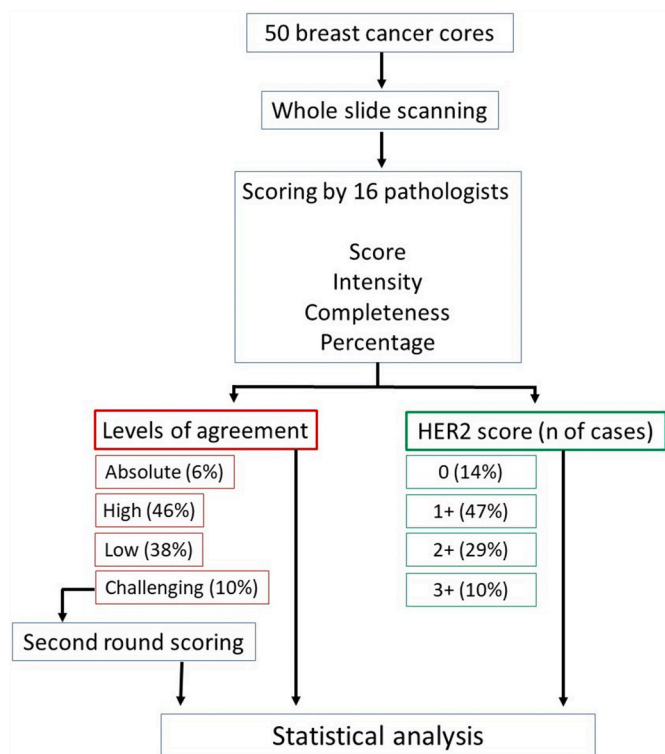


Fig. 1. Flow chart of the study design.

Table 1
HER2 scoring parameters.

Parameter	Score	Description
HER2 Final score	0	Negative
	1+	Negative (1+)
	2+	Equivocal
	3+	Positive
Intensity	0	None
	1	Faint
	2	Weak
	3	Moderate
Completeness of membrane staining	4	Strong
	0	Incomplete
	1	Complete
Percentage (%)	0	<10
	1	≥10

Table 2
Levels of agreement.

Number of concordant scorers (total =16)	Percentage of scorers agreeing	Level of agreement
16	100%	Absolute
12-15	75-100%	High
9-11	>50-75%	Low
≤8	≤50%	Poor/Challenging

HER2 stained tumours designated as challenging with agreement only among 8 pathologists (50%) were studied and re-scored by all 16 pathologists. Since the level of agreement does not necessarily reflect clinical significance in management of the patients, if the scores were all on the positive or the negative sides, only scores which were significantly different, resulting in potential over or under-treatment, were designated as significant minority scoring.

Fleiss multiple-rater kappa (κ) statistics of inter-observer agreement

for scoring cases (as 0, 1+, 2+ and 3+) was calculated. Fleiss' kappa, κ is a measure of inter-rater agreement used to determine the level of agreement between two or more raters when the method of assessment, known as the response variable, is measured on a categorical scale. Fleiss multi-rater kappa was used to measure overall agreement (OA) and agreement on scoring categories, among all the pathologists. Cohen's weighted Kappa was calculated to assess inter-observer agreement between each two of the scoring pathologists as well as between each pathologist and the consensus score. Cohen's weighted kappa was analysed both in the whole cohort and in the HER2-low group. Kappa results are interpreted as follows: values ≤ 0 indicate no agreement and 0.01–0.20: none to slight, 0.21–0.40: fair, 0.41–0.60: moderate, 0.61–0.80: substantial and 0.81–1.00: almost perfect agreement.

For statistical analysis, the consensus score was considered the reference score, upon which cases were classified into HER2 negative, ultralow, low and positive categories.

4. Results

4.1. Cohort characteristics

The cohort comprised 50 breast tumours. The median age of patients was 58.5 years (IQR 50.75-76.25). The invasive carcinomas were predominantly grade 2 (29/58; 58%). Invasive carcinoma of no special type (NST) was the commonest (41/50; 82%), followed by invasive lobular carcinoma (4/50; 8%).

The original HER2 scoring, designated by the pathologist at the time of reporting, divided tumours into score 0 (20/50; 40%), 1+ (13/50; 26%), score 2+ 'equivocal' (12/50; 24%) and 3+ (5/50; 10%). Equivocal cases (score 2+) were subsequently categorised by FISH into non-amplified (10/50; 20%) and amplified (2/50; 4%), Table 3.

4.2. Consensus scoring

The 16 pathologists scored 50 cases, with a total of 800 scores recorded. The most common category was 1+ (323/800; 40.40%), followed by 2+ (259/800; 32.38%), 0 (136/800; 17%) and 3+ (82/800; 10.25%). When the consensus score (score agreed by the majority of raters) was considered, the distribution of scores was comparable, where score 1+ was the mostly agreed score (23.5; 47%), followed by 2+ (14.5; 29%), 0 (7; 14%) and 3+ (5; 10%). According to consensus/majority scoring, the 7 tumours scored as 0, were subsequently divided into HER2 negative (1/50; 2%) and HER2 ultralow (6/50; 12%), the latter defined as evident focal staining in less than 10% of cells. Tumours scored as 2+, were divided according to the results of prior FISH testing where available (results not known to the participating pathologists/raters) into FISH amplified (n = 2) and FISH non-amplified (n = 12.5) including a case with no agreement with half the raters scoring it as 2+ and the other half as 1+. The latter group, in addition to cases scored as 1+ (n = 23.5), represented the HER2-low group (n = 36). FISH amplified cases, in addition to cases scored as 3+, represented the HER2 positive tumour group cases (n = 7). The distribution of the 800 scores, submitted by the pathologists and the consensus scores are shown in Table 4.

4.3. Levels of agreement

The percentage of complete agreement (absolute agreement), in which all 16 pathologists assigned tumours to the same IHC score category (0, 1+, 2+ or 3+), was achieved in only 6% (3 of 50), all of which were scored as 3+, Fig. 2A. No absolute agreement was achieved in tumours scored as 0, 1+ or 2+ (Table 5).

When HER2 scores were combined, the level of agreement increased, the percentage of tumours with absolute and high agreement both increased. The highest level of absolute and high agreement (86%) was achieved when tumours were divided into just 2 categories (score 0 versus others). This was followed by using the combined category

Table 3
Clinicopathological characteristics of the studied cohort.

Age (yrs)	Median (IQR)	58.50 (50.75; 76.25)		
Grade	1	10/50; 20%		
	2	29/50; 58%		
	3	11/50; 22%		
Histologic Type	NST	41/50; 82%		
	Lobular	4/50; 8%		
	Tubular	2/50; 4%		
	Mucinous	1/50; 2%		
	Metaplastic	1/50; 2%		
	Apocrine	1/50; 2%		
Molecular Type	Luminal	37/50; 74%		
	Her2 positive	8/50; 16%		
	Triple negative	5/50; 10%		
ER Status	Negative	7/50; 14%	Absent	7/50; 14%
	Positive	43/50; 86%	Weak	3/50; 6%
PR Status	Negative	15/49; 31%	Moderate	0/50; 0%
			Strong	40/50; 80%
	Positive	34/49; 69%	Weak	1/49; 0.02%
HER2 Original Score	0	1/50; 2%	No staining	Negative
			Staining in <10%	Negative (HER2 ultralow)
	1+	13/50; 26%	FISH Non-Amplified	Her2-Low
	2+	10/50; 20%	FISH Amplified	HER2 Positive
	3+	5/50; 10%		HER2 Positive

Table 4
Distribution of the scores among raters.

Pathologist	HER2 Scoring Categories			
	0	1+	2+	3+
P1	19	13	13	5
P2	6	18	22	4
P3	3	12	31	4
P4	4	27	13	6
P5	7	14	25	4
P6	5	33	7	5
P7	7	19	19	5
P8	20	17	9	4
P9	15	20	10	5
P10	1	20	20	9
P11	9	22	14	5
P12	19	16	10	5
P13	5	34	7	4
P14	6	24	16	4
P15	2	10	32	6
P16	8	24	11	7
Total (n); %	136	323 40.4	259 32.4	82
Consensus (n); %	17			10.2
	7	23.5	14.5	5
	14	47	29	10
HER2 Categories (n)	N	UL	LOW	POS
	1	6	36	7

N: Her2 negative; UL: HER2 ultralow; LOW: HER2-low; POS: HER2 positive.

(1+/2+) where high and absolute agreement was achieved in 84% of tumours. For the HER2-low category alone, combining scores 1+ and 2+ resulted in an absolute and high agreement in 92% of cases.

5. Challenging cases

Overall agreement above 50% (8/16) was seen in the majority of tumours (45/50, 90%) while the remaining tumours (5/50; 10%) had poor agreement. These challenging tumours were all HER2 negative, of luminal molecular type and belonged to the HER2-low (n = 4) and HER2 ultralow (n = 1) groups, Table 6. Using the current HER2 scoring algorithms and clinical management pathways, these tumours would all be classified as HER2 negative and no adverse clinical impact expected. Reasons for discordance included heterogeneous HER2 expression with some focal staining bordering the 10% cut-off leading to difficulties in discerning 0 from the 1+ category. Cytoplasmic staining was also a major contributing factor particularly when prominent and mimicking faint/weak membrane staining, leading to overestimation of HER2 staining Figs. 2B,C, 2D. One case was focally blurred on the digital scanned image resulting in difficulties in scoring.

The 5 challenging cases were subsequently re-scored (supplementary spreadsheet 1) to assess intra-observer concordance. The distribution of raters' scores in both rounds is shown in Table 6. The second round scoring resulted in better level of agreement in 4 cases, while agreement remained the same in the fifth case. The level of agreement improved from poor to low and high in 3/5 and 1/5 cases respectively, while remained poor in the fifth case. This latter case was a mucinous carcinoma.

5.1. Minority scores of adverse clinical significance

Tumours which showed a difference between each rater's score and the consensus score, potentially resulting in either under or over-treatment and hence inappropriate management, were considered scores of adverse clinical significance. Eight tumours (8/50; 16%) showed such significant minority scoring and were submitted by a total of 6 pathologists (Table 7). These 8 tumours included those designated as negative (n = 3) while the consensus was HER2 positive, and tumours classified as positive (n = 5) with the consensus being negative. One tumour (case 8, Table 7) showed two distinct clones (clustered heterogeneity); a minority 3+ clone and a majority 0 clone with variable interpretation of results (Two pathologists reported as 3+, three reported as 1+ and one as 0, while majority (n = 10) as 2+). Figs. 2E and 2F.

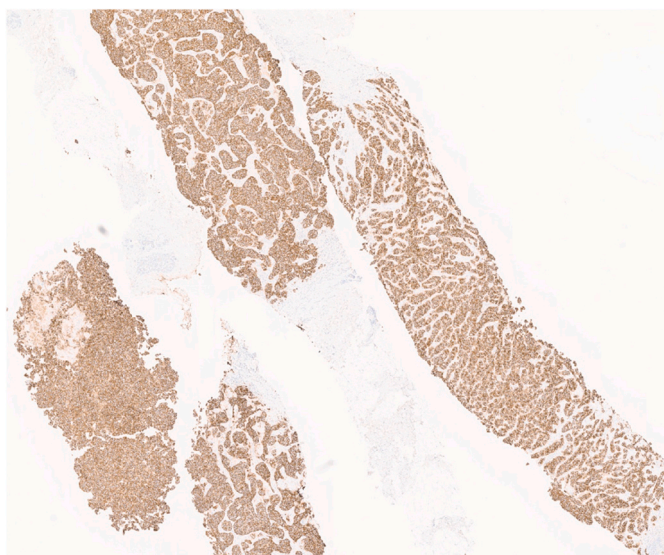
Fleiss Multi-rater kappa of overall agreement (OA) and among scoring categories was calculated, Table 8. Improvement in kappa of OA agreement and individual agreement was obtained with clustering scores 1+ and 2+ categories (1/2), Table 8.

5.2. Inter-observer and consensus observers' agreement

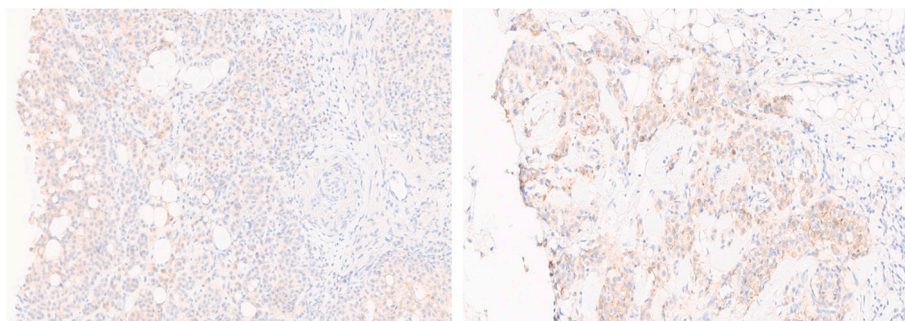
Cohen's Kappa was calculated to assess inter-observer agreement between each two of the scoring pathologists as well as between each pathologist and the consensus score. This has been investigated in the whole cohort and in the HER2-low group (Table 9). In the whole cohort, inter-observer agreement was predominantly moderate (46%) to substantial (41%), while it was fair (40.7%) to moderate (30%) in the HER2-low group. Similarly, consensus-observer agreement was substantial (50%) to almost perfect (19%) in the whole cohort, while it was moderate (43.75%) to substantial (25%) in the HER2-low group (Table 10).

6. Discussion

This is the first set of data from the UK and Ireland on the consistency of pathologists reporting of the HER2-low and ultralow categories of



A. Breast carcinoma showing absolute (100%) agreement on scoring of 3+.



B, C. Two fields of challenging case 4 with a consensus score of 1+ (HER2-low).

Fig. 2a. Examples of HER2 immunohistochemical staining and interpretation in the studied cohort. A. Breast carcinoma showing absolute (100%) agreement on scoring of 3+. B, C. Two fields of challenging case 4 with a consensus score of 1+ (HER2-low).

Table 5
Levels of agreement across standard & clustered categories.

Levels of Agreement	Individual categories				Combined categories				
	0	1+	2+	3+	0	1/2/3	0	1/2	3+
Absolute	3; 6%				22; 44%		15; 30%		
	0	0	0	3					
High	23; 46%				21; 42%		27; 54%		
	3	9	10	1					
Low	19; 38%				6; 12%		7; 14%		
	3	11	4	1					
Challenging	5; 10%				1; 2%		1; 2%		
	1	3.5	0.5						
Cases/category	7	23.5	14.5	5	7	43	7	38	5

invasive breast carcinomas. In addition to the categorisation of the HER2 IHC into the standard scores of 0, 1+, 2+ and 3+, pathologists detailed the parameters they based their scores on, including whether the membranous staining was circumferential or not, intensity of staining and the percentage of membranous staining expressed as < or ≥10%. These parameters follow the recently updated UK HER2 reporting guidelines that also recognize HER2-low tumours as a new entity

[7]. The majority of tumours (90%) were accurately classified. However, a small proportion (10%) remained challenging. Reasons for discordance were predominantly due to difficulties in estimating percentage expression around the 10% cut off value. HER2 cytoplasmic staining rendered the assessment of faint membrane positivity rather tricky and heterogeneity of HER2 expression resulted in difficulties in estimating the percentage of expression. Cohen Kappa statistics showed lower inter-observer and observer-consensus agreement in the HER2-low group compared with the whole four standard HER2 IHC scoring categories. This lower consistency for diagnosing the HER2-low group requires further work to refine the categorisation and improve concordance.

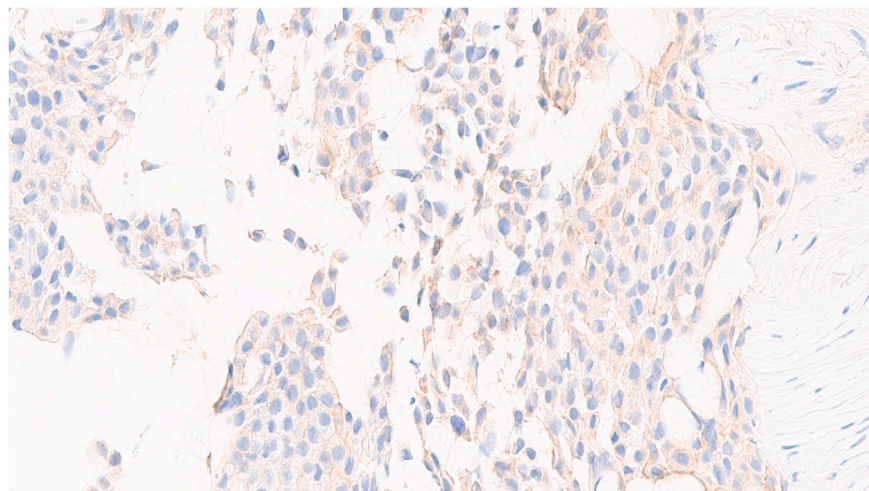
Tumour heterogeneity has been highlighted as an important issue causing difficulties in HER2 IHC and FISH interpretation and leading to uncertainty on treatment decisions. Three patterns of heterogeneity have been described including clustered (two distinct patterns of positive and negative expression), mosaic, and scattered heterogeneity [12]. One of the tumours with discordant scoring among the non-majority diagnoses showed a clustered pattern of expression (small 3+ focus within a majority negative score 0 tumour), Figs. 2E and F. It is of note that further sampling from a metastatic axillary lymph node showed only a HER2 0 profile. Following multidisciplinary discussion, the tumour was regarded and managed as HER2 negative (TNBC) in view of the very focal expression and the patient received neoadjuvant

Table 6

Challenging cases consensus first and second round scoring.

(1*) Lobular carcinoma; (5*+) Mucinous carcinoma; rest of the cases were of the NST histological type. On the left side of the table; grey-shaded cells represent the number of raters agreeing on scoring (8/16 or lower; 50% denote poor level of agreement). On the right side of the table; pink-shaded cells display number of raters agreeing on a particular score, after re-scoring (8/16 or lower; 50%).

ID	First round				Consensus Score	Level of Agreement	Second round				Consensus Score	Level of Agreement
	Number of pathologists						Number of pathologists					
	0	1+	2+	3+			0	1+	2+	3+		
1*	0	8	8	0	1+/2+	50%; Poor	0	5	11	0	2+	%69; Low
2	1	8	7	0	1+	50%; Poor	1	4	11	0	2+	%69; Low
3	6	8	2	0	1+	50%; Poor	1	12	3	0	1+	%75; High
4	6	8	2	0	1+	50%; Poor	3	10	3	0	1+	%63; Low
5**	8	7	1	0	0	50%; Poor	5	8	3	0	1+	50%; Poor



D. Challenging case 5 showing faint membranous staining and background staining making it difficult to assess. Consensus was 0 in round 1 and 1+ in round 2.

Fig. 2b. D. Challenging case 5 showing faint membranous staining and background staining making it difficult to assess. Consensus was 0 in round 1 and 1+ in round 2.

Table 7

Clinicopathological features of cases with minority scores of adverse clinical significance

POS = positive, NEG = negative

*Tumour showing clustered heterogeneity (3+ and 0 areas) Cells shaded in grey highlights the number of incorrect minority scorings resulting in clinically significant change of HER2 status from positive to negative and vice versa.

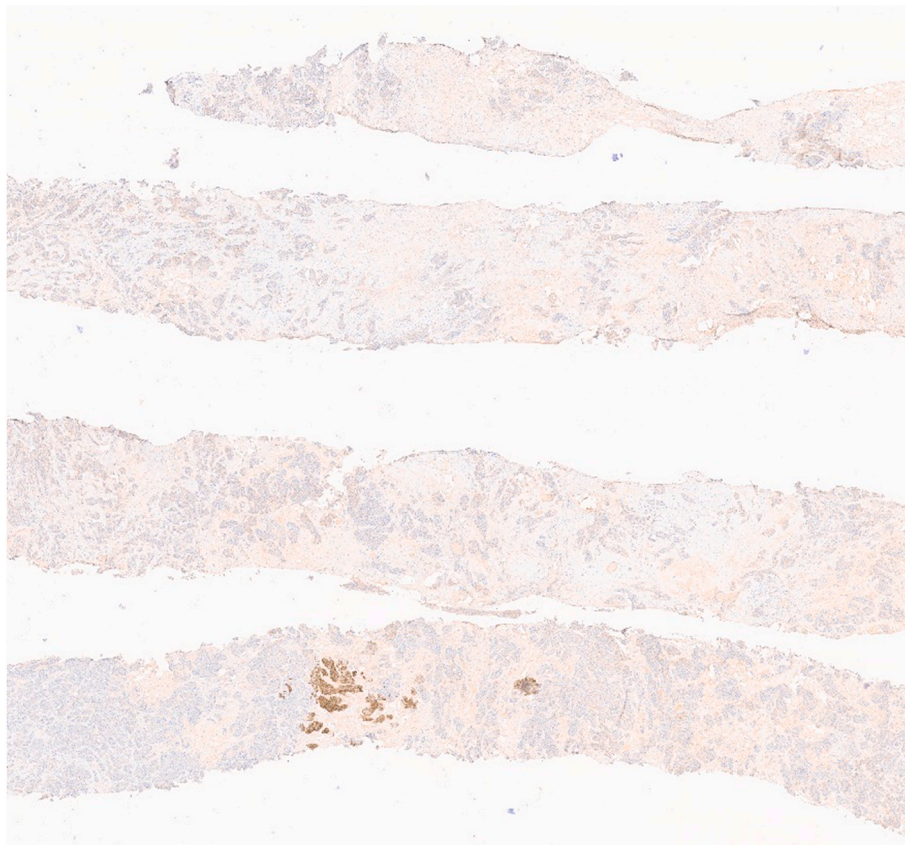
ID	HER2 Status	DISTRIBUTION of SCORES				CONSENSUS SCORE	INCORRECT SCORING	GRADE
		0	+1	+2	+3			
1	POS	0	1	15	0	2+ (+FISH)	POS > NEG	3
2	POS	0	1	14	1	2+ (+FISH)	POS > NEG	3
3	POS	0	1	0	15	3+	POS > NEG	2
4	NEG	0	0	13	3	2+ (-FISH)	NEG > POS	2
5	NEG	0	0	15	1	2+ (-FISH)	NEG > POS	1
6	NEG	0	0	14	2	2+ (-FISH)	NEG > POS	2
7	NEG	0	6	9	1	2+ (-FISH)	NEG > POS	2
8	NEG*	1	3	10	2	2+ (-FISH)	NEG > POS	3

Table 7: Clinicopathological features of cases with minority scores of adverse clinical significance

POS= positive, NEG = negative

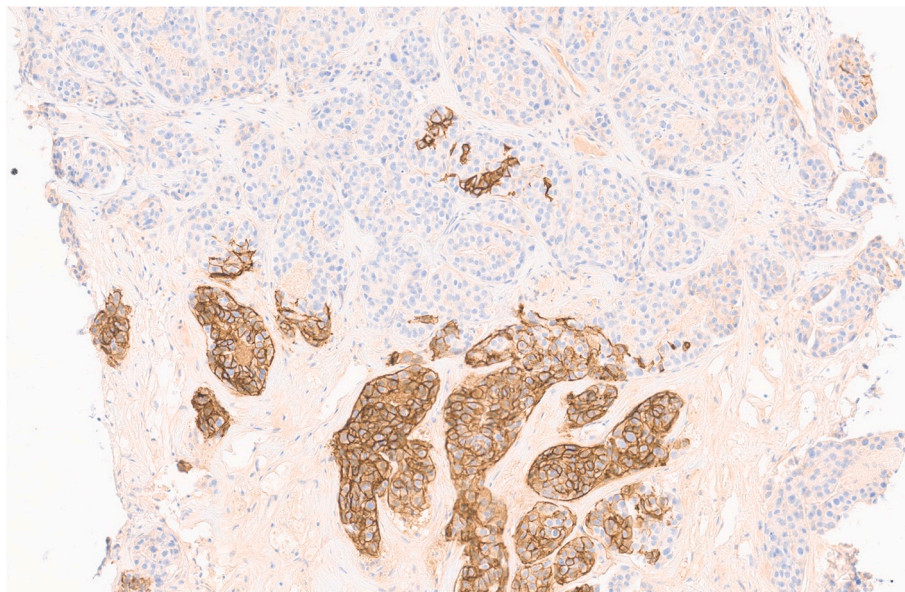
*Tumour showing clustered heterogeneity (3+ and 0 areas)

Cells shaded in grey highlights the number of incorrect minority scorings resulting in clinically significant change of HER2 status from positive to negative and vice versa.



E. Low power overview of breast carcinoma cores showing clustered heterogeneous expression of HER2. Consensus score was 2+.

Fig. 2c. E. Low power overview of breast carcinoma cores showing clustered heterogeneous expression of HER2. Consensus score was 2+.



F. high power view of the same case showing a clone with strong complete membranous staining.

Fig. 2d. F. high power view of the same case showing a clone with strong complete membranous staining.

chemotherapy (NACT). Testing of the residual carcinoma post NACT revealed TNBC with a HER2 IHC score of 0. HER2 heterogeneity is known to be more prevalent in the borderline and negative groups and

hence will be represented at a higher frequency in the HER2-low tumours. This adds to the complexity of HER2-low scoring and may influence the correct categorisation of such tumours.

Table 8

Fleiss multi-rater kappa of agreement, among individual and combined categories.

OA: overall agreement.

Scoring Categories	
Individual	Combined
OA (0.433)	OA (0.548)
0 (0.437)	
1+ (0.292)	1+/2+ (0.509)
2+ (0.431)	
3+ (0.803)	

The type of specimen, i.e. biopsy versus excision, has recently been shown to be relevant in the assessment of HER2 low cancer. In a Chinese series of 5610 paired specimens, a discordance rate of 17.63% was found between core and excision HER2 low assessment. This was largely presented by cases diagnosed as HER2-low on surgical excision following a core biopsy assessment as negative, score 0 (n = 530, 9.4%). The reverse also occurred with 387 cases (6.9%) switching from a HER2 0 to HER2-low status [13]. In routine practice, HER2 is tested primarily on diagnostic core biopsies and hence the assessment of breast cores in the present study.

The HER2-low group represents a considerable proportion (40–50%) of all breast cancers and includes both hormone receptor positive and negative carcinomas with more representation of the former [14–16]. A real world UK and Ireland web based biomarker audit showed that HER2 1+ tumours accounted for 32.2% of all breast cancers, and the HER2 equivocal (2+) category for 12.9%, with 72.1% of the latter FISH non-amplified. This corresponds to a rate of 41.5% HER2-low breast

cancers [17]. In a study of 281 consecutive breast cancers, the HER2-low group represented 31% of tumours and was more common in the ER positive group compared with ER negative tumours (33.6% vs 15%, $p = 0.017$). Those tumours were predominantly of ‘ductal’ type [18]. Following on from the National Comprehensive Cancer Network (NCCN) guidelines version 3.2023 [19], other national and international guidelines are likely to include the HER2-low category in management pathways and therefore standardised diagnostic criteria should be available for pathologists to ensure consistency of diagnosis.

Concordance studies of HER2-low scoring among pathologists were recently identified as an unmet research need by Baez-Navarro and colleagues [20]. Awareness of the recent advances in therapeutic options and the requirement to correctly identify the HER2-low group of breast tumours is particularly relevant as the boundaries between 0 and 1+ scores, as highlighted in the current study, are less clear. In the current state of knowledge, it is not known whether pathologists will also be required to identify tumours with HER2 ultralow expression. Initial data from the DAISY trial (Study of DS-8201a, an Antibody Drug Conjugate for Advanced Breast Cancer Patients, With Biomarkers Analysis (DAISY) (ClinicalTrials.gov Identifier: NCT04132960), suggest that patients with tumours that display very low levels of HER2 expression (currently regarded as score 0) may still benefit from T-DXd therapy. If this is proven, reversion to the ASCO/CAP 2007 criteria of IHC score 0 confined to tumours with complete absence of HER2 expression (rather than the current 10% cut off) would be appropriate [20]. We have shown that combining scoring categories resulted in a higher concordance among pathologists and will meet the clinical need. This will undoubtedly be guided by the clinical requirements for determining patient eligibility for treatment with anti-HER2 ADCs and awaits evidence from current ongoing trials.

Table 9

Cohen’s weighted kappa for pairwise agreement between each rater’s score and the consensus score, in all HER2 IHC categories and in the HER2-low group.

The bottom left triangular half below the equatorial blank cells represents values of the agreement among the whole cohort of cases (n = 50), while the opposite top right triangle represents the HER2-low cases (n = 36). (C-A): Consensus score of the whole cohort; (C-L): Consensus score of the HER2-low cohort. Cell shading colours reflect kappa levels of agreement as follows; yellow for almost perfect agreement (0.81–1), green for substantial agreement (0.61–0.8), orange for moderate agreement (0.41–0.6), blue for fair agreement (0.21–0.4), and red for poor agreement (0–0.2).

	C-A	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16
C-L		0.484	0.533	0.265	0.655	0.503	0.390	0.493	0.171	0.422	0.415	0.690	0.362	0.409	0.716	0.167	0.635
P1	0.71		0.267	0.144	0.345	0.271	0.327	0.299	0.434	0.548	0.248	0.541	0.568	0.363	0.501	0.100	0.532
P2	0.742	0.555		0.484	0.459	0.494	0.239	0.528	0.146	0.285	0.615	0.459	0.247	0.274	0.562	0.328	0.375
P3	0.524	0.39	0.702		0.229	0.239	0.085	0.284	0.084	0.113	0.308	0.143	0.100	0.100	0.278	0.400	0.189
P4	0.757	0.567	0.654	0.511		0.344	0.398	0.418	0.091	0.339	0.379	0.497	0.214	0.329	0.618	0.150	0.580
P5	0.749	0.564	0.77	0.599	0.577		0.178	0.272	0.163	0.220	0.552	0.438	0.193	0.206	0.424	0.192	0.363
P6	0.68	0.604	0.586	0.432	0.724	0.52		0.272	0.314	0.483	0.156	0.517	0.394	0.348	0.415	0.056	0.383
P7	0.748	0.596	0.776	0.579	0.662	0.648	0.637		0.152	0.308	0.415	0.416	0.261	0.310	0.412	0.077	0.422
P8	0.5	0.646	0.463	0.325	0.384	0.479	0.562	0.464		0.484	0.095	0.211	0.504	0.198	0.249	0.058	0.214
P9	0.709	0.758	0.588	0.382	0.592	0.563	0.716	0.63	0.689		0.199	0.495	0.557	0.330	0.354	0.077	0.564
P10	0.557	0.427	0.652	0.562	0.588	0.61	0.445	0.579	0.284	0.41		0.384	0.177	0.255	0.457	0.250	0.390
P11	0.854	0.757	0.719	0.463	0.688	0.727	0.745	0.725	0.54	0.761	0.54		0.488	0.441	0.618	0.077	0.580
P12	0.651	0.763	0.537	0.35	0.503	0.513	0.653	0.576	0.693	0.772	0.38	0.735		0.270	0.376	0.069	0.340
P13	0.58	0.515	0.485	0.339	0.563	0.469	0.599	0.491	0.421	0.529	0.395	0.596	0.477		0.460	0.063	0.424
P14	0.824	0.694	0.768	0.531	0.737	0.73	0.653	0.684	0.55	0.65	0.569	0.799	0.632	0.644		0.186	0.601
P15	0.412	0.309	0.561	0.643	0.436	0.505	0.336	0.406	0.258	0.301	0.457	0.357	0.277	0.29	0.457		0.191
P16	0.815	0.72	0.682	0.476	0.735	0.651	0.7	0.73	0.511	0.765	0.544	0.792	0.628	0.551	0.759	0.412	

Table 10

Cohen’s weighted kappa for pairwise agreement between.

Cohen’s Kappa / Level of agreement	Inter-rater		Consensus-Rater	
	All cases	HER2-low	All cases	HER2-low
0.81 - 1 / Almost perfect	0	0	3/16; 19%	0
0.61- 0.8 / Substantial	49/120; 41%	4/120; 3.3%	8/16; 50%	4/16; 25%
0.41-0.6 / Moderate	55/120; 46%	36/120; 30%	5/16; 31%	7/16; 43.75%
0.21-0.4 / Fair	16/120; 13%	49/120; 40.7%	0	3/16; 18.75%
0-0.2 / Slight	0	31/120; 26%	0	2/16; 12.5%

Recently, Baez -Navarro et al. assessed the concordance of HER2 scoring among 16 pathologists in two rounds. In round one, absolute agreement on all categories was reached in a minority of cases (4.7%). This compares with 6% in the current study. Best absolute agreement (74.3%) was reached when the HER2 0 category (using the 2007 ASCO/CAP guidelines) was compared to a combined group of HER2 1+/2+ (i.e., no staining vs any staining that is not 3+) [9]. Our study confirmed the above data where clustering of score 0 against the rest of the IHC scores combined resulted in increased both absolute and high agreement to 86% (compared to 52%). Combining scores 1+ and 2+ also resulted in improved concordance with higher percentages of absolute and high agreement to 84%. While the use of the term HER2-low by pathologists may not be currently required as per the recent ESMO consensus [21], it is essential that pathologists distinguish between HER2 scores 0 and 1+ in their report. An unqualified HER2 negative statement, that was previously acceptable for primarily identifying the HER2 positive cancers, is no longer justified. Data from the current study also supports that the use of the 2007 guidelines for the definition of HER2 0 category as complete absence of membranous staining would improve the consistency of pathologists scoring. This group lacking any evidence of HER2 expression has been recognized and referred to as “HER2-null”. The decision as to whether this latter term should be used in conjunction with the HER2-ultralow category awaits the results of the DB-06 trial [21].

The current study cohort of breast tumours was stained using the 4B5 Ventana IHC assay. Differences among HER2 staining platforms exist and the proportion of tumours regarded as HER2-low may vary according to the assay used. A recent comparison of 119 breast cancers stained by the Dako polyclonal HercepTest and the Ventana monoclonal 4B5 assay revealed that the former classified more cases as HER2 2+ [22], including tumours with HER2 amplification and tumours that were non-amplified (HER2-low). However, an earlier study of 500 primary and metastatic breast cancer tested the concordance of the two tests and showed the proportion of HER2 1+/2+ scores to be 28% and 11.6% by the 4B5 and HercepTest respectively, with several tumours designated as HER2 0 by the HercepTest, classified as 1+/2+ (i.e., HER2-low) by the 4B5 assay. All tumours that scored 0 by the HercepTest were also scored 0 by the 4B5 assay [10]. In a series of 205, mostly HER2 score 2+ tumours, stained by the HercepTest at 14 Spanish centres and tested centrally by FISH, agreement was only low to moderate [23]. The number of tumours that were FISH negative (i.e., HER2-low) increased with a decrease in the number of coincident observers but the opposite was found for the FISH positive cases. Possible reasons for disagreement, proposed by the authors, included subjective assessment, pre-analytical conditions of samples, criteria used and observer experience [23].

Data on the prognostic significance of HER2-low tumours are conflicting. Xu et al. analysed 777 non-HER2 positive breast cancers and found no prognostic differences between the HER2-2 low and the HER2 negative group. The hormone receptor HER2-low group had a better prognosis [24]. PAM50 data from 3689 HER2-low and HER2 score 0 breast cancers revealed that the HER2-low tumours were predominantly of the luminal type compared with the TNBC (65.4% VS 36.6% respectively) [25].

Similarly, a large retrospective study of breast tumours from 5235 patients, that included 2917 HER2-low tumours, did not identify any specific histopathological features associated with this HER2 profile¹⁷. The expression of hormone receptors was significantly higher in HER2-low compared with the HER2 0 tumours ($p < 0.001$) [26] but there was no significant difference in disease-free or overall survival between hormone receptor positive and hormone receptor negative HER2 low tumours. These findings imply that HER2-low breast cancer is not a distinct biological entity. Others, however, have proposed that, particularly in the neoadjuvant chemotherapy setting, that HER2-low tumours can be considered a “new subgroup of breast cancer by standardised IHC, distinct from HER2-zero tumours” highlighting a lower

pathological response rate in hormone receptor positive lesions compared to HER2 negative lesions and differences in prognosis [27]. Clearly this potential group of invasive breast cancers requires further research. A recent study of 351 patients including 189 with HER2 low breast cancers, reported that the latter had better survival compared with the HER2 0 group [28].

Strengths of the current study include that participants, Professional Clinical Advisors for the NHS Breast Screening Programme, are experts in breast pathology morphological diagnosis and HER2 immunohistochemistry assessment. Their roles include supporting delivery of high quality of breast cancer service and providing guidance to pathologists in various regions of the UK and Ireland. These data will inform education and training of breast and molecular pathologists in both countries. Another strength is the robust IHC staining and FISH protocols performed at a large UK molecular pathology service which is also a regional referral centre for HER2 immunohistochemistry and FISH testing. In addition, all tumours were stained using the Ventana PATHWAY anti-HER2/neu (4B5) Rabbit Monoclonal Primary Antibody test. It is the most commonly used HER2 IHC assay in the UK, the test used for the DESTINY-Breast 04 trial [4] and an approved companion diagnostic for HER2-low assessment by the U.S. Food and Drug Administration (FDA). The use of digital pathology is a strength that ensured that all scorers assessed the identical slides but can also be a limitation for pathologists who are less familiar with using digital platforms for scoring. A limitation of this study is the relatively small number of tumours (50) included. Whilst all these cases were core biopsy specimens, this is the specimen that is assessed for HER2 testing in the UK. This may not reflect the examination of HER2 status on breast cancer excisions. Of particular note, the cases in this study were enriched for HER2-low expressors and selected to predominantly assess the concordance at the lower end of the HER2 staining spectrum; in other words, it must be noted that the concordance presented here does not necessarily represent the agreement for overall HER2 testing for all cases.

This current study highlights that a proportion breast cancers can be challenging to designate as immunohistochemistry score 1+ versus 0 even for expert breast pathologists. We recommend double scoring of those cases at the border between the two categories and/or discussion with colleagues to reach a consensus. Pre-analytical factors such as appropriate fixation, careful attention to the quality and sensitivity of HER2 immunohistochemically staining and use of appropriate controls are also essential. This will be facilitated by the enrolment in relevant laboratory quality assurance programmes. Recently, the United Kingdom National External Quality Assessment Service (UK NEQAS) has launched an educational HER2-low technical and interpretive pilot scheme and presented findings of the first round. Results have shown poor concordance (<50%) in HER2-low laboratory staining and scores when compared with expert pathologists using a well validated assay [29]. Continuous education and raising awareness among pathologists will improve concordance and reporting accuracy.

Digital pathology and artificial intelligence (AI) algorithms have been suggested as promising applications for HER2 IHC and FISH assessment [30]. AI algorithms for HER2 scoring are promising but can only be implemented if the gold standard pathologist reporting can be standardised for an AI algorithm calibration. Furthermore, other techniques such as quantitative immunofluorescence [31] and *in vitro* diagnostic RNA based tests for the assessment of HER2 status, particularly at the low expression level, have been developed [32,33]. These may provide more accurate and reproducible categorisation of the HER2-low tumours compared with IHC or used as confirmatory tests in challenging cases. Further studies comparing various methods of HER2 assessment against tumour response are required to inform decision-making and therapeutic strategies.

Declaration of competing interest

EP has received honoraria from Roche, Novartis and AstraZeneca for lectures and participating in advisory committees. SP has received honoraria as speaker and advisory board member for Exact Sciences. AMS has received honoraria as speaker and advisory board member for Exact Sciences, Veracyte, Roche, Hologic, Diaceutics and AstraZeneca.

Acknowledgements

AMS was funded by Birmingham CR-UK Centre grant no. C17422/A25154.

References

- [1] Wu Y, Zhong R, Ma F. HER2-low breast cancer: novel detections and treatment advances. *Crit Rev Oncol Hematol* 2023;181:103883.
- [2] Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol* : official journal of the American Society of Clinical Oncology 2013;31:3997–4013.
- [3] Wolff AC, Hammond MEH, Allison KH, Harvey BE, Mangu PB, Bartlett JMS, et al. Human epidermal growth factor receptor 2 testing in breast cancer: American society of clinical oncology/college of American pathologists clinical practice guideline focused update. *Arch Pathol Lab Med* 2018;142:1364–82.
- [4] Modi S, Jacot W, Yamashita T, Sohn J, Vidal M, Tokunaga E, et al. Trastuzumab deruxtecan in previously treated HER2-low advanced breast cancer. *N Engl J Med* 2022;387:9–20.
- [5] European Medicines Agency. Enhertu (trastuzumab deruxtecan) an overview of Enhertu and why it is authorised in the EU. 2023.
- [6] NifHaC Excellence. Trastuzumab deruxtecan for treating HER2-positive unresectable or metastatic breast cancer after 1 or more anti-HER2 treatments, TA862. Technology appraisal guidance; 2023.
- [7] Rakha EA, Tan PH, Quinn C, Provenzano E, Shaaban AM, Deb R, et al. UK recommendations for HER2 assessment in breast cancer: an update. *J Clin Pathol* 2022;76(4):217–27.
- [8] Wolff AC, Somerfield MR, Dowsett M, Hammond MEH, Hayes DF, McShane LM, et al. Human epidermal growth factor receptor 2 testing in breast cancer: ASCO-Collge of American pathologists guideline update. *J Clin Oncol* : official journal of the American Society of Clinical Oncology 2023;JCO2202864.
- [9] Baez-Navarro X, van Bockstal MR, Nawawi D, Broeckx G, Colpaert C, Doebar SC, et al. Interobserver variation in the assessment of immunohistochemistry expression levels in HER2-negative breast cancer: can we improve the identification of low levels of HER2 expression by adjusting the criteria? An international interobserver study. *Mod Pathol* : an official journal of the United States and Canadian Academy of Pathology, Inc. 2023;36:100009.
- [10] Scott M, Vandenberghe ME, Scorer P, Boothman A-M, Barker C. Prevalence of HER2 low in breast cancer subtypes using the VENTANA anti-HER2/neu (4B5) assay. *J Clin Oncol* 2021;39.
- [11] UK National External Quality Assurance (UK NEQAS). UK national external quality assurance (UK NEQAS) ICC and ISH assessment. 2018.
- [12] Marchio C, Annaratone L, Marques A, Casorzo L, Berrino E, Sapino A. Evolving concepts in HER2 evaluation in breast cancer: heterogeneity, HER2-low carcinomas and beyond. *Semin Cancer Biol* 2021;72:123–35.
- [13] Lu Y, Zhu S, Tong Y, Fei X, Jiang W, Shen K, et al. HER2-Low status is not accurate in breast cancer core needle biopsy samples: an analysis of 5610 consecutive patients. *Cancers* 2022;14.
- [14] Giuliani S, Ciniselli CM, Leonardi E, Polla E, Decarli N, Luchini C, et al. In a cohort of breast cancer screened patients the proportion of HER2 positive cases is lower than that earlier reported and pathological characteristics differ between HER2 3+ and HER2 2+/Her2 amplified cases. *Virchows Arch* : an international journal of pathology 2016;469:45–50.
- [15] Miglietta F, Griguolo G, Bottosso M, Giarratano T, Lo Mele M, Fassan M, et al. Evolution of HER2-low expression from primary to recurrent breast cancer. *NPJ breast cancer* 2021;7:137.
- [16] Holthuis EI, Vondeling GT, Kuiper JG, Dezentje V, Rosenlund M, Overbeek JA, et al. Real-world data of HER2-low metastatic breast cancer: a population based cohort study. *Breast* 2022;66:278–84.
- [17] Dodson A, Parry S, Ibrahim M, Bartlett JM, Pinder S, Dowsett M, et al. Breast cancer biomarkers in clinical testing: analysis of a UK national external quality assessment scheme for immunocytochemistry and in situ hybridisation database containing results from 199 300 patients. *The journal of pathology Clinical research* 2018;4:262–73.
- [18] Zhang H, Katerji H, Turner BM, Audeh W, Hicks DG. HER2-low breast cancers: incidence, HER2 staining patterns, clinicopathologic features, MammaPrint and Blueprint genomic profiles. *Mod Pathol* : an official journal of the United States and Canadian Academy of Pathology, Inc. 2022;35:1075–82.
- [19] NCCN (. NCCN. NCCN guidelines version 3.2023 invasive breast cancer. 2023.
- [20] Baez-Navarro X, Salgado R, Denkert C, Lennerz JK, Penault-Llorca F, Viale G, et al. Selecting patients with HER2-low breast cancer: getting out of the tangle. *Eur J Cancer* 2022;175:187–92.
- [21] Tarantino P, Viale G, Press MF, Hu X, Llorca FP, Bardia A, et al. ESMO Expert Consensus Statements (ECS) on the definition, diagnosis, and management of HER2-low breast cancer. *Ann Oncol* : Offic J Eur Soc Med Oncol/ ESMO 2023. <https://doi.org/10.1016/j.annonc.2023.05.008>.
- [22] Ruschoff J, Friedrich M, Nagelmeier I, Kirchner M, Andresen LM, Salomon K, et al. Comparison of HercepTest mAb pharmDx (Dako Omnis, GE001) with Ventana PATHWAY anti-HER-2/neu (4B5) in breast cancer: correlation with HER2 amplification and HER2 low status. *Virchows Arch* : an international journal of pathology 2022;481:685–94.
- [23] Castera Redal C, Bernet Vegue L. "HER2 immunohistochemistry inter-observer reproducibility in 205 cases of invasive breast carcinoma additionally tested by ISH" Answer to the statistical issue to avoid misinterpretation. *Ann Diagn Pathol* 2020;48:151566.
- [24] Xu H, Han Y, Wu Y, Wang Y, Li Q, Zhang P, et al. Clinicopathological characteristics and prognosis of HER2-low early-stage breast cancer: a single-institution experience. *Front Oncol* 2022;12:906011.
- [25] Schettini F, Chic N, Braso-Maristany F, Pare L, Pascual T, Conte B, et al. Clinical, pathological, and PAM50 gene expression features of HER2-low breast cancer. *NPJ breast cancer* 2021;7:1.
- [26] Tarantino P, Jin Q, Tayob N, Jeselsohn RM, Schnitt SJ, Vinciguilla J, et al. Prognostic and biologic significance of ERBB2-low expression in early-stage breast cancer. *JAMA Oncol* 2022;8:1177–83.
- [27] Denkert C, Seither F, Schneeweiss A, Link T, Ju Blohmer, Just M, et al. Clinical and molecular characteristics of HER2-low-positive breast cancer: pooled analysis of individual patient data from four prospective, neoadjuvant clinical trials. *Lancet Oncol* 2021;22:1151–61.
- [28] Almstedt K, Heimes AS, Kappenberg F, Battista MJ, Lehr HA, Krajnak S, et al. Long-term prognostic significance of HER2-low and HER2-zero in node-negative breast cancer. *Eur J Cancer* 2022;173:10–9.
- [29] Parry S, Shaaban A, Jasani B, Dodson A. 54P Proficiency testing for HER2-low expression in breast cancer: results of the UK NEQAS ICC & ISH programme examining technical and interpretive accuracy. *ESMO Open* 2023;8.
- [30] Wu S, Yue M, Zhang J, Li X, Li Z, Zhang H, et al. The role of artificial intelligence in accurate interpretation of HER2 immunohistochemical scores 0 and 1+ in breast cancer. *Mod Pathol* : an official journal of the United States and Canadian Academy of Pathology, Inc. 2023;36:100054.
- [31] Moutafi M, Robbins CJ, Yaghoobi V, Fernandez AI, Martinez-Morilla S, Xirou V, et al. Quantitative measurement of HER2 expression to subclassify ERBB2 unamplified breast cancer. *Lab Invest* 2022;102:1101–8.
- [32] Varga Z, Lebeau A, Bu H, Hartmann A, Penault-Llorca F, Guerini-Rocco E, et al. An international reproducibility study validating quantitative determination of ERBB2, ESR1, PGR, and MKI67 mRNA in breast cancer using MammaTyper(R). *Breast cancer research. BCR* 2017;19:55.
- [33] Atallah NM, Toss MS, Green AR, Mongan NP, Ball G, Rakha EA. Refining the definition of HER2-low class in invasive breast cancer. *Histopathology* 2022;81:770–85.