Poster 1455

Comparison of digital and computational algorithms for quantifying human epidermal growth factor receptor 2 (HER2) protein expression in metastatic breast cancer (mBC) from clinical samples

Savitri Krishnamurthy, MD[1]; Dhanrajan Tiruchinapalli, PhD, MBA[2]; Clara Lam, PhD, MPH[2]; Simon M. Collin, PhD[3]; Rosemary Taylor, MSc[4]; Linlin Luo, MSc[2]; Anupriya Dutta, PhD[2]; Ehab El-Gabry, MD[2,*]; Michele Sue-Ann Woo, PhD[5]; Grace Kwon, PharmD, MBA[5]; Robert Egger, PhD[6]; Jennifer Hipp, MD, PhD[6]; Lauren Brunner, PhD[6]; Jeppe Thagaard, PhD[7]; Thomas W. Ramsing, MSc[7]; Henrik Hoeg, MSc[7]; Wonkyung Jung, MD[8]; Heon Song, MSc[8]; Chang Ho Ahn, MD, PhD[8]; Vladimir Kravtsov, MSc[8]; Patrick Frey, PhD[9]; Ralf Banisch, PhD[9]; Stella Redpath, PhD[9]

[1]Department of Pathology, Division of Pathology and Laboratory Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; [2]Medical Affairs, AstraZeneca Pharmaceuticals LP, Gaithersburg, MD, USA; [3]Global Medical Affairs, AstraZeneca, Cambridge, UK; [4]Oncology Biometrics, AstraZeneca, Macclesfield, UK; [5]US Medical Affairs, Daiichi Sankyo, Inc., Basking Ridge, NJ, USA; [6]PathAI, Inc., Boston, MA, USA; [7]Computational Pathology, VisioPharm, Hersholm, Denmark; [8]Lunit, Seoul, Republic of Korea; [9]MindPeak GmbH, Hamburg, Germany
*At the time of the study.

Objectives

- To evaluate the concordance of 4 standalone computational pathology-assisted (CPa) tools with pathologists' manual consensus scoring of digitized mBC whole-slide images (WSIs) for quantification of the lower end of the HER2 immunohistochemistry (IHC) spectrum (IHC 0 absent membrane staining, IHC 0 with membrane staining [HER2-ultralow], and IHC 1+ [HER2-low][1-3])
- To evaluate whether these CPa tools can maintain acceptable agreement across HER2 IHC categories while reducing review time

Conclusions

- Pathologist manual consensus showed only low–moderate agreement for HER2 IHC 0 to 1+ cases, underscoring the inherent variability in scoring of HER2-low and HER2-ultralow specimens
- Standalone CPa algorithms demonstrated fair–moderate concordance with manual consensus across the HER2 IHC spectrum while reducing median review time by ~65%, suggesting potential efficiency gains if they were to be incorporated into digital pathology workflows
- Despite the remaining challenges (including variability in sample quality, dataset generalizability, and interpretability), standalone CPa tools aligned relatively well with pathologist scoring in HER2-low/-ultralow cases and may assist pathologists and support more consistent, efficient evaluation when used within ''human-in-the-loop'' workflows
- These findings support the potential integration of CPa tools as assistive technologies in HER2 IHC workflows, with further validation needed using clinically representative mBC datasets to help address some of the remaining challenges

Plain language summary

Why did we perform this research?

Some breast cancers have very low levels of a protein called HER2.[1] New treatments may help patients whose tumors have low or very low HER2 levels, but it can be difficult for doctors to identify them consistently using standard microscope review alone.[2-7] Computer-based tools that analyze digital images of tumor samples may help support pathologists with this process, but their performance needs to be carefully evaluated.[7]

How did we perform this research?

This study investigated computer images of tumor samples from patients with metastatic breast cancer that were previously scored by pathologist as having no HER2 expression or low levels of HER2 expression. Three pathologists reviewed each image to determine the HER2 score, and their consensus score was used as the reference standard. These manually rescored digitized images were also analyzed using 4 computer-based tools, and we compared the results and the time needed for review with computer-based tools.

What were the findings of this research?

The standalone performance of the 4 computer-based tools showed different levels of agreement with the pathologists' assessments when classifying HER2 levels, with 3 out of 4 tools matching the pathologists' results quite closely (around 70% overall agreement). In addition, computer-based analysis generally required less review time than manual assessment by pathologists.

What are the implications of this research?

Computer-based tools may help reduce the time required to review and report HER2 test results, but agreement with pathologists' assessments remains variable (fair–moderate). These digital tools are still being developed and are likely to improve over time. Further research is needed to evaluate their performance when used by pathologists and implemented in clinical practice.

References

1. Wolff AC, et al. *Arch Pathol Lab Med.* 2018;142(11):1364-1382.
2. Schettini F, et al. *NPJ Breast Cancer.* 2021;7(1):1.
3. Bardia A, et al. *N Engl J Med.* 2024;391(22):2110-2122.
4. Venetis K, et al. *Front Mol Biosci.* 2022;9:834651.
5. Viale G, et al. *ESMO Open.* 2023;8(4):101615.
6. Modi S, et al. *N Engl J Med.* 2022;387(1):9-20.
7. Krishnamurthy S, et al. Poster presented at the 2024 San Antonio Breast Cancer Symposium (SABCS); December 10-13, 2024; San Antonio, TX, USA.

This study was sponsored and led by AstraZeneca. In March 2019, AstraZeneca entered into a global development and commercialization collaboration agreement with Daiichi Sankyo for trastuzumab deruxtecan (T-DXd; DS-8201). Poster presented at the 2026 American Association for Cancer Research (AACR) Annual Meeting; April 17-22, 2026; San Diego, CA, USA, by Dr. Savitri Krishnamurthy.
Corresponding author email address: skrishna@mdanderson.org

Introduction

- HER2 expression is a key biomarker in breast cancer, and HER2 status is routinely determined by IHC and/or in situ hybridization (ISH). According to the classical binary classification defined by the ASCO-CAP 2018 guidelines, tumors are categorized as HER2-positive (IHC 3+ or IHC 2+/ISH-positive) or HER2-negative (IHC 2+/ISH-negative, IHC 1+, or IHC 0)[2,3]
- Accurate and timely identification of HER2 IHC status by pathologists is key for treatment selection in breast cancer. This is particularly relevant in the metastatic setting where patients with HER2-expressing tumors may be eligible for therapies such as T-DXd (a HER2-directed antibody-drug conjugate)[4]
- Clinical benefit of T-DXd has expanded the clinical relevance of low levels of HER2 expression, with approvals based on the DESTINY-Breast04 and DESTINY-Breast06 clinical trials for patients with HER2-low (IHC 1+ or IHC 2+/ISH-negative) and HER2-ultralow (IHC 0 with faint or incomplete membrane staining in ≤10% of tumor cells) mBC[5-10]
- Analyses of mBC samples originally scored as HER2 IHC 0 and IHC 1+ (HER2-negative) have shown that a substantial proportion of tumors, particularly those initially scored as IHC 0, may be rescored as IHC 0 with membrane staining or HER2 IHC 1+, and that inter-and intra-observer concordance among pathologists is variable, notably at the lower levels of the HER2 IHC expression spectrum[6,8,10,11]
- As the clinical relevance of HER2-low and HER2-ultralow has increased, accurate identification of very low HER2 expression has become increasingly important. Supporting pathologists' confidence in interpreting these cases highlights the value for standardized protocols, additional training, and digital tools to aid HER2 IHC scoring decisions[11]
- Advances in the analysis of WSIs have transformed breast pathology by enabling workflows that support computational algorithms; CPa tools can help enhance diagnostic precision, reproducibility, and clinical decision-making[12]
- This study builds on previous real-world findings[11] of manual scoring by 3 pathologists and compares 4 standalone CPa tools versus pathologist-derived consensus scoring

Methods

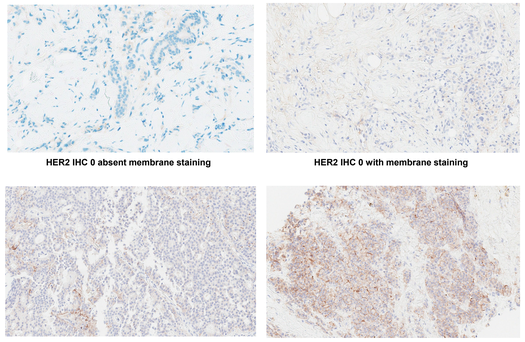
- This retrospective real-world evidence study evaluated WSIs from mBC biopsy samples collected between 2020 and 2023 and stained using the PATHWAY HER2 (4B5) assay (Roche Diagnostic Solutions, Tucson, AZ, USA)
- A total of 384 WSIs (scanned at 20× magnification) originally scored as HER2 IHC 0 (n = 246) or IHC 1+ (n = 138) were included in the analysis
- WSIs were rescored into the following HER2 IHC categories: IHC 0 absent membrane staining, IHC 0 with membrane staining in ≤10% of tumor cells, IHC 1+, IHC 2+, and IHC 3+ according to the CAP 2025 Biomarker Reporting Template[1,3]
- The WSIs were independently rescored manually by 3 board-certified pathologists and independently by 4 CPa tools under development, with vendor identities being blinded with code names (KL84Q, RV73X, ZX19P, and MQ52G)
- Each pathologist performed 2 blinded rescoring assessments per WSI using 2023 ASCO/CAP guidelines, with randomized slide order and a washout period of ≥2 weeks between readings
- Manual consensus was defined as agreement by at least 2 of 3 pathologists
- Concordance between digital pathology-assisted outputs and manual consensus scores across HER2 IHC categories was assessed using overall percentage agreement (OPA) and Cohen's κ statistics
- Pathologist and CPa review times were recorded to evaluate potential differences in efficiency between manual and CPa-assisted workflows

Results

Manual Consensus HER2 Reclassification

- Of 384 WSIs assessed, 375 had aligned HER2 IHC scores following pathologist review; 9 WSIs showed discordant scores (all 3 pathologists assigned different scores) and were excluded from consensus analyses
  - Agreement between 2 out of 3 pathologists occurred in 154 cases (41.1%), while agreement between all 3 pathologists occurred in 221 cases (58.9%)
- Representative HER2 IHC–stained WSIs illustrating reclassifications across HER2 IHC categories are shown in **Figure 1**

Figure 1. Representative HER2 IHC WSI by Consensus Classification (20x Power of Magnification)



HER2 IHC 0 absent membrane staining | HER2 IHC 0 with membrane staining | HER2 IHC 1+ | HER2 IHC 2+

HER2 IHC Reclassification by CPa Tools

- 81 of 375 WSIs (21.6%) were manually rescored as HER2 IHC 0 absent membrane staining, 85 (22.7%) as HER2 IHC 0 with membrane staining, 203 (54.1%) as HER2 IHC 1+, and 6 (1.6%) as HER2 IHC 2+ (**Table 1**)
- HER2 IHC reclassification generated by each of the 4 CPa tools showed variability when compared with manual consensus scoring across HER2 expression categories (**Table 1**)

Table 1. HER2 IHC Rescores by Manual Consensus and CPa Tools

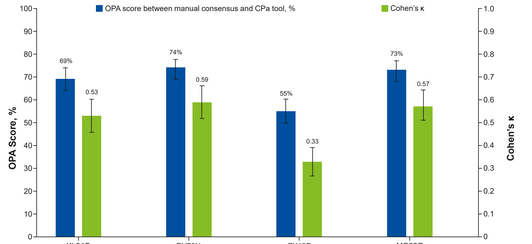
HER2 IHC Rescore, n (%)	Manual Consensus	KL84Q	RV73X	ZX19P	MQ52G
	n = 375	n = 375	n = 375	n = 375	n = 375
HER2 IHC 0 absent membrane staining	81 (21.6)	82 (21.9)	72 (19.2)	27 (7.2)	77 (20.5)
HER2 IHC 0 with membrane staining	85 (22.7)	137 (36.5)	100 (26.7)	201 (53.6)	126 (33.6)
HER2 IHC 1+	203 (54.1)	145 (38.7)	176 (46.9)	130 (34.7)	168 (44.8)
HER2 IHC 2+	6 (1.6)	10 (2.7)	5 (1.3)	6 (1.6)	4 (1.1)
HER2 IHC 3+		1 (0.3)		3 (0.8)	
Missing[a]			22 (5.9)	8 (2.1)	
Coverage %	-	100	94.1	97.9	100

OPA and Cohen's κ calculations were based on the biopsy sample only by each CPa algorithm reported.
[a]A few of the CPa tools have in-built QC algorithms that prevented processing of certain images.

Concordance Between CPa Tools and Manual Consensus

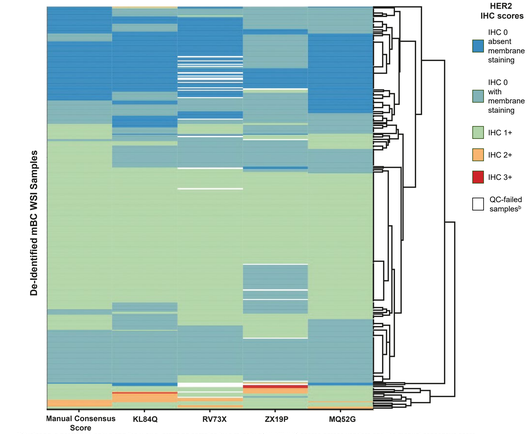
- Concordance with manual consensus varied across CPa tools (**Figure 2**):
  - OPA ranged from 55% to 74% across CPa tools
  - Cohen's κ ranged from 0.33 to 0.59 across CPa tools

Figure 2. OPA and Cohen's κ Between Manual Consensus and CPa Tools



- Concordance visualization using heatmaps further demonstrated differences in classification alignment between CPa tools and manual consensus, including variability in assignment of cases across the HER2 IHC 0 absent membrane staining, IHC 0 with membrane staining, and HER2 IHC 1+ categories (**Figure 3**)

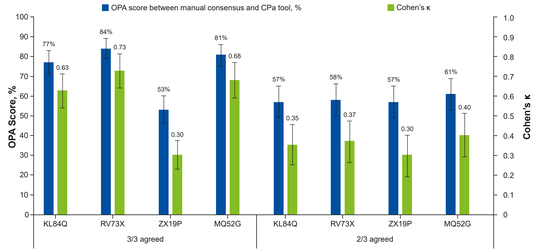
Figure 3. Heatmap of Concordance Between Manual Consensus and CPa Tools[a]



[a]Clustering is performed only along the y-axis, grouping samples by similarity of HER2 IHC scores across vendors and the manual consensus score. X-axis clustering was disabled; therefore, relationships between vendors cannot be inferred.
[b]Samples failed QC due to pre-analytical issues. OPA and Cohen's κ calculations were based on the biopsy sample only by each CPa algorithm reported.

- Agreement between standalone CPa tools and manual consensus scoring ranged from 53% to 84% OPA (Cohen's κ: 0.30-0.73) in the 3/3 consensus subset and 57% to 61% OPA (Cohen's κ: 0.30-0.40) in the 2/3 consensus subset (**Figure 4**)

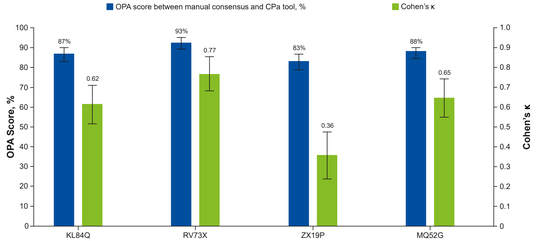
Figure 4. OPA and Cohen's κ by Consensus Tier (3/3 vs 2/3) and CPa Tool



Primary analyses use 2-of-3 consensus (N = 375). The 3-of-3 subset, if identified, is for robustness illustration only.

- In binary classification (HER2 IHC 0 absent membrane staining vs HER2 IHC 0 with membrane staining/ IHC 1+/ IHC 2+), high overall concordance with CPa tools was observed with manual consensus (83% to 93%) and moderate agreement by Cohen's κ (0.36-0.77), suggesting reliable identification of true HER2 IHC 0 absent membrane staining cases versus those with any detectable membrane staining (**Figure 5**)

Figure 5. Binary Agreement: IHC 0 Absent Membrane Staining vs IHC 0 With Membrane Staining, IHC 1+, IHC 2+



Comparison of Review Time for Manual and CPa Assessment

- Median review time per WSI was shorter using CPa tools compared with manual pathologist review (**Table 2**):
  - The median review time for manual consensus review was 6.7 minutes (range, 3.0-11.5 minutes)
  - Median CPa tool review times ranged from 0.7 to 2.4 minutes

Table 2. Median and Mean Review Time[a] by Manual Consensus and CPa Tools

Review Time, minutes	Manual Consensus	KL84Q	RV73X	ZX19P	MQ52G
	n = 375	n = 375	n = 375	n = 375	n = 375
Median (range)	6.7 (3.0-11.5)	2.4 (0.5-25.8)	2.1 (0.3-50.2)	Not available	0.7 (0.1-8.5)
Mean (SD)	6.6 (1.3)	3.3 (2.9)	3.2 (4.4)	2.5[b] (Not available)	1.0 (0.9)

Review times are not normalized across computer environments and include hardware specifications for each vendor.
[a]Review times are for the CPa tools to run and do not include pathologist review time.
[b]ZX19P can only provide the average review time from the total samples (n = 600; including non-biopsy samples).

Abbreviations

ASCO, American Society of Clinical Oncology; **CAP**, College of American Pathologists; **CPa**, computational pathology-assisted; **HER2**, human epidermal growth factor receptor 2; **IHC**, immunohistochemistry; **ISH**, in situ hybridization; **mBC**, metastatic breast cancer; **OPA**, overall percentage agreement; **QC**, quality control; **SD**, standard deviation; **T-DXd**, trastuzumab deruxtecan; **WSI**, whole-slide image.

Acknowledgments

Medical writing support, under the direction of the authors, was provided by Ilias Boltsis, PhD, Selene Jarrett, PhD, and Jill Shults, PhD, of ApotheCom, and was funded by AstraZeneca in accordance with Good Publication Practice (GPP) guidelines (http://www.ismpp.org/gpp-2022).

Disclosures

Dr. Savitri Krishnamurthy serves on scientific advisory boards for Daiichi Sankyo and AstraZeneca.

References

1. College of American Pathologists. Reporting Template for Reporting Results of Biomarker Testing of Specimens from Patients with Carcinoma of the Breast (Version 1.6.0.0). Accessed March 2026.
2. Wolff AC, et al. *Arch Pathol Lab Med.* 2018;142(11):1364-1382.
3. Wolff AC, et al. *J Clin Oncol.* 2023;41(22):3867-3872.
4. Tozbikian G, et al. *Histopathology.* 2024;85(3):489-502.
5. Schettini F, et al. *NPJ Breast Cancer.* 2021;7(1):1.
6. Bardia A, et al. *N Engl J Med.* 2024;391(22):2110-2122.
7. Venetis K, et al. *Front Mol Biosci.* 2022;9:834651.
8. Viale G, et al. *ESMO Open.* 2023;8(4):101615.
9. ENHERTU (fam-trastuzumab deruxtecan-nxki) for injection, for intravenous use. Prescribing information. Daiichi Sankyo Inc.; 2025.
10. Modi S, et al. *N Engl J Med.* 2022;387(1):9-20.
11. Krishnamurthy S, et al. Poster presented at the 2024 San Antonio Breast Cancer Symposium (SABCS); December 10-13, 2024; San Antonio, TX, USA.
12. Datwani S, et al. *Human Pathol.* 2025;162:105819.