Poster 177P

HER2 testing in multiple solid tumors: concordance between 3 scoring algorithms

Wentao Yang,[1] Josef Rüschoff,[2] Michelle Shiller,[3,4] Jessica Baumann,[5*] George Kumar,[6] Adam Brown,[7] Linlin Luo,[8] Helen Bridge,[9] Sunil Badve[10]

[1]Department of Pathology, Fudan University Shanghai Cancer Center, China; [2]Discovery Life Sciences Biomarker GmbH, Kassel, Germany; [3]PathGroup, Brentwood, TN, US; [4]Baylor University Medical Center, Dallas, TX, US; [5]Global Oncology Diagnostics, AstraZeneca, Gaithersburg, MD, US; [6]Medical Affairs, AstraZeneca, Seattle, WA, US; [7]Global Oncology Medical Affairs, Daiichi Sankyo, Basking Ridge, NJ, US; [8]Oncology Medical Above Brand, AstraZeneca, Gaithersburg, MD, US; [9]Oncology R&D, AstraZeneca, Macclesfield, UK; [10]Emory University School of Medicine, Atlanta, GA, US

*Current affiliation: NEOPATHOLOGY CORP., Frederick, MD, US

Objective

- Describe concordance between 3 human epidermal growth factor receptor 2 (HER2) immunohistochemistry (IHC) scoring algorithms across multiple solid tumor types

Conclusions

- In this study, the American Society Clinical Oncology (ASCO) / College of American Pathology (CAP) scoring algorithms for gastric and breast cancer were comparable in their identification of HER2 IHC 3+
  - Lower concordance was observed when identifying IHC 2+ and IHC 1+
- Concordance between the gastric and endometrial algorithms was low across all HER2 expression levels, including IHC 3+
- The low concordance between 3 independent non-academic HER2 IHC-experienced pathologists highlights a real-world issue of inter-pathologist variability and emphasizes the need for greater awareness on best scoring practices and additional education across different tumor types to ensure reliable identification of patients likely to benefit from treatment with trastuzumab deruxtecan (T-DXd)

Plain language summary

Why did we perform this research?
Human epidermal growth factor receptor 2 (known as HER2) is a protein found at higher-than-normal levels on the cell surface of various cancers.[1] The level of HER2 can be used to help identify patients who may benefit from HER2-targeted treatment.[2] Trastuzumab deruxtecan (T-DXd) is a recommended treatment in the US for adults with solid tumors that have the highest level of HER2 (also known as immunohistochemistry [IHC] 3+) and who have received previous treatment(s) and have no other treatment options.[3] There is no agreement on the best way to determine HER2 levels in tumors other than gastric and breast. This analysis compared 3 different methods for measuring HER2 levels (known as scoring algorithms) in different types of solid tumors.

How did we perform this research?
This study used images of tissue across different tumor types (biliary tract, bladder, cervical, endometrial, non-small cell lung cancer, ovarian, pancreatic, salivary gland, and other) from the DESTINY-PanTumor02 study and a commercially provided set. The tumor tissue had been stained to show HER2 protein levels. Three pathologists looked at the images and individually evaluated the HER2 levels (IHC 3+/2+/1+/0) using 1 algorithm usually used for breast cancer, 1 algorithm usually used for gastric cancer, and 1 algorithm used in the Fader et al clinical study evaluating endometrial cancer.[4] The primary outcome of the study was to compare the results from the different scoring algorithms to see if they produced similar results. An additional outcome was to compare the results between pathologists to see if they scored the images in the same way.

What were the findings of this research?
When results from the breast scoring algorithm were compared with the gastric scoring algorithm, agreement on the level of HER2 was higher for images identified as IHC 3+ and IHC 0, and lower for those identified as IHC 2+ and IHC 1+. Findings were similar for all tumor types. When results from the endometrial scoring algorithm were compared with the gastric scoring algorithm, agreement on the level of HER2 was low across all HER2 levels. When results from 1 pathologist were compared with results from another, agreement was higher for images identified as IHC 3+ and IHC 0, and lower for images identified as IHC 2+ and IHC 1+.

What are the implications of this research?
This study shows that the gastric and breast scoring algorithms are similar in their identification of the highest level of HER2 (IHC 3+). The gastric and endometrial scoring algorithms showed low agreement across all HER2 levels, including the highest level of HER2 (IHC 3+). The lack of agreement between the pathologists' scores shows there is a need for more awareness on the best processes to follow when scoring HER2 levels in different types of solid tumors to make sure patients who may benefit from treatment with T-DXd are correctly identified.

Where can I access more information?
For information about DESTINY-PanTumor02, please visit https://clinicaltrials.gov/study/NCT04482309, or see primary data published in the *Journal of Clinical Oncology* here. Please also reach out to Prof. Yang at yangwt2000@163.com.

1. Uzunparmak B, et al. *Ann Oncol*. 2023;34:1035–1046; 2. Zhang H, et al. *Histopathology*. 2024;85:3–19; 3. US Food and Drug Administration (FDA). Resources for information. Available from: https://www.fda.gov/drugs/resources-information-approved-drugs/fda-grants-accelerated-approval-fam-trastuzumab-deruxtecan-nxki-unresectable-or-metastatic-her2 (Accessed June 13, 2024); 4. Fader AN, et al. *J Clin Oncol*. 2018:36:2044–2051

Poster

Please scan this quick response (QR) code with your smartphone camera or app to obtain a copy of these materials. Alternatively, please click on the link below.

https://bit.ly/3yFO3tN

Poster presented at ESMO 2024, September 13–17, Barcelona, Spain, by Prof. Yang. Corresponding author email address: yangwt2000@163.com

Introduction

- HER2 expression is seen in a wide range of solid tumor types and is associated with a biologically aggressive phenotype and poor prognosis[1–5]
- In April 2024, based on the results from the DESTINY-PanTumor02 (DP-02), DESTINY-Lung01, and DESTINY-CRC02 studies, T-DXd was granted accelerated approval in the US for adult patients with unresectable or metastatic HER2-positive (IHC 3+) solid tumors that have progressed after prior treatment and have no alternative therapies[6–9]
- Although HER2 expression was assessed according to the gastric-specific criteria in the studies contributing to T-DXd's tumor-agnostic approval,[7–9] there is currently no consensus on the best scoring practices for non-breast and non-gastric solid tumors
  - Studies evaluating T-DXd outside of breast cancer, including those contributing to the T-DXd tumor-agnostic approval, have utilized the ASCO/CAP gastric scoring guidelines[7–11]
  - Use of an appropriate scoring algorithm is essential for accurate and reliable patient identification to support clinical decision making[12]
- This observational analysis reports concordance between 3 HER2 IHC scoring algorithms across multiple solid tumor types

Methods

- This observational study utilized images of stained tissue across solid tumor types from the DP-02 study (NCT04482309) and a commercial set
  - Images from DP-02 were blinded to all study participants and rescored in this analysis
- HER2 expression status was assessed using different antibodies depending on tumor type
  - HercepTest (DAKO Autostainer Polyclonal)-stained tissue was used to assess biliary tract, bladder, cervical, endometrial, ovarian, pancreatic, and other (including salivary gland) tumors from DP-02 and the commercial set
  - VENTANA HER2 4B5 (Roche)-stained tissue was used to assess non-small cell lung cancer (NSCLC) tumors from the commercial set
- Images were assessed in random sequence by 3 independent HER2 IHC-experienced non-academic pathologists and scored according to:*
  - The ASCO/CAP gastric HER2 scoring algorithm (used in DP-02[7] and as the reference scoring algorithm in this analysis)[11†]
  - The ASCO/CAP breast HER2 scoring algorithm[13†]
  - The endometrial clinical trial HER2 scoring algorithm (endometrial tumors only)[14,15]
  - There was a 2-week washout period between assessments; pathologists were blinded to previous scores
- Concordance was assessed according to positive percent agreement (PPA), negative percent agreement (NPA), and Cohen's kappa coefficient
- PPA is the percentage of images assigned a specific IHC status by the reference method in which the same status was assigned with the test method. NPA is the percentage of images not assigned a specific IHC status by the reference method in which the same status was not assigned with the test method
  - To evaluate concordance of the HER2 scoring algorithms, the gastric scoring algorithm was assigned as the reference method
  - To evaluate inter-pathologist concordance, 1 pathologist in each pair was arbitrarily assigned as the reference
- Cohen's kappa coefficient was used to assess categorical agreement between 2 pathologists, accounting for the expected agreement by chance
- All analyses performed were exploratory

*As per guidelines, no prespecified magnification was used to analyze images; †in situ hybridization testing is part of the ASCO/CAP guidelines and so was not carried out for equivocal samples

Primary outcome

- Concordance between HER2 scoring algorithms by tumor type and by pathologist

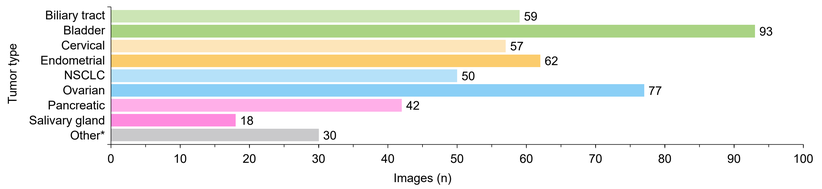
Secondary outcome

- Concordance between pathologists by tumor type and by scoring algorithm

Results and interpretation

- A total of 488 images were assessed. Numbers of images assessed per solid tumor type are presented in **Figure 1**

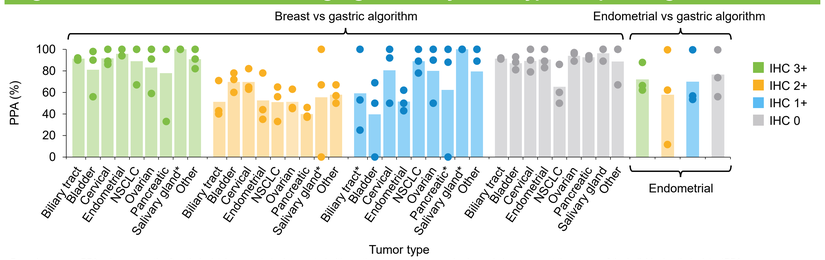
Figure 1. Number of images assessed per tumor type



*Other tumor cohort from DESTINY-PanTumor02, including adenocarcinoid tumor of the appendix, adenoid cystic carcinoma, salivary gland cancer, extramammary Paget disease, head and neck cancer, lip and/or oral cavity cancer, oropharyngeal neoplasm, intestinal adenocarcinoma, malignant neoplasm of unknown primary site, cutaneous melanoma, esophageal adenocarcinoma, esophageal squamous cell carcinoma, testis cancer, and vulval cancer.[7] NSCLC, non-small cell lung cancer

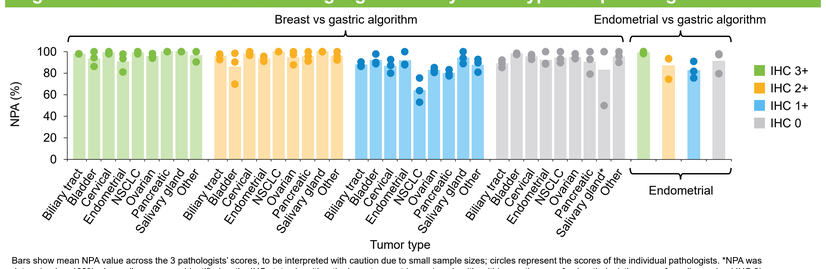
- Between the breast and gastric algorithms, PPA was greater when scoring IHC 3+ and IHC 0 compared with scoring IHC 2+ and IHC 1+ (**Figure 2**)
- Between the endometrial and gastric algorithms, PPA lacked consistency across pathologists for all HER2 expression levels (**Figure 2**)
- NPA was greatest when scoring IHC 3+ and IHC 2+ compared with scoring IHC 1+ and IHC 0 (**Figure 3**)

Figure 2. PPA between HER2 scoring algorithms by tumor type and pathologist



Bars show mean PPA value across the 3 pathologists' scores, to be interpreted with caution due to small sample sizes; circles represent the scores of the individual pathologists. *PPA was determined as 100% where there were no cases identified as the IHC status by the breast or gastric scoring algorithm (this was the case for all 3 pathologists' scores for salivary gland IHC 3+, and for 1 pathologist's scores for each salivary gland IHC 2+ and 1+, biliary tract IHC 1+, and pancreatic IHC 1+). HER2, human epidermal growth factor receptor 2; IHC, immunohistochemistry; NSCLC, non-small cell lung cancer; PPA, positive percent agreement

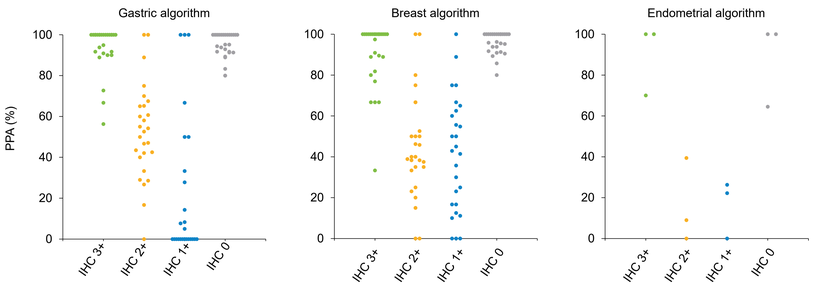
Figure 3. NPA between HER2 scoring algorithms by tumor type and pathologist



Bars show mean NPA value across the 3 pathologists' scores, to be interpreted with caution due to small sample sizes; circles represent the scores of the individual pathologists. *NPA was determined as 100% where all cases were identified as the IHC status by either the breast or gastric scoring algorithm (this was the case for 1 pathologist's scores for salivary gland IHC 0). HER2, human epidermal growth factor receptor 2; IHC, immunohistochemistry; NPA, negative percent agreement; NSCLC, non-small cell lung cancer

- Across the HER2 expression levels, inter-pathologist PPA was greatest when scoring IHC 3+ and IHC 0 for all algorithms; substantial inter-pathologist variability was observed across tumor types when scoring IHC 2+ and IHC 1+ for all algorithms (**Figure 4**)

Figure 4. Inter-pathologist PPA per HER2 scoring algorithm



Circles represent individual inter-pathologist pairwise PPA scores and may overlap owing to proximity of scores. There are 3 PPA scores per IHC status for the endometrial algorithm (3 pathologist pairings × 1 tumor type) and 27 PPA scores per IHC status for the gastric algorithm and breast algorithm (3 pathologist pairings × 9 tumor types). PPA result depends on which pathologist in each pair is arbitrarily assigned as the reference; therefore, figure should be interpreted with caution. HER2, human epidermal growth factor receptor 2; IHC, immunohistochemistry; PPA, positive percent agreement

- Across tumor types, when images were scored according to the gastric and breast algorithms, the majority of inter-pathologist pairwise comparisons had Cohen's kappa coefficient values of >0.4, indicating at least moderate agreement (**Table 1**)
- Using the endometrial algorithm, inter-pathologist pairwise comparisons for endometrial tumors had Cohen's kappa coefficients (95% confidence interval) of 0.43 (0.28, 0.58), 0.17 (0.04, 0.29), and 0.33 (0.18, 0.48), indicating moderate, slight, and fair agreement, respectively[16]

Table 1. Inter-pathologist Cohen's kappa coefficient by HER2 scoring algorithm and tumor type

Tumor type, Cohen's kappa value (95% CI)	Gastric algorithm			Breast algorithm		
	Pathologist A vs B	Pathologist A vs C	Pathologist B vs C	Pathologist A vs B	Pathologist A vs C	Pathologist B vs C
Biliary tract	0.60 (0.44, 0.75)	0.47 (0.33, 0.60)	0.52 (0.38, 0.65)	0.56 (0.41, 0.71)	0.65 (0.50, 0.79)	0.59 (0.43, 0.75)
Bladder	0.63 (0.50, 0.75)	0.57 (0.45, 0.69)	0.64 (0.53, 0.76)	0.43 (0.31, 0.55)	0.52 (0.40, 0.64)	0.60 (0.48, 0.72)
Cervical	0.65 (0.50, 0.81)	0.52 (0.38, 0.67)	0.64 (0.49, 0.79)	0.55 (0.40, 0.70)	0.56 (0.42, 0.70)	0.55 (0.39, 0.71)
Endometrial	0.64 (0.50, 0.78)	0.41 (0.26, 0.57)	0.49 (0.34, 0.63)	0.55 (0.40, 0.69)	0.59 (0.45, 0.73)	0.71 (0.58, 0.85)
NSCLC	0.54 (0.36, 0.71)	0.36 (0.21, 0.51)	0.46 (0.28, 0.64)	0.34 (0.15, 0.52)	0.27 (0.08, 0.46)	0.57 (0.40, 0.75)
Ovarian	0.66 (0.53, 0.80)	0.45 (0.32, 0.58)	0.56 (0.43, 0.68)	0.46 (0.32, 0.61)	0.52 (0.39, 0.65)	0.58 (0.44, 0.71)
Pancreatic	0.52 (0.30, 0.74)	0.53 (0.36, 0.71)	0.51 (0.30, 0.72)	0.64 (0.44, 0.84)	0.63 (0.45, 0.81)	0.57 (0.37, 0.76)
Salivary gland	0.44 (0.03, 0.85)	0 (0.0, 0.0)	0 (0.0, 0.0)	0.14 (−0.08, 0.37)	0.44 (0.03, 0.85)	0.64 (0.00, 1.00)
Other	0.63 (0.42, 0.83)	0.44 (0.26, 0.63)	0.42 (0.22, 0.62)	0.51 (0.31, 0.71)	0.48 (0.27, 0.69)	0.57 (0.37, 0.76)

Agreement key:[16] 0 to 0.2 (slight) >0.2 to 0.4 (fair) >0.4 to 0.6 (moderate) >0.6 to 0.8 (substantial)

CI, confidence interval; HER2, human epidermal growth factor receptor 2; NSCLC, non-small cell lung cancer

Study limitations

- The tissue staining assays used were not validated for use in the tumor types analyzed; further validation of the staining protocols would be required to ensure accurate comparison of assay performance
- The sample numbers for each tumor type were relatively small and only 3 pathologists assessed images, both limiting interpretation of the findings

Acknowledgments

Thank you to the patients and their caregivers for their participation, and the study site staff for their contributions to the DESTINY-PanTumor02 study. We also thank Eun-Ang Raiber-Moreau and Flavia Michelini (AstraZeneca employees) for interpretation of the data and scientific review of the poster, and the pathologists from Avaden Bio, Seattle, WA, US, for assessing and scoring images for this analysis. DESTINY-PanTumor02 and this observational analysis were supported and sponsored by AstraZeneca in collaboration with Daiichi Sankyo. In March 2019, AstraZeneca entered into a global development and commercialization collaboration agreement with Daiichi Sankyo for trastuzumab deruxtecan (T-DXd; DS-8201). Under the guidance of the authors and in accordance with Good Publication Practice, medical writing and editorial support was provided by Zoë Hine, BSc, of Helios Medical Communications, part of Helios Global Group, and was funded by AstraZeneca.

Disclosures

Prof. Yang declares no conflicts of interest.

References

1. Yan M, et al. *Cancer Metastasis Rev*. 2015;34:157–164
2. Li Z, et al. *EBioMedicine*. 2020;62:103074
3. Uzunparmak B, et al. *Ann Oncol*. 2023;34:1035–1046
4. Xing F, et al. *Mol Cancer*. 2023;22:6
5. Halle MK, et al. *Br J Cancer*. 2018;118:378–387
6. Enhertu (fam-trastuzumab deruxtecan-nxki): highlights of prescribing information. 2024. Available from: www.accessdata.fda.gov/drugsatfda_docs/label/2024/761139s028lbl.pdf (Accessed June 13, 2024)
7. Meric-Bernstam F, et al. *J Clin Oncol*. 2024;42:47–58
8. Smit EF, et al. *Lancet Oncol*. 2024;25:439–454
9. Raghav K, et al. *Lancet Oncol*. 2024;doi 10.1016/S1470-2045(24)00380-2: Aug 05 [Epub ahead of print]
10. Li BT, et al. *Lancet Oncol*. 2024;25:707–719
11. Hofmann M, et al. *Histopathology*. 2008;52:797–805
12. Zhang H, et al. *Histopathology*. 2024;85:3–19
13. Wolff AC, et al. *Arch Pathol Lab Med*. 2023;147:993–1000
14. Fader AN, et al. *J Clin Oncol*. 2018:36:2044–2051
15. Wolff AC, et al. *Arch Pathol Lab Med*. 2007;131:18–43
16. Landis JR, Koch GG. *Biometrics*. 1977;33:159–174