Accurate quantification of slide-level HER2 scores in breast cancer using a machine-learning model, AIM-HER2 Breast Cancer

Poster #P02-14-12

STUDY BACKGROUND

- HER2 expression level is a key factor in determining the optimal treatment course for breast cancer patients. Roughly 15% of breast cancers are HER2+, and determination of HER2 status is routinely assessed by immunohistochemistry (IHC). Accurate assessment of the HER2 IHC score (0, 1+, 2+, 3+) by pathologists is therefore critical, especially in light of novel therapeutic approaches demonstrating efficacy in the HER2-low setting (IHC scores 1+, and 2+/FISH-)[1,2].
- To assist pathologists with the consistent provision of reproducible and accurate scores across the entire HER2 scoring range, we developed a machine-learning algorithm ("AIM-HER2") to generate accurate, slide-level HER2 scores aligned with ASCO-CAP guidelines in clinical breast cancer HER2 IHC specimens.

AIM-HER2 BREAST MODEL

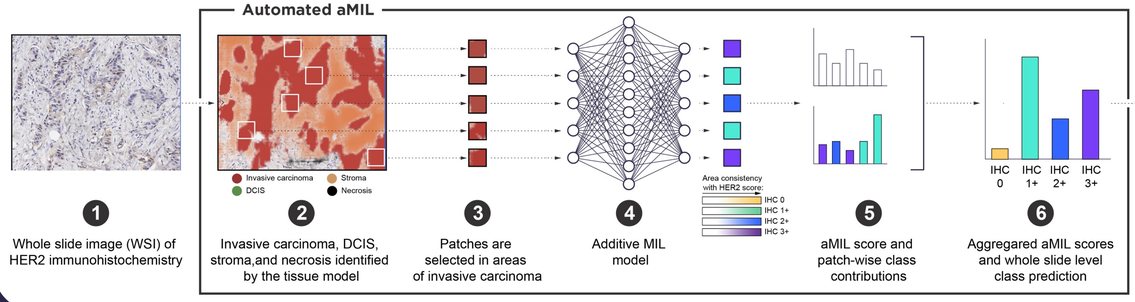


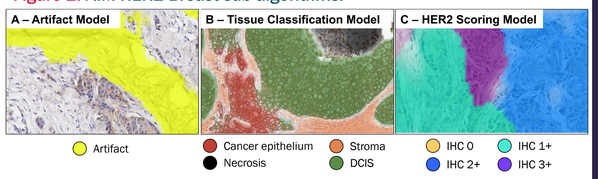
Figure 1. AIM-HER2 Breast additive multiple instance learning (aMIL) model. From WSI of HER2 IHC, image patches are selected in model-predicted areas of cancer and stroma. The aMIL model predicts the likelihood of each patch as being a certain HER2 score (IHC 0, 1+, 2+, or 3+, consistent with ASCO/CAP guidelines). From the model output, a slide-level HER2 score can be predicted, and image overlays show the relative contribution of areas of tissue to the model-predicted score.

METHODS

Model Development. AIM-HER2 was developed using whole-slide images (WSI; N=4261) from clinical and commercial sources. WSI were split into training (N=2694, 63%) and optimization (N=1567, 37%) sets. An additive multiple instance learning (aMIL) model[4] was trained to predict HER2 scores directly from WSI and create interpretable heatmaps that depict HER2 predictions in tissue images (Fig. 1). The model was trained on >157,000 annotations and approximately 12,000 slide-level HER2 scores from over 65 board-certified pathologists; labels were collected from three pathologists per slide. Image artifacts and in situ carcinomas were identified using artifact and tissue segmentation models and were excluded, leaving only regions of invasive carcinoma to be analyzed. AIM-HER2 Breast makes use of three sub-algorithms: 1) Artifact Model, 2) Tissue Model, and 3) HER2 Scoring Model (Fig. 2). Model outputs include: HER2 score (0, 1+, 2+, 3+), area of invasive carcinoma, and aMIL density heatmap overlays (Fig. 3).

Evaluation of AIM-HER2 Breast. AIM-HER2 performance assessed on HER2 IHC slides obtained from five academic or commercial sources (N=804 total, 770 marked evaluable by pathologists). Two sources, comprising 223 slides (30.3% of total), were held out and not seen during training. Board-certified pathologists (N=52) with relevant experience provided manual HER2 scores based on ASCO-CAP guidelines. Nested pairwise non-inferiority analysis[5] was used to compare model performance to that of pathologists (N=3 pathologists per slide). In the nested pairwise framework, agreement among pathologists was compared to agreement between AIM-HER2 and pathologists via linear kappa, so that summary metrics account for inter-pathologist variability.

Figure 2. AIM-HER2 Breast sub-algorithms.



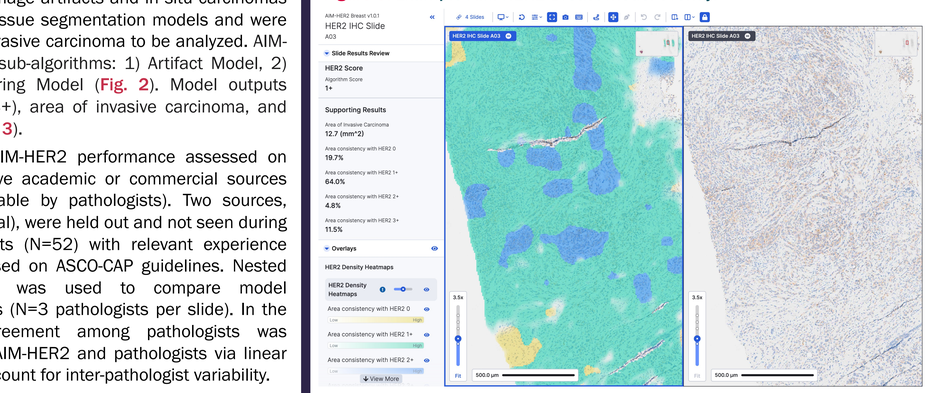
AIM-HER2 sub-algorithms **A)** detect and remove all artifact (e.g. tissue folds, damaged tissue, blur), **B)** identify and classify regions of stroma, necrosis, invasive cancer, and ductal in situ carcinoma (DCIS), and **C)** calculate slide-level HER2 scores according to ASCO/CAP guidelines.

RESULTS

Key Results:

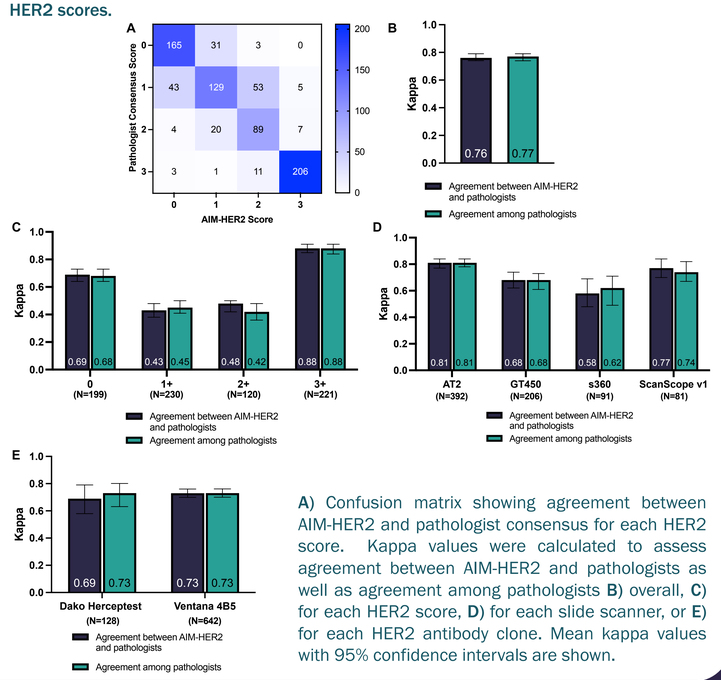
- High concordance was observed between AIM-HER2-predicted and pathologist-labeled slide-level HER2 scores, overall (Figure 4A, 4B).
- Similar results were observed when assessing AIM-HER2 performance
  - For each HER2 scoring level (Figure 4C)
  - For multiple slide scanners (Figure 4D)
  - With IHC with multiple HER2 IHC antibody clones (Figure 4E).

Figure 3. Example AIM-HER2 Breast model overlay.



A representative breast cancer WSI is shown after AIM-HER2 deployment. HER2 IHC is shown at the right, while the model overlay is shown on the left. Model overlays indicate tissue regions predicted to be IHC 0, IHC 1+, IHC 2+, or IHC 3+, with additional metrics (area of invasive carcinoma and area consistency with each HER2 score) listed in the side panel.

Figure 4. Concordance between AIM-HER2 predicted and pathologist-labeled slide-level HER2 scores.



A) Confusion matrix showing agreement between AIM-HER2 and pathologist consensus for each HER2 score. Kappa values were calculated to assess agreement between AIM-HER2 and pathologists as well as agreement among pathologists **B)** overall, **C)** for each HER2 score, **D)** for each slide scanner, or **E)** for each HER2 antibody clone. Mean kappa values with 95% confidence intervals are shown.

CONCLUSIONS

We developed AIM-HER2, a novel aMIL-based approach for predicting slide-level HER2 IHC scores. AIM-HER2 has similar levels of agreement with pathologists as pathologists have with each other for determining HER2 score. This result is upheld when slides imaged using multiple scanning platforms and stained using multiple HER2 antibody clones. The performance of AIM-HER2 on multiple scanners and after multiple assays supports broad applicability of this algorithm in clinical laboratories, including for the identification of HER2-low cases. Work is ongoing to perform similar analyses in an independent, real-world dataset.

AUTHORS

Zahil Shanis[1], Ryan Cabeen[1,#], Shreya Chakraborty[1, #], John Shamshoian[1], Marc Thibault[1], Blake Martin[1], Harshith Pagidela[1], Dinkar Juyal[1], Syed Ashar Javed[1], William Qian[1, #], Juhyun Kim[1], Beckett Rucker[1], Jacqueline Brosnan-Cashman[1], Harsha Pokkalla[1], Jimish Mehta[1, #], Amaro Taylor-Weiner[1, #], Ben Glass[1], Santhosh Balasubramanian[1]

[1]PathAI, Boston, MA
Employed at PathAI at time of study

CONTACT

santhosh.bala@pathai.com
publications@pathai.com

REFERENCES

1) Fernandez, AI, et al. *JAMA Oncol*. 2022 8(4):1-4.
2) Modi, S et al. *N Engl J Med*. 2022 387:9-20.
3) Wolff, AC, et al. *J Clin Oncol*. 2018 36:2105-22.
4) Javed, SA, et al. *Adv Neural Inf Process Syst*. 2022 35: 20689-702.
5) Gerardin, Y, et al. 2023 arXiv:2306.04709

ACKNOWLEDGMENTS

We would like to thank the members of PathAI Pathology Network who contributed towards this work.

This poster template was developed by SciStories LLC. https://scistories.com/

An electronic copy of this poster can be found here:

San Antonio Breast Cancer Symposium, December 5-9, 2023, San Antonio, TX