Prediction of Pathologic Complete Response from Histopathology Images of HER2+ Breast Cancer Using an AI Foundation Model

Reva Basho[1], Edwin A. Heredia[1], Heather McArthur[2], Zihang Fang[1], Alexis LeVee[3], Valerie Narumi[1], Ethan Hamid[1], Lara Papasian[1], Thomas Jonsson[1], Farnaz Dadmanesh[4], Shikha Bose[4], Shiva Patre[1], Thomas Keller[1], John Paine[1], Brian Tran[1], Brianna Simental[1], Olga Castellanos[1], Jerry Lee[1], Stephen Shiao[4], David Agus[1], Naim Matasci[1]

[1]Ellison Medical Institute; Los Angeles, CA; [2]UT Southwestern; Dallas, TX; [3]University of California Los Angeles (UCLA); Los Angeles, CA; [4]Cedars-Sinai Medical Center; Los Angeles, CA

Abstract: 1001 PS#-04-01

Background

- Neoadjuvant taxane-based chemotherapy plus anti-HER2 therapy is standard for locally advanced, operable HER2+ breast cancer.[1]
- Pathologic complete response (pCR) after neoadjuvant therapy is associated with improved event-free survival.[2]
- pCR occurs less frequently in patients with HER2+, hormone receptor-positive (HR+) disease compared to patients with HER2+, HR-negative (HR-) disease (45-55% vs 65-75%).[3]
- Improved prediction of pCR outcome for patients treated with standard neoadjuvant therapy can facilitate optimal patient selection for trials of novel therapeutic approaches in both HR+ and HR- subsets.

Objectives

- Develop and validate an artificial intelligence/machine learning (AI/ML) model to predict pCR status after standard neoadjuvant chemotherapy consisting of taxane-based chemotherapy in combination with anti-HER2 therapy for HER2+/HR+ and HER2+/HR- patients using H&E-stained baseline biopsy slides.
- Apply a pathology foundation model (CanvOI[4]) to enhance pCR prediction.

Materials and Methods

Data Sources

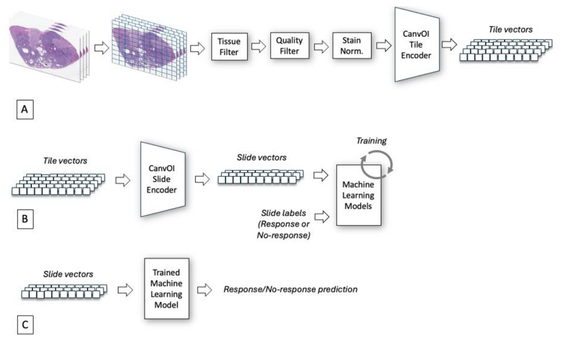
- Patients with operable HER2+ breast cancer treated neoadjuvantly with taxane-based chemotherapy plus anti-HER2 therapy.
- Patients pooled from Cedars-Sinai Medical Center under an IRB-approved protocol (Cohort 1) and 2 publicly available datasets: IMPRESS[5] (Cohort 2) and Yale[6] (Cohort 3).
- N=231 patients; N=114 HER2+, HR+ (classified as ER and/or PR > 10%) and N=117 HER2+, HR- (classified as ER and PR ≤ 10%).

Methods

- H&E-stained baseline biopsy whole-slide images (WSIs) digitized at 20× magnification.
- WSIs divided into 380×380-pixel tiles.
- CanvOI foundation model applied (Figure 1):
  - Tile encoder → vector embeddings
  - Slide encoder → aggregate slide-level vectors
- Logistic regression and other ML classifiers trained on vectors for pCR classification within HER2+/HR+ and HER2+/HR- subsets.
- 4x4 cross-validation with random partitions.
- Data balancing (trimming dominant class) to mitigate pCR imbalance within HER2+/HR+ and HER2+/HR- subsets.
- Multivariable analysis using logistic regression and 5-fold cross validation with clinical variables associated with pCR.

Results

Figure 1: WSI Processing Workflow



(A) Whole-slide images are tiled and converted into vectors.
(B) Tile vectors are aggregated into slide-level vectors for model training.
(C) Vectors from test images are used with the trained model to assess pCR prediction accuracy.

Patient Characteristics

Table 1: Baseline Patient Characteristics by Cohort

N (%)		Cohort 1	Cohort 2	Cohort 3
Total Number		86	62	83
Age, median (range)		49 (25–87)	56 (30–76)	
Histology	Ductal	81 (94)	61 (98)	
	Lobular	3 (4)	1 (2)	
	Mixed	2 (2)		
Grade	1		1 (2)	
	2	28 (33)	27 (44)	
	3	57 (66)	34 (55)	
	†	1 (1)	0 (0)	83 (100)
HER2 IHC	1+	0 (0.0)	2 (3)	
	2+	18 (21)	7 (11)	30 (36)
	3+	68 (79)	53 (85)	53 (64)
HR	Positive	39 (45)	28 (45)	47 (57)
	Negative	47 (55)	34 (55)	36 (43)
T	T1	12 (14)		
	T2	50 (58)		
	T3	21 (24)		
	T4	3 (4)		
Node Status	N0	37 (43)		
	N+	47 (55)		
	†	2 (2)	62 (100)	83 (100)
pCR		48 (56)	38 (61)	35 (42)

Table 2: Clinical Variables for HER2+/HR+ Patients (Cohorts 1 & 2)

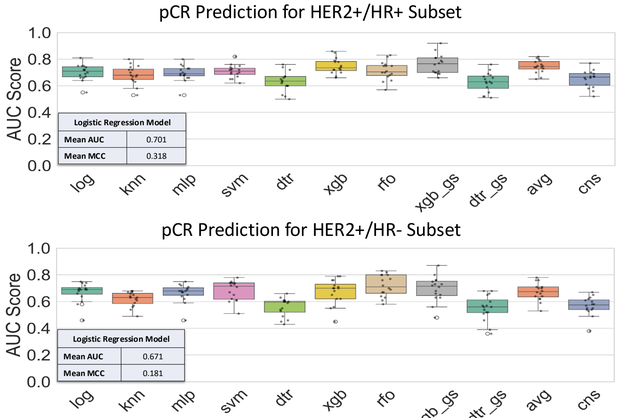
N (%)		Overall	pCR	Non-pCR	P-value*
Total Number		67	27 (40)	40 (60)	
HER2 IHC	1–2+	15 (22)	2 (7)	13 (32)	0.034
	3+	52 (78)	25 (93)	27 (68)	
Grade	1–2	33 (49)	13 (48)	20 (50)	1.000
	3	33 (49)	14 (52)	19 (48)	
	†	1 (1)	0 (0)	1 (2)	

Table 3: Clinical Variables for HER2+/HR- Patients (Cohorts 1 & 2)

N (%)		Overall	pCR	Non-pCR	P-value*
Total Number		81	59 (73)	22 (27)	
HER2 IHC	1–2+	12 (15)	2 (3)	10 (45)	<0.001
	3+	69 (85)	57 (97)	12 (55)	
Grade	1–2	23 (28)	18 (31)	5 (23)	0.679
	3	58 (72)	41 (69)	17 (77)	

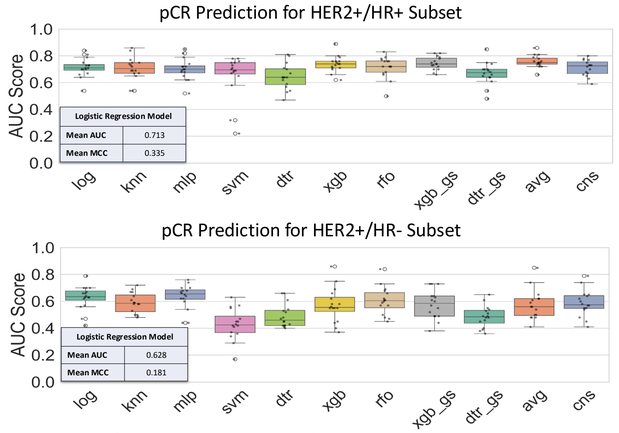
† Data not available
* P-values calculated using chi-square test of homogeneity

Figure 2: Prediction of pCR by HR Subset



Area under the curve (AUC) and Matthews correlation coefficient (MCC) scores for prediction of pCR using CanvOI foundation model adapted by various machine leaning methods in the HER2+/HR+ subset (top) and HER2+/HR- subset (bottom).

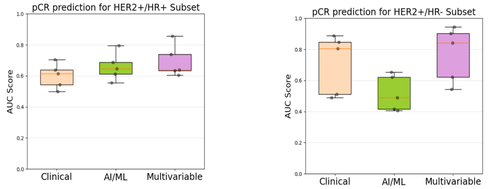
Figure 3: Prediction of pCR by HR Subset after Data Balancing



AUC and MCC scores for prediction of pCR using CanvOI foundation model adapted by various machine leaning methods in the HER2+/HR+ subset (top) and HER2+/HR- subset (bottom) after data balancing by trimming. Logistic regression AUC score is not significantly improved after data balancing in either the HER2+/HR+ (p=0.591) or HER2+/HR- (p=0.154) subsets. P-values were calculated using Welch's T-test.

Results (continued)

Figure 4: Multivariable Analysis



HER2+/HR+	Mean AUC	P-value
Clinical	0.600	--
AI/ML	0.659	0.306
Multivariable	0.694	0.150

HER2+/HR-	Mean AUC	P-value
Clinical	0.708	--
AI/ML	0.518	0.102
Multivariable	0.772	0.603

Multivariable model using logistic regression of clinical variables (HER2 IHC and grade) and the AI/ML model for pCR prediction in both HER2+/HR+ and HER2+/HR- subsets. The model is limited to Cohorts 1 and 2 due to availability of clinical data. P-values were calculated using Welch's T-test, comparing the AI/ML model and multivariable model to clinical variables.

Conclusions and Future Directions

- A digital pathology-based AI/ML model applied on baseline biopsy slides from patients with operable HER2+ breast cancer may enhance prediction of sensitivity to standard neoadjuvant therapy.
- A multivariable model of the AI/ML model and clinical variables resulted in a trend toward improved prediction of pCR in both HER2+/HR+ and HER2+/HR- subsets but did not reach statistical significance when compared to clinical variables alone in this limited dataset.
- Data balancing by data trimming did not significantly improve the model.
- Such a model may help guide therapeutic decisions and clinical trial selection criteria for patients with HER2+ early breast cancer.
- For patients with HER2+, HR+ breast cancer, rates of pCR with current standard therapy remain low and improved novel approaches driven by biomarker selection are needed.
- Larger datasets are needed to resolve variations in data sources and identify new biomarkers based on histological patterns.

References and Acknowledgements

References

1. A. Schneeweiss et al., Long-term efficacy analysis of the randomised, phase II TRYPHAENA cardiac safety study: Evaluating pertuzumab and trastuzumab plus standard neoadjuvant anthracycline-containing and anthracycline-free chemotherapy regimens in patients with HER2-positive early breast cancer. Eur J Cancer 2018.
2. C. Yau et al., Residual cancer burden after neoadjuvant chemotherapy and long-term survival outcomes in breast cancer: a multicentre pooled analysis of 5161 patients. Lancet Oncol 2022.
3. N. Harbeck et al., Neoadjuvant and adjuvant treatment of patients with HER2-positive early breast cancer. Breast 2022.
4. J. Zalach, J. et al., CanvOI, an Oncology Intelligence Foundation Model: Scaling FLOPS Differently. Preprint arXiv:2409.02885 2024.
5. Z. Huang et al., Artificial intelligence reveals features associated with breast cancer neoadjuvant chemotherapy responses from multi-stain histopathologic images. npj Precis Onc 2023.
6. S. Farahmand et al., Deep learning trained on hematoxylin and eosin tumor region of Interest predicts HER2 status and trastuzumab treatment response in HER2+ breast cancer. Mod Pathol 2022.

Acknowledgements

This research was partially supported by a grant from the Breast Cancer Research Foundation (BCRF-2020-002).