Multimodal AI for predicting treatment response to breast cancer neoadjuvant chemotherapy

Aryan Kumar[1], Jefferey Wang[2], Rina Hui[2,3], Hoyeon Lee[3,4]
[1]Department of Computer Science, Faculty of Computing and Data Science, University of Hong Kong
[2]Department of Medicine, Queen Mary Hospital
[3]Centre of Cancer Medicine, LKS Faculty of Medicine, University of Hong Kong
[4]Department of Diagnostic Radiology, LKS Faculty of Medicine, University of Hong Kong

INTRODUCTION

- Breast cancer is the most common cancer among women worldwide [1].
- A standard treatment approach is neoadjuvant chemotherapy (NACT), which integrates chemotherapy and surgical resection[2].
- The treatment response of NACT can be assessed through pathological analysis, and several studies have demonstrated that absence of residual cancer, referred to as pathological complete response (pCR) is strongly correlated with favorable clinical outcomes.
- Accurate prediction for pCR provides valuable information clinical decision-making [3,4].
- While mammograms (MMGs) are commonly available in clinical practice, most existing studies on pCR prediction studies primarily focused on MRI, which is less widely available and more expensive [5,6].
- Furthermore, most previous studies have relied primarily on imaging information to predict pCR [7].
- Therefore, this study aims to develop a multimodal AI to predict pCR using MMGs and clinical information.

MATERIALS AND METHODS

- Post-NACT mammograms and clinicopathological data from 151 patients with HER2-positive or triple-positive breast cancer were retrospectively collected at Queen Mary Hospital, Hong Kong (IRB Ref. No.: UW-25-447).
- Clinicopathological features included age, molecular subtype, hormone-receptor status, and chemotherapy regimen.
- We developed a multimodal AI as shown in Figure 1.
- The proposed model utilizes 2-views of MMGs and clinicopathological data, i.e., age, treatment regimen, HER2 status, and HR status, as inputs.
- Features from MMGs were extracted using Mammo-FM, a foundation model for MMGs [8].
- Subsequently, the extracted imaging features were conditioned on the molecular subtypes of the tumor and concatenated with clinical features extracted using a multilayer perceptron (MLP).
- To compare the performance of the proposed multimodal AI, we implemented 2 additional AI models, which utilize clinical data only (Tabular) and imaging data only (Image-only).
- Five-fold cross-validation and compared the pooled out-of-fold (OOF) area under the receiver operating characteristic curve (AUROC) to evaluate the performances of the 3 models.

Table 1. Patients' demographics

Age	pCR	Molecular subtypes	Treatment regimen
56 (33 – 84)	Positive: 83 Negative:68	Triple positive: 81 HER2 positive: 70	DDPH: 120 TCPh: 19 Others:12

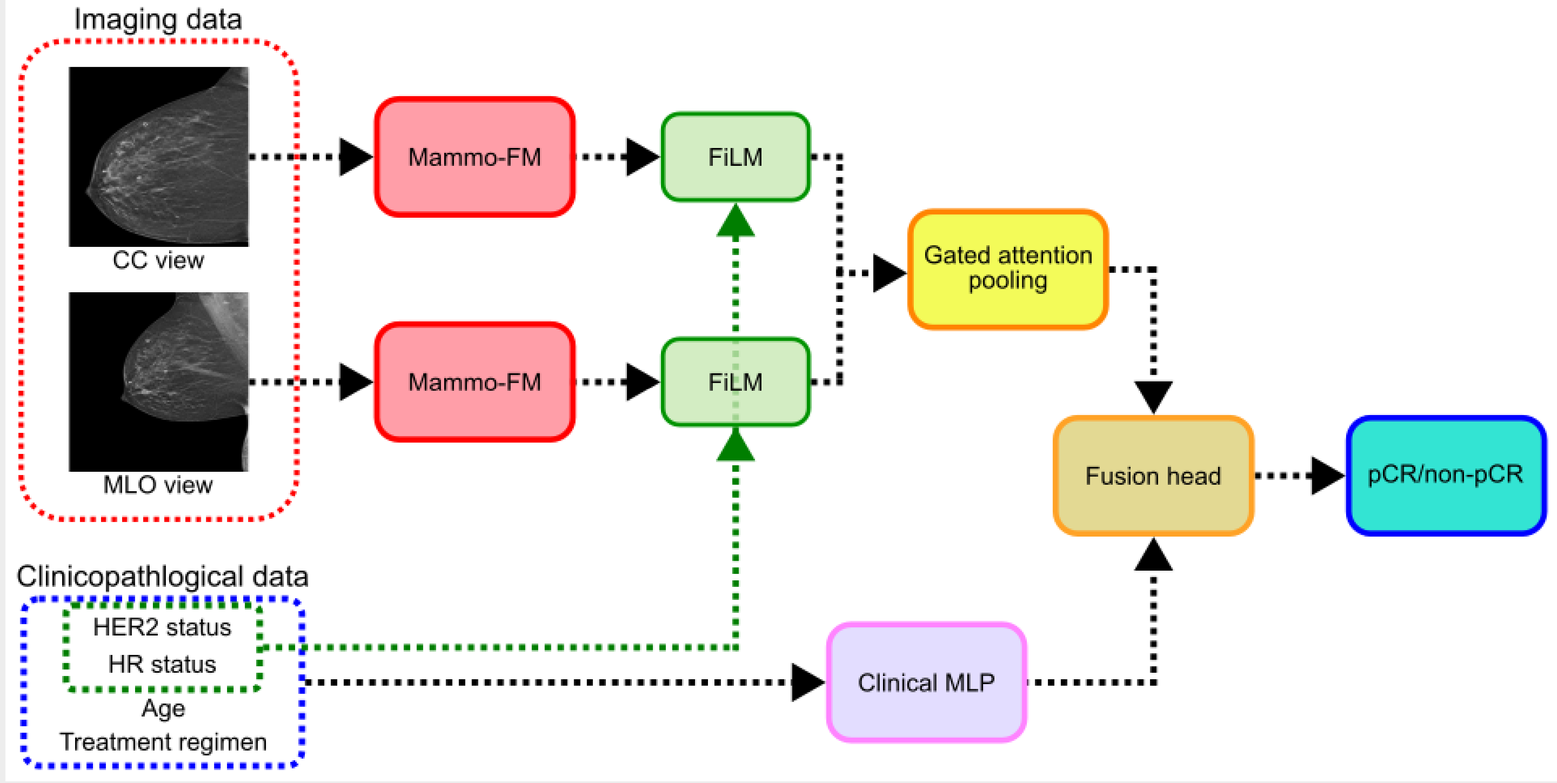

Figure 1. Proposed multimodal AI for pCR prediction

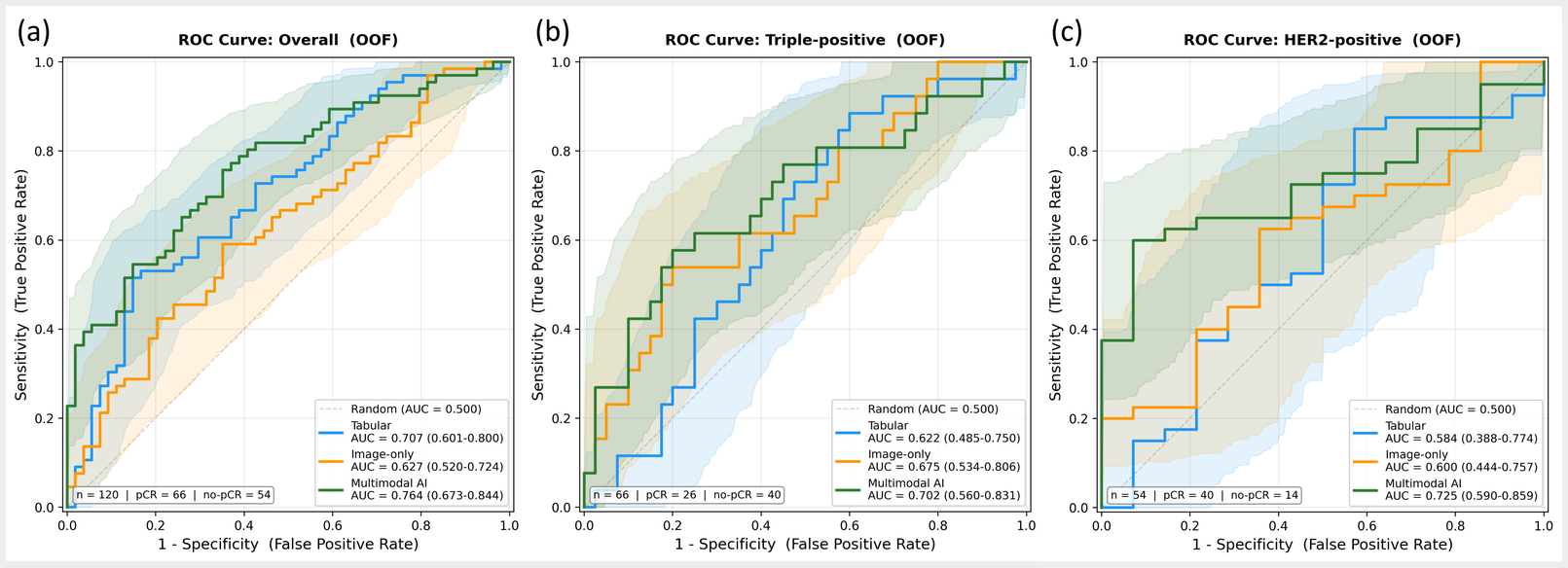

Figure 2. Receiver Operating Characteristics (ROC) curves for (a) overall dataset, (b) Triple-positive cases, and (c) HER2-positive cases.

RESULTS

- Patient demographics are summarized in Table 1.
- As shown in Figure 2 (a), the proposed multimodal AI achieved a higher AUROC compared with the Tabular and Image-only models.
- Considering the differences in the pCR rates between the molecular subtypes (74% vs 39%), we also compared the prediction performance between different molecular subtypes.
- The ROC curves of demonstrate that the proposed multimodal AI achieved consistently better performance than the other models across the different molecular subtypes.
- These results demonstrate that integrating imaging and clinical data can capture complementary patient-specific information and improve prediction performance.

DISCUSSION AND CONCLUSION

- In this study, we proposed a multimodal AI to predict pCR following NACT for breast cancer.
- The multimodal AI utilizes MMGs, which are commonly available in clinical practice, and clinical information to predict the pCR.
- These results demonstrate that MMGs can be utilized to predict pCR following NACT for breast cancer and that integrating imaging and clinical information could improve the robustness of pCR prediction.
- Despite the promising results, this study analyzed data from a small cohort data collected from a single institution.
- To further evaluate the clinical impact of the proposed multimodal AI, we will conduct a multicenter study and prospective clinical trials.

REFERENCES

[1] Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I., & Jemal, A. (2024). Global Cancer Statistics 2022: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, *74*(3), 229–263. https://doi.org/10.3322/caac.21834
[2] Korde, L. A., Somerfield, M. R., Carey, L. A., Crews, J. R., Denduluri, N., Hwang, E. S., Khan, S. A., Loibl, S., Morris, E. A., Perez, A., Regan, M. M., Spears, P. A., Sudheendra, P. K., Symmans, W. F., Yung, R. L., Harvey, B. E., & Hershman, D. L. (2021). Neoadjuvant Chemotherapy, Endocrine Therapy, and Targeted Therapy for Breast Cancer: ASCO Guideline. *Journal of Clinical Oncology*, 39(13), JCO.20.03399. https://doi.org/10.1200/jco.20.03399
[3] Cortazar, P., Zhang, L., Untch, M., Mehta, K., Costantino, J. P., Wolmark, N., Bonnefoi, H., Cameron, D., Gianni, L., Valagussa, P., Swain, S. M., Prowell, T., Loibl, S., Wickerham, D. L., Bogaerts, J., Baselga, J., Perou, C., Blumenthal, G., Blohmer, J., von Minckwitz, G. (2014). Pathological complete response and long-term clinical benefit in breast cancer: The CTNeoBC pooled analysis. *The Lancet*, 384(9938), 164–172. https://doi.org/10.1016/s0140-6736(13)62422-8
[4] Spring, L. M., Fell, G., Arfe, A., Sharma, C., Greenup, R., Reynolds, K. L., Smith, B. L., Alexander, B., Moy, B., Isakoff, S. J., Parmigiani, G., Trippa, L., & Bardia, A. (2020). Pathologic Complete Response after Neoadjuvant Chemotherapy and Impact on Breast Cancer Recurrence and Survival: A Comprehensive Meta-analysis. *Clinical Cancer Research*, 26(12), 2838–2848. https://doi.org/10.1158/1078-0432.ccr-19-3492
[5] Liang, X., Yu, X., & Gao, T. (2022). Machine learning with magnetic resonance imaging for prediction of response to neoadjuvant chemotherapy in breast cancer: A systematic review and meta-analysis. *European Journal of Radiology*, 150, 110247. https://doi.org/https://doi.org/10.1016/j.ejrad.2022.110247
[6] Krasniqi, E., Filomeno, L., Arcuri, T., Ferretti, G., Gasparro, S., Fulvi, A., Roselli, A., D'Onofrio, L., Pizzuti, L., Barba, M., Maugeri-Saccà, M., Botti, C., Graziano, F., Puccica, I., Cappelli, S., Pelle, F., Cavicchi, F., Villanucci, A., Paris, I., … Vici, P. (2025). Multimodal deep learning for predicting neoadjuvant treatment outcomes in breast cancer: A systematic review. *Biology Direct*, 20(1). https://doi.org/10.1186/s13062-025-00661-8
[7] Joo, S., Ko, E. S., Kwon, S., Jeon, E., Jung, H., Kim, J.-Y., Chung, M. J., & Im, Y.-H. (2021). Multimodal deep learning models for the prediction of pathologic response to neoadjuvant chemotherapy in breast cancer. *Scientific Reports*, 11(1). https://doi.org/10.1038/s41598-021-98408-8
[8] Ghosh, S., Parthesh, J. V., Rayan, S., Budhraja, P., Aya, K., Morrison, K. C., Tang, A., Cheung Aiden, W. H., Varshney, A., Basak, P., Dai, W., Gichoya, J. W., Hari M, T., Banerjee, I., Visweswaran, S., Poynton, C. B., & Batmanghelich, K. (2025). Mammo-fm: Breast-specific foundational model for integrated mammographic diagnosis, prognosis, and reporting. *arXiv*. https://arxiv.org/abs/2512.00198
[9] Ilse, M., Tomczak, J. M., & Welling, M. (2018). Attention-based deep multiple instance learning. *arXiv*. https://doi.org/10.48550/arXiv.1802.04712

ACKNOWLEDGMENT

This research was conducted under the HKU Laidlaw Scholars Programme, supported by the Laidlaw Foundation.

CONTACT

Aryan Kumar (aryank@connect.hku.hk)
Hoyeon Lee (hlee01@hku.hk)