10th International Breast Density & Cancer Risk Assessment Workshop. June 7-9, 2023 | Kailua-Kona, Hawai`i

Check out our other research!

Image-Based Models for Predicting Advanced Breast Cancer Risk

Lambert Leong[1,2], Thomas Wolfgruber[1], Brandon Quon[1], **Arianna Bunnell[1,2]**, Dustin Valdez[1], Jami Fukui[1], Brenda Hernandez[1], Yurii Shvetsov[1], Karla Kerlikowske[3], John Shepherd[1,2]

University of Hawaiʻi Cancer Center, Honolulu, HI, USA[1], University of Hawaiʻi at Mānoa, Honolulu, HI, USA[2], University of California San Francisco[3]

ARTIFICIAL INTELLIGENCE PRECISION HEALTH INSTITUTE
UNIVERSITY OF HAWAIʻI

Introduction

- Advanced-stage breast cancer (stages III and IV) rates in the U.S.-Affiliated Pacific Islands (USAPI) are much higher than in the continental United States. Examples include Guam (60% advanced-stage cancer rate), American Samoa (74%), and the Federated States of Micronesia (81%) [1]. Reducing advanced-stage cancer involves cancer detection and risk assessment.
- Kerlikowske et al [2] found obesity, breast density BI-RADS category C and D, and proliferative atypia to be associated to advanced stage cancer risk.
- Density is a measurement which is obtained from imaging and it is a highly compressed representation of the total afforded information.
- Image-based ML/DL models have shown that imaging contains information related to cancer risk that is orthogonal to breast density [3].
- We hypothesis that images may contain information beyond breast density related to advanced cancer risk

Methods

Study Design: Prospective study of all women who received screening diagnostic mammograms from 2009 to 2021 at clinical sites participating in the Hawaiʻi and Pacific Islands Mammography Registry (NIH R01CA263491 and U54CA143728)
Participant Selection: Participants had to have a prior negative screening mammography, all four mammographic views, images had to have been acquired on Hologic systems, staging had to be known, and had to have no prior tumors. Cases were excluded if laterality or staging was unknown, if diagnosis was not within two years of ipsilateral mammogram acquisition, and if the contralateral mammogram was obtained outside of 1 year of the ipsilateral mammogram. Controls were all BI-RADS diagnostic scores 2 or lower.
Deep Feature Extraction: A DenseNet121 pretrained with ImageNet weights was used to derive the deep radiomic image features. Mammograms were fed into the locked DenseNet model which output a 1024 feature vector. The feature vectors of the left and right views of the breast were concatenated and used to derive predictive models. **Figure 1** provides an illustrative diagram.
Modeling: A total of 5 logistic models were trained to predict binary outcomes of advanced stage cancer and not advance stage cancer. The first two models used clinical and deep learning derived breast density, respectively, as inputs into the model. The third model only used the deep radiomic features as inputs. The fourth and fifth models combined both density and deep radiomics as input into the models. All models which used breast density as an input were age adjusted. **Figure 2** shows all model configurations with their respective inputs.
Evaluation and Statistics: 5-fold cross validation was used to evaluate the performance of each logistic model. During each fold, all models were trained and evaluated using the same split for fair performance evaluation. Splits were constrained to preserve the ratio of advanced cancers to non-advanced cancers and they were performed on patient ID to avoid leakage as well as to ensure that multiple observations for the same patient remained within only one of the splits. Area Under the receiver operating characteristic Curve (AUCs) were calculated on the test set for each fold and the mean of all folds is reflected in the results.

References

[1] Cancer in the U.S. Affiliated Pacific Islands 2007-2018. Pacific Regional Cancer Registry, 2021.
[2] Kerlikowske K, Chen S, Golmakani MK, Sprague BL, Tice JA, Tosteson ANA, Rauscher GH, Henderson LM, Buist DSM, Lee JM, Gard CC, Miglioretti DL. Cumulative Advanced Breast Cancer Risk Prediction Model Developed in a Screening Mammography Population. JNCI: Journal of the National Cancer Institute. 2022;114(5):676-85. doi: 10.1093/jnci/djac008.
[3] Zhu X, Wolfgruber TK, Leong L, Jensen M, Scott C, Winham S, Sadowski P, Vachon C, Kerlikowske K, Shepherd JA. Deep Learning Predicts Interval and Screening-detected Cancer from Screening Mammograms: A Case-Case-Control Study in 6369 Women. Radiology. 2021;301(3):550-8. doi: 10.1148/radiol.2021203758. PubMed PMID: 34491131.

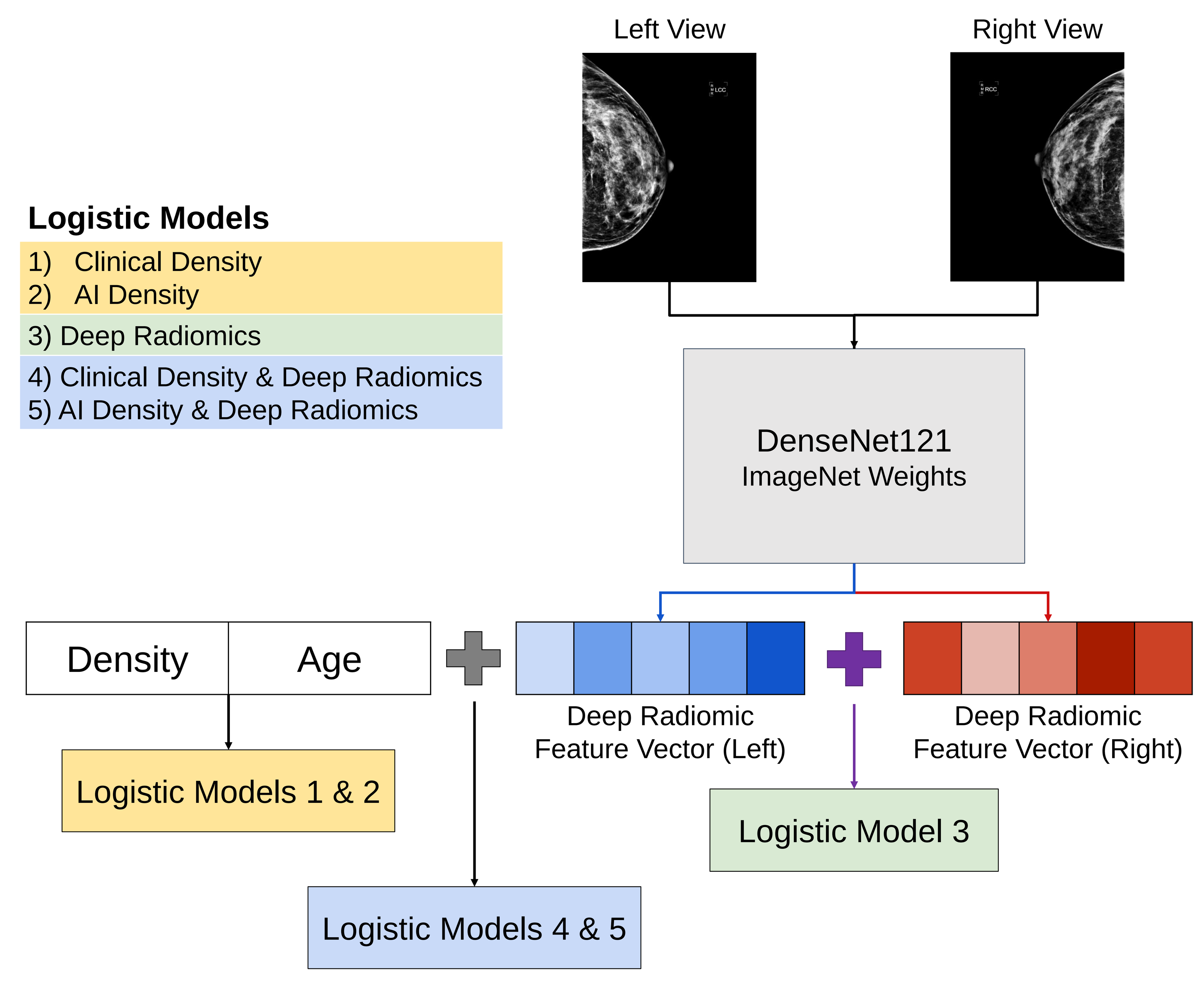

Figure 2: Model design for all 5 logistic models constructed in this work. All models predict the binary outcome of advanced cancer/not advanced cancer. All models which include either clinical or AI-derived breast density are adjusted for age.

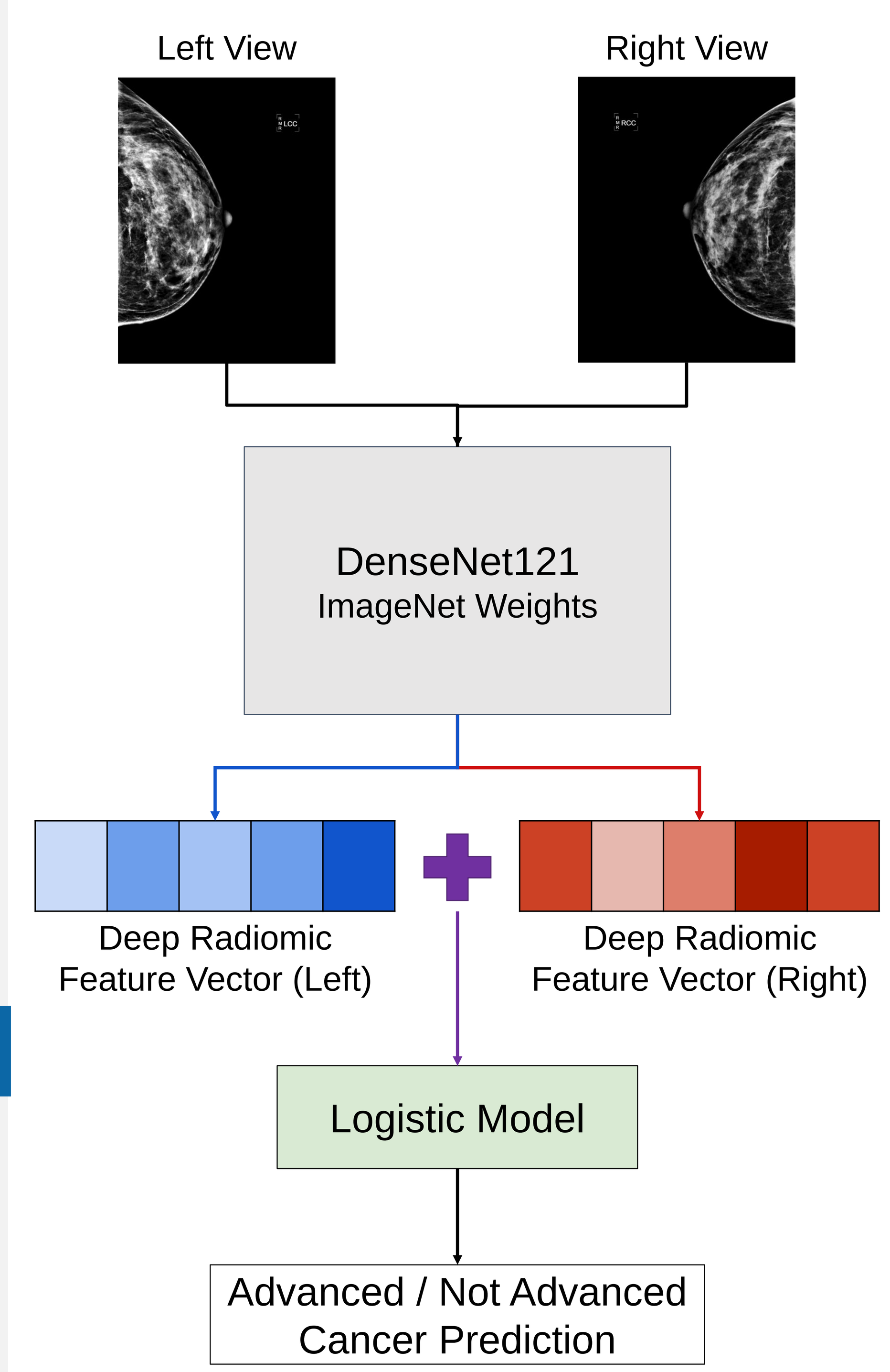

Figure 1: Model design for deep radiomics/image feature-only model.

Table 1: Population descriptive statistics for women included in the dataset sourced from the HIPIMR.

		Population	No Advanced Cancer	Advanced Cancer
N Total (%)		61,472 (100)	61,315 (99.25)	157 (0.26)
Age		Mean (SD)		
		59.02 (11.99)	59.01 (11.99)	60.73 (12.50)
		Median [Min, Max]		
		58 [20, 100]	58 [20, 100]	62 [31, 93]
Clinical Density Total (%)	A	4,164 (6.77)	4,154 (6.77)	10 (6.37)
	B	24,945 (40.58)	24,883 (40.58)	62 (39.49)
	C	29,120 (47.37)	29,039 (47.36)	81 (51.59)
	D	3,243 (5.28)	3,239 (5.28)	4 (2.55)
AI Density Total (%)	A	1,636 (2.66)	1,632 (2.66)	4 (2.55)
	B	30,578 (49.74)	30,516 (49.77)	62 (39.49)
	C	26,955 (43.85)	2,6870 (43.82)	85 (54.14)
	D	2,303 (3.75)	2,297 (3.75)	6 (5.28)

Results

A total of 61,472 women were used in this analysis. The dataset consisted of 463 total cancers of which 157 were advance stage. Comparing clinical density to DL derived density resulted in a Kappa-score of 0.53. **Table 1** provides detailed counts.
Models using clinical density and DL density resulted in AUC of 0.57 (95% CI: 0.49, 0.65) and 0.56 (95% CI: 0.50, 0.65), respectively while image only models resulted in an AUC of 0.63 (95% CI: 0.58, 0.69). Combining imaging predictions with clinical density resulted in an AUC of 0.69 (95% CI: 0.62, 0.73) and combining imaging predictions with DL density resulted in an AUC of 0.67 (95% CI: 0.62, 0.72) See **Table 2** and **Figure 3** for performance breakdown.

Table 2: Area Under the receiver operating characteristic Curve (AUC) for all 5 model configurations and odds ratios (OR) to the reference category (scattered/B density). Image feature ORs are per decile.

	Single Input Models		Combined Image and Density Models	
	OR (95% CI)	AUC (95% CI)	OR (95% CI)	AUC (95% CI)
Clinical Density		0.57 (0.49, 0.65)		0.69 (0.62, 0.73)
A	0.50 (0.17, 0.81)		0.48 (0.11, 0.77)	
B	1.00 (ref)		1.00 (ref)	
C & D	1.33 (0.71, 1.72)		1.63 (1.11, 2.01)	
Image Features	N/A		2.12 (1.31, 3.23)	
DL Density		0.56 (0.50, 0.65)		0.67 (0.62, 0.72)
A	0.46 (0.15, 0.84)		0.48 (0.15, 0.88)	
B	1.00 (ref)		1.00 (ref)	
C & D	1.23 (0.88, 1.47)		1.51 (0.90, 1.65)	
Image Features	N/A		2.05 (1.05, 3.19)	
Image Features	1.89 (1.02, 3.05)	0.63 (0.58, 0.69)	N/A	N/A

Conclusion

Imaging contains predictive information related to risk of advanced cancer, unique to breast density. Models which use this information have the potential to better identify high risk women which better enables early and effective intervention.

Limitations: This data export had a low number of cases and not all the standard risk factor variables were available. Also, there were a small number of density category D. Further validation is needed to confirm these findings.

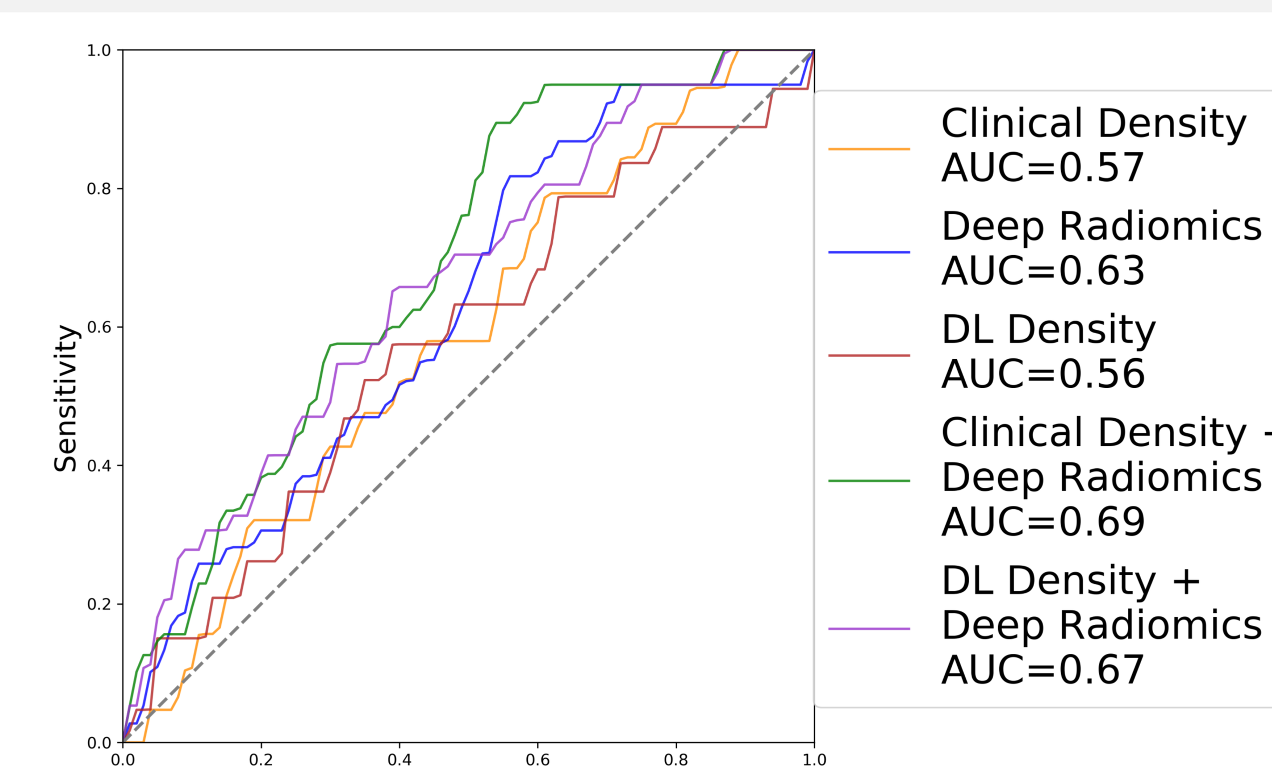

Figure 3: Receiver operating characteristic curves for all 5 model configurations. From top to bottom in the legend: Model 1, Model 3, Model 2, Model 4, Model 5.