Clinical Validation of an Artificial-Intelligent (AI) Enabled Digital Test Using the Patient's Diagnostic Breast Biopsy to Predict Invasive Breast Cancer Recurrence within 6 years.

Gerardo Fernandez, Jack Zeineh, Abishek Sainath Madduri, Richard Scott, Marcel Prastawa, Michael J. Donovan.

Background

Invasive Breast cancer (IBC) has surpassed lung cancer as the leading diagnosed cancer worldwide, representing some 15.5% of all cancer deaths. There remains an outstanding need to improve the current standard of care at diagnosis, including a reproducible and quantitative assessment of histologic grade and biological phenotype. We previously validated a digital laboratory-developed test to predict breast cancer recurrence using the surgical resection specimen. We now present the same approach for the diagnostic biopsy, providing a risk of recurrence earlier in the treatment planning and decision process.

Methods

1559 patients from 2004-2016 (Mount Sinai Health System, NY, NY, USA) with 6-year median follow-up divided 3:1, training (surgical cohort, 14% event rate) and validation (biopsy cohort, 13% event rate). H&E Whole slide Images (WSI), 40X magnification (Philips, Netherlands) were deconstructed with an AI-generated, precision medicine 'morphology feature array' (MFA) designed to extract tumor cell and tissue architectural features. Both cohorts were predominantly early-stage, ER/PR+ve, Her2-ve. Only age at diagnosis was utilized for biopsy model development. Recurrence events were classified as locoregional, distant metastasis and overall survival. C-index / AUC curves, Kaplan-Meier, hazards ratio, sensitivity, specificity, NPV, and PPV were used to assess risk discrimination.

Results

Surgical training model (n=1559), age (mean 60 years) (**Table 1**) combined with 7 imaging features representing an AI-(grade) yielded a C-index of 0.75 (95% CI, 0.73-0.77) vs. clinical (age) 0.62 (95% CI, 0.59-0.65) (**Table 2; Figure 1**). A risk score of 59.25 (scale 0-100) stratified patients as low- or high-risk, HR 4.9, P-value <0.001, with sensitivity 0.71, specificity 0.71, NPV 0.94, and PPV 0.27 (**Table 3, Figure 2**) for predicting BC recurrence within six years. In the diagnostic biopsy validation cohort (n=570, **Table 1)**, the model produced a C-index of 0.76 (95% CI, 0.72-0.80) vs. age only 0.65 (95%CI, 0.59-0.71) (**Table 2, Figure 3**). When patients were stratified by a risk score of 59.25, the HR was 4.9, P value <0.001), sensitivity 0.76, specificity 0.67, NPV 0.96, and PPV 0.22 for predicting BCR (**Table 4, Figure 4**). Examples of low and high-risk test patients are represented in **Figures 5 and 6, respectively.**

Conclusion

We developed and validated a breast biopsy AI-enabled digital platform that successfully predicted early-stage BC recurrence within six years using only the H&E-stained image and age at diagnosis. The test is designed to assist in characterizing clinical risk and the overall management of patients at the time of diagnosis. Additional studies are underway to refine the impact on treatment selection further.

	Train	Test
N	1559	570
Race/Ethnicity		
asian	9 (0.58%)	4 (0.7%)
black	81 (5.2%)	27 (4.74%)
latino	22 (1.41%)	8 (1.4%)
other	186 (11.93%)	86 (15.09%)
unknown	408 (26.17%)	124 (21.75%)
white	853 (54.71%)	321 (56.32%)
Median age (range)	60.0 [24, 90]	60.0 [28 90]
ER*		
0	204 (13.09%)	29 (5.09%)
1	1355 (86.91%)	220 (38.6%)
PR*		
0	292 (18.73%)	35 (6.14%)
1	1267 (81.27%)	213 (37.37%)
HER2*		
0	1362 (87.36%)	230 (40.35%)
1	197 (12.64%)	18 (3.16%)
Tumor size (cm)	1.54±1.12 [0.1, 17.0]	N/A
Stage		
Stage1	1055 (67.67%)	N/A
Stage 2	386 (24.76%)	N/A
Stage IIIA/B	81 (5.2%)	N/A
Stage IIC	36 (2.31%)	N/A
Stage IV	1 (0.06%)	N/A
LN		
posLN=0	1075 (68.95%)	N/A
microLN or isolatedLN and posLN=0	127 (8.15%)	N/A
1<=posLN<=3	239 (15.33%)	N/A
posLN>3	118 (7.57%)	N/A
grade		
1	290 (18.6%)	89 (16%)
2	649 (41.63%)	248 (44%)
3	620 (39.77%)	233 (41%)
total events		
0	1339 (85.89%)	494 (86.67%)
1	220 (14.11%)	76 (13.33%)
time to event (months)	75.28 [-16.0, 68.0, 200.0]	81.26 [0.0, 75.0, 194.0]

Table 1. Demographics of the PreciseDx Breast Biopsy Train and Test Cohorts. Only ~44% of biopsy cases had ER/PR or HER2 status available. However, they are proportionately balanced with the training cohort.

Training C-Index on surgical cohort (1559)	**0.753654**
Training Sensitivity/Specificity Threshold:	**59.25**
Train Sensitivity:	**0.715328**
Train Specificity:	**0.713661**
Test CI on biopsy cohort (570)	**0.760138**
Test Sensitivity:	**0.76**
Test Specificity:	**0.670025**
Feature	**Weight in Final Model**
Proliferative Activity	**-23.5534**
Gland Architecture_1	**-3.33142**
Nuclear Pleomorphism	**-27.7339**
Age at Diagnosis	**-34.7084**
Tumor Infiltrating Lymphocytes	**20.8371**
Gland Architecture_2	**7.7395**
Gland Architecture_3	**3.3563**
Gland Architecture_4	**55.6735**

Table 2. PreciseDx Breast Biopsy C-Index (CI) Train and Test models with AI-imaging features combined with age at diagnosis.

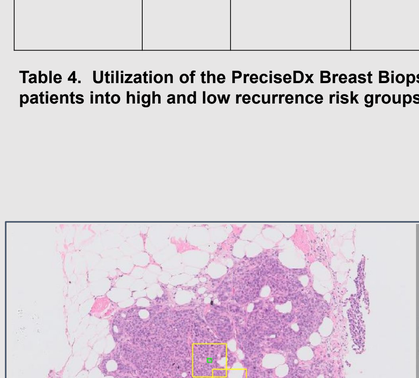

Figure 3. AUC-ROC curve of Biopsy Validation Model (n=570, blue line) utilizing biopsy specimens and age at diagnosis: CI 0.76 (95% 0.72-0.80) vs. Clinical only model (orange line) using only age at diagnosis: CI 0.65 (95%CI 0.59-0.71).

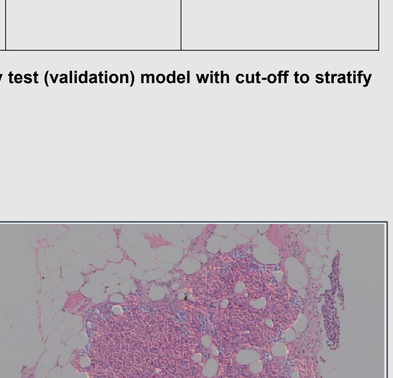

Figure 4. Kaplan-Meier Breast Biopsy Validation model vs clinical only age model applying a <59.25 or >/= 59.25 cut-off as low or high risk for breast cancer recurrence. HR 4.9 (95%CI, 2.9-8.06, p<0.001).

Test	Events	Censored	Total	Performance	
Risk Score ≥ 59.25	55	158	213	**Sensitivity**	0.76
Risk Score < 59.25	21	336	357	**Specificity**	0.67
				PPV	0.224852
				NPV	0.956835
				CI	0.7605 (0 .72253, 0.80355)

Table 4. Utilization of the PreciseDx Breast Biopsy test (validation) model with cut-off to stratify patients into high and low recurrence risk groups.

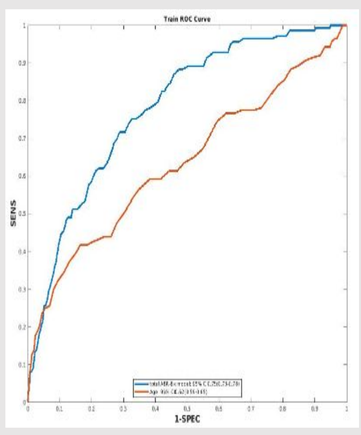

Figure 1. AUC - ROC curve of Training Model (n=1559, blue line) utilizing excision specimens and age at diagnosis: CI 0.75 (95% 0.73-0.77), HR4.9, p<0.001 vs. Clinical only model (orange line) using only age at diagnosis: CI 0.62 (95%CI 0.59-0.65), HR 2.27 p<0.001).

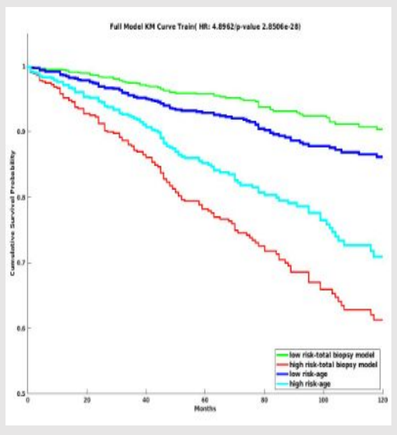

Figure 2. Kaplan-Meier of training model performance with a 59.25 cut-off distributing patients as high or low risk for breast cancer recurrence vs. age at diagnosis clinical model. HR= 4.9 (95%CI, CI,3.7-6.5, p<0.001).

Train	Events	Censored	Total	Performance	
Risk Score ≥ 59.25	148	370	518	**Sensitivity**	0.7153
Risk Score < 59.25	72	969	1041	**Specificity**	0.7137
				PPV	0.2722
				NPV	0.9436
				CI 95%	0.7536 (0.7292, 0.7751)

Table 3. PreciseDx Breast Biopsy training model with a cut-off to stratify patients into high and low-recurrence risk groups.

Figure 5 Is a 44 y/o low risk woman(PreciseDx Breast Risk Score 32), clinically assigned Bloom-Richardson grade 2. Left panel shows the patient's H&E-stained biopsy with the mitotic figure detector (none detected). The right panel shows the same region of H&E-stained biopsy tissue with a gland morphology overlay with orange representing high grade gland morphology and blue representing lower grade morphology.

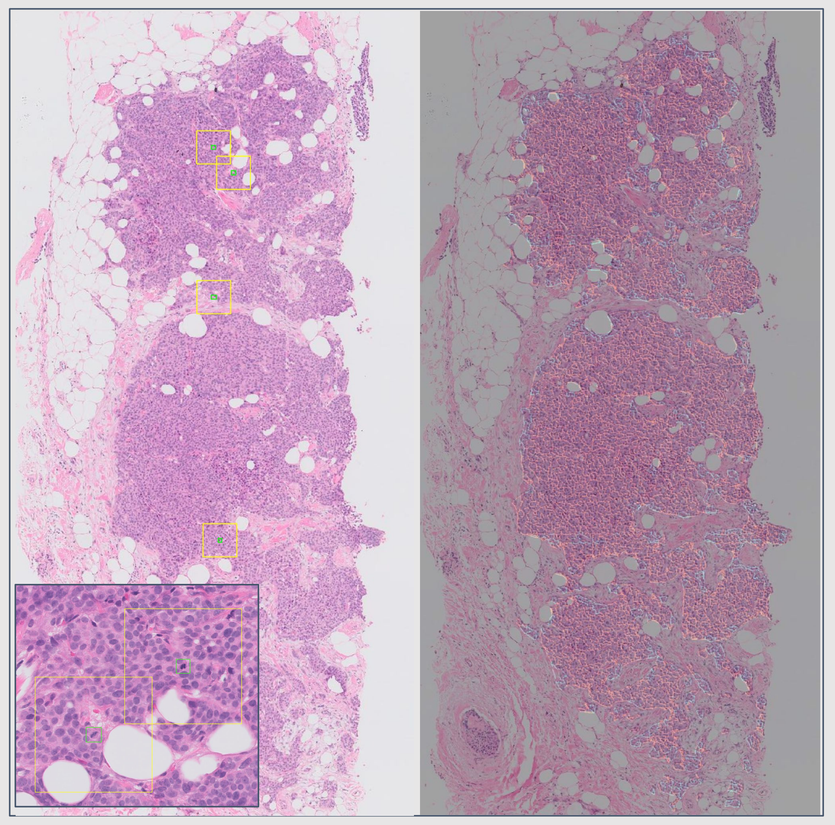

Figure 6. Is an 82 y/o high risk woman(PreciseDx Breast Risk Score 74), clinically assigned Bloom-Richardson grade 2. Left panel shows the patient's H&E-stained biopsy with the mitotic figure detector overlay (yellow box with green box insert) and a higher magnification insert of detected mitotic figures. The right panel shows the same region of H&E-stained biopsy tissue and a gland morphology overlay with orange representing high grade gland morphology and blue representing low grade morphology.

References

1. Role of the Surgical Pathologist in the Diagnosis and Management of the Cancer Patient—Holland-Frei Cancer Medicine—NCBI Bookshelf. Available from: https://www.ncbi.nlm.nih.gov/books/NBK13237/.
2. Rakha EA, Reis-Filho JS, Baehner F, Dabbs DJ, Decker T, Eusebi V, et al. Breast cancer prognostic classification in the molecular era: the role of histological grade. Breast Cancer Res. 2010;12:1–12.
3. Rakha EA, Aleskandarany MA, Toss MS, Mongan NP, ElSayed ME, Green AR, et al. Impact of breast cancer grade discordance on prediction of outcome. Histopathology. 2018;73:904–15.
4. Jiang Y, Yang M, Wang S, Li X, Sun Y. Emerging role of deep learning-based artificial intelligence in tumor pathology. Cancer Commun. 2020;40:154–66.