Deep Learning Model Translates Imaging Biomarkers to Predict Future Breast Cancer Risk, Surpassing Traditional Methods of Risk Assessment

[1]Lamb L MD MSc, [2]Yala A MEng, [2]Mikael P BS,
[2*]Barzilay R PhD, [1*]Lehman CD MD PhD

[1]Breast Imaging Division
Department of Radiology
Massachusetts General Hospital
[2]Massachusetts Institute of Technology

*Joint senior authors

Disclosures

Leslie Lamb, MGH

- No disclosures

Adam Yala, MIT

- No disclosures

Peter Mikael, MIT

- No disclosures

Regina Barzilay, MIT

- Susan G Komen, Breast Cancer Research Foundation, Quanta Computing, MIT J-Clinic

Connie Lehman, MGH

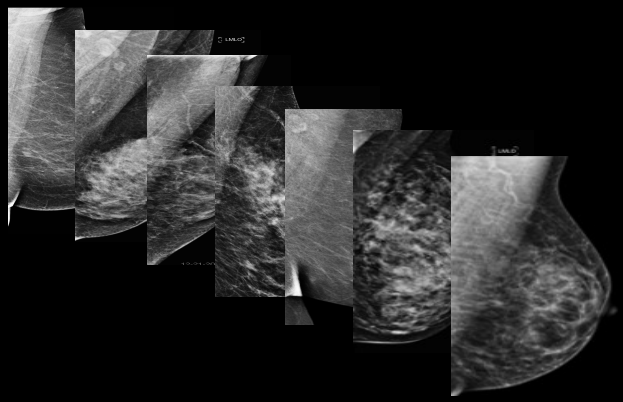
- Breast Cancer Research Foundation, Institutional GE Research Support

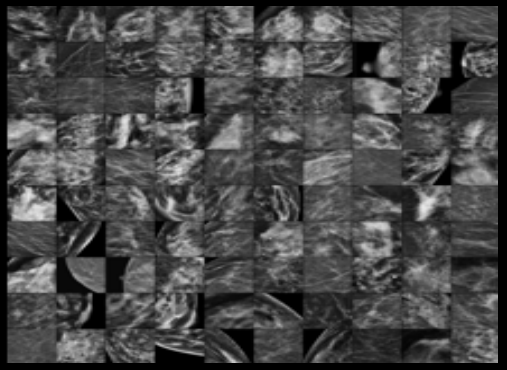
Background

- Risk-based screening has been elusive due to inability to accurately predict risk
  - Current screening guidelines leverage risk models
- Risk assessment models include Tyrer-Cuzick (TC)/IBIS, Gail/NCI BCRAT, Claus, BOADICEA, BRCAPRO, BCSC
  - AUCs[1-3]: TC 0.62 & Gail 0.59
- Current risk models incorporate a small fraction of available patient data (ex. family history, prior biopsies, hormonal/reproductive history, breast density)

Background

- With the power of artificial intelligence, why limit ourselves?





- Previous deep learning (DL) model (DL1) was developed that combined imaging and traditional risk factors[4]
  - AUCs: 0.68 image-only model, 0.70 for image + risk factor (RF) model vs 0.51 TC version 7 (TC7), 0.62 TC version 8 (TC8)
- A new DL algorithm (DL2) was designed to predict a patient's risk of developing breast cancer at multiple time points using mammographic image biomarkers alone

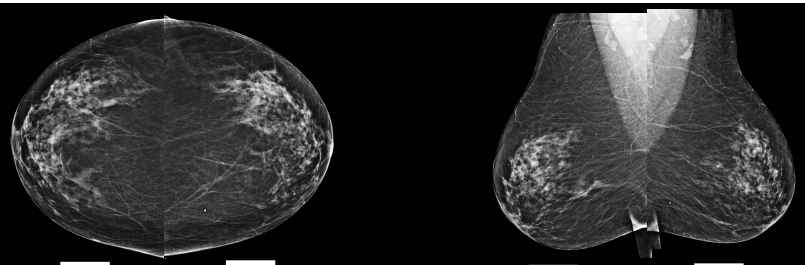


Image Encoder

1. Pass each image through Image Encoder

Image Aggregator

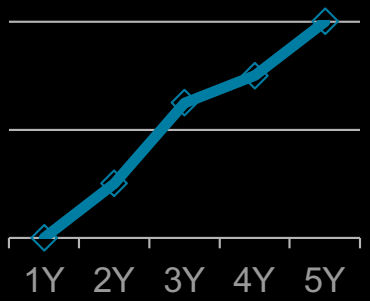
2. Combine information across views

3. Predict RFs

Risk Factor Predictor

Additive Hazard Layer

4. Predict patient's risk



Objective

- To compare the predictive accuracy of the DL2 (image-only) model to that of TC8, a traditional risk model that incorporates patient history and breast density to predict future breast cancer risk

Materials and Methods

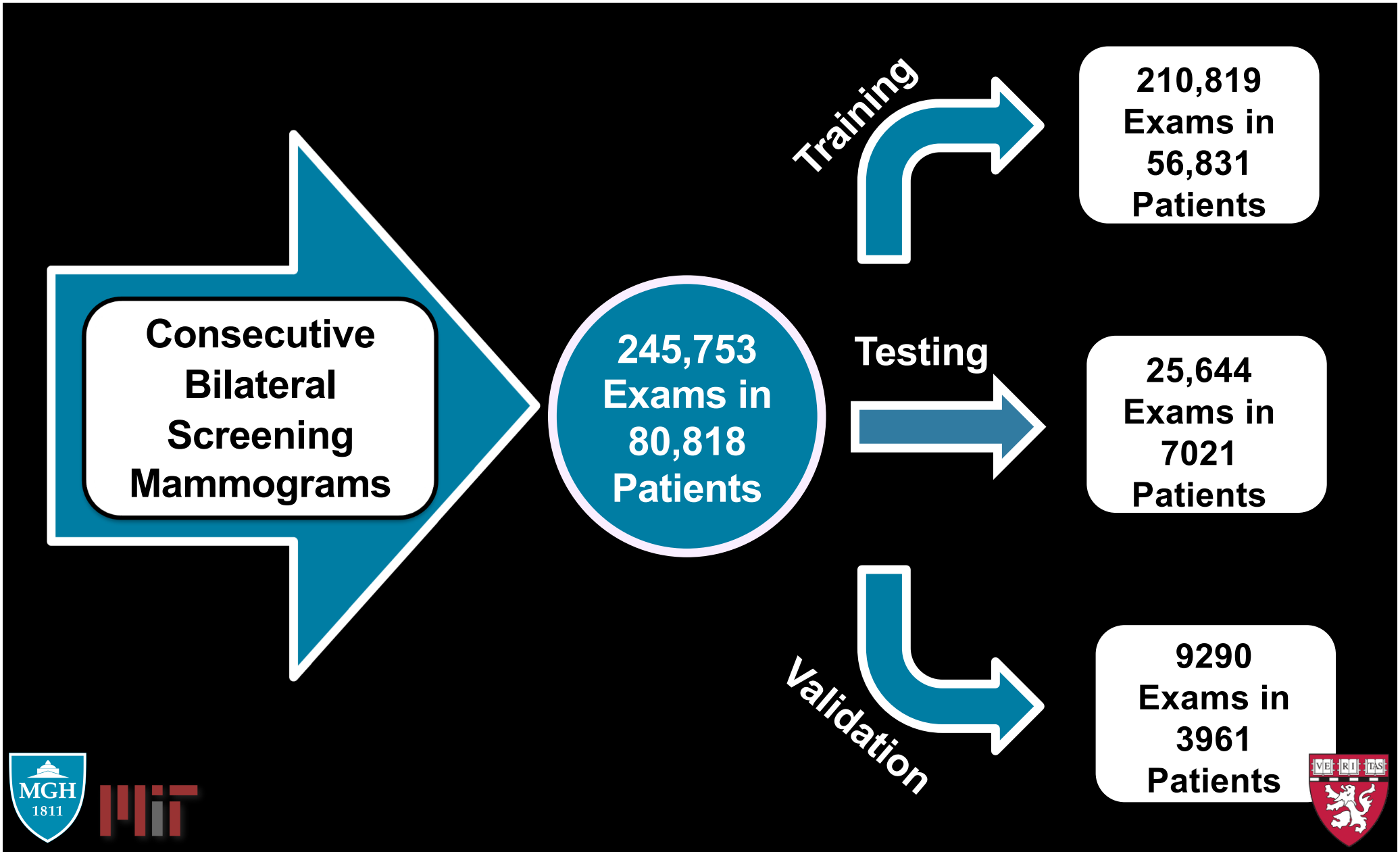
Study Design

- Institutional Review Board approved
- HIPAA compliant
- Retrospective
- Single tertiary academic institution with 5 screening sites
- Cancer outcomes linked to regional tumor registry

Study Population

- **Inclusions:**
  - 2009 to 2016
  - Consecutive patients with bilateral screening mammography
  - Personal history of breast cancer (except test set), implants, all races, all biopsy results
- **Exclusions:**
  - Unilateral
  - Non-Hologic mammography unit
  - Personal history of breast cancer (test set)
  - Other cancers in the breast (ex. sarcoma)
  - Lacked 1y screening follow-up
  - Developed breast cancer within 3m of screening

Model Development



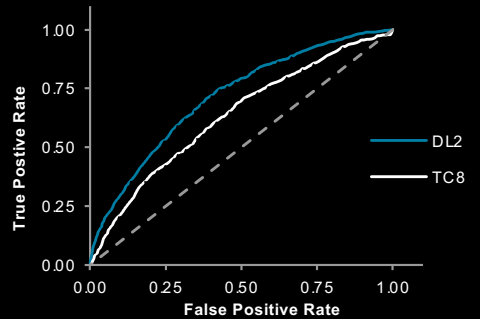
Materials and Methods

Primary Outcome

- Development of breast cancer within 5 years of index mammogram

Statistical Analysis

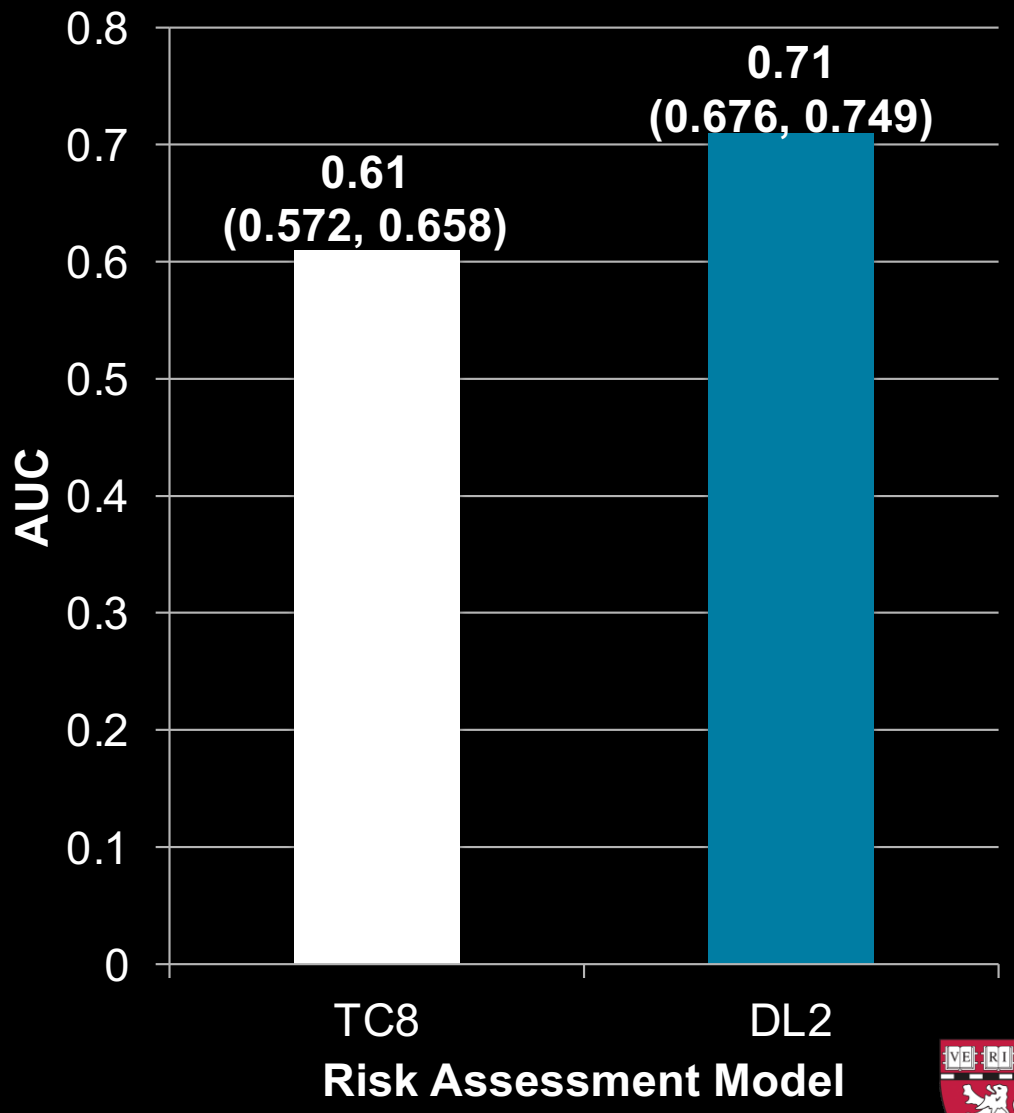
- DL2 vs TC8 model performance was compared using areas under the receiver operating characteristic (ROC) curve (AUCs) with DeLong test
- *p*-values <0.05 considered significant



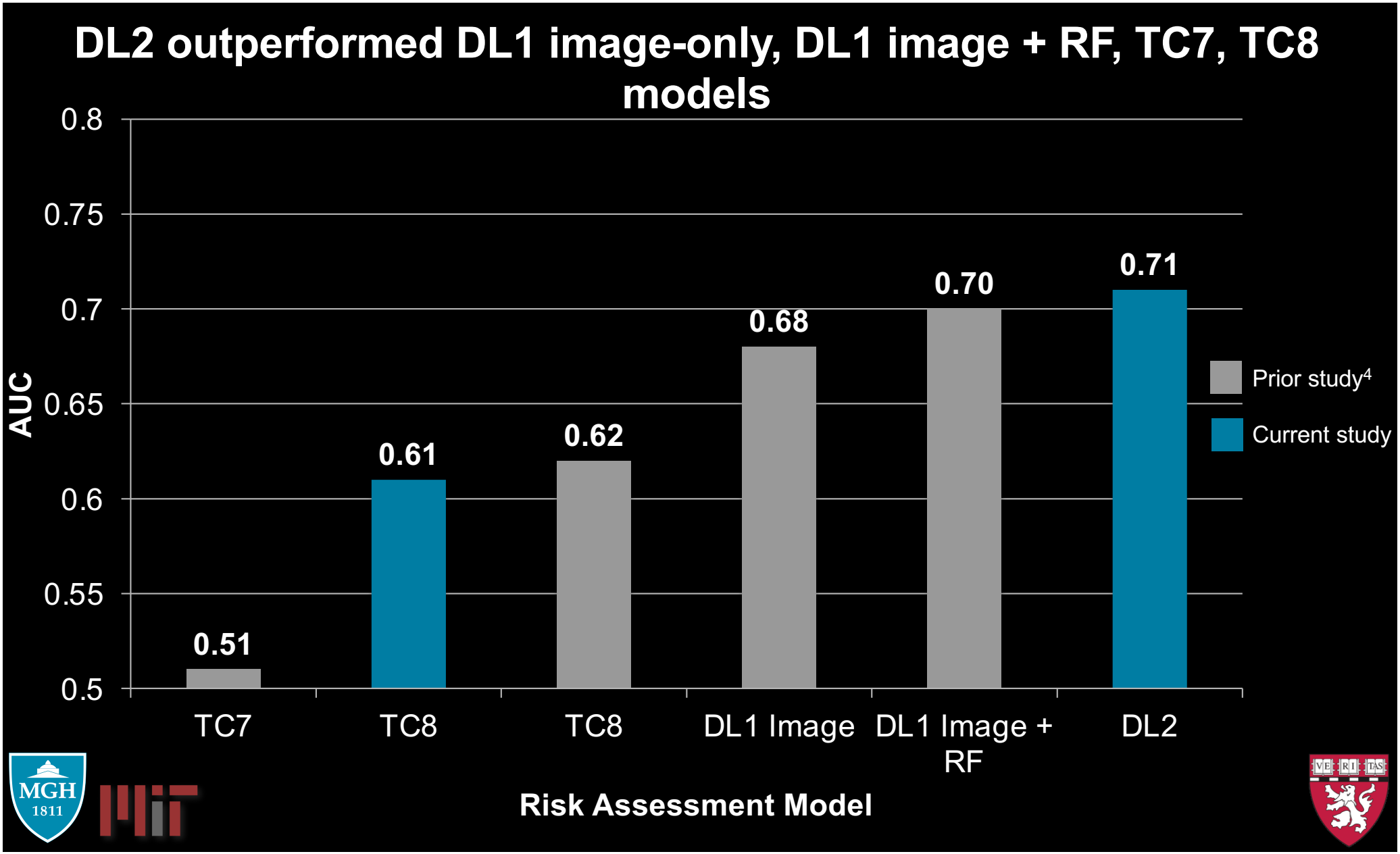
Results

Demographics

- Age: 56.4y (range: 35-91y)
- Pre-menopausal: 29.5% vs post-menopausal: 70.5% ($p<0.001$)
- Non-dense: 55.7% vs dense: 44.3% ($p<0.001$)
- White: 81.4%, African American: 4.8%, Asian/Pacific Islander: 4.8%, Other: 9.1% ($p<0.001$)



Discussion


DL2 outperformed DL1 image-only, DL1 image + RF, TC7, TC8 models
AUC
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.51
0.61
0.62
0.68
0.70
0.71
TC7
TC8
TC8
DL1 Image
DL1 Image + RF
DL2
Prior study[4]
Current study
Risk Assessment Model

Strengths

- Large cohort
- DL model is inclusive – consecutive screening patients (personal history of breast cancer, all races, implants, prior biopsies)
- Feasibility – time and staffing resources

Limitations

- Predominantly white population (81.4%)
  - Small sub-group numbers limits analysis
- Five screening sites, part of single academic institution

Future Directions

- Validation:
  - External validation has been since performed at Karolinska Institute, Sweden and Chang Gung Memorial Hospital, Taiwan
  - Further validation in larger African American and minority subgroups required
- Further model development
- Clinical implementation

ML Density: B - Scattered fibroglandular densities
Prior Breast Density: B - Scattered fibroglandular densities
Accession #: E16505034

AI1: 0.179	AI2: 0.480	AI5: 1.493
TC8 Life: 6.5	TC8 10y: 2.70	
NCI Life: -2.00	NCI 5y: -2.00	
BRCA Life: 9.40		

Conclusions

- Mammograms contain highly predictive biomarkers of future cancer risk, not identified by traditional risk models
- A DL model using screening mammography alone can improve risk discriminatory accuracy compared to traditional modern risk models which rely on clinical history and mammographic breast density

Traditional risk models can be time-consuming to acquire and rely on inconsistent or missing data. A DL image-only risk model can provide increased access to more accurate, less costly risk assessment

References

1. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Statistics in Medicine* 2004;23(7):1111-1130
2. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *JNCI: Journal of the National Cancer Institute* 1989;81(24):1879-1886
3. Brentnall AR, Harkness EF, Astley SM, et al. Mammographic density adds accuracy to both the Tyrer-Cuzick and Gail breast cancer risk models in a prospective UK screening cohort. *Breast Cancer Research* 2015;17(1): 147
4. Yala A, Lehman C, Schuster T, Portnoi T, Barzilay R. A deep learning mammography-based model for improved breast cancer risk prediction. *Radiology* 2019;292(1):60-66

Deep Learning Model Translates Imaging Biomarkers to Predict Future Breast Cancer Risk, Surpassing Traditional Methods of Risk Assessment

THANK YOU!

Questions/Comments?

LRLAMB@mgh.harvard.edu