DANA-FARBER/BRIGHAM AND WOMEN'S CANCER CENTER

PathAI

BRIGHAM HEALTH
BRIGHAM AND WOMEN'S Program in Precision Medicine

HARVARD MEDICAL SCHOOL

Artificial Intelligence powered predictive analysis of atypical ductal hyperplasia from digitized pathology images

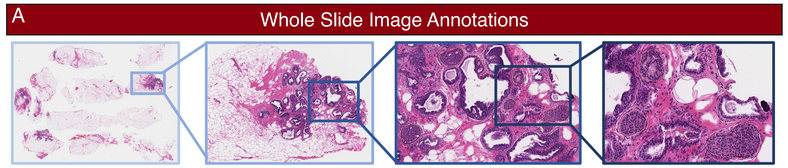
J. K. Kerner[1], A. Cleary[2], S. Jain[1], H. Pokkalla[1], B. Glass[1], S. Grossmith[3,4], M. Harary[4], E. Mittendorf[3,4], A. H. Beck[1], A. Khosla[1], S. J. Schnitt[2,3], I. Wapinski[1], T.A. King[3,4]

[1]PathAI, Boston, MA, [2] Department of Pathology, Brigham and Women's Hospital, 3 Breast Oncology Program, Dana Farber / Brigham and Women's Cancer Center [4] Division of Breast Surgery, Brigham and Women's Hospital, Boston, MA

BACKGROUND

- Approximately 15-25% of patients with atypical ductal hyperplasia (ADH) on breast core needle biopsy (CNB) are upgraded to ductal carcinoma in situ (DCIS) or invasive carcinoma (IC) on surgical excision.
- We hypothesized that a machine learning approach could be utilized to train models to recognize ADH on digitized pathology images and to identify cases of ADH more likely to be upgraded to DCIS or IC at excision.
- Here we demonstrate the accuracy of the machine learning approach to identify ADH.

OVERVIEW



B Region Annotation Counts

Cribriform ADH	2198	Lobular Neoplasia	746	Apocrine metaplasia	438
Micropapillary ADH	178	DCIS	27	Microcyst	363
Solid ADH	176	Normal TDLU	2154	Artifact	167
Pagetoid ADH	35	Adipose	1947	Lumen	127
ADH Stroma	363	Non-atypical breast tissue	909	Sclerosing adenosis	86
FEA	1476	Columnar cell change	816	Papilloma	12
Calcifications	1186	Usual ductal hyperplasia	641	Background	44

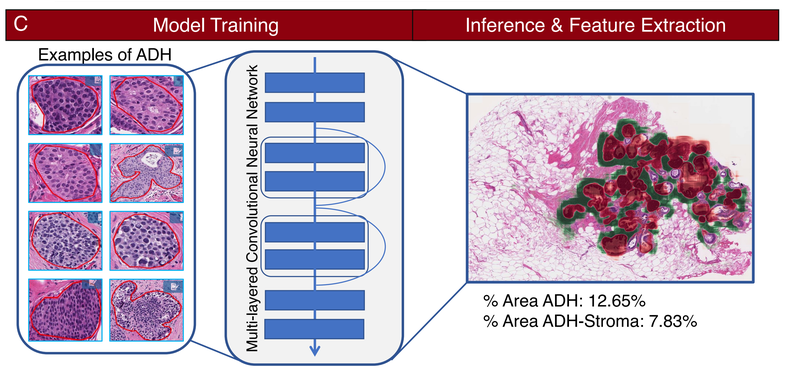


Figure 1. Overview of whole slide image (WSI) annotations and model training regimen. WSIs are annotated by breast pathologists for each region of interest on the PathAI research platform (counts shown in the table). Examples of each substance are presented to multi-layer CNN, which learn to recognize morphological patterns within images. Once trained, models are deployed on test set images and features are summarized to characterize each slide.

METHODS

- From 11/2004-3/2018 we identified 286 cases with a core bx diagnosis of ADH; from these cases we retrieved 673 digitized whole slide images (WSI).
- Clinical risk factors such as patient race, tobacco smoking, pregnancy history, age at menarche, breast composition, family history, etc. were also documented for each patient as metadata.
- Independent histologic review by two breast pathologists identified slides with and without ADH.
- 39 board certified pathologists with experience in evaluation of breast biopsies were employed for tissue region annotation on the PathAI research platform, yielding 14,116 tissue region annotations (Figure 1, panel B).
- Annotations were used to train a convolutional neural network (CNN) with 35 layers and approximately 9 million parameters to identify ADH (Figure 1).
- Data were split into training and testing sets, representing 61% and 39% of the data respectively. The distribution of ADH slides and number of ADH upgrade cases was balanced between the two.
- Morphological tissue features were derived from the output of the CNN model, which were then used to train a logistic regression classifier to predict slide level ADH label.

RESULTS

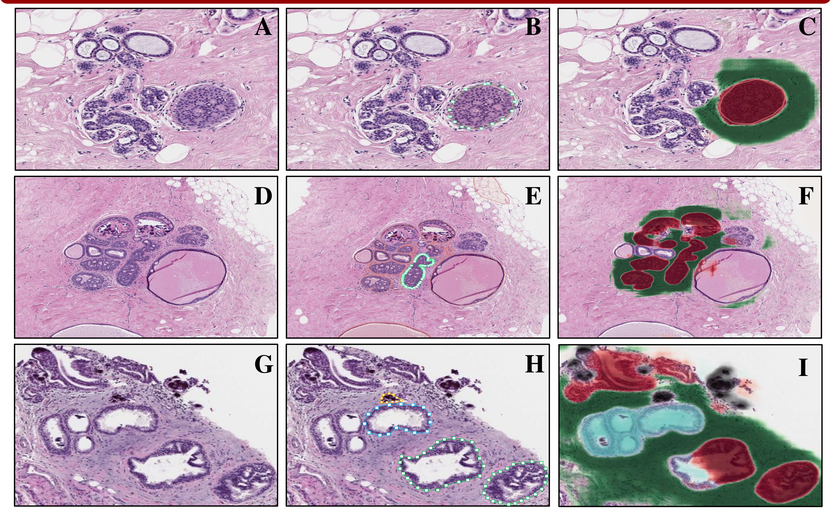
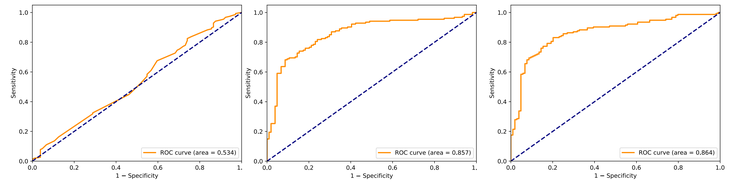


Figure 2. A, D, G. Sample images without annotations or overlays. **B, E, H.** Sample images with annotations from contributing pathologists. Annotations in **[B]** include ADH (green annotations with white boxes), in **[E]** include ADH (green annotations with white boxes), stroma adjacent to ADH, FEA, microcyst, calcification, and adipose tissue (orange annotations) and in **[H]** include ADH (green annotations with white boxes), FEA (blue annotations with white boxes) and calcifications (orange annotations with white boxes). **C, F, I.** Sample images with overlays showing ADH predictions (red), stroma adjacent to ADH (green), FEA (cyan) and calcifications (black). Areas without an overlay are predicted to be normal. Focal false positives ADH predictions are seen in **[F]** and FEA false positives are seen in **[I]**.

RESULTS



(A) Metadata features (B) Tissue features (C) Metadata and Tissue features

Figure 3. ROC and AUC for ADH detection with metadata (A), tissue features (B) and combined features (C). Model prediction compared to ground truth of histologic examination by pathologists (slide-level ADH label).

- AI model trained with only patient metadata (clinical risk factors) was not able to predict ADH diagnosis, resulting in a slide-level area under the receiver operator curve (AUC) of 0.534 (Figure 3A).
- AI model trained with morphological tissue features extracted automatically from the slide was able to predict the diagnosis of ADH with an AUC of 0.857 (Figure 3B).
- AI model trained with a combination of metadata (patient clinical risk factors and tissue features) improves the performance and was able to predict the diagnosis of ADH with 79% sensitivity (122 of 154 images within the test set) and 82% specificity (85 of 104 images within the test set). The slide-level AUC for this model was 0.864 (Figure 3C).
- Area of ADH and ADH Stroma were the most important morphological features for slide level ADH diagnosis.

CONCLUSIONS

- A deep learning-based classifier was developed for the identification of ADH. Morphological features extracted from this system showed strong performance for predicting slide level ADH labels.
- Future analyses will focus on determining if morphologic features of ADH extracted by the deep learning system can be used to predict upgrade to DCIS and IC. This approach may help stratify patients with ADH on CNB into those who require surgical excision and those who can be followed with active surveillance.

REFERENCES

Linsk A, Mehta TS, Dialani V, Brook A, Chadashvili T, Houlihan MJ, Sharma R (2018). Surgical upgrade rate of breast atypia to malignancy: An academic center's experience and validation of a predictive model. *Breast J.*, 24(2):115-119. doi: 10.1111/tbj.12885

Mesurolle B, Perez JC, Azzumea F, et al. Atypical ductal hyperplasia diagnosed at sonographically guided core needle biopsy: frequency, final surgical outcome, and factors associated with underestimation. *Am J Roentgenol*. 2014; **202**: 1389- 1394.

Mooney KL, Bassett LW, Apple SK (2016). Upgrade rates of high-risk breast lesions diagnosed on core needle biopsy: a single-institution experience and literature review. *Mod Pathol*, 29(12):1471-1484. doi: 10.1038/modpathol.2016.127

Weinfurtner RJ, Patel B, Laronga C, Lee MC, Falcon SL, Mooney BP, Yue B, Drukteinis JS (2015). Magnetic resonance imaging-guided core needle breast biopsies resulting in high-risk histopathologic findings: upstage frequency and lesion characteristics. *Clin Breast Cancer*, 15(3):234-9. doi: 10.1016/j.clbc.2014.12.005