

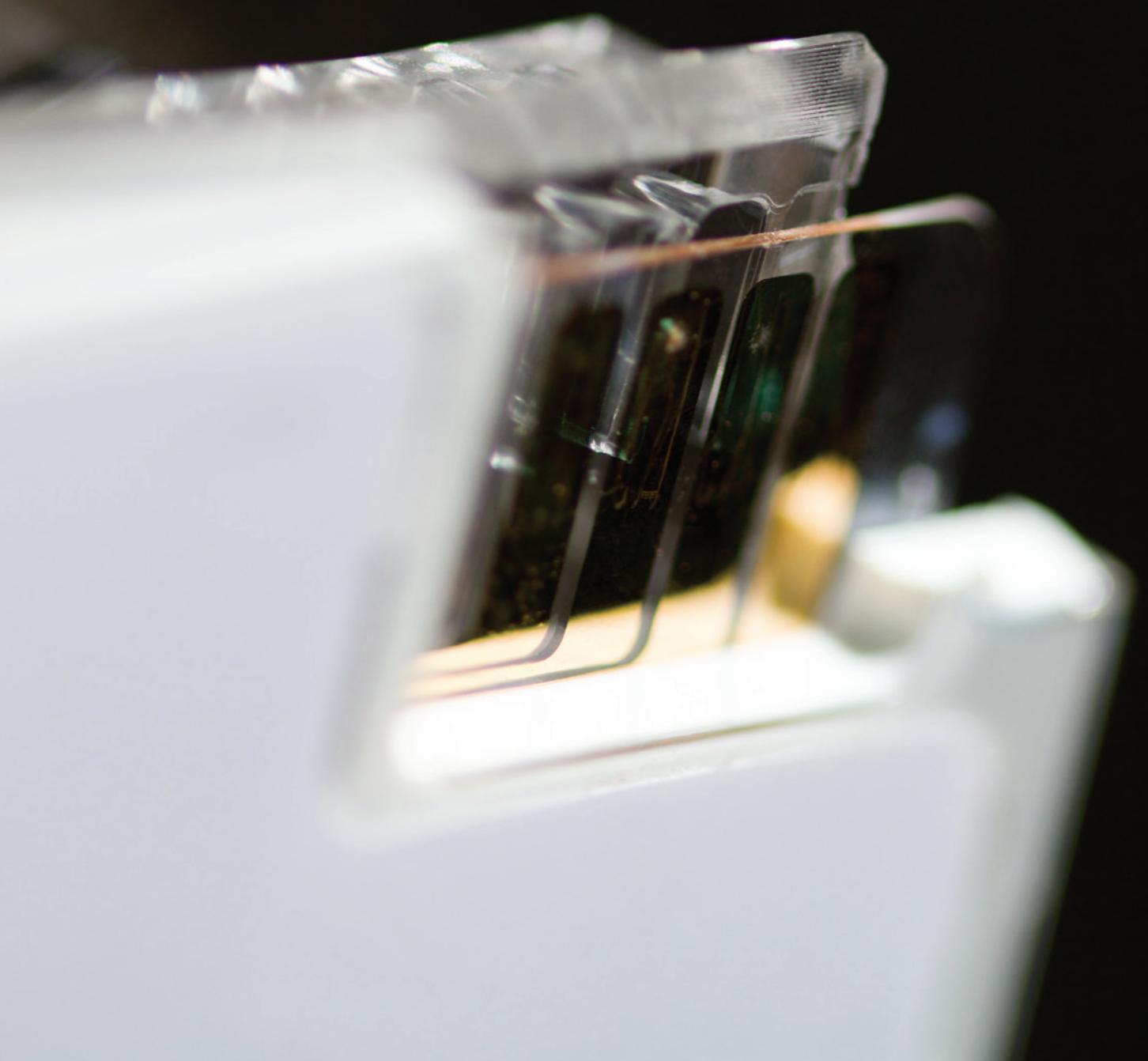


The logo consists of a dark blue, stylized, rounded font spelling "IDG" above the word "PROGRAM". A thin, dark blue curved line arches over the top of the letters. Below the main text is the year "2019" in a smaller, light blue serif font.

IDG
PROGRAM

WORLD MEDICAL
INNOVATION
FORUM™





The Innovation Discovery Grants (IDG) program aims to enhance the commercial outcomes of the Partners HealthCare community and increase its innovative potential. It is designed to stimulate new inventive concepts, identify areas of commercially significant scientific strength and accelerate commercialization of PHS intellectual assets. In 2018, the IDG program funded projects that focused on the advancement of opportunities where artificial intelligence can be used to improve clinical care.



2019
artificial intelligence
investigators

Innovation Discovery
Grant Awardee Presentations

Eleven clinical AI teams culled through the Innovation Discovery Grant program present their work illustrating how AI can be used to improve patient health and healthcare delivery. This session is designed for investors, entrepreneurs, investigators, and others who are interested in commercializing AI opportunities that are currently in development with support from the Innovation Office.

- 6** Peter Dunn, MD
- 7** Kevin Elias, MD
- 8** Alexandra Golby, MD
- 9** Jayashree Kalpathy-Cramer, PhD
- 10** Jochen Lennerz, MD, PhD
- 11** Bruno Madore, PhD
- 12** Jinsong Ouyang, PhD
- 13** David Papke, MD, PhD
- 14** Martin Teicher, MD, PhD
- 15** Christian Webb, PhD
- 16** Brandon Westover, MD, PhD

Note: Speakers and content are subject to change.

BH Brigham and Women's Hospital | **HMS** Harvard Medical School | **MGH** Massachusetts General Hospital

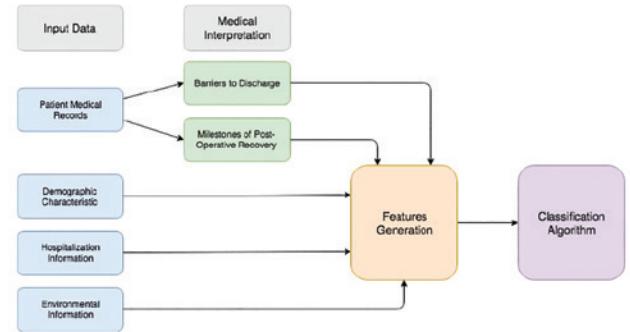


Figure 1. Discharge Prediction Model Building Blocks

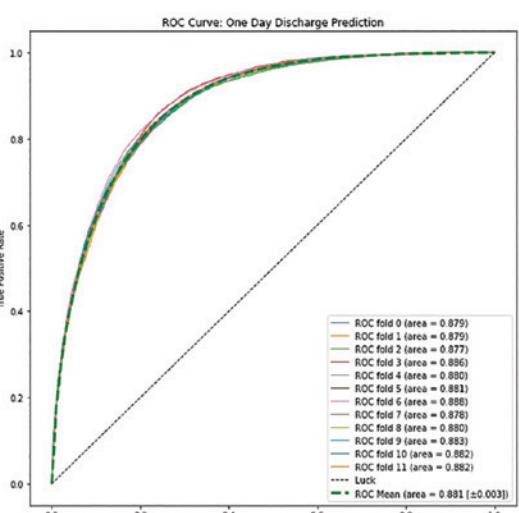


Figure 2. ROC Curve for 80%-20% Random Training/Testing Sets

Using Deep Learning to Optimize Hospital Capacity Management

Peter Dunn, MD

Vice President, Perioperative Services and Healthcare System Engineering, MGH; Assistant Professor, Anesthesia, HMS
pdunn@partners.org

MGH and many US hospitals operate near maximum bed capacity. Efficient capacity management is critical to their missions. They face the daily task of matching high bed demand with unpredictable bed supply, with little predictability or transparency around discharges. Currently, hospitals ask clinical teams to identify who will be discharged each day and single out candidates for early discharge. This task is time-consuming, manual, and complex. The consequences are: (1) lack of ability to proactively manage patient flow, which usually results in suboptimal bed allocation decisions; (2) late discharges both inter-day and intra-day, and (3) upstream congestion in the Emergency Department, Intensive Care Units, and Operating Rooms.

MGH Healthcare Systems Engineering, in collaboration with faculty and students at MIT, has created a Discharge Prediction Tool that each morning predicts which surgical patients will leave the hospital within 24 hours. More specifically, the Tool: (1) ranks patients in order of their likelihood of discharge that day, (2) predicts the total number of patients that will be discharged, and (3) provides a comprehensive list of discharge barriers for each patient. The prediction model is based on a neural

network with 2 layers and 20 nodes. The model tracks patients' clinical milestones and barriers to discharge based on over 900 input variables extracted from the hospital's electronic medical record. It has been trained on over 20,000 patients (AUC of 0.88) and is being piloted at MGH.

The Tool enables a timely, automatic, and transparent process to identify patients who can be discharged that day that is available to all stakeholders. This gives clinical teams the necessary time to work through potential barriers, prepare patients for early discharge and, more generally, they can be freed up to focus on patient care. Reducing unnecessary hospital days will free up beds and increase access to hospital services. Finally, the Tool systematically collects data on obstacles to discharge to target specific issues and improve efficiency.



Screening for Cancer Using Serum miRNA Neural Networks

Kevin Elias, MD

Director, Gynecologic Oncology Research Laboratory, BH
Assistant Professor, HMS
kelias@bwh.harvard.edu

Diagnosis of cancers at early stage is associated with improved overall survival, lower costs, and reduced morbidity for patients. However, the current state of cancer screening remains highly fragmented and incomplete. Patients need to visit multiple providers and undergo multiple procedures. Moreover, many common and highly lethal cancers, such as ovarian and pancreatic cancer, do not have approved screening tests. Consolidating cancer screening for several different types of cancer into a single assay would save costs, improve compliance, and increase the number of cancers amenable to early diagnosis.

We have developed a strategy for screening multiple cancers simultaneously using serum microRNAs (miRNAs). Our technology combines advances in detection of rare circulating miRNA transcripts with a neural network machine learning approach. In our original paper, we showed that a model incorporating serum levels of 14 miRNAs had a high degree of accuracy (AUC 0.93, 95%CI 0.81-1.00) for distinguishing cases of invasive ovarian cancer from non-invasive lesions, benign tumors, or healthy controls. When applied to an independent dataset of 454 patients, the model had 100% specificity for identifying ovarian cancers versus 12 competing diagnoses. We have shown that the miRNAs are derived from the tumors themselves and can identify even microscopic lesions. We have since validated this approach in mouse and non-human primate models and applied the same technique to breast, colon, pancreatic, and prostate cancers. miRNA profiles are tissue-specific and can distinguish one cancer type from another. Our laboratory has built a high throughput pipeline for serum miRNA analysis and a cloud computing platform to streamline the machine learning analysis. We can now screen for up to 10 cancers simultaneously among 40 individual patients in less than 4 hours. Data can also be batch uploaded from reference laboratories and the software will return individualized risk reports for each patient. Our goal is to expand this into an analytics service which can be licensed to clinical laboratories.

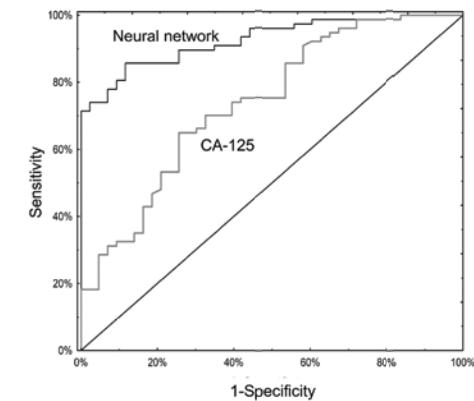


Figure 1. Receiver operating characteristic curve comparing the miRNA neural network to CA125 for diagnosing ovarian cancer.

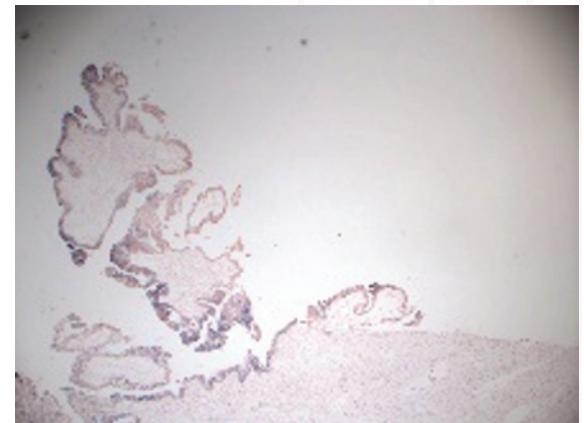


Figure 2. In situ hybridization showing miRNA expression in a pre-invasive ovarian cancer lesion.



Figure 3. Example of miRNA-based risk report for ovarian cancer generated by the cloud computing platform.

Using Machine Learning to Optimize Optical Image Guidance for Brain Tumor Surgery



Pablo Valdes, MD, PhD

Neurosurgery Resident, BH

pvaldesquevedo@bwh.harvard.edu

Alexandra Golby, MD (pictured)

Director, Image-Guided Neurosurgery, BH

Professor, Neurosurgery Radiology, HMS

agolby@bwh.harvard.edu

Extent of resection is the most significant modifiable predictor of quality of life and survival in patients with intracranial tumors including low- and high-grade gliomas, meningiomas, and metastases. Current technologies are limited due to interruption to the surgical workflow, high costs, limited tumor-tissue contrast, lack of immediate feedback, and non-quantitative assessments of tumor presence. To address the limitations of current intraoperative technologies, we proposed to develop a novel intraoperative technology which harnesses the power of optical imaging to quantitatively measure multiple tumor biomarkers at the tissue level including fluorescent tumor markers like protoporphyrin IX (PpIX), oxy- and deoxy-hemoglobin and scattering of tissue across the full surgical field of view. We call this novel system a quantitative optical imaging (qOI) system for multiplexed biomarker quantification. We are developing technology that integrates with modern surgical microscopes to collect fluorescence and non-fluorescence optical data from the surgical field to ultimately quantify multiple biomarkers. These biomarkers can be used in machine learning algorithms (e.g., support vector machines, neural networks) to help improve delineation of tumor tissue. Development of the technology involves both pre-clinical testing and validation of the qOI system in phantoms; and a clinical implementation for acquisition of clinical data. As a multi-stage technology development proposal, we are optimizing machine learning algorithms in a clinical intraoperative spectroscopy dataset of >50 patients that includes multiple tumor types including low- and high-grade gliomas, meningiomas, and metastases. In parallel, we have developed the hardware and software for system control and data acquisition of the qOI system. We are currently in the process of pre-clinical testing of our system with quantification accuracies >95%. We expect to transition to first clinical use of the qOI system by end of 2019. At the conclusion of this project, we would have produced a qOI system for multiplexed quantification of optical biomarkers, integrating our optimized spectroscopy machine learning algorithm for use with the qOI system for immediate, real-time intraoperative feedback.

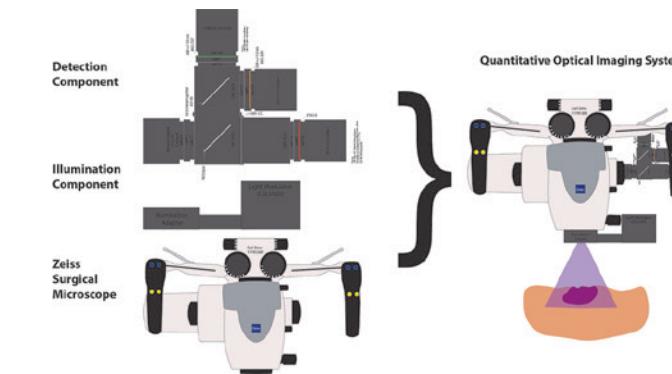


Figure 1. Light Transport Schematics: vFI vs qOI. Intraoperative optical imaging in tissue with different optical properties (right and left bottom panels) but identical concentration of fluorescent marker, PpIX. Top panels demonstrate difference in fluorescence intensity images for vFI due to the different optical properties, but equal and accurate, fluorescence intensity images using qOI

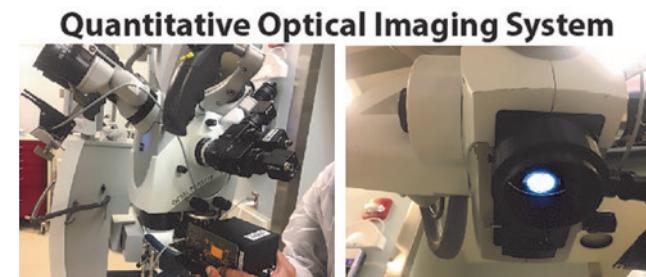


Figure 2. Schematics of qOI system. (Left) Components of the intraoperative system including the detection and illumination system and commercial microscope system. (Right) The qOI system integrated with the surgical microscope for intraoperative use

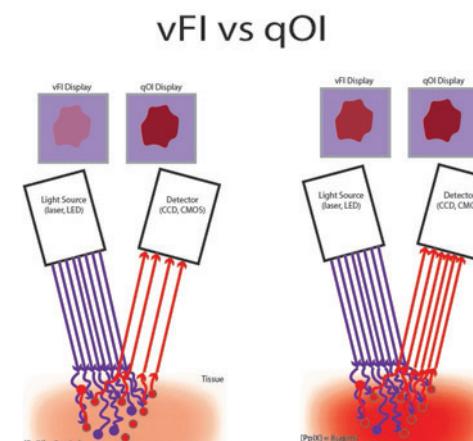


Figure 3. Live images of intraoperative qOI system in the neurosurgical OR. Images show the detection and illumination components integrated with the surgical microscope



DeepROP: Point-of-Care System for Diagnosis of Plus Disease in Retinopathy of Prematurity

James Brown, PhD

Senior Lecturer, Computer Science,

University of Lincoln, UK

jamesbrown@lincoln.ac.uk

Jayashree Kalpathy-Cramer, PhD (pictured)

Director, QTIM Lab, MGH

Associate Professor, Radiology, HMS

kalpathy@nmr.mgh.harvard.edu

Retinopathy of prematurity (ROP) is a retinal vascular disease that affects approximately two-thirds of premature infants weighing less than 1250 grams at birth. It is one of the leading causes of preventable childhood blindness. Although most cases are mild, 5-10% of cases (in the US) progress to severe disease that can lead to retinal detachment and permanent blindness if untreated. It is estimated that there are 40,000 cases of ROP per year in the US resulting in blindness in about 600 infants. The estimated economic impact is \$40-60M/year (1993 dollars). Ironically, advancements in neonatology worldwide are leading to an increased incidence of ROP. However, the number of ophthalmologists willing and able to manage ROP in the US and worldwide is insufficient, resulting in a great unmet need for screening and clinical decision support tools in this domain. Plus disease, characterized by retinal vessel tortuosity and dilation, is a key hallmark of treatment-requiring ROP. Several major NIH-funded studies and clinical trials have shown that severe ROP with plus disease may be effectively treated with laser photocoagulation or intravitreal injection of pharmacological agents such as bevacizumab. Therefore, it is essential to diagnose plus disease in an accurate and timely manner. Our solution is an AI system for the automated diagnosis of plus disease from retinal images. The software consists of an image pre-processing pipeline and classifier based on convolutional neural networks. The system has been trained and evaluated on a multi-institutional retinal image dataset, collected as part of i-ROP cohort study. Additionally, we have developed a disease severity scale that enables the monitoring of disease progression and response to therapy, with better granularity and predictive ability. Our software achieves diagnostic accuracies that are comparable with ROP experts. The near-real-time analysis making it highly tenable for clinical use. Deployment opportunities for the software include integration with the camera system including low-cost mobile phone attachments, integration with the electronic health record system, mobile phone applications and a web-based tool.



Figure 1. Example fundus photographs demonstrating the normal and abnormal vasculature in ROP

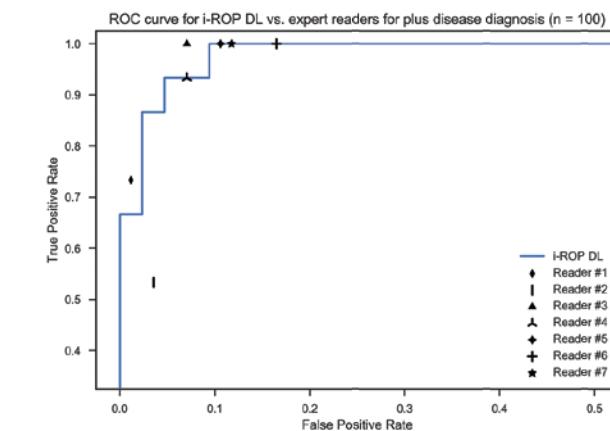


Figure 2. Diagnostic performance of the algorithm and 8 ROP experts when compared to the reference standard diagnosis

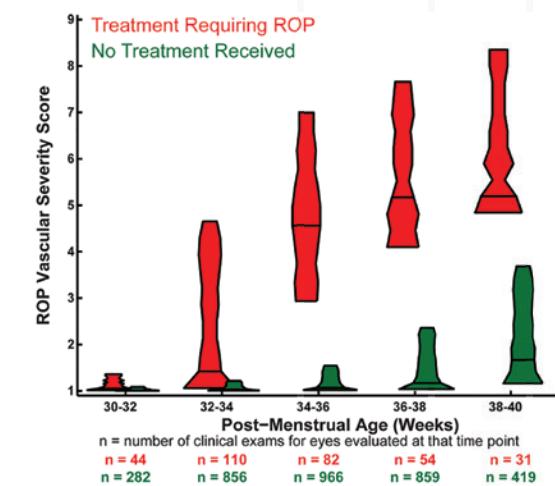


Figure 3. Severity score generated by the system is a biomarker for treatment requiring disease

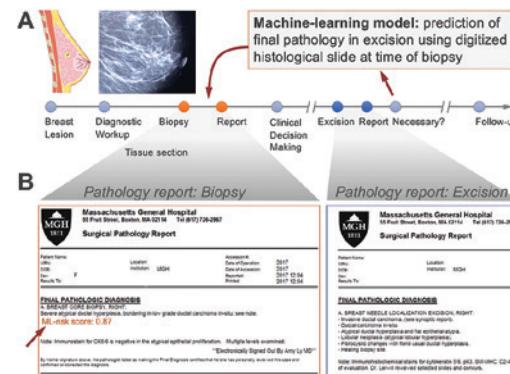


Figure 1. Unmet need and proposed solution. A. Timeline in the diagnostic workup of a breast lesion. The biopsy result is a central component of the clinical decision making. Only after excision it is clear whether the surgery was necessary or not. We plan to implement a machine-learning model that predicts the excision results at the time of biopsy using data from a digitized histopathological slide B. Proposed clinical implementation as a machine-learning (ML-) score in the report. The machine-learning result will be an output ranging from 0-1 (0=no risk; 1=high-risk for malignancy in a subsequent excision). This ML-score will be implemented in the final pathology report.

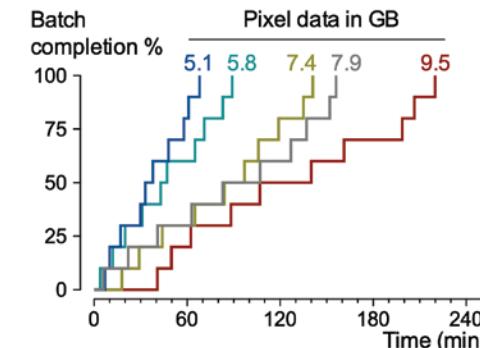


Figure 2. Pixel data acquisition times for 5 representative batches of 10 slides.

Predicting Unnecessary Surgeries in High-Risk Breast Lesions



Jochen Lennerz, MD, PhD

Associate Director, Center for Integrated Diagnostics, MGH; Assistant Professor, HMS

jlennerz@partners.org

Early detection of breast cancer can reduce mortality and suspicious lesions identified on mammography are biopsied to clarify the nature of the lesion. Pathologists evaluate the biopsies, render diagnoses, and determine whether surgical excision (for high-risk lesions) or clinical follow-up (for low-risk lesions) is indicated. In the US, ~10% of the 2,000,000 breast biopsies done each year yield high-risk lesions. The current standard of care for the 200,000 women with high-risk lesions is surgical excision. Notably, most (>90%) of high-risk lesions are ultimately benign on final surgical excision. Unnecessary surgeries in patients with high-risk breast lesions have numerous healthcare, aesthetic, safety, legal, and financial implications. We proposed to integrate a machine-learning (ML) algorithm for evaluating histologic sections and incorporate these digital data into a ML model that predicts cancer in the subsequent excision - with the explicit goal of reducing unnecessary breast surgeries for benign lesions.

To approach a fully-integrated clinical solution we implemented the following critical components:

- (1) An interdisciplinary working group, currently consisting of >10 representatives from 5 disciplines.
- (2) The model, by Drs. Lehman (MGH radiology) and Dr. Barzilay (MIT) has been published¹ and is theoretically able to reduce unnecessary breast surgeries by 30%.
- (3) The initial dataset of 1006 lesions in 986 patients with full clinical, radiologic and electronic medical record derived annotations surmounts to ~20,000 data elements.

- (4) The reporting infrastructure –portrayed within the timeline of the diagnostic workup– of a high-risk breast lesion has been outlined (Figure 1). Physicians involved in management of high-risk breast lesions will have access to the ML-score for malignancy at time of biopsy (Figure 1).
- (5) A vendor-neutral environment for efficient slide scanning in clinical production (Figure 2). We have currently generated pixel and pixel-related metadata for 542 patients.
- (6) Although the primary outcome predicted by the model is determined using pathology diagnoses, the ML model currently does not include histologic slides because a unified data model and slide scanning environment in clinical production is currently missing. In collaboration with the MGH&BH Center for Clinical Data Science, we have implemented DICOM for digital pathology².

Generating the data for machine learning models is the key limiting factor for adoption of ML in digital pathology; we anticipate that the developed infrastructure serves as a blueprint to overcome this significant hurdle to integrate ML into patient care.

References

1. Bahl M, Barzilay R, Yedidia AB, Locascio NJ, Yu L, Lehman CD. High-Risk Breast Lesions: A Machine Learning Model to Predict Pathologic Upgrade and Reduce Unnecessary Surgical Excision. *Radiology*. 2018 Mar;286(3):810-818.
2. Herrmann MD, Clunie DA, Fedorov A, Doyle SW, Pieper S, Klepeis V, Le LP, Mutter GL, Milstone DS, Schultz TJ, Kikinis R, Kotekas G, Hwang DH, Andriole KP, Iafrate AJ, Brink JA, Boland GW, Dreyer KJ, Michalski M, Golden JA, Louis DN, Lennerz JK. Implementing the DICOM Standard for Digital Pathology. *J Pathol Inform*. 2018 Nov 2;9:37.



Figure 1. Sensors consist of an MR-compatible transducer contained in a 3D-printed capsule of our own design, easily attached to the skin.

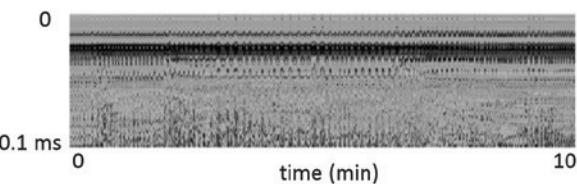


Figure 2. Sensor signals are sensitive to internal motion, breathing motion here. The sensors accompany patients to their imaging exams, leading to parallel streams of data that are then combined through machine learning.

Sensor Technology for Enhanced Medical Imaging



Bruno Madore, PhD

Associate Professor, Radiology, BH, HMS

bruno@bwh.harvard.edu

In the United States, medical imaging is an expensive component of a particularly-costly healthcare system. Acquiring more clinically-relevant information in a shorter amount of time would add value to imaging exams. Our approach revolves around the use of sensors to supplement traditional imaging by generating parallel streams of data where only one (the traditional scanner) would normally exist. The area of skin closest to the anatomy of interest is thought of as prime real-estate and sensors were designed here to best exploit this spatial resource.

More specifically, we designed ultrasound-based sensors meant to accompany patients to MRI, PET or ultrasound imaging exams. The sensors can easily be fixed to the skin and they are relatively small, about 3x3x1 cm in size. They emit unfocused ultrasound beams, the signals they collect are exquisitely sensitive to internal motion, and they can be employed up to four at a time for distributed sensing. Machine learning algorithms translate sensor signals into readily-useable forms, such as synthetic MRI images, synthetic electrocardiogram signals and/or synthetic optical tracking signals. In the latter application, sensors were employed in a receive-only mode, in a manner reminiscent of passive sonar, to track the position and orientation of a clinical ultrasound imaging probe placed on the skin.

The main application pursued so far has been to carry the power of MRI outside of the imaging suite. MRI is performed with sensors in place while machine learning discovers correlations between images and sensor signals; later on, based on sensor signals and known correlations, time series of MRI-like images can be obtained even after the subject has left the MRI suite. This functionality can be used to help bridge separate scans from different modalities in a manner that accounts for internal motion, and applications to image-guided therapies are considered. A PCT application was filed on the sensor technology. The 3D-printed capsules are meant to be single-use, and as such could play a key role in future revenue streams.

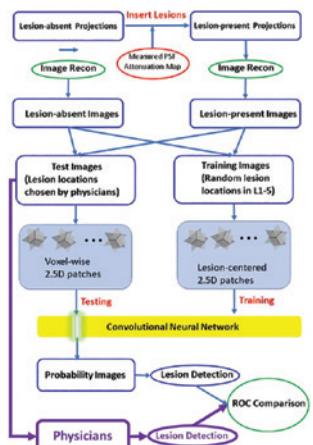


Figure 1. Lesion detection using a neural network trained on artificially inserted lesions. Realistic artificial lesions are inserted in human data in the projection space using measured point spread function and subjects' attenuation map. Lesion-absent/present data are reconstructed. The resulting images are used to train a convolutional neural network. The performance of lesion detection is evaluated by receiver operating characteristic curves (ROC) using human observer results as the reference.

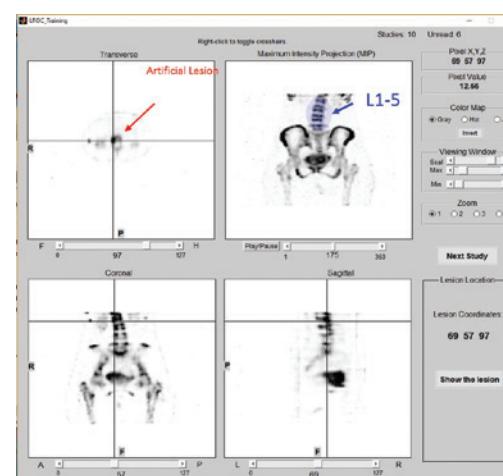


Figure 2. ^{99m}Tc -MDP bone images of a healthy subject in transverse, coronal, sagittal, and maximum intensity projection views. An artificial lesion is inserted at the crosshairs, which is located on the posterior aspect of the lumbar spine.

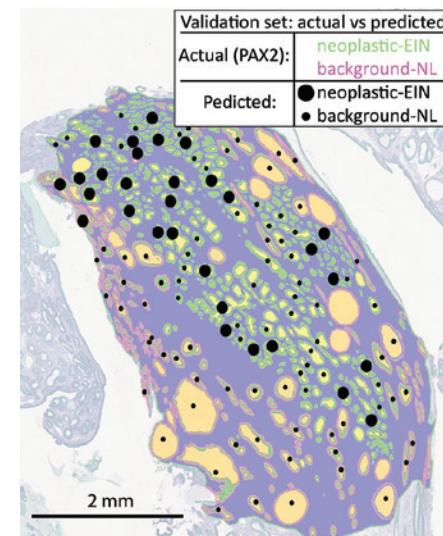


Figure 1. Use of data clouds facilitates human visual integration of gland classifications to identify regions of diagnostic interest. Large black circles identify glands predicted to be neoplastic, and small black circles identify glands predicted to be non-neoplastic. Green glands are neoplastic based on biomarker status (loss of PAX2 expression), while purple glands are non-neoplastic. A human user can readily identify the regions with high concentrations of large circles, which, given the 89% positive predictive value of the algorithm, are very likely to contain neoplastic glands.



Augmented Digital Microscopy for Diagnosis of Endometrial Neoplasia

David Papke, MD, PhD (pictured)

Resident, Surgical Pathology, BH; Clinical Fellow, HMS

dpapke@partners.org

George Mutter, MD

Medical Director, Reproductive Endocrinology Laboratory, BH; Professor of Pathology, HMS

Training a Neural Network to Detect Lesions



Jinsong Ouyang, PhD

Physicist, MGH; Associate Professor, HMS

ouyang.jinsong@mgh.harvard.edu

Our goal is to use neural network for lesion detection in medical imaging. The use of machine learning in medical imaging requires large number of labeled clinical data sets. There is also no guarantee that the diagnostic information associated with each clinical data set is accurate. Moreover, the behavior of neural network changes due to data shift, i.e., the difference between training and test data distributions, partly because of the variation on imaging protocols including scanners, procedures, and image reconstruction parameters. We propose to artificially insert lesions into human imaging data in the projection space to build the training data for a neural network. This makes it possible to use a small number of human studies to build almost unlimited number of training data sets. As a result, we can train the neural network with each type of imaging protocol to avoid the data shift. Also, the ground truth of each inserted lesion is known. We are now focusing on bone lesion detection using ^{99m}Tc -MDP single photon emission computed tomography, which is the most challenging lesion detection task in clinical setting. We collected projection data sets on 65 subjects. Using the lesion-insertion/image reconstruction software we

developed, we inserted realistic artificial bone lesions into projection data and created both training and test data sets for both the neural network and human observer studies. We evaluated different convolutional neural networks (CNN) for lesion detection. We are deriving receiver operating characteristic curves (ROC) using our trained CNN. We will evaluate the performance of our method by comparing the resulting ROC curves with the ones obtained from human observer studies. Our method can be applied to other clinical applications, such as tumor detection using positron emission tomography with ^{18}F -FDG.

Variability in subjective human diagnoses is an unsolved problem in anatomic pathology, and there is need for auxiliary techniques to improve diagnostic replicability in several challenging yet common clinical scenarios. One such scenario is the identification of pre-malignant neoplasia in endometrial biopsies ("endometrial intraepithelial neoplasia", or EIN), which has poor reproducibility among human pathologists. Here, we outline a novel algorithmic approach to localize and classify glandular lesions, using diagnosis of EIN in endometrial biopsies as a model system.

We obtained archival endometrial biopsy slides from 108 patients (34 EIN, 74 normal), immunostained them for pan-keratin, and scanned all slides to generate digital whole slide images. The keratin stain was used to direct automated segmentation of tissue into epithelial and stromal compartments. Individual glands in the validation set were identified as background or neoplastic based on positive or negative staining status, respectively, of PAX2 (a transcription factor lost in neoplastic EIN glands). Gland location in the PAX2 stain was mapped to companion image analysis results from the keratin stain to enable automatic labelling of glands as neoplastic (component

of EIN) or background (NL). Voronoi diagrams were generated for each gland based on nuclear coordinates, and the graph parameters were used to train random forest algorithms to classify individual glands as neoplastic-EIN or NL.

We hypothesized that a classification accuracy of approximately 70% would suffice to generate a "data cloud" overlay that a human user could visually integrate to identify a region of interest. We also hypothesized that small glands are uninformative, and that excluding small glands from the training set would improve classification accuracy. We found that exclusion of small glands improved classification accuracy in the validation set from 59.3% to up to 77.9%. This accuracy was sufficient to allow our algorithm to identify EIN lesion epicenters as dense clouds of individual glands classified as neoplastic-EIN (Figure 1).

This work demonstrates the utility of sequential enhancement of classification by filtering uninformative glands and leveraging human visual integration of data clouds. This approach is generalizable to any diagnostic problem involving localizing glandular lesions, such as those commonly seen in breast and prostate biopsies.

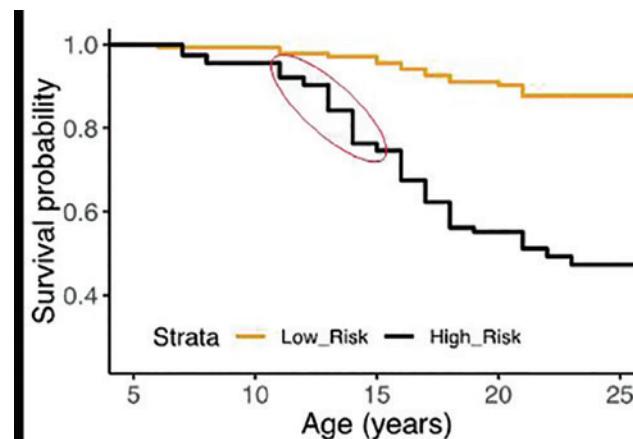


Figure 1. Kaplan-Meier survival analysis indicating time course until onset of first MDD episode in community participants with AI predicted low versus high risk for MDD. Red ellipse highlights period of rapid emergence of MDD shortly after age 12.

Poly-Exposure Risk Scores for Psychiatric Disorders



Martin Teicher, MD, PhD

Director, Developmental Biopsychiatry Research Program, McLean
Associate Professor, Psychiatry, HMS
mteicher@mclean.harvard.edu

The key problem that this technology is intended to address is the accurate prediction of risk for major depression (MDD) and anxiety disorders. These are the most prevalent psychiatric disorders and are leading causes of disability adjusted life years worldwide. Psychiatry has lagged behind other branches of medicine in prevention strategies, due in part to an inability to identify individuals at an early stage with sufficient risk to motivate cost-effective preventative treatment trials. This invention is designed to close this gap. The long-term benefit would be to modify psychiatric practice to focus on prevention/preemption of common disorders. Considerable effort is being devoted to developing Polygenic Risk Scores for schizophrenia and bipolar disorder, but the disorders we are focused on are less heritable and early exposure to adversity plays a critical role in their emergence. We have discovered that there are sensitive periods when exposure to specific types of abuse are the most important predictors of risk and developed a highly reliable instrument for assessing type and timing of exposure (MACE scale). For these evaluations we used two data sets: an interviewed sample ($N=667$) recruited from the community but

enriched to have a greater percentage of participants with maltreatment and an online sample ($N=2092$) recruited from the community. We first established that Random Forest Classification (RFC) and Model Averaged Neural Networks provided the best predictive accuracy results and we used RFC throughout due to their superior computational speed. Briefly, we developed predictive models that, based on exposure data up to age 12, plus family history, could identify a high-risk subgroup (~20% of sample) with a 52% risk for developing major depression and a low risk subset (~20%) with an 13% risk (\log_{10} likelihood $p < 10-11$). These high-risk individuals are at sufficient risk to warrant enrollment in prevention trials. PERS prediction greatly exceeded published results based on GWAS for MDD (16% of the variance versus 0.5 – 1%). Prospective longitudinal studies are needed to verify these models.

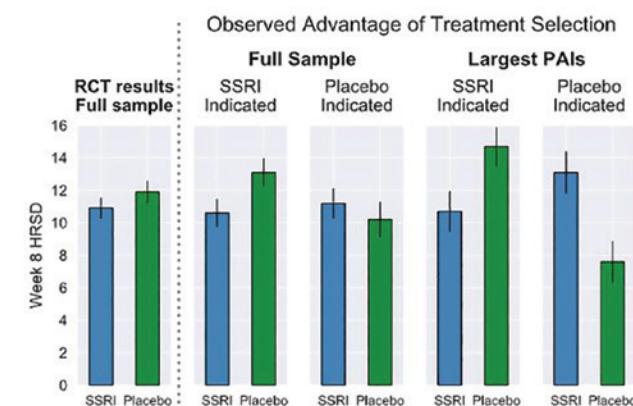


Figure 1. Comparison of post-treatment depression scores for patients randomized to their algorithm-indicated vs. non-indicated treatment condition.

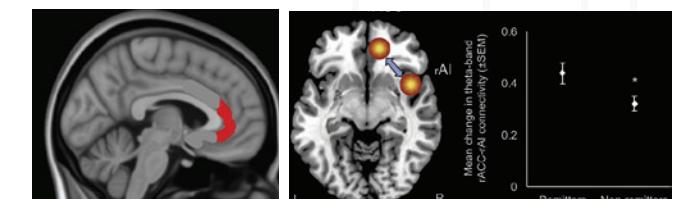


Figure 2. Rostral anterior cingulate cortex (rACC) activity (left panel) and rACC-insula connectivity (right panel) predicts depression remission.

Leveraging Machine Learning to Match Depressed Patients to the Optimal Treatment



Christian Webb, PhD

Director, Treatment and Etiology of Depression, Youth Lab, McLean
Assistant Professor, Psychiatry, HMS
cwebb@mclean.harvard.edu

Major Depressive Disorder (MDD) is a prevalent and recurrent condition associated with substantial disability, staggering economic costs, and a high suicide rate. Despite significant effort, MDD remains very challenging to treat. Treatment recommendations in current clinical practice are primarily guided by trial-and-error (i.e., trying different antidepressant medications, alone or in combination) and clinician preference (e.g., cognitive behavioral therapy vs. antidepressants). Clinically useful algorithms predicting outcomes for individual patients are already in use in various medical fields including oncology, cardiology, endocrinology, and critical care, where quantifying patient-specific risk can be critical for informing recommendations for preventive interventions and treatment planning. However, psychiatry has lagged behind and such algorithms are unfortunately not available for data-driven treatment recommendation for depressed individuals. There is a critical need for more efficient, data-driven methods to predict treatment outcome and optimize treatment selection for individuals suffering from depression. As 50-70% of depressed individuals do not respond to the first antidepressant they receive, the bulk of patients must suffer the distress

of prolonged depressive episodes before the suitable treatment is identified. Moreover, these inefficiencies and prolonged lengths of treatment are ultimately costly to payers and tax the limited resources (e.g., beds, clinical staff) of psychiatric clinics. Research aimed at identifying predictors of antidepressant response may ultimately help inform optimal treatment assignment. Our recent work demonstrates the potential for machine learning to improve individual patient outcomes through algorithm-guided treatment recommendations. Specifically, we recently developed a machine learning algorithm to predict optimal treatment assignment for depressed individuals on the basis of pre-treatment patient characteristics, including neural (8-minute electroencephalography [EEG] recording), clinical (e.g., symptom profiles, depression history), demographic (e.g., age, gender, marital and employment status), and behavioral (e.g., cognitive control task) data. We are currently adapting and testing the utility of this algorithm within a clinical unit at McLean Hospital.

Deep Learning to Diagnose Epilepsy



Brandon Westover, MD, PhD

Executive Director, Clinical Data Animation Center, MGH
Associate Professor, Neurology, HMS

mwestover@mgh.harvard.edu

Diagnosing epilepsy depends critically on the ability of specially trained clinical experts to detect interictal epileptiform discharges (IEDs) in electroencephalograms (EEGs) recordings. However, experts are in short supply, and the quality of EEG interpretations vary. Before now, lack of a sufficiently large and representative annotated dataset has precluded development of automated IED detection algorithms. In this talk I will describe how we developed the world's largest annotated IED dataset and used it to train deep neural networks that detect IEDs better than typical human experts.

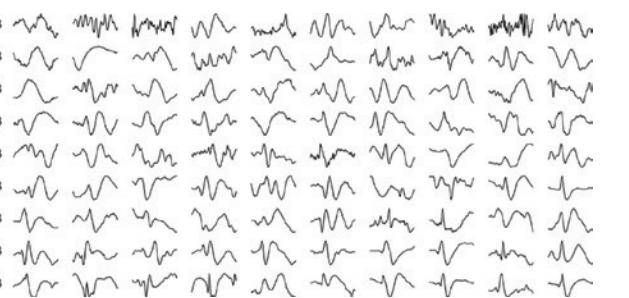


Figure 1. Examples of epileptic spikes, the hallmark of epilepsy. The numbers left of each row show the number of experts who voted that examples in that row were epileptic. These examples provide some sense of the diversity of waveform shapes that can occur in a large population of patients with and without epilepsy, and the challenge of interpreting EEGs.

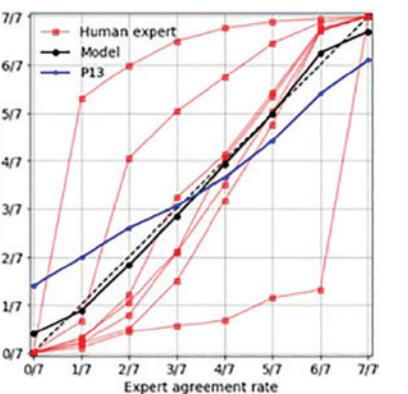


Figure 2. Calibration curves for humans (red), the "industry standard" algorithm (black), and our deep neural network. The industry standard overcalls low-probability events, and undercalls high probability events. Our neural network performs better than human experts.

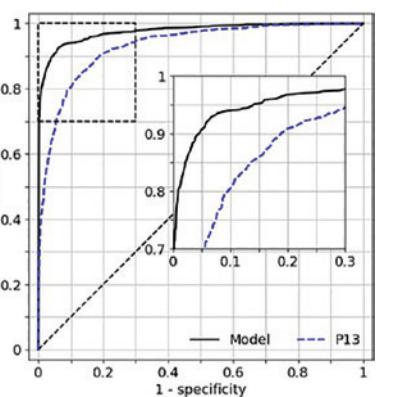


Figure 3. ROC curve for our deep neural network (blue curve) and the industry standard (black curve) in summarizing an entire EEG (as opposed to individual spikes) as epileptic vs not epileptic.

Notes

WORLD MEDICAL INNOVATION FORUM S P O N S O R S

The 2019 World Medical Innovation Forum was made possible through the generosity of the following sponsors.

presenting



stakeholder



strategic



collaborator



Partners HealthCare I N N O V A T I O N

The World Medical Innovation Forum is brought to you by Partners HealthCare Innovation, the 125-person business development unit responsible for the worldwide commercial application of the capabilities and discoveries of Partners' 74,000 employees.



DEPARTMENTS

- Boston Biomedical Innovation Center
- Business Development
- Innovation Operations
- Licensing and Innovation Management
- Market Sectors
- Open Innovation – Alliances and Fellows
- Partners HealthCare International
- Partners Innovation Fund

FUNCTIONS

- Business development
- Company creation
- Industry collaborations
- Innovation management
- International consulting
- Licensing
- Research translation strategy and funding
- Technology marketing
- Venture investing
- Workforce capacity building



215 First Street, Suite 500, Cambridge, MA 02142
857-307-2400 | innovation.partners.org



215 First Street, Suite 500 | Cambridge, MA 02142
857-307-2400 | innovation.partners.org



BRIGHAM HEALTH



BRIGHAM AND
WOMEN'S HOSPITAL



MASSACHUSETTS
GENERAL HOSPITAL



MASSACHUSETTS
EYE AND EAR



McLean HOSPITAL
HARVARD MEDICAL SCHOOL AFFILIATE



NEWTON-WELLESLEY
HOSPITAL



SPAULDING.
REHABILITATION NETWORK