First Look

The Next Wave of AI Breakthroughs in Health Care

2019

WORLD MEDICAL
INNOVATION
FORUM℠

2018 First Look | The Next Wave of Artificial Intelligence Breakthroughs

Leveraging AI and Biological Networks to Functionally Interpret Vast Genetic Datasets

Kasper Lage, PhD
Director of Bioinformatics, MGH; Associate Professor, Surgery, HMS

Early career Harvard Medical School investigators kick-off the 2019 World Medical Innovation Forum with rapid fire presentations of their high-potential new technologies. Eighteen rising stars from Brigham Health, Massachusetts General Hospital, Massachusetts Eye and Ear Infirmary, McLean Hospital and Spaulding Rehabilitation Hospital will give ten-minute presentations highlighting their discoveries and insights that will disrupt the field of artificial intelligence. This session is designed for investors, leaders, donors, entrepreneurs, investigators and others who share a passion for identifying emerging high-impact technologies.

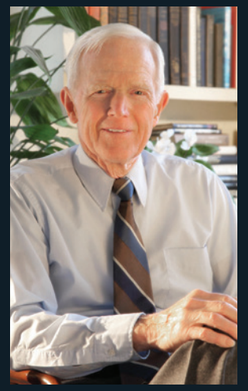
Two of the top presenters will be awarded the Peter K. Ranney Innovation Award. The prize carries a $10,000 award.

First Look

peter k. ranney innovation award

The Peter K. Ranney Innovation Award will be given to honor the top two First Look presenters who embody the innovative, entrepreneurial and visionary spirit that the World Medical Innovation Forum was established to recognize. The two $10,000 awards will be granted based on overall presentation quality, innovativeness, commercial potential, caliber of disruption, and market need.

The Award will be judged throughout the morning session on April 8th with winners announced at the annual Innovator's Dinner on Wednesday, April 10, 2019.



2019 artificial intelligence **investigators**

6 Henry Chueh, MD
7 Synho Do, PhD
8 Laura Germine, PhD
9 Satrajit Ghosh, PhD
10 Xudong Huang, PhD
11 Tina Kapur, PhD
12 Bharti Khurana, MD
13 Vesela Kovacheva, MD. PhD
14 Constance Lehman, MD, PhD
15 Lisa Nickerson, PhD
16 Federico Parisi, PhD
17 Stuart Pomerantz, MD
18 Sandro Santagata, MD, PhD
19 Joseph Schwab, MD
20 Chris Sidey-Gibbons, PhD
21 Hiroyuki Yoshida, PhD
22 Nazlee Zebardast, MD
23 Li Zhou, MD, PhD

Note: *Speakers and content are subject to change.*
BH *Brigham and Women's Hospital* | ***HMS*** *Harvard Medical School*
MEE *Mass. Eye and Ear* | ***MGH*** *Massachusetts General Hospital*
SRN *Spaulding Rehabilitation Network*

Dxplain: Expanding Diagnostic Horizons

Henry Chueh, MD
Director, MGH Lab of Computer Science, MGH
Assistant Professor, Medicine, HMS
chueh.henry@mgh.harvard.edu

Getting a wrong diagnosis can lead to unnecessary treatment, bad outcomes, and even death. It's tempting to believe that in our age of many resources and technologies that this is a rare event – but that would be false. Studies have estimated that as many as 1 in 6 new diagnoses are wrong, leading to up to 36% of preventable adverse events. In a review of malpractice claims, almost half of these errors were due to failure to consider the correct diagnosis at the time of decision-making. In fact, the leading cause of wrong diagnosis appears to be the tendency of even expert clinicians to settle too early on a diagnosis.

Dxplain (Fig. 1) is a technology created by the MGH Lab of Computer Science that can help doctors make the right diagnosis the first time. Dxplain consists of a massive knowledge base of over 2500 diseases, 5500 clinical findings and more than 1 million quantitative relationships between them. When paired with algorithms of probabilistic inference, Dxplain can show a clinician the possible common, rare, and urgent conditions that should be considered, given a set of clinical findings. Unlike some other technologies, Dxplain can explain why it arrived at a particular diagnosis (Fig. 2). The Dxplain technology has had decades of validation through curation, independent studies, and clinical use. Recently the MGH LCS implemented Dxplain as a robust set of application programming interfaces hosted through cloud-based microservices on Amazon Web Services. This platform is already being used to support tens of millions of hits monthly from prominent symptoms checker applications. A mobile iPhone app has also been created (Fig. 3).

The capabilities of Dxplain have barely been tapped, and numerous market opportunities exist: feeding notes to Dxplain to identify potentially missed diagnoses after visits, establishing tools that malpractice providers will set as standard of care for preferred rates, consumer-based self-assessment and triage, inclusion into electronic records or ambient systems to suggest diagnoses in real-time. Since data and algorithms are at the core of Dxplain, the possibilities are endless, and it remains a unique commercialization opportunity.

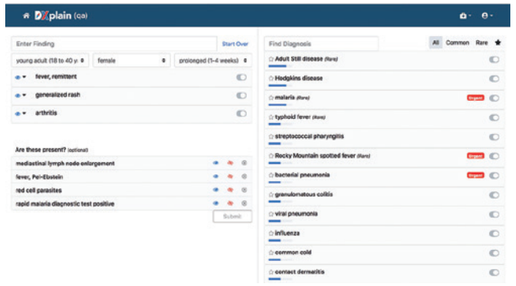


Figure 1. Screenshot of a clinical case in Dxplain showing a differential diagnosis ranked by likelihood, given a set of three clinical findings. Rare and urgent diagnoses are highlighted.

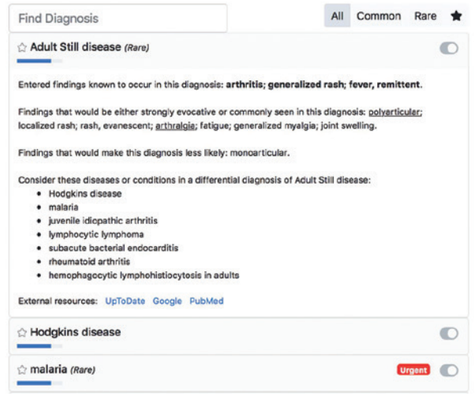


Figure 2. Screenshot of a diagnosis in Dxplain expanded, showing how it can explain why the diagnosis is included, and how it can also suggest other findings that would support or diminish the likelihood of the diagnosis. This can enhance trust in the results.

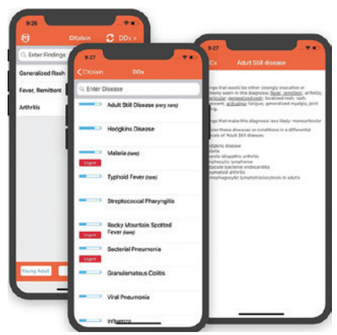


Figure 3. Screenshots of a clinical case in Dxplain mobile, which is powered by the same cloud-based microservices as the Web version.

Leveraging a Deep-Learning Algorithm for the Detection of Acute Intracranial Hemorrhage

Synho Do, PhD
Director, Laboratory of Medical Imaging and Computation (LMIC), MGH; Assistant Professor, HMS
sdo@mgh.harvard.edu

With the rapid progress of machine learning, deep-learning algorithms have the potential to change medicine's landscape. Specifically, advances in image recognition could increase diagnostic accuracy and speed and enhance physician workflow. However, there are still obstacles hindering the translation of deep-learning systems into clinical environments. This includes the necessary access to large datasets from which to "train" machine-learning, which can be costly and time-consuming to accumulate. An additional obstacle is the inability for users to understand the algorithm's decision-making process. For example, even if an algorithm correctly identifies a certain diagnosis, how can we understand its justification? In our research, we addressed both challenges by using a small dataset to construct an explainable, deep-learning algorithm for the image detection of acute intracranial haemorrhage (ICH). While using a small, imbalanced dataset of less than 1,000 images, we emphasized the standard of quality. Rather than having general radiologists simply label the presence or absence of ICH in each image, we recruited five specialty neuro-radiologists not only to label the presence of ICH, but also to label its specific subset out of five options. Furthermore, we adjusted the system's image processing to mimic radiologists' own workflow. We found that even with a small dataset, enhancing the quality of the data and paralleling the algorithm's processing to clinical work enabled a system performance similar to that of expert radiologists. Beyond optimizing performance, we made the algorithm explainable, having it create an atlas from the training set which in turn illuminated its decision-making. The "explainability" of an algorithm is essential not only for understanding the system's predictions, but also for continuing improvement and optimization.

By providing a reliable, accurate second opinion in diagnosing brain hemorrhages, the implementation of this system has the potential to enhance patient care, empower patients and cut costs. The benefits of deep-learning systems extend beyond neuroradiology, and by constructing an explainable deep-learning algorithm from small datasets, our research helps address challenges traditionally hindering their implementation.

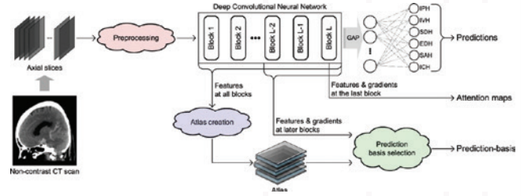


Figure 1. A diagram of the explainable deep learning system for ICH detection and classification.

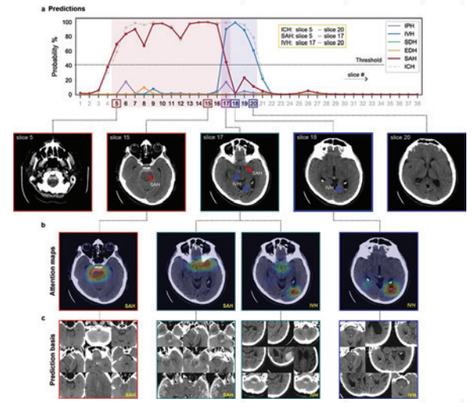


Figure 2. An output summary of the explainable deep learning algorithm. a. Probabilities for the presence of each type of ICH. b. For each positive case, the system generates a color-coded attention map, c. A set of prediction bases that are most relevant to each positive image

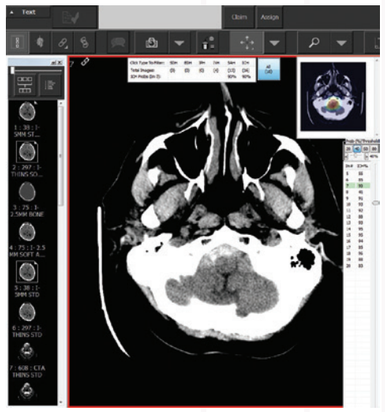


Figure 3. A screenshot of ICH detection AI implementation in the clinical PACS(Picture Archiving and Communication System).

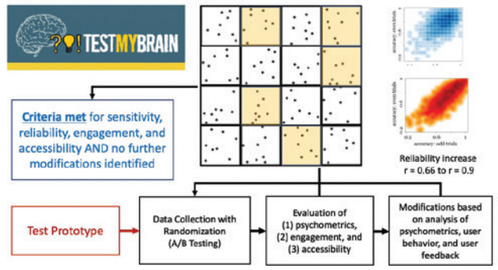


Figure 1. Schematic of Iterative Test Development Procedure Based on Structured A/B Testing and Automated Parameter Selection

The Next Generation of Cognitive and Behavioral Assessment

Laura Germine, PhD
Director, Laboratory for Brain and Cognitive Health Technology, McLean
Assistant Professor, Psychiatry, HMS
lgermine@mclean.harvard.edu

The future of cognitive and behavioral assessment will be digital. New approaches to assessment development are needed to (i) exponentially increase the cycle of innovation and validation, (ii) go beyond print publication models of intellectual property, and (iii) provide continued and long-term commercialization opportunities. Here, we describe the MAIAD approach (Machine-Assisted Iterative Assessment Development), an iterative and high throughput method to assessment development based on the low-cost collection of large structured datasets and algorithmic optimization of assessment parameters. Rather than providing a specific assessment, algorithm, or technology, MAIAD serves as an engine for the development of novel digital, algorithmic assessments for quantifying cognitive, behavioral, and neuropsychological functions. MAIAD is enabled by two unique capabilities: (1) A structured and participant engagement driven approach to data collection that allows us to capture structured data across massive samples (2.2 million thusfar through our digital research platform TestMyBrain.org), and (2) A validated iterative development framework that allows us to rapidly identify tests, items, and item parameters that best capture a particular characteristic, phenotype, or outcome (see Figure 1). New assessments and algorithms can be evaluated in hours or days, rather than the typical months or years. As a proof of concept, we show a straightforward open loop application of MAIAD for capturing visuospatial capability. Using this approach, we were able to develop a measure that uniquely predicts SAT score based on automated identification of item parameters. Cognitive measures developed by this approach are already being incorporated into large scale assessment initiatives, supported by funding across four NIH institutes.

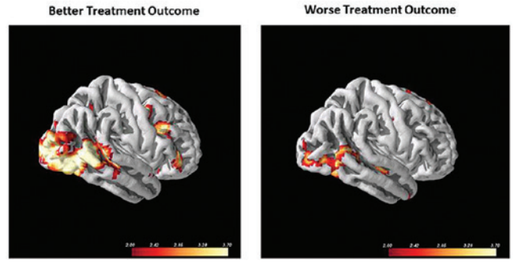


Figure 1. Brain imaging shows differences in activity that are predictive of treatment outcome in social anxiety disorder.

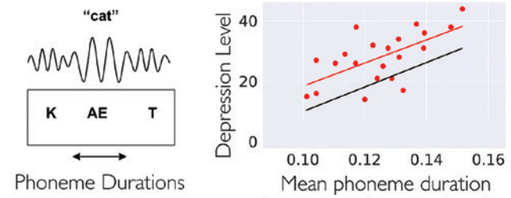


Figure 2. Speech characteristics in individuals with depression are indicative of severity.

Assistive Intelligent Technologies for Brain Health

Satrajit Ghosh, PhD
Research Associate, MEE; Principal Research Scientist, MIT
Assistant Professor, Otolaryngology, HMS
satra@mit.edu

The creation of the Web catalyzed a societal transformation. It enabled discussion forums, instant messaging, smartphones, and wearables. These technologies allow us to communicate better, to interact as a collective, to disseminate knowledge and information, and, perhaps most importantly, they act as sensors to measure the pulse of the world and of individuals. While specific areas of healthcare have seen significant advances, our ability to measure, track, treat, and predict treatment outcome of brain-related disorders, and specifically mental health, remains limited. Many existing pharmacological and behavioral therapies remain ineffective. Current approaches rely on intermittent assessments and self or caregiver reports, which are subjective and often unquantifiable.

We and our collaborators are changing this by combining modern sensing using brain imaging and smartphones with advances in AI technologies to improve assessments and treatments for mental health and neurological disorders. We have used brain scans to help predict treatment outcome in disorders such as social anxiety and major depression and demonstrated that speech recordings provide viable markers for tracking depression and Parkinson disease. We have developed methods for detecting meningioma tumors and helped advance AI algorithms along the way. In developing these technologies, we now know that more data can help make these methods more robust, but also that we do not really know when and how these models will fail.

To address this, we need a fundamental shift away from algorithms that simply learn from more data to algorithmic models that better understand the process. These models will link our targets of interest (e.g., depression level, treatment outcome) to the variations in our sensors (e.g., voice, brain images) and, to the extent possible, to the underlying neurobiology. Over the next decade, we envision using these models in a system that is continuously learning from sensors across individuals, relating information to life and health outcomes, and guiding individuals and caregivers to a more proactive version of care for brain health, and healthcare more broadly.

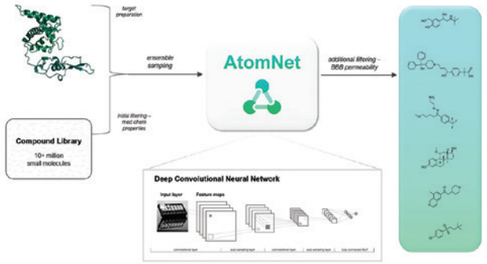


Figure 1. DCNN-based Virtual Drug Discovery Platform- AtomNet™

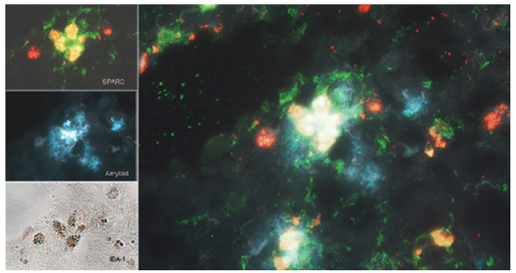


Figure 2. Increased SPARC protein is associated with Aβ amyloid plaques and microglia in Alzheimer's cortical tissues

Leveraging Artificial Intelligence for Brain Drug Discovery

Xudong Huang, PhD
Co-Director, Neurochemistry Laboratory; MGH; Associate Professor, Psychiatry, HMS
Huang.Xudong@mgh.harvard.edu

Drug discovery process is incredibly capital-intensive and time-consuming, and it is one of key factors that are driving up the astronomical healthcare costs. This is largely owing to the inefficiency in identifying suitable early leads and optimizing them into drug candidates, which typically involves the synthesis and testing of large numbers of small molecules. Teamed up with Atomwise, Inc, we use AI, i.e. deep convolutional neural networks (DCNN) (Fig. 1) to identify molecules that "fit" sites on target proteins- secreted protein acidic and rich in cysteine (SPARC) (Fig. 2) and three tryptophan-catabolizing enzymes in tryptophan-kynurenine pathway such as Indoleamine 2,3-dioxygenase 1 and 2 (IDO1/2) and Tryptophan 2,3-dioxygenase (TDO). These proteins are potential druggable targets for human diseases such as cancer, Alzheimer's disease (AD), and depression.

Unlike conventional molecular mechanics/quantum chemistry-based molecular modeling which attempts to explicitly parameterize the underlying H-bonding, van der Waals, electrostatics, and hydrophobic interactions (which are notoriously difficult to model accurately), this deep learning approach automatically generates a set of predictive parameters, often non-intuitive and abstract in nature, without human adjudication.

With the technology, we can rapidly screen millions of diverse molecules in silico to find ones that bind to a target protein specifically. We can also undertake mechanism of action screens to virtually identify proteins that bind compounds identified as hits in phenotypic screens, and to optimize lead molecules for bioavailability, including in silico prediction of blood-brain barrier (BBB) permeation if needed. This provides researchers with a small set of compounds that not only have a high probability to bind a target protein but also have elevated drug-like properties. Because the initial set of candidate compounds is small, they can often be experimentally validated quickly and inexpensively - using assays that researchers often already have in their lab without the need of developing a new medium or high throughput screening assay. Upon identification of early hits, we can also undertake hit-expansion exercises through in silico enumeration and evaluation, thereby minimizing the need for expensive chemical synthesis and expansive in vitro/in vivo testing. This AI-powered virtual drug screening technology will hopefully shorten the time from initial hit to clinical trial and ultimately reduce the cost of early drug development for unmet medical needs.

Using AI to Better Visualize Needles in Ultrasound-Guided Liver Biopsies

Tina Kapur, PhD
Executive Director, Image-Guided Therapy, BH
Assistant Professor, Radiology, HMS
tkapur@bwh.harvard.edu

Nearly a million image-guided liver biopsies are performed every year worldwide, and the numbers are projected to grow at an annual rate of 4%. At our institution, typically 30% are non-targeted liver parenchymal biopsies to evaluate diffuse disease, while 70% are performed to target a specific lesion or mass in the liver. Focal liver biopsies are not performed using ultrasound guidance when lesion visibility is expected to be poor with ultrasound imaging; half of our focal liver mass biopsies are performed using ultrasound guidance, with most of the other half performed using CT guidance. In some ultrasound frames, as shown in the top left image, the needle is clearly visible relative to the target (and critical blood vessels), therefore, the physician can proceed with the biopsy rapidly. At other times, as shown in the top right image, only experienced physicians can deduce where the needle is in the image, and whether the tip of the needle is outside the ultrasound image plane. We have previously developed AI-based software solutions for enhanced visualization of needles in MRI-guided prostate biopsies in which we trained deep neural networks on data from our institution and achieved accuracy in results that are within the range of inter-expert concordance. We propose an AI-powered software solution that shows the liver biopsy needle continually and in its full length on the ultrasound display, with the tip easily recognized as being in plane or out of plane (as shown in the bottom two images). The proposed solution is a high impact low-cost innovation for liver biopsies, which are low-reimbursement procedures with little room for the added time, complexity, or incremental cost of available commercial products. A software-only, AI-powered needle and tip visualization capability, integrated in the real-time ultrasound display, will provide better accuracy in hitting target lesions and avoiding critical structures, while affording shorter procedure times and shorter learning curves for physicians-in-training. AI can increase the utilization and cost-effectiveness of US-guided liver biopsies.

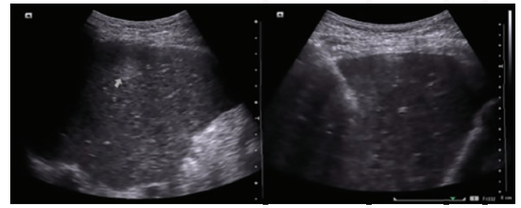


Figure 1. Ultrasound images during liver biopsy. In the left image the white arrow shows the faintly visible needle tip, and right image shows the length of the needle more clearly.

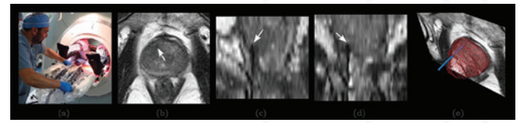


Figure 2. Finding Needles in transperineal in-gantry MRI-targeted prostate biopsy. (a) The patient is placed in the supine position in the MRI gantry, and his legs are elevated to allow for transperineal access. The skin of the perineum is prepared and draped in a sterile manner, and the needle guidance template is positioned. (b), (c) and (d): Axial, sagittal and coronal views of intraprocedural T2-W MRI with needle tip marked by white arrow. (e) 3D rendering of the needle (blue), segmented by our method, and visualized relative to the prostate gland (purple), and an MRI cross-section that is orthogonal to the plane containing the needle tip.

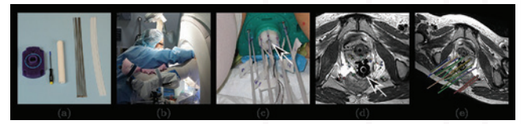


Figure 3. Finding Needles in MRI-guided high dose rate brachytherapy. (a) interstitial brachytherapy applicator which includes a Syed-Neblett template, a plastic obturator, tungsten alloy needles to help push in the catheters, and hollow air-filled plastic catheters. (b) a physician inserting catheters through the template in the bore of a 3T MRI scanner, (c) a close-up view of the inserted catheters, (d) a T2-weighted MRI axial cross-section with catheters segmented and color-coded, (e) a rotated view of the axial plan to illustrate the catheters in 3D.

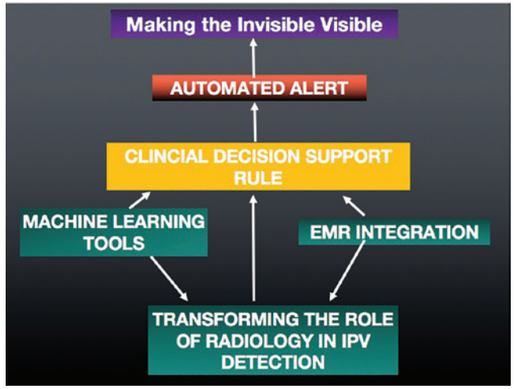


Figure 1. Automated detection of Intimate Partner Violence

Making the Invisible Visible: Bringing Intimate Partner Violence into Focus

Bharti Khurana, MD
Director, Emergency Musculoskeletal Radiology, BH
Assistant Professor, HMS
bkhurana@bwh.harvard.edu

On Nov 25th, 2018, the United Nations chillingly reported that the most dangerous place for women is inside their own homes. No fewer than 58% of female homicides are committed by current or former intimate partners or family members. Intimate Partner Violence (IPV) is defined as physical, sexual or emotional violence between partners or former partners. One in four women have reported IPV during their lifetime. The CDC estimated the cost of IPV exceeding $5.8 billion dollars ($9.3 billion in 2017 dollars). Despite the high prevalence and fatality of this critical public health issue, IPV continues to be profoundly underdiagnosed due to under-reporting by the victim. The lack of identification of IPV as the primary cause reduces the ability to offer early preventive services and may lead to further violence with each physical injury increasing the likelihood of sustaining a life-threatening injury. Currently an IPV screening questionnaire is a core component of every health care visit but the proportion of identifiable IPV cases to date only represents the tip of the iceberg. Imaging has made substantial contributions to the detection of nonaccidental trauma but sadly, the role of radiology in identifying adult victims has never been explored before.

We believe that rapid advances in imaging, information technology and machine learning present an opportunity to achieve a breakthrough in identifying IPV victims and thereby, activating timely intervention. Our pilot study on "Radiological findings in IPV victims" has allowed us to create an exhaustive list of imaging and clinical findings that are associated with IPV. We are now developing an integrated, multi-dimensional clinical decision support tool that uses patterns derived from expert analysis of historical radiological and clinical data, classification models, statistical evidence and alert system that classifies injuries for their likelihood of being as a result of IPV, and automatically alerts clinicians if a patient's injuries have a high or low risk probability for IPV. Our goal is to first validate and then implement this alert system locally, nationally, and globally.

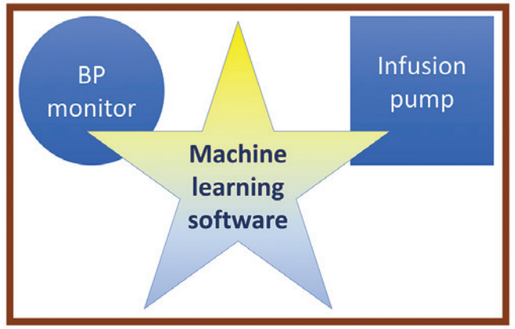


Figure 1. Anesthesia Automation System

Harnessing the Power of Machine Learning to Automate Drug Infusions in the OR and ICU

Vesela Kovacheva, MD, PhD
Attending Anesthesiologist, BH; Instructor, Anesthesiology, HMS
vkovacheva@bwh.harvard.edu

Machine learning holds the promise to automate and personalize drug administration in order to achieve optimal health outcome. Administration and titration of medications to a monitored parameter is one of the major sources for medication errors and risk patient safety. We are developing a machine learning automation system to maintain normal maternal blood pressure during cesarean delivery.

Low maternal blood pressure occurs in up to 74.1% of cesarean deliveries. In the US, 900,000 mothers annually experience complications associated with low maternal blood pressure like nausea, lightheadedness, and, infrequently, stroke; resulting in a prolonged recovery. In the baby, low maternal blood pressure can cause acidosis, hypoxia and low Apgar scores, which can lead to neonatal ICU admission and can correlate with poor developmental outcome. All these complications incur more than $100M costs to the healthcare system annually.

Currently, maintaining blood pressure is done by monitoring the vital signs and manually adjusting the rate of drug infusion every minute. We are developing an anesthesia automation system (AAS), which will utilize real-time blood pressure input from the patient, calculate the needs for vasopressor medication using a deep learning algorithm and deliver it directly to the patient via commercially available infusion pump. The software will calculate the amount of drug needed based not only on the current value of the blood pressure, but also on the blood pressure trend for that particular patient, the time course and the pharmacokinetics of the drug. While semi-autonomous, our system is designed to work with the physician present at all times and responsible for the patient care.

Our system would increase the safety and well-being of millions of mothers and babies annually (estimated $600M market for this application alone). As we expand to include medications like propofol and insulin, we envision that AAS will be invaluable also for any complex sedation, general anesthesia or ICU case. In the long run, AAS would augment the workflow of the anesthesiologists, increase patient safety, decrease cost of care and, ultimately, transform the way we perform anesthesia.

AI-Based Care Delivery: A New Paradigm for Curing Cancer

Constance Lehman, MD, PhD
Chief, Breast Imaging Division, MGH
Professor of Radiology, HMS
clehman@partners.org

The global burden of cancer in lives lost and financial costs is expanding exponentially. Every year over two million women are diagnosed with, and over 600,000 women die from, breast cancer worldwide. Current approaches that treat late stage malignant disease are costly and fail to cure. Breast cancer is cured when detected early, yet the vast majority of women do not have access to the benefits of early detection due to 1) lack of accurate risk prediction models to support effective prevention and screening strategies and 2) dearth of specialized radiologists to interpret screening mammograms. The cure for cancer, not only breast cancer but the full diversity of solid tumor cancers, lies in our ability to identify patients at increased risk and to implement effective programs for disease prevention and early detection/treatment.

Our AI-based paradigm delivers high quality, cost effective care by providing two immediate applications for clinical implementation. First, our innovative methods of risk assessment leverage the strength of Artificial Intelligence to identify women at risk for breast cancer. Current breast cancer risk models incorporate only a small fraction of patient data available and have failed to accurately predict future risk in individual women. We have developed a deep-learning (DL) model that operates over a full resolution digital mammogram image with traditional risk factor data to predict a patient's future breast cancer risk. Rather than manually identifying discriminative image patterns, we rely on our machine learning model to discover these patterns directly from the data. Unlike traditional models, our DL model performs equally well across diverse races, ages, and family histories. Second, we have developed a DL model that can provide interpretation of mammograms approximating the level of specialized human readers, transitioning screening mammography from a costly test, highly dependent on subspecialized radiologist expertise, to an inexpensive test that can be read by machines. Our lead scientists from MIT (Dr. Regina Barzilay) and MGH (Dr. Connie Lehman) are uniquely positioned to combine the strength of AI, the wealth of our large curated, modern quality databases, and our expertise in effective clinical implementation to rapidly integrate advanced AI risk models and AI image interpretation into clinical practice.

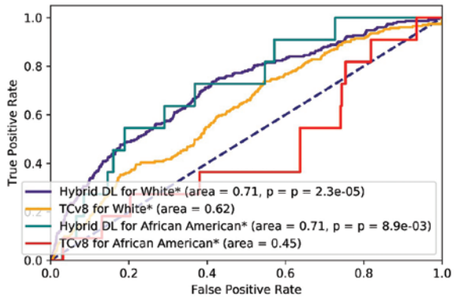


Figure 1. Deep learning model significantly better at predicting future risk of cancer in diverse races than current "best practice" advanced risk model Tyrer-Cuzick version 8.

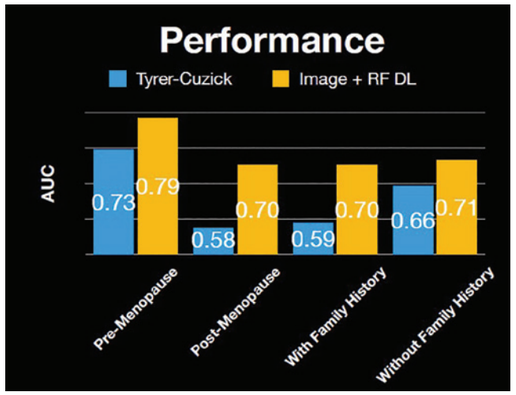


Figure 2. Deep learning model significantly better at predicting future risk of cancer across diverse ages and histories compared to "best practice" advanced risk model Tyrer-Cuzick version 8.

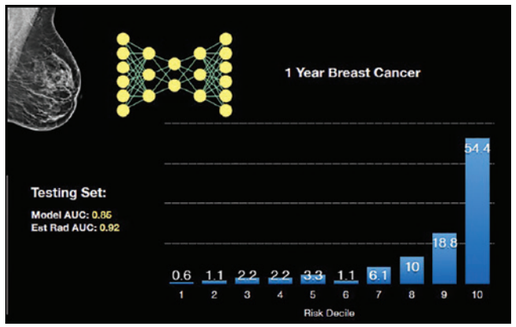


Figure 3. Deep learning model detects cancers on modern digital screening mammograms

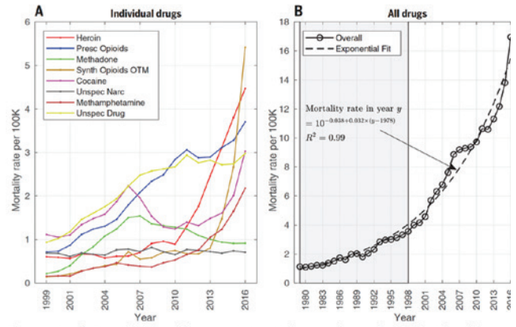


Figure 1. A) Death rates from unintential overdoses for individual drugs from 1999-2016. B) Overdoses from all drugs from 1979-2016 shows drug overdose deaths have been increasing exponentially for the past 40 years. Figure from Jalal et al. Science, 2018.

Figure 2. Features extracted from smartphone and wearables data are analyzed using prediction functions and classifiers from previous training of machine learning algorithms to forecast and detect drug use and overdose.

Using Digital Phenotyping and Machine Learning to Forecast, Detect, and Prevent Drug Overdose Deaths

Lisa Nickerson, PhD
Director, Applied Neuroimaging Statistics Lab, McLean; Assistant Professor, HMS
lisa_nickerson@hms.harvard.edu

Drug overdose deaths have grown exponentially for the past 40 years, with more than 700,000 deaths since the 1990's (Figure 1) and 70,000 in 2017 alone. Astounding increases in opioid-related deaths have lead the US government to declare a national emergency. By 2025, drug overdose deaths will exceed 100,000 each year and nearly half a million people are projected to die from drug overdoses in the next decade. While opioids were involved in nearly 50,000 deaths in 2017, we estimate that there were also nearly a million non-fatal opioid overdoses in 2017. Non-fatal overdoses cause significant financial and medical burdens and greatly increase risk of future overdose. Creative solutions are needed to combat the growth in drug overdose death rates and to slow the opioid epidemic.

Our team is focusing on "digital phenotyping" using personal smartphones and wearables to develop a mobile health ("mHealth") tool for forecasting and detecting drug use, and preventing drug overdose deaths. Personal devices employ numerous internal sensors to make moment-by-moment measurements in a person's natural environment. These data can be used to decode daily activities, such as walking, sitting, standing, eating, and drinking. Digital phenotyping and machine learning can also be used to detect drug use and prevent overdose deaths (Figure 2). mHealth apps for preventing opioid overdose deaths are available, but they don't leverage digital phenotyping and machine learning with commonplace personal devices. For example, some apps trigger an emergency response if the drug user does not stop a timer after they are high that they started before using drugs, while others crowdsource naloxone.

We have identified a great need for an evidence-based mHealth tool that provides useful forecasting, feedback, and interventions for preventing overdose deaths that patients will actually use in their natural environment. Our goal is to develop a commercial product based on comprehensive digital phenotyping in patients, application of machine learning and predictive analytics, design and testing in patients, with clinician input, to meet this need.

Mobile Health Technologies for Monitoring Motor Fluctuations in Patients with Parkinson's Disease

Federico Parisi, PhD
Research Fellow, Wyss Institute for Biologically Inspired Engineering, SRN
fparisi@partners.org

Parkinson's Disease (PD) is a neurodegenerative condition affecting more than 10 million people worldwide. Typical motor symptoms include tremor, bradykinesia, and rigidity. In the late stage of the disease, patients eventually develop motor complications, such as dyskinesias (i.e., hyperkinetic involuntary movements), and start experimenting fluctuations in symptoms' severity. Accurate titration of medications is crucial to minimize the impact of motor complications while maintaining effective the management of PD symptoms.

Current clinical tools for monitoring motor fluctuations rely on sporadic visits in the clinic and patients' self-reports, which provide clinicians only with a fragmented and unreliable picture of the subject's condition. The symbiosis between wearables, mobile devices, wireless technologies, and artificial intelligence, often identified as Mobile Health (mHealth), has the potential to impact the future of healthcare considerably. At the Motion Analysis Lab, we investigate this potential in the context of PD, among others, since almost two decades, collaborating with both academia and industry. For example, we demonstrated that it is possible to use wearables and machine learning to accurately estimate clinical scores for tremor, bradykinesia, and dyskinesia during the performance of standardized tasks and to deploy these models in a web-based platform for remote and longitudinal monitoring of PD subjects in the home setting. Our latest research efforts in this context aim at overcoming the limitations of current approaches, such as the dependency on standardized motor tasks or the insufficient time resolution in the sampling of symptoms' severity across the day.

We recently acquired a new dataset, where we collected data from wearable sensors during a complete medication cycle in both the laboratory and a simulated apartment (equipped with cameras). Our goal is to exploit this dataset, together with the latest AI techniques, to train robust models for continuous and objective tracking of the severity of motor complications in subjects with PD and to validate the developed algorithms using the data collected in the apartment (using video recordings and clinical evaluations as ground truth).

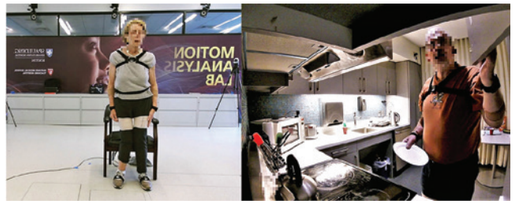


Figure 1. Experimental set-up during a data collection with subjects with PD in the laboratory (right) and the simulated apartment (left) at the SRH.

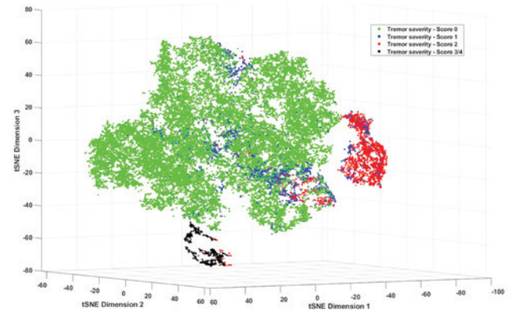


Figure 2. Visualization of severity clusters for resting tremor estimates in a reduced-dimensionality features space.

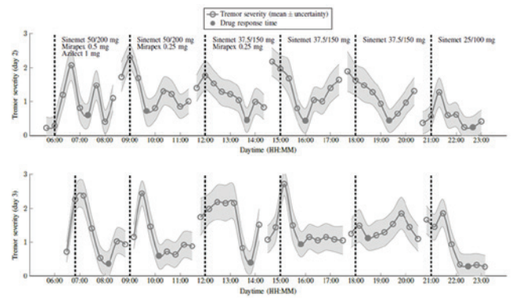


Figure 3. Tremor severity dynamics across different medication cycles for a single subject in the home setting.

AI-Powered Diagnostic Reporting Tool for Spinal MRI of Degenerative Disease

Stuart Pomerantz, MD
Director, Neuro-CT, Neuroradiology, MGH
Instructor, HMS
spomerantz@mgh.harvard.edu

Spinal degeneration is a highly prevalent condition with large societal costs, much of which comes from medical imaging. Spinal MRI, the study of choice, is among the most expensive of imaging procedures yet heavily utilized for both initial diagnosis and longitudinal evaluation. Its interpretation is challenging and time-consuming even for those with sub-specialty expertise.

Due to this complexity and a lack of universally-accepted grading standards, spinal MRI reporting still exhibits a large degree of inter-reader variability frustrating both referring clinicians and their patients, who increasingly demand direct access to their care documentation. The recent advances in convolutional neural-network machine learning methodologies and the availability of the requisite GPU capability has allowed us to fully leverage the immense value in our institutional image and reporting archive to train an AI-based solution for these workflow and reporting issues. DeepSPINE is an automated and accurate system for lumbar spinal MRI analysis and report generation. Its processing pipeline is comprised of deep-learning algorithm components for automated vertebral segmentation, disc-level labelling, optimized reorienting of image slice angulation, and level-by-level grading of central canal and foraminal stenosis with report text generation. Our efforts have achieved state-of-the-art performance in the machine-learning literature (Proceedings of Machine Learning Research 85:1-16, 2018) but algorithmic performance is only one component of a successful solution. Efficient and elegant integration into pre-existing workflows is essential for both rapid deployment and clinician acceptance and thus has been an equally important focus of our efforts.

As an AI-powered reporting tool, DeepSPINE populates algorithmic outcomes into standardized report templates for more consistent grading terminology. Further development will encompass additional features of spinal degeneration and eventually integrate longitudinal imaging analysis and additional patient data elements from the electronic healthcare record to provide value beyond the traditional radiology report such as workflow-prioritization and predictive outcome modelling to aid surgical planning and other therapeutic approaches.

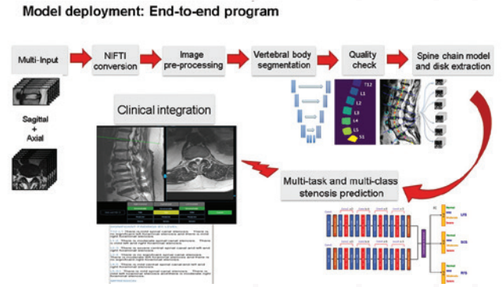


Figure 1. DeepSPINE end-to-end pipeline for automated lumbar spine MRI stenosis grading: multi-input (T2-Sag, T2-axial MRI Series) series extraction, vertebral body segmentation, disc-level labelling, optimized per-level series angle reformatting, multi-task (central and bilateral foraminal) and multi-class (normal, mild, moderate, severe) stenosis grading, and clinical integration with a dynamic user-interface for image and algorithmic report correlation and report text generation and injection into standard departmental dictation platform.

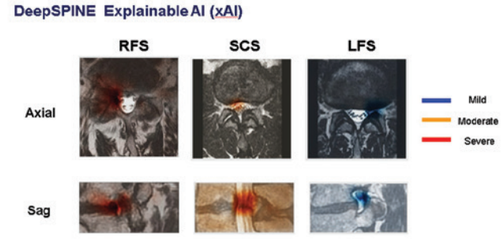


Figure 2. Saliency map generation highlights localized severity of stenosis to aid user validation of DeepSPINE algorithm grading inferences. Most striking is that the highly specific localization is achieved from an algorithm trained from stenosis labels extracted by natural-language-processing from archival reporting without any manual image annotation of anatomic structures or areas of degenerative disease.

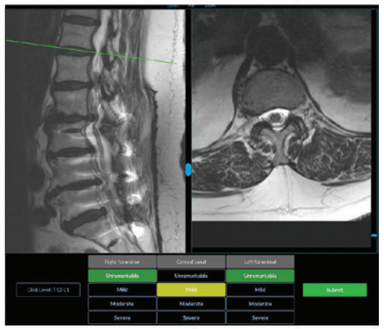


Figure 3. DeepSPINE dynamic user-interface for image analysis and validation of algorithm grading outputs.

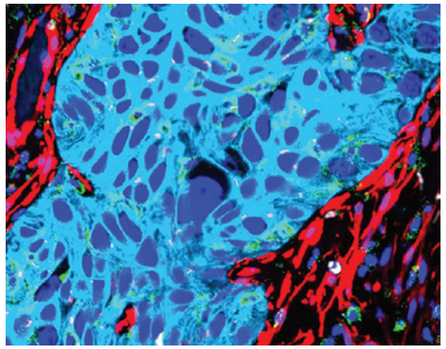


Figure 1. t-CyCIF image of a human squamous cell carcinoma metastasis to the brain.

Multiplexed Tissue Imaging and Quantitative Pathology for Discovery and Translational Medicine

Sandro Santagata MD, PhD
Assistant Professor, Pathology, BH, HMS
ssantagata@bics.bwh.harvard.edu

Pathologists contribute to the care of cancer patients by providing important information relevant to diagnosis, patient prognosis and the prediction of response to treatment. This information is gleaned from standard histopathology review of tissue samples and increasingly from the integration of genomic information. Despite numerous cases of remarkable responses to a new generation of targeted therapies, many tumors remain refractory to treatment and new approaches are needed which may provide critical information for patient stratification, drug target discovery, and precision medicine. Dissociative techniques (e.g. single-cell RNA sequencing), functional techniques (e.g. BH3 profiling or single-cell mass measurements) and spatially-resolved imaging techniques (e.g. mass spectrometry imaging, multiplexed antibody imaging and spatial transcriptomics) are being increasingly deployed in the characterization of tumor samples. Multiplexed antibody imaging has received significant attention as of late, due to the explosion in interest in immuno-oncology and the need to characterize the cells residing within the tumor microenvironment.

Our group has implemented a tissue cyclic immunofluorescence method (t-CyCIF) that is an extension of older approaches to sequentially assemble high-plex images using fluorophore inactivation and/or antibody stripping. t-CyCIF constructs high-dimensional images from clinical archival formalin-fixed, paraffin-embedded (FFPE) tissues by sequential 4-6 color immunofluorescence imaging on a conventional microscope and can be extended to at least 60 antigens. We are developing an information resource at https://www.cycif.org/ comprising a list of reagents, images, caveats and best practices; the site also describes the data analysis, visualization and management software that will be used for the development of pre-cancer and cancer atlases for the Human Tumor Atlas Network (HTAN), an initiative funded by the NCI Biden Moonshot. Spatially-resolved optical imaging of clinical fixed tissues has the potential to provide unprecedented insights into tumor cell biology and to generate new approaches for biomarker discovery and clinical testing.

Artificial Intelligence for Diagnosis and Management in Spine Surgery

Joseph Schwab, MD
Chief, Orthopaedic Spine Surgery, MGH
Associate Professor, HMS
jhschwab@mgh.harvard.edu

In the United States, annual direct costs of degenerative spinal pathologies are more than $100 billion. National inpatient expenses for spinal fusion surgery increased from $10 billion in to over $50 billion in the last two decades (2% of total national health care expenses). Total hospital charges for these patients are heavily driven by postoperative care: complications can double or triple the cost of spinal fusion surgery. Given a lower bound of 5% complication rate, 450,000 fusion surgeries per year, and surgeries with complications costing twice the amount without complications, the addressable market is at least $1.8 billion in the United States (of the $50+ billion) that can be preoperatively predicted and adequately planned to avoid unnecessary costs and poor patient outcomes. Data science is an emerging area in spine care but existing solutions have failed to adapt to innovations in predictive analytics. Existing tools are based on rough estimations from small studies of less than 100-200 patients. The tools are fragmented, lack peer-reviewed publications and external validation, but are the only options for spinal care providers. To address this tremendous need in spine care, we have developed algorithms by using machine learning, natural language processing and deep learning to prevent adverse events and aid decision making. The most rigorous standards for clinical prediction models have been followed in developing these models. Our algorithms have been published by the leading journals in spine care (The Spine Journal, Neurosurgery).

External validation of these algorithms has borne out their utility in diverse populations. The algorithms have been made available for providers as web applications and we have presented our work at the national organizations for spine care. Our usage patterns indicate that these algorithms are accessed daily to guide decision making on topics ranging from prediction of postoperative opioid dependence to determination of expected postoperative survival in spinal metastatic disease. We are actively working on integrating these algorithms into electronic health records to provide a seamless clinical workflow. Additionally, we are exploring the use of these algorithms by both payors and providers to establish risk-adjusted reimbursement plans in population health management.

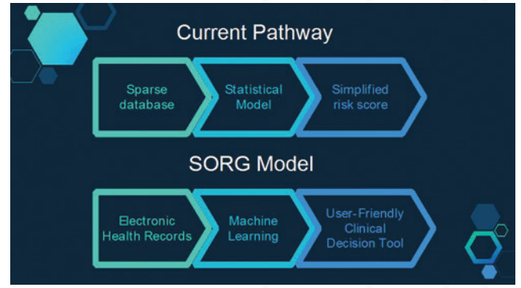


Figure 1. Comparison of SORG to existing tools in spine care

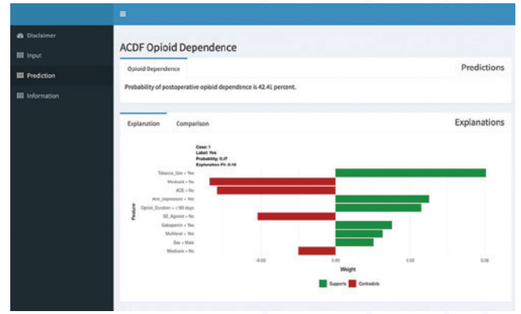


Figure 2. SORG algorithm for prediction of opioid dependence after anterior cervical discectomy and fusion providing both predictions and explanations

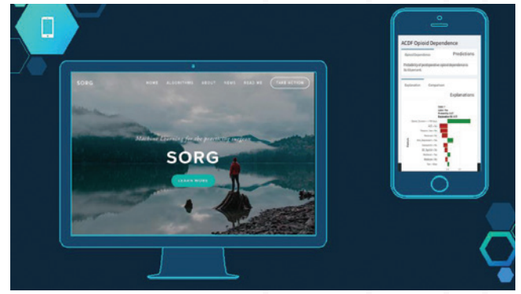


Figure 3. User-interfaces for SORG algorithms, desktop (left), smartphone (right)

Three Computational Techniques and One Tool to Bring the Patient Voice into Care

Chris Sidey-Gibbons, PhD
Co-Director, PROVE Center, BH; Member of Faculty, HMS
cgibbons2@bwh.harvard.edu

Data collected directly from patients has the potential to generate unique insights into the process and outcomes of care. These data are usually collected using structured questionnaires, referred to as patient-reported outcome and experience measures (PROMs and PREMs). Despite their apparent simplicity, both PROMs and PREMs are effective at improving both processes and outcomes of care.

Recent notable examples include increased identification of chemotherapy side-effects which translated into significantly improved survival in studies conducted in the US and Europe. Patient experience data can also be used to identify poorer outcomes including unplanned admissions and reoperations. Reliance on static fixed-length questionnaires has led to several issues with the collection and interpretation of patient-reported data. Many PROMs, designed originally to monitor outcomes in clinical trials, are lengthy and do not translate well to clinical practice. In addition to being lengthy, it is often unclear what clinical actions should be taken based on the information collected. Likewise, patient-reported experience measures also suffer from a lack of actionability, as little insight into potential quality improvement mechanisms can be gleaned from a single numeric score. Research conducted within the Brigham and Women's PROVE Centre has demonstrated that the limitations of patient-reported data tools can be overcome using novel computational tools including computerized adaptive testing and machine learning.

This talk will introduce the current state of the art in the assessment, analysis and feedback of patient-reported assessments by demonstrating three technologies and one tool to facilitate implementation. Specifically, this talk will introduce three machine learning techniques to individually tailor patient-reported assessments, predict individual outcomes for patients undergoing breast reconstruction, and automatically make sense of open-text feedback about doctor's performance. Finally, I will introduce Concerto, an open-source platform to facilitate the development and deployment of secure patient-reported assessments which combine cutting-edge AI techniques and flexible user interfaces.

Gibbons, C., Bower, P., Lovell, K., Valderas, J., & Skevington, S. (2016). Electronic quality of life assessment using computer-adaptive testing. Journal of Medical Internet Research, 18(9).

Gibbons, C., Richards, S., Valderas, J. M., & Campbell, J. (2017). Supervised machine learning algorithms can classify open-text feedback of doctor performance with human-level accuracy. Journal of Medical Internet Research, 19(3).

Figure 1. Overview of the Concerto System as it is implemented for patient-reported outcomes and experience assessment.

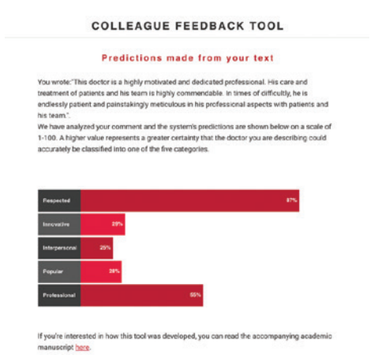


Figure 2. Individual tailored patient feedback with geo-relevant referral links (see Gibbons et al., 2016)

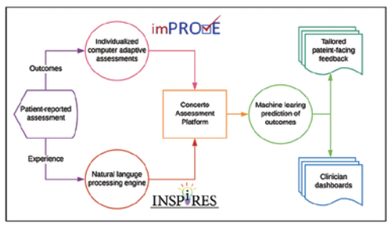


Figure 3. Machine learning classifications of open-text reports of doctor performance (see. Gibbons et al., 2017)

AI-Imaging for Patient-friendly Colon Cancer Screening

Hiroyuki Yoshida, PhD
Director, 3D Imaging Research, MGH
Associate Professor, Radiology, HMS
yoshida.hiro@mgh.harvard.edu

With more than 50,000 annual deaths in the United States, colon cancer is the second leading cause of cancer deaths for men and women in the United States. These deaths could be prevented by early detection and removal of precursor polyps. However, only 62% of adults aged ≥50 years adhere to colorectal screening guidelines today. Laxative (cathartic) bowel preparation, a mandatory pre-exam preparation for optical colonoscopy (OC), has been identified as the single most important barrier to patient adherence to colorectal examinations, especially for the Medicare population, a prime target for colorectal screening, who are fragile to laxative preparation.

Computed tomographic colonography (CTC), also known as virtual colonoscopy, is an alternative complete colon cancer screening method rated as "A" by the US Preventive Services Task Force for the detection of polyps and cancers. Also CTC is usually performed with laxative preparation, whereas we previously developed a computer-assisted laxative-free CTC scheme to eliminate the need for laxatives in colorectal examinations. This could substantially increase the capacity, safety, accuracy, and patient adherence to colorectal examinations. Our multi-center clinical trial showed that laxative-free bowel preparation was easy to tolerate for patients and enabled the detection of large polyps at a sensitivity comparable to that of colonoscopy. However, small polyps were difficult to detect and identify due to a large amount of distracting fecal residue that either hides polyps or mimics their appearance.

We thus employed state-of-the-art deep-learning methods to develop AI colonography, where AI performs virtual bowel cleansing of laxative-free CTC cases by removing the fecal residue from CTC images and then automatically detects and identifies the polyps that would have otherwise disappeared among the abundant distracting fecal residues. Successful deployment of AI colonography is expected to make CTC a patient-friendly yet highly accurate and cost-effective option for large populations, especially for Medicare population, thereby promoting early diagnosis and ultimately reducing mortality due to colorectal cancer.

Deep Learning for Glaucoma Detection

Nazlee Zebardast, MD
Instructor, Ophthalmology, MEE, HMS
nazlee_zebardast@meei.harvard.edu

Glaucoma is the leading cause of irreversible blindness worldwide, with an estimated global prevalence of 3.5% in persons aged 40 to 80 years, and affecting approximately 64 million people worldwide as of 2013.[1] Nearly half of glaucoma cases in the US are undiagnosed with many individuals with significant visual field loss at the time of diagnosis. Glaucoma is an ideal disease to screen for as early treatment of glaucoma, by reduction of intraocular pressure (IOP), has been shown to delay glaucoma progression.[2] Current screening strategies rely on clinical examination and physician-interpreted testing and are expensive, time intensive, and lack adequate specificity.[3] Successful community screening must be simple and scalable. Automated classification of fundus photographs and optical coherence tomography (OCT) images using deep learning (DL) algorithms has the potential for improving screening accuracy, lowering cost and improving accessibility to eye care.

DL has been used previously for detection of glaucoma using fundus photographs. [4,5] In these prior studies, glaucoma cases have been defined by clinical opinion based on ophthalmoscopically-apparent characteristic changes to the optic nerve head. This method of data labeling is inherently subjective and of poor sensitivity and specificity. As any DL algorithm can only attempt to match its reference standard, for improved screening there is a need to improve this standard. It may be possible to utilize objective measures to identify individuals at high risk for glaucoma. There is significant evidence for genetic risk factors that influence glaucoma susceptibility.[6] Additionally, numerous epidemiologic studies have shown that high IOP is a major risk factor for glaucoma[7]. We propose to construct a DL model to classify fundus photographs and OCTs in order to identify individuals with high IOP and high genetic risk for glaucoma, and therefore high risk of glaucomatous optic neuropathy. Additionally, we aim to detect fundus photography and macular OCT biomarkers of elevated IOP and high genetic risk for glaucoma that can be used for screening or prognostication purposes. We hypothesize that these algorithms can be harnessed to find new imaging features predictive of glaucomatous optic neuropathy as well as future risk of the disease.

Ultimately we aim to use disease-predicting fundus image features identified in this study to construct multi-modal models using multiple objective modalities such as IOP, genetic risk, and risk derived from fundus images to improve our detection and prediction rates for glaucoma.

1. Tham YC et al. Ophthalmology. 2014;121(11):2081-2090
2. The Advanced Glaucoma Intervention Study (AGIS). Am J Ophthalmol. 2000;130(4):429-440.
3. Tatemichi M et al. Am J Ophthalmol. 2002;134(4):529-537
4. Li Z et al. Ophthalmology 2018; 125 (8): 1199-1206.
5. Ting DSW et al. JAMA Ophthal 2017; 318 (22): 2211-2223.
6. Wiggs JL, Pasquale LR. Hum Mol Genet. 2017;26(R1):R21-R27.
7. Sommer A. Am J Ophthalmol. 1989;107(2):186-188

Machine Learning and NLP to Track Disease Progression and Predict Health Outcomes

Li Zhou, MD, PhD
Associate Professor/Lead Investigator, BH
Associate Professor, HMS
lzhou@bwh.harvard.edu

Predicting patients' clinical trajectories and providing personalized care interventions at an appropriate time drive high-value care. However, knowing which patients will benefit from which intervention and when is difficult. Our group seeks to develop innovative solutions using machine learning and natural language processing (NLP), leveraging the vast information in longitudinal clinical notes from electronic health records (EHRs).

One of our projects aims to address multiple unmet needs within the palliative care domain for patients with dementia. To discover and understand the illness and care trajectories, we conducted automatic topic modeling to capture trends of various themes (e.g., cognitive function, nutritional status) in a large volume of clinical notes over the last two years of a patient's life. We then developed and validated a deep learning algorithm using clinical notes and demographics to predict mortality risk for use as a proxy indicator to identify dementia patients with unmet palliative care needs. Our 1- and 2-year models reached AUCs of 0.956 and 0.943, respectively. We further validated the algorithm using a set of 266 patients who were screened by clinicians for palliative care interventions using the "surprise question" ("Would you be surprised if this patient died in the next 2 years?"); the model outperformed the clinician screening. In addition, the model can trace and visualize clinically meaningful note topics used in each patient's prediction.

Our work demonstrates that deep learning predictive models hold promise in patient stratification for clinical practice. Next, we will collaborate with our population health team to put our tool into practice and investigate the feasibility of applying this method in clinical settings for identifying patients who will benefit from early palliative care interventions. Our methods are generalizable to a wide-range of applications and could be used in other clinical areas to predict other important outcomes, such as hospital readmission. We have also developed similar approaches and applied our NLP solution, MTERMS, for early detection of clinical deterioration in hospitalized patients using nursing notes.

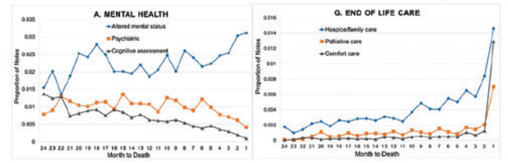


Figure 1. Example figures demonstrating how topic modeling can support descriptive analysis to understand the illness and healthcare trajectories for patient with dementia. Figure 1.A shows that the documentation of altered mental status increased towards the end of life, while the topic trends for cognitive assessment and psychiatric disorder decreased. Figure 1.G. shows that all categories related to end of life care were trending up in the last months of life. The topics about family/hospice care began to increase around a year before death, but documentation of palliative care and comfort care topics did not begin to rise until the last 2 months of life. (Wang L, et al, AMIA 2018 Symposium)

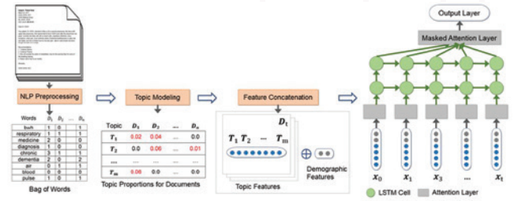


Figure 2. Overview of the predictive modeling using longitudinal clinical notes and demographics

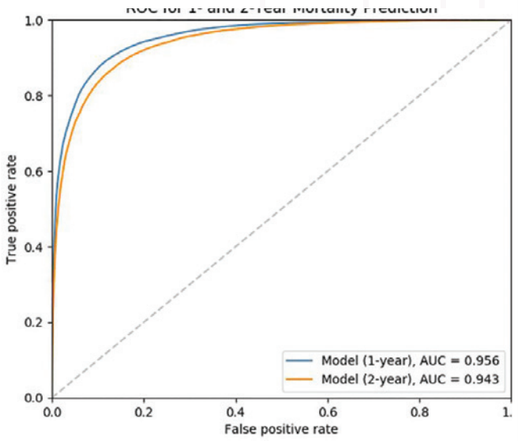


Figure 3. Receiver Operating Characteristics (ROC) of the models in predicting 1-year and 2-year mortality

Notes

Notes

WORLD MEDICAL INNOVATION FORUM

sponsors

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

presenting

BAYER

stakeholder

Bristol-Myers Squibb

GE Healthcare

NUANCE

NVIDIA

strategic

AMGEN

Biogen

SIEMENS Healthineers

Wolters Kluwer

collaborator

Boston Scientific
Advancing science for life™

Canon

MGH & BWH CENTER FOR CLINICAL DATA SCIENCE

FUJIFILM
Value from Innovation

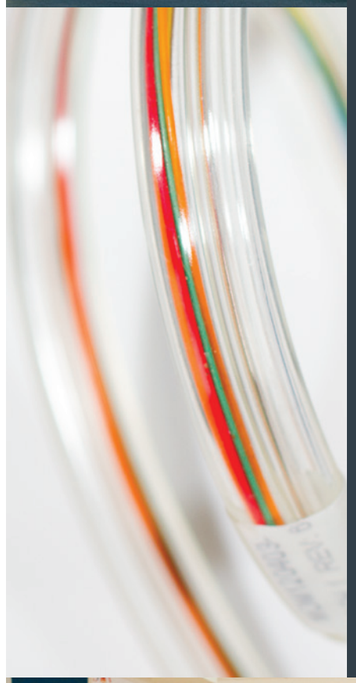
MINTZ

北极光创投 northern light VENTURE CAPITAL

PERSISTENT

PURESTORAGE

REDHILL 丹麓资本 CAPITAL



Partners HealthCare

INNOVATION

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

PARTNERS HEALTHCARE | INNOVATION

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

PARTNERS HEALTHCARE | INNOVATION

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

PARTNERS HEALTHCARE | FOUNDED BY BRIGHAM AND WOMEN'S HOSPITAL AND MASSACHUSETTS GENERAL HOSPITAL

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