

Pittsburgh Health Data Alliance

Carnegie Mellon University University of Pittsburgh UPMC

Center for Machine Learning and Health
Carnegie Mellon University

Project Overviews

Table of Contents

Current Project Overviews Pages 2-8

Machine Learning based Localization of Focal Epilepsy from Magnetoencephalography	Page 2
Scaling Remote Extracorporeal Membrane Oxygenation (ECMO) Training with Intelligent Agents	Page 3
Smart Sensors for Evaluating and Improving Cochlear Implants	Page 4
Improving Breast Cancer Diagnosis with Interpretable Machine Learning	Page 5
Evaluating the Predictive Capability of Machine Learning Algorithms in Adult Cardiac Surgery	Page 6
Integrating Deep Learning with High Throughput Materials Engineering for Detecting Noroviruses	Page 7
Towards Automated Multimodal Behavioral Screening for Depression	Page 8

Completed Project Overviews Pages 9-26

Brain Tsunamis	Page 9
Developing a Biomarker for Alzheimer's Disease Using Machine Learning and Immune Cell Epigenomics	Page 10
Non-Invasive Intracranial Pressure Monitoring	Page 11
Sepsis Phenotyping from Electronic Health Records	Page 12
Diagnosis Coding Engine for Electronic Health Records	Page 13
Clinical Genomics Modeling Platform	Page 14
Deep Learning for Placental Pathology	Page 15
Diagnosis Coding Engine for Electronic Health Records	Page 16
In-Home Movement Therapy Data Collection	Page 17
Ingestible Impedance Sensors to Acquire Large-Scale Data Sets from Patients with Eosinophilic Esophagitis	Page 18
Machine Learning Algorithms for Advanced Manufacturing of High-Fidelity 3D Printed Biomaterials	Page 19
Real-Time Tool to Predict Clinical Outcomes After Cardiac Arrest	Page 20
MyHealthyPregnancy	Page 21
Phylogenetic Models for Predicting Cancer Progression	Page 22
Computational Modeling of Behavioral Rhythms to Predict Readmissions	Page 23
Programming Framework for Managing Patient Privacy	Page 24
Healthcare Trails and Care Coordination Using Admissions, Discharges, and Transfers Data	Page 25
Detecting Intestinal Activity by Analyzing Gut Sounds	Page 26



Machine Learning based Localization of Focal Epilepsy from Magnetoencephalography

PROJECT SUMMARY

Epilepsy affects over 65 million people worldwide. For approximately 30% of the 3.4 million Americans with epilepsy, the seizures are not controlled by medical therapy alone. Focal epilepsy, which are seizures that begin in a focal region of the brain, represents the most common type of drug resistant epilepsy. Currently, treatment options for these patients are limited to epilepsy surgery, vagus nerve stimulation, or enrollment in experimental protocols, such as brain stimulation. Epilepsy surgery has the best chance of curing epilepsy but is only an option if the brain region generating seizures can be accurately localized and safely removed. Unfortunately, many patients are not candidates for epilepsy surgery because the brain region responsible for generating seizures cannot be localized and mapped, or seizures originate from the eloquent cortex that cannot be removed without neurological deficits. These results highlight the need for improved diagnostic tools and therapeutic options for drug resistant epilepsy.

Researchers for this project includes Bin He, PhD, Trustee Professor of Biomedical Engineering at Carnegie Mellon University and Anto Bagic, MD, Professor of Neurology and Chief of Epilepsy Division at the University of Pittsburgh Medical School.

The goal of this proposed research project is to develop a novel machine learning (ML)-based technology that can image and localize epileptogenic brain from noninvasive magnetoencephalography (MEG) measurements to aid surgical planning in patients with drug resistant epilepsy.

Determining the epileptogenic brain is of crucial significance for the successful surgical treatment of intractable epilepsy. The successful completion of the proposed research promises to lead to establishment of a disruptive technology, which can significantly advance state-of-the-art management of intractable epilepsy, by means of an innovative ML approach. The development of such technology to localize and image the epileptogenic zone in drug resistant epilepsy patients will benefit numerous patients and the healthcare system.

RESEARCHERS

Bin He, PhD

Trustee Professor of Biomedical Engineering
Carnegie Mellon University

Anto Bagic, MD

Professor of Neurology and Chief of
Epilepsy Division at the University of
Pittsburgh Medical School
University of Pittsburgh

Scaling Remote Extracorporeal Membrane Oxygenation (ECMO) Training with Intelligent Agents

PROJECT SUMMARY

The success of surgical training relies on the close connection between instructors and trainees. In case-based simulations for extracorporeal membrane oxygenation (ECMO), for example, trainees perform tasks such as cannulation while instructors observe, guide, and correct. Such on-site training, however, does not scale to large numbers of trainees, and instructors must travel to training sites, limiting the availability of the training. Remote training using Augmented Reality and Virtual Reality has the potential to enable scalable and safe training. Current approaches, however, do not provide high quality training in a scalable manner.

In the proposed research, CMU's team of David Lindlbauer, PhD, Assistant Professor at the Human-Computer Interaction Institute, and Jean Oh, PhD, Associate Research Professor at the Robotics Institute, in collaboration with Raj Ramanan, MD, Assistant Professor of Critical Care Medicine at UPMC and Medical Director UPMC Procirca ECMO Education, aims to improve the quality and scale of ECMO training by enabling a tight connection between instructors and trainees in a remote setting. This requires instructors to track which trainee requires assistance and who shows a high level of proficiency. To address this challenge, they propose to pair a novel virtual remote training environment with novel intelligent agents that observe, track, guide, and assist instructors and trainees during the training. As a result, instructors can deliver training to a larger number of trainees without sacrificing the quality of learning experience, enabling highly scalable remote training. The proposed work is a vital step toward training medical experts in a scalable and efficient way.

The team proposes to address the challenge of providing a scalable, immersive, and tightly connected training experience to instructors and trainees. The technological innovation lies in combining a novel advanced remote training environment with an intelligent agent that supports instructors when working with dozens of trainees simultaneously. This work is the first to focus on the close connection between instructors and trainees, while providing ECMO training in a remote manner. This will allow for training more clinical professionals in ECMO and provide recurrent rather than one-time training.

The researchers see their work as a first step toward improving clinical training in general. The software of the remote training environment will enable them to incorporate other training scenarios such as the closely related extracorporeal cardiopulmonary resuscitation (ECPR) or other parts of clinical practice such as minimally invasive procedures, all simply by designing new environments.

RESEARCHERS

David Lindlbauer, PhD

Assistant Professor at the Human-Computer Interaction Institute
Carnegie Mellon University

Jean Oh, PhD

Associate Research Professor at the Robotics Institute
Carnegie Mellon University

Raj Ramanan, MD

Assistant Professor of Critical Care Medicine at UPMC and Medical Director UPMC Procirca ECMO Education
UPMC

PROJECT SUMMARY

Hearing loss is one of the leading causes of disability worldwide, responsible for an estimated global cost of almost \$1 trillion annually. Hearing loss is also personally devastating, as it causes social isolation and is a large risk factor for dementia. When hearing loss progresses beyond the point where hearing aids are effective, acoustic hearing can be augmented with electronic hearing via a cochlear implant (CI). These life changing devices are increasingly seeing use amongst non-deaf patients who have residual acoustic hearing. In these cases, preserving this hearing during implantation yields better hearing outcomes by allowing simultaneous use of a hearing aid and CI. In addition, the acoustic hearing can be lifesaving in emergency situations, such as hearing a fire alarm while the implant is turned off.

Currently, suboptimal cochlear implantation results in 10-15 dB reduction of acoustic hearing on average after the surgery, although some patients experience no detectable loss in hearing due to implantation and others are left completely deaf. In all, only 40% of patients have any residual hearing preservation after surgery. The substantial risk posed to residual hearing due to the implantation limits the eligible population for these devices and is a major barrier to adoption amongst the current eligible population.

Carnegie Mellon University's Maysam Chamanzar, PhD, Professor of Electrical and Computer Engineering and Jay Reddy, Postdoctoral Research Associate of Electrical and Computer Engineering are working on this project. They propose to integrate smart sensors onto the CI electrode arrays, for the first time, to provide real-time quantitative feedback during implantation surgery. Adjusting the insertion technique to minimize these quantitative metrics will aid residual hearing preservation. This will, in turn, expand the eligible patient population for CIs and combat hearing loss by reducing the risk of the implantation procedure. The long-term goal of this project is to develop a novel smart sensor array that attaches to current CI electrode array designs to give feedback to the surgeon during implantation, resulting in better hearing for more people.

RESEARCHERS

Maysam Chamanzar, PhD

Professor of Electrical and Computer Engineering
Carnegie Mellon University

Jay Reddy

Postdoctoral Research Associate of Electrical and Computer Engineering
Carnegie Mellon University

PROJECT SUMMARY

Our goal aims to increase the effectiveness of breast cancer screenings by using AI to enhance current processes for continually training radiologists using clinical data. To motivate one possible avenue to improve care, we note the well-documented and significant inconsistencies in how various doctors evaluate mammogram screenings.

Recall rates (the fraction of patients asked to return for more targeted diagnostic testing) vary widely across doctors. In current quality initiatives, doctors periodically review past cases, often focusing on those that have been retrospectively confirmed to be miscalled. We believe that deep learning can enhance the selection process for review in the paraclinical setting, helping to increase cancer detection rates and reduce false positives, leading to a more effectively treated patient population.

By using hundreds of thousands of mammogram screenings and modeling not only ground truth labels, but also the tendencies of each radiologist, our AI solution will identify ideal training cases for review, e.g., by selecting those instances where either specific doctors are predicted to disagree with the ground truth, or cases where specific doctors are predicted to disagree with each other. Our solution will utilize state-of-the-art computer vision techniques together with human-centered visualizations and design principles, to make the results of the AI interpretable to medical experts with limited AI expertise.

RESEARCHERS

Pulkit Grover, PhD

Assistant Professor of Electrical and Computer Engineering
Carnegie Mellon University

Shawn Kelly, PhD

Senior Systems Scientist, Institute for Complex Engineered Systems
Carnegie Mellon University

Marlene Behrmann, PhD

Professor, Psychology
Carnegie Mellon University

Michael Tarr, PhD

Department Head and Professor,
Psychology Carnegie Mellon University

Jonathan Elmer, MD

Assistant Professor of Critical Care Medicine and Emergency Medicine
University of Pittsburgh

Lori A. Shutter, MD

Vice Chair of Education, Professor of Critical Care Medicine, Neurology and Neurosurgery
University of Pittsburgh

PROJECT SUMMARY

A team from UPMC/Pitt and CMU is investigating the application of machine learning to clinical cardiac surgery in ways that could have profound implications in clinical trials, therapy selection, patient prognostication and counseling, and surgeon and hospital evaluations.

According to the American Heart Association, one in three Americans have heart disease. The cost of treating heart disease is expected to triple from \$273 billion in 2010 to \$818 billion in 2030. This is largely attributable to the aging population. In addition, a shortage of cardiac surgeons may be at hand, with the number of active cardiothoracic surgeons expected to decline from 3,708 to 3,411 between 2010 and 2025. Taken together, these trends have profound implications for cardiac surgery as a field. The increased demand, coupled with decreased supply, will underscore the need for cost-effective, personalized care that optimizes outcomes for patients with surgical heart disease.

Risk modeling has played a vital role for many years in adult cardiac surgery – with the Society of Thoracic Surgeons (STS) database created in 1989 as an initiative for quality improvement and patient safety. The team of researchers from UPMC/Pitt and CMU is investigating how machine learning might enhance risk modeling. There are opportunities to apply machine learning to data that is already used for risk modeling, and to incorporate vast amounts of additional data that are routinely collected but not used in current models. The ultimate goals of this research are to improve the accuracy of risk models in adult cardiac surgery, introduce new and potentially predictive data, and develop clinical platforms and technology that can provide real-time guidance to physicians in clinical management of cardiac surgery patients to improve outcomes.

RESEARCHERS

Kyle Miller

Project Scientist, Chemical Engineering
Carnegie Mellon University

Arman Kilic, MD

Assistant Professor, Cardiothoracic Surgery
and Director, Surgical Quality and Analytics
for Cardiac Surgery; Co-Director, Center for
Cardiovascular Outcomes and Innovation
UPMC

Integrating Deep Learning with High Throughput Materials Engineering for Detecting Noroviruses

PROJECT SUMMARY

Researchers from CMU's Mechanical Engineering Department are developing technologies for rapidly detecting if surfaces are contaminated with norovirus, a major cause of epidemic gastroenteritis.

Norovirus, sometimes referred to as the winter vomiting bug, results in about 685 million cases of disease and 200,000 deaths globally each year. About half of food-borne disease outbreaks in the United States are caused by norovirus.

Noroviruses are very contagious, and less than 20 virus particles can cause an infection. One of the challenges faced by healthcare facilities, like hospitals, is preventing contamination by and transmission of noroviruses. The spread of the virus is of particular concern during winter, when ventilation of buildings may be reduced.

Mechanical Engineering's B. Reeja Jayan and Amir Barati Farimani are introducing highly sensitive and specific sensors that can rapidly detect noroviruses and significantly reduce the chance of these infections and epidemics, especially in healthcare facilities. Their research combines deep learning techniques with molecular-scale materials engineering to allow for quick, sensitive detection of norovirus.

RESEARCHERS

Reeja Jayan, PhD

Assistant Professor, Mechanical Engineering,
Materials Science and Engineering
Carnegie Mellon University

Amir Barati Farimani, PhD

Assistant Professor, Mechanical Engineering
Carnegie Mellon University

PROJECT SUMMARY

Depression is a leading cause of disability worldwide. Effective, evidence-based treatments for depression exist but many individuals suffering from depression go undetected and therefore untreated. Efforts to increase the accuracy, efficiency, and adoption of depression screening thus have the potential to minimize human suffering and even save lives.

Recent advances in computer sensing technologies provide exciting new opportunities to improve depression screening, especially in terms of their objectivity, scalability, and accessibility. Professor Morency and Dr. Szigethy are collaborating to develop sensing technologies to automatically measure subtle changes in individuals' behavior that are related to affective, cognitive, and psychosocial functioning. Their goal is to develop and refine computational tools that automatically measure depression-related behavioral biomarkers and to evaluate the clinical utility of these measurements.

RESEARCHERS

Louis-Philippe Morency, PhD

Finmeccanica Associate Professor,
Language Technologies
Carnegie Mellon University

Eva Szigethy, MD, PhD

Professor, Psychiatry, Medicine and
Pediatrics; Director, Behavioral Health with
the Chief Medical and Scientific Office
University of Pittsburgh

Completed Project Overviews

Brain Tsunamis

PROJECT SUMMARY

The last decade of research has established that Cortical Spreading Depolarizations (CSDs), or "Brain Tsunamis," play an important role in many disorders, including Traumatic Brain Injury (TBI), stroke, hemorrhage, and migraine, that collectively affect more than a billion people worldwide and are major causes of death and disability. CSDs are waves of neurochemical changes that propagate across the brain surface and are thought to mediate secondary brain damage following an injury or event.

The only reliable way to measure CSDs now is an invasive method using electrodes; it is costly, risky, and inappropriate for patients with certain contraindications. In addition, detection is slow and requires a physician to visually inspect the data. Diagnosis and treatment in real-time thus is not realistic in most clinical settings.

Our solution is a Noninvasive Platform for Automated CSD Detection and Suppression. Our multidisciplinary team will develop algorithms and techniques to noninvasively suppress CSDs by using online detection and ensuing stimulation. This is the critical first step for discovery of novel therapeutic strategies and would eliminate the need for invasive monitoring. Real-time detection would allow rapidly iterated, individualized care, timely intervention, and enhanced patient outcomes.

CURRENT STATUS

The research team is continuing to carry out commercially orientated research on this topic.

RESEARCHERS

Pulkit Grover, PhD

Assistant Professor of Electrical and Computer Engineering
Carnegie Mellon University

Shawn Kelly, PhD

Senior Systems Scientist, Institute for Complex Engineered Systems
Carnegie Mellon University

Marlene Behrmann, PhD

Professor, Psychology
Carnegie Mellon University

Michael Tarr, PhD

Department Head and Professor,
Psychology Carnegie Mellon University

Jonathan Elmer, MD

Assistant Professor of Critical Care Medicine and Emergency Medicine
University of Pittsburgh

Lori A. Shutter, MD

Vice Chair of Education, Professor of Critical Care Medicine, Neurology and Neurosurgery
University of Pittsburgh

Developing a Biomarker for Alzheimer's Disease Using Machine Learning and Immune Cell Epigenomics

PROJECT SUMMARY

Researchers from CMU's School of Computer Science and the University of Pittsburgh School of Medicine are combining areas of expertise and techniques that span machine learning, immunology, and genomics to develop new methods of diagnosing Alzheimer's disease (AD) that are less expensive than current standards and that may help researchers discover possible treatments.

Currently 5.7 million Americans have AD dementia, including approximately one out of every 10 people over the age of 65. As the population ages, it is projected that 13.8 million Americans will have the disease by 2050. The estimated cost of Alzheimer's disease is \$277 billion annually. Unfortunately, there are still no highly effective treatments or cures for AD.

This team is developing biomarkers that would enable AD to be diagnosed with a blood test, rather than by use of expensive brain imaging techniques. CMU's Andreas Pfenning has proposed a solution that leverages research in epigenetics and knowledge of the immune mechanisms underlying AD predisposition to develop a novel biomarker. The team will also continue to investigate the impact of genetic changes on AD predisposition and progression of the disease. The team hopes to leverage these models to inform future research and the design of effective therapeutics for the treatment of AD.

CURRENT STATUS

The research team is continuing to carry out commercially orientated research on this topic.

RESEARCHERS

Andreas R. Pfenning

Assistant Professor, Computational Biology
Carnegie Mellon University

Joshua C. Cyktor

Research Assistant Professor, Associate
Director, Virology Laboratory Division of
Infectious Diseases
University of Pittsburgh

John W. Mellors

Distinguished Professor of Medicine; Chief,
Division of Infectious Diseases
University of Pittsburgh Alzheimer's Disease
Research Center

Non-Invasive Intracranial Pressure Monitoring

PROJECT SUMMARY

The healthy brain maintains a relatively constant blood flow even when there are changes in blood pressure or intracranial pressure (ICP). The mechanism preserving blood flow is called cerebral autoregulation (CA), which is known to be impaired in a variety of diseases, such as diabetes, Parkinson's disease, stroke, and Traumatic Brain Injury (TBI).

ICP currently can only be measured by placing an invasive pressure sensor inside the brain.

This project seeks to establish an optical technique called near-infrared spectroscopy (NIRS) as a noninvasive method to monitor ICP in humans. Development of a noninvasive ICP sensor would optimize patient treatment in cases where invasive ICP is not possible or could be dangerous to the patient.

CURRENT STATUS

The team received follow-on funding after the completion of their PHDA project and continues research toward creating a commercial solution.

RESEARCHERS

Jana Kainerstorfer, PhD

Assistant Professor, Biomedical Engineering
Carnegie Mellon University

Elizabeth Tyler-Kabara, MD, PhD

Associate Professor, Neurological Surgery,
Bioengineering, Physical medicine and
Rehabilitation
UPMC Children's Hospital of Pittsburgh

Sepsis Phenotyping from Electronic Health Records

PROJECT SUMMARY

Sepsis is a life-threatening organ dysfunction caused by a dysregulated host response to infection that has a high mortality rate and is a large burden on the healthcare industry, constituting over \$24 billion annually.

The team will use machine learning and artificial intelligence methods to develop analytic tools to identify sepsis earlier and define subgroups of patients who are at high risk for sepsis and share other traits. Successful completion of this research will advance efforts to identify subgroups with sepsis for whom particular treatments are more effective, thereby reducing morbidity and mortality. It will also support the development of physician-competitive, health records-based risk scores that can be used for risk stratification and for early warning clinical decision support. This proposal leverages an established collaboration between experts in machine learning and sepsis at CMU, Pitt, and UPMC.

CURRENT STATUS

The research team has wrapped up this project and is pursuing new directions and applications for their innovative approaches.

RESEARCHERS

Jeremy Weiss, MD, PhD

Assistant Professor, Health Informatics
Carnegie Mellon University

Christopher Seymour, MD

Associate Professor, Critical Care and
Emergency Medicine
University of Pittsburgh

PROJECT SUMMARY

Medical diagnostic errors impact 12 million adults each year in the U.S. The number of deaths due to medical diagnostic errors is the third leading cause of death, equivalent to the crash of a large aircraft every day, based on estimates in the U.S. alone. A key reason why diagnostic errors are made – even by the best clinicians in highly reliable organizations – is the increasing complexity of the diagnostic process, with over 10,000 diseases and 5,000 laboratory tests to choose from.

This project focuses specifically on preventing coding and billing errors. To address this cognitively complex problem, the team is developing an engine that will predict likely diagnosis codes based on information available in a patient's electronic health record. Specifically, the solution will review both structured and unstructured data, such as clinical notes, and apply a machine learning-based mapping from these data to specific diagnosis codes.

When the diagnosis code suggestions are acknowledged and accepted, these can then also flow into billing codes through electronic medical records or other systems. The solution will thus aid clinicians in their medical management, decision support, accurate documentation, billing, and quality improvement.

CURRENT STATUS

The research team is continuing to carry out commercially orientated research on this topic.

RESEARCHERS

Tamar Krishnamurti, PhD

Assistant Research Professor,
Engineering and Public Policy
Carnegie Mellon University
Assistant Professor, Medicine
University of Pittsburgh

Hyagriv Simhan, MD, MS

Professor, Obstetrics, Gynecology and
Reproductive Science; Division Chief,
Maternal Fetal Medicine
Magee Women's Hospital

Alexander Davis, PhD

Assistant Professor, Engineering
and Public Policy
Carnegie Mellon University

PROJECT SUMMARY

The Clinical Genomics Modeling Platform is an engine for easily building precision-medicine models for various diseases and populations. Triage algorithms, for instance, might help to determine if patients with a certain disease should be sent home with monitoring or sent to the intensive care unit.

This project focuses on the understanding of the relationship between genetic variation and medical outcomes in a large population, which is key to realizing the vision of personalized medicine. Efforts are now underway to obtain the genome sequences of thousands of individuals. As cost of sequencing continues to drop, it will become routine to sequence patients in a medical setting. However, a number of computational and practical challenges remain in the way of using genomic sequencing for clinical decision making. In this project, Dr. Carl Kingsford and Dr. Christopher Langmead are aiming to increase the usage of predictive systems based on machine learning techniques and genomics through the development and commercialization of a computational system that will:

- Make predictive models easy to build for clinical researchers
- Make predictive models easy to share (sell) and apply
- Make results of models easy to understand

CURRENT STATUS

The research team is continuing to carry out commercially orientated research on this topic.

RESEARCHERS

Carl Kingsford, PhD

Herbert A. Simon Professor,
Computer Science
Carnegie Mellon University

Christopher Langmead, PhD

Associate Professor, Computational Biology
Carnegie Mellon University

PROJECT SUMMARY

A collaborative team from CMU, Pitt, and UPMC is developing image analysis software to examine slide images of placentas following delivery and identify abnormalities that require review by a pathologist. Abnormalities found in these images can provide critical information about the health of the infant and mother.

Though the placenta receives a routine examination in the delivery room, experienced pathologists – especially those with perinatal subspecialty expertise – are not always available to handle the volume of incoming placentas requiring examination. The software being developed in this project thus would reduce the risk that physicians would miss detecting otherwise avoidable complications.

The researchers will apply deep learning techniques to create an algorithm that will support high-throughput examination of whole-slide images of placentas and identification of placentas that most need a pathologist's attention. This approach will be enabled by using the digitized whole slide imaging dataset of placenta slides at UPMC, with the current rich dataset available from the Magee Obstetric Maternal & Infant (MOMI) Biobank of placenta images, coupled with pathology metadata, clinical data regarding the delivery, and long-term data about the infant's health.

CURRENT STATUS

The research team is continuing to carry out commercially orientated research on this topic.

RESEARCHERS

Jonathan Cagan, PhD

George Tallman and Florence Barrett
Ladd Professor, Mechanical Engineering
Carnegie Mellon University

Philip R. LeDuc, PhD

William J. Brown Professor,
Mechanical Engineering, Biomedical
Engineering, Computational Biology,
and Biological Sciences
Carnegie Mellon University

Janet M. Catov, PhD, MS

Associate Professor, Department of
Obstetrics, Gynecology and Reproductive
Sciences and Department of Epidemiology
University of Pittsburgh

Liron Pantanowitz, MD

Vice Chairman for Pathology Informatics
and Director of Cytopathology
UPMC Shadyside

PROJECT SUMMARY

Medical diagnostic errors impact 12 million adults each year in the U.S. The number of deaths due to medical diagnostic errors is the third leading cause of death, equivalent to the crash of a large aircraft every day, based on estimates in the U.S. alone. A key reason why diagnostic errors are made – even by the best clinicians in highly reliable organizations – is the increasing complexity of the diagnostic process, with over 10,000 diseases and 5,000 laboratory tests to choose from.

This project focuses specifically on preventing coding and billing errors. To address this cognitively complex problem, the team is developing an engine that will predict likely diagnosis codes based on information available in a patient's electronic health record. Specifically, the solution will review both structured and unstructured data, such as clinical notes, and apply a machine learning-based mapping from these data to specific diagnosis codes.

When the diagnosis code suggestions are acknowledged and accepted, these can then also flow into billing codes through electronic medical records or other systems. The solution will thus aid clinicians in their medical management, decision support, accurate documentation, billing, and quality improvement.

CURRENT STATUS

The research team is continuing to carry out commercially orientated research on this topic.

RESEARCHERS

Tamar Krishnamurti, PhD

Assistant Research Professor,
Engineering and Public Policy
Carnegie Mellon University
Assistant Professor, Medicine
University of Pittsburgh

Hyagriv Simhan, MD, MS

Professor, Obstetrics, Gynecology and
Reproductive Science; Division Chief,
Maternal Fetal Medicine
Magee Women's Hospital

Alexander Davis, PhD

Assistant Professor, Engineering
and Public Policy
Carnegie Mellon University

PROJECT SUMMARY

A team from CMU, Pitt, and UPMC is working on a project to evaluate the ability of cameras, inertial measurement units, and other sensors, in combination with machine learning (ML) algorithms, to assess patients' movement therapy exercises in the home. The long-term goal is to increase the quality and efficacy of physical therapy (PT) by providing patients with automated, in-home monitoring, near real-time feedback on exercise performance, and feedback to providers when issues arise outside of the PT setting. This will enable providers to prioritize clinic time for patients whose recovery has stalled while avoiding unnecessary appointments for those who are progressing satisfactorily.

CURRENT STATUS

The research team is continuing to carry out commercially orientated research on this topic.

RESEARCHERS

Jessica Hodgins, PhD

Professor, Robotics Institute
Carnegie Mellon University

Dan Siewiorek, PhD

Buhl University Professor,
Electrical and Computer Engineering
and Computer Science
Carnegie Mellon University

Asim Smailagic, PhD

Research Professor, Institute for Complex
Engineered Systems and Director,
Laboratory for Interactive Computer
Systems and Wearable Computers
Carnegie Mellon University

Andrew Whitford, PhD

Senior Research Associate, Robotics
Institute
Carnegie Mellon University

Keelan Enseki, MS, PT

Director, Clinical Practice Innovation
UPMC Center for Rehab Services

Adam Popchak, PhD

Research Assistant Professor,
Department of Physical Therapy
University of Pittsburgh

Ingestible Impedance Sensors to Acquire Large-Scale Data Sets from Patients with Eosinophilic Esophagitis

PROJECT SUMMARY

An ingestible sensor may provide a faster, more convenient alternative to biopsies for diagnosing esophageal disorders, such as eosinophilic esophagitis (EoE). EoE is an inflammatory disorder of the gastrointestinal tract that can cause persistent feeding problems, vomiting, and abdominal pain.

Christopher J. Bettinger of CMU's Biomedical Engineering Department is developing an ingestible, flexible sensor that could be used to detect and monitor EoE and other esophageal disorders. EoE affects more than 150,000 Americans each year and the incidence of EoE is increasing rapidly, especially in Western Pennsylvania. It is thought to arise from the patient's immune response to specific foods and, when properly diagnosed, can be treated by identifying and eliminating those foods. But timely diagnosis is challenging because it currently relies on biopsies, which are slow, expensive, and highly invasive.

If EoE could be reliably monitored on a more frequent basis, the resulting data could be analyzed in conjunction with a patient's diet using machine-learning algorithms to identify and eliminate specific foods that may be responsible for causing EoE on a case-by-case basis.

CURRENT STATUS

The research team is continuing to carry out commercially orientated research on this topic.

RESEARCHERS

Christopher Bettinger, PhD
Associate Professor, Biomedical
Engineering
Carnegie Mellon University

Machine Learning Algorithms for Advanced Manufacturing of High-Fidelity 3D Printed Biomaterials

PROJECT SUMMARY

CMU's Newell Washburn, Philip R. LeDuc, and Jennifer Bone are designing a tool that would enable the 3D printing of biomaterials, removing one of the obstacles to using additive manufacturing to produce surgical implants and organ transplants.

The need for bioprinted implants and organs is significant. At any given time, nearly 3,500 – 4,000 people are waiting for a heart or heart-lung transplant, and every 10 minutes a new person is added to the national transplant waiting list. Furthermore, the prevalence of rejection and immunosuppression currently impact the success of transplantation and highlight the need for patient-specific transplants that will increase the rate of survival. Due to this high demand, the market for patient-specific implants is projected to reach \$1 billion by 2020.

Based on hierarchical machine learning (HML), this work will enable the manufacturing of high-fidelity 3D constructs from various starting materials, including biological components. The tool will be used by designers, manufacturers, and others who require the ability to integrate information from various models in order to leverage 3D bioprinting for efficient and accurate manufacturing.

CURRENT STATUS

The research team is continuing to carry out commercially orientated research on this topic.

RESEARCHERS

Newell Washburn, PhD

Associate Professor, Chemistry,
Biomedical Engineering
Carnegie Mellon University

Philip LeDuc, PhD

William J. Brown Professor,
Mechanical Engineering,
Biomedical Engineering, Computational
Biology, and Biological Sciences
Carnegie Mellon University

Jennifer Bone

PhD Candidate, Biomedical Engineering
Carnegie Mellon University

Real-Time Tool to Predict Clinical Outcomes After Cardiac Arrest

PROJECT SUMMARY

Over 150,000 Americans are treated in hospitals each year after cardiac arrest, most of whom are initially comatose. Once their hearts are revived, most such patients eventually die in the hospital from brain injury, but a sizable minority will awaken and have a good neurological recovery. With current technology, no sign, symptom, or test result in the first 72 hours after cardiac arrest is sufficient to preclude a favorable recovery. Prognosis often remains uncertain for days or even weeks thereafter. The project team is developing technology that will use advanced signal processing and modeling methods to allow accurate neurological prognostication sooner than currently possible.

CURRENT STATUS

The research team is continuing to carry out commercially orientated research on this topic.

For more information, reach out to the PHDA at healthdataalliance.com/contact.

RESEARCHERS

Daniel Nagin, PhD

Teresa and H. John Heinz III Professor,
Public Policy and Statistics
Carnegie Mellon University

Jonathan Elmer, PhD

Assistant Professor of Critical Care Medicine
and Emergency Medicine
University of Pittsburgh

Cliff Callaway, MD, PhD

Executive Vice Chair, Emergency Medicine
University of Pittsburgh Medical School

PROJECT SUMMARY

The costs of preterm births to the healthcare system and society are exceedingly large, and the incidence of preterm births is on the rise despite medical advances. This has created a need to provide reliable, personalized information to pregnant women about how they can reduce their risk of premature births. In response, researchers have created the "MyHealthyPregnancy" smartphone app.

The app applies statistical machine learning techniques to comprehensive pregnancy data sets to improve the app's patient-specific risk predictions of adverse pregnancy outcomes. It also employs decision science techniques to extend the app's ability to provide postpartum education and behavioral feedback to minimize infant mortality outcomes. The team, which has performed pilot studies of the concept, expects this work will result in a product that can be tailored to both the UPMC health system and developed into an off-the-shelf application for the general population of pregnant women.

CURRENT STATUS

MyHealthyPregnancy formed a startup called Naima Health.

*UPMC has a financial interest in Naima Health.

For more information, reach out to the PHDA at healthdataalliance.com/contact.

RESEARCHERS

Tamar Krishnamurti, PhD

Assistant Research Professor,
Engineering and Public Policy
Carnegie Mellon University
Assistant Professor, Medicine
University of Pittsburgh

Hyagriv Simhan, MD, MS

Professor, Obstetrics, Gynecology and
Reproductive Science; Division Chief,
Maternal Fetal Medicine
Magee Women's Hospital

Alexander Davis, PhD

Assistant Professor, Engineering
and Public Policy
Carnegie Mellon University

PROJECT SUMMARY

Screening technologies and public health efforts have improved the detection of early-stage cancers. However, methods to reliably predict which of these cancers will progress to more aggressive stages, metastasis, and ultimately result in patient mortality are lacking.

The project team is working on a suite of software tools that clinicians will use to improve cancer diagnosis and therapeutics based on the molecular signatures of the patient's tumor genome.

The goal of the project is to develop novel models, algorithms, and software tools to better understand the origin and evolution of tumor cells and how a patient's tumors are likely to progress. This information would contribute to personalized, precision treatment of cancer.

CURRENT STATUS

The research team is continuing to carry out commercially orientated research on this topic.

For more information, reach out to the PHDA at healthdataalliance.com/contact.

RESEARCHERS

Russell Schwartz, PhD

Professor, Biological Sciences and
Computational Biology
Carnegie Mellon University

Jian Ma

Associate Professor, Computational Biology
Carnegie Mellon University

PROJECT SUMMARY

Readmission after complex cancer surgery is common, with studies reporting between 15% and 50% of patients being readmitted within 30 days of discharge from the hospital. Readmissions are associated with increased healthcare costs, poor clinical outcomes including increased risk of infection and early mortality, and patient and family stress and suffering. Prior research has identified demographic and clinical predictors of readmission, but the role of patient-centered behavioral processes remains relatively unexplored. Identifying behavioral factors associated with readmission, which might include physical activity levels, sleep habits, and social contacts, could highlight ways to prevent readmissions and empower patients to take a more active role in their recovery.

The project team is collaborating to take a generalizable and scalable approach that holistically looks at patients' behavior before and after surgery and identifies routines in this behavior. These routines will form the basis of their understanding of the predictive factors for readmissions, particularly while patients are still in the hospital after surgery or other treatments.

Harnessing technology to monitor these and other behavioral risks before surgery, during inpatient recovery, and during the critical transition from hospital to home will advance the field in a number of ways:

- Reliably assessing and computationally modeling behavioral risks associated with readmissions to improve risk stratification
- Identifying optimal targets and timing for behavioral intervention to reduce preventable readmissions
- Providing data that may be helpful and motivating for patients and that could inform clinical decision making to improve quality of care

CURRENT STATUS

The research team is evaluating the potential clinical value of their models and continues to develop new analytical techniques related to this study.

For more information, reach out to the PHDA at healthdataalliance.com/contact.

RESEARCHERS

Anind Dey, PhD

Charles M. Geschke Director, Human-Computer Interaction Institute
Carnegie Mellon University

Carissa Low, PhD

Assistant Professor, Hematology/Oncology,
Center for Behavioral Health and Smart
Technology
University of Pittsburgh

PROJECT SUMMARY

As health data is increasingly digitized, the need to protect patient privacy is unprecedented. The Programming Framework for Managing Patient Privacy team is collaborating on an approach to protect patients' health records and computations involving sensitive patient data from being leaked in a breach.

Project researchers propose a programming model and software framework for managing the privacy of health data. The programming model allows programs to attach privacy policies directly to the data, rather than requiring the programmer to implement the correct privacy checks across code that uses the data. Attaching privacy policies to the data enables programs to be policy-agnostic and to be updated independently of the specification and implementation of privacy policies. The system tracks the flow of data, as well as the policies, so the programmer need not do so. This approach facilitates auditing, as policies may now be centralized and to be implemented only once, instead of as checks across the program. Auditors may also leverage the fact that the framework, rather than the programmer, is now responsible for correctly implementing the checks.

This project is the first step in a long-term plan to improve patient care and promote scientific discovery by allowing sensitive health data to be shared safely. This plan involves 1) applying prior work on policy-agnostic programming for information flow policies in the context of patient health records, 2) extending and combining prior work to support more expressive policies, and 3) building a practical framework that can be used to safely share sensitive data.

CURRENT STATUS

The research team is completing analyses from their project and exploring next steps.

For more information, reach out to the PHDA at healthdataalliance.com/contact.

RESEARCHERS

Jean Yang, PhD

Assistant Professor, Computer Science
Carnegie Mellon University

Matt Fredrikson, PhD

Assistant Professor, Computer Science and
Institute for Software Research
Carnegie Mellon University

Jian Ma

Associate Professor, Computational Biology
Carnegie Mellon University

PROJECT SUMMARY

This project is determining if coordinated care decisions could be improved by analyzing and monitoring how patients move through the healthcare system.

The team will be focusing on healthcare trials – time-ordered sequences in which a patient obtains services such as dialysis, blood work, or psychological counseling. These trials will be analyzed to identify those that are commonly used by patients with similar conditions and those that most often improve patient conditions. Factors such as a patient's age, healthcare coverage, and medical condition can affect the flow of these services, and differences in flow can be associated with different patient outcomes.

The goal is to develop, test, and assess algorithms and interfaces for analyzing hospital admissions data and for providing feedback to doctors and caregivers through automated patient tracking and notification systems. By analyzing the timing of admits, discharges, and transfers, the application could help inform caregivers about how a patient's care is being managed and detect when a patient might be at risk based on their healthcare trial.

CURRENT STATUS

The Health Care Trails research team is examining the outputs from this project and potential future directions.

For more information, reach out to the PHDA at healthdataalliance.com/contact.

RESEARCHERS

Kathleen Carley, PhD

Professor, Institute for Software Research
Carnegie Mellon University

L. Richard Carley, PhD

Professor, Electrical and Computer
Engineering
Carnegie Mellon University

Detecting Intestinal Activity by Analyzing Gut Sounds

PROJECT SUMMARY

A noninvasive device that analyzes sounds from the intestinal tract could become a powerful new tool to help physicians diagnose and monitor a variety of gastric illnesses, such as acute pancreatitis, bowel obstructions, irritable bowel syndrome, inflammatory bowel disease, and Crohn's disease.

CMU researchers, collaborating with UPMC gastroenterologists, are developing, testing, and performing clinical research on a wearable sensor array to detect intestinal sounds – a new vital sign for the "gut." Its impact could be substantial: In the U.S. alone, digestive diseases afflict 60-70 million people annually with treatment costs totaling more than \$100 billion.

Using machine learning techniques, the researchers will develop a means of interpreting the acoustic signals to examine the determinants of gut activity and activity suppression, the ability to identify gastric disorders on the basis of sounds, and the device's therapeutic potential for enabling gut regulation via biofeedback. The project will also develop classification procedures that use these features to distinguish between the various types of gastrointestinal activity.

CURRENT STATUS

The research team has wrapped up this project and is pursuing new directions and applications for their innovative approaches.

For more information, reach out to the PHDA at healthdataalliance.com/contact.

RESEARCHERS

George Loewenstein, PhD

Herbert A. Simon University Professor,
Economics and Psychology
Carnegie Mellon University

Ali Momeni, PhD

Associate Professor, Art
Carnegie Mellon University

Rich Stern, PhD

Professor, Electrical and Computer
Engineering
Carnegie Mellon University

Max Gsell, PhD

Associate Professor, Statistics
Carnegie Mellon University

Valerie Ventura, PhD

Associate Professor, Statistics
Carnegie Mellon University.