

THE DUKE INSTITUTE FOR HEALTH INNOVATION



INTRAPRENEURS

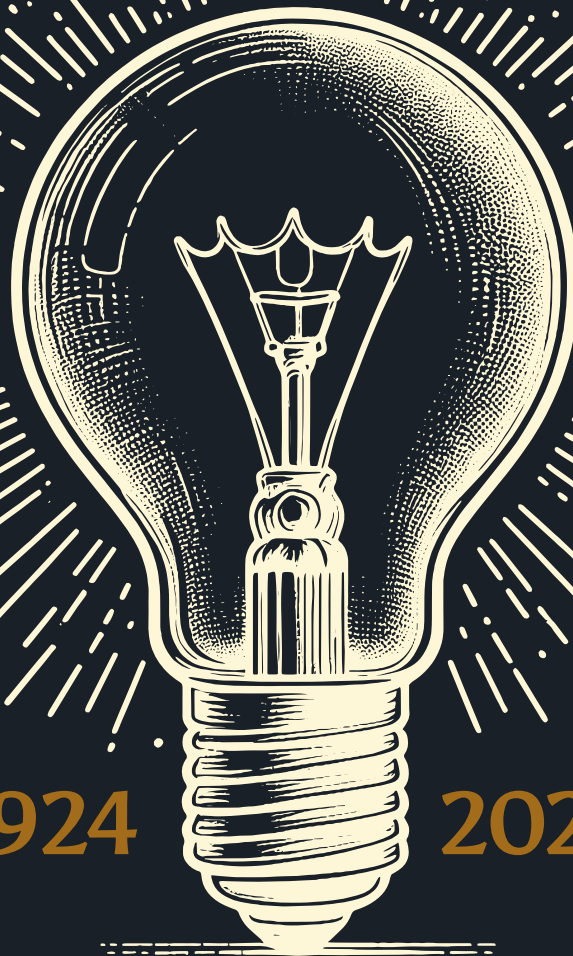
HEALTH INNOVATION TEAM

10th

100th

1924

2024



QUALITY, ACCESS,
EQUITY AND IMPACT

PATENTS
INVENTIONS
DEVICES
GEN AI

VOLUME 25

2025

LEARNING
NETWORK

CATALYZING
TRANSFORMATIVE
INNOVATION
AND RESEARCH

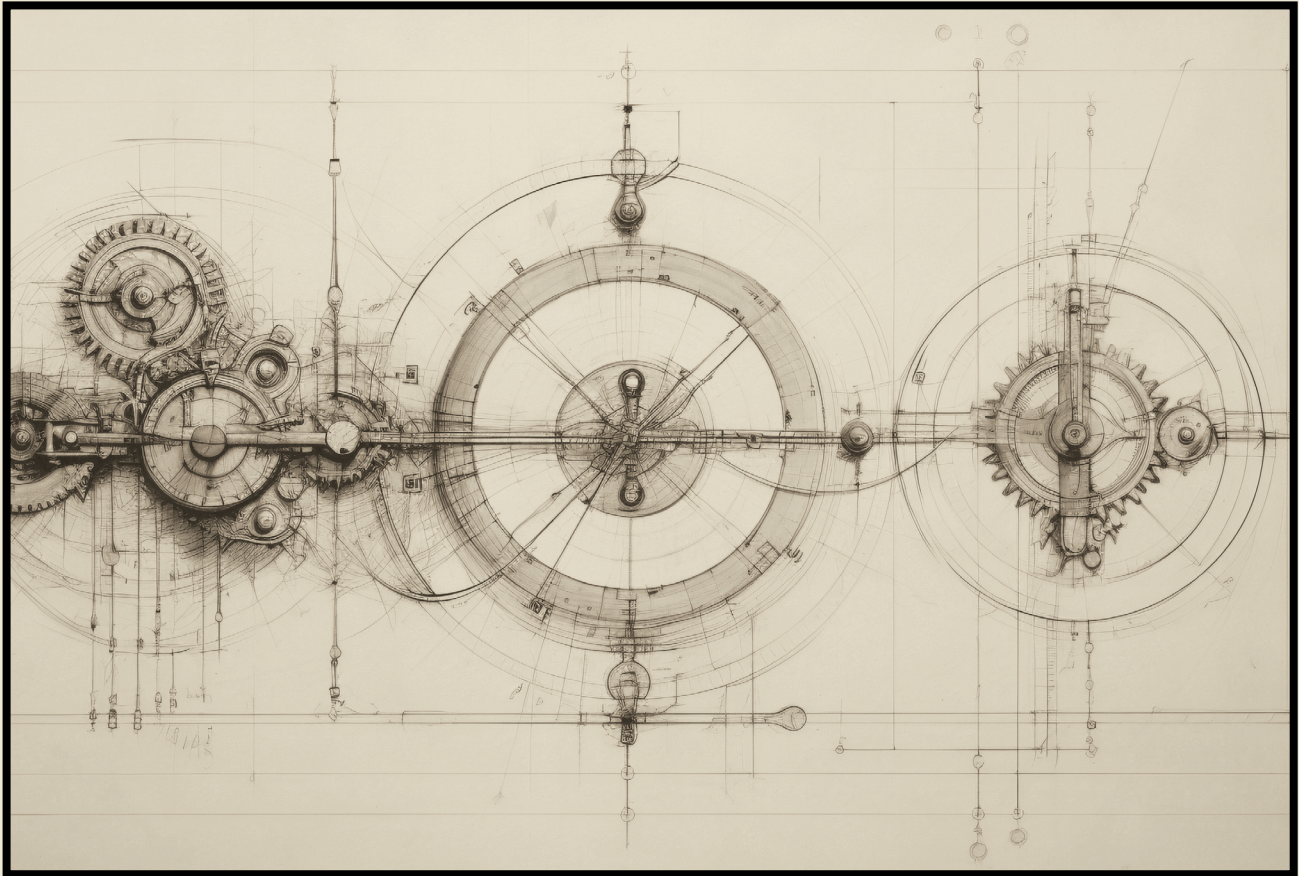
About the Artwork

This year marks a special milestone as we celebrate our 10th Anniversary and Duke University's centennial. This volume's artwork was inspired by 1890-1920 medical almanacs and the hand-drawn designs of that era's innovators, symbolizing our commitment to building on Duke's rich 100-year legacy while embracing modern technologies and innovative solutions. Trinity College became Duke University during a transformative era marked by significant advancements in automation, communication technology, and creative arts. This pioneering spirit resonates with DIHI's mission to tackle complex healthcare challenges. Our dedication to enhancing healthcare quality, access, equity, and value reflects the progressive ideals of the 1920s, such as the expansion of women's voting rights and the establishment of public health resources in schools. Our focus on workflow improvements and impact measurement mirrors the industrialization ideals that defined the era. Just as the 1920s brought about breakthroughs in radio, automobiles, and household appliances, we are now witnessing exciting advancements in generative AI and large language models. These innovations inspire us to envision a future where work is less burdensome and more fulfilling. We reflect on a time when social interactions were primarily face-to-face, fostering tight-knit communities, and we yearn for unstructured time filled with imagination and creativity—where information overload and multitasking are minimized. With "Eruditio et Religio" (Duke's original motto), we can pave the way for a healthier, more rewarding life.

All artwork is from Duke, DIHI, or AI-generated. Generated artworks are original creations inspired by detailed prompts, without intentional replication of copyrighted works.

THE DUKE INSTITUTE FOR HEALTH INNOVATION

VOLUME 25



About DIHI

The Duke Institute for Health Innovation (DIHI) catalyzes transformative innovations in health and healthcare through implementation of high-impact innovations, leadership development, and cultivation of a community of entrepreneurship. We bring innovative solutions to the most pressing challenges in health and healthcare through multidisciplinary teamwork across Duke University and Duke Health and by fostering collaborations with national and international thought leaders.

Contents

A Decade of Influential Innovation

Executive Summary	04
Impact Infographic	05
A Year of AI and Human Connection at DIHI	06

RFA Impact

An Automated Post-Hospitalization Track for Patients Discharged on Outpatient Parenteral Antimicrobial Therapy	08
Making it EASY to Schedule Telehealth Visits in Rheumatology	10
Automating CMS Quality Measure Curation	12
Use of Machine Learning to Support Triage of Patients with Lower Back Pain	17
Duke Raleigh Hospital Automated Staff and Bed Assignment	20

Expanding the Portfolio

Health Record Optimization to Identify Candidates for Pre-exposure Prophylaxis	22
Cardiogenic Shock	25
Federated Learning: Revolutionizing Healthcare AI with Owkin	27
Understanding and Managing Clinical Metadata with Poseidon	29

AI Governance & Ethics

Framework for Assessing the Impact of AI on Health Equity: Health Equity Across the AI Lifecycle (HEAAL)	31
Creating an AI Ethics Training Program for Physicians in Collaboration With the American Medical Association	34
Establishing an AI Governance System to Facilitate the Safe and Responsible Adoption of AI Software in a Canadian Context	35
Building a One-of-a-Kind AI Product Evaluation Rubric Tool	37

Large Language Models

Leveraging a Large Language Model to Deploy a Tool for Accurate Care Engagement Classification Among Patients with HIV (PWH) Within the Duke University Health System	38
Inbasket Messaging With LLMs	41
Duke Health Policy Assistant – The RAG Journey	44
Keep LLMs Local	47
The Large Language Model (LLM) Frontier in 2024	49

Our Team’s Take

Intrapreneurial Healthcare Strategy	51
Commercializing Innovation: An Interview with Suresh Balu and the OTC	56
Utilizing Statistical Process Control to Track the Impact of Our Health Innovations: Lessons from a Pediatric Project	59
Perspectivation in Healthcare: Driving Innovation Through Diverse Perspectives	62



Legacy Innovations: Reflecting on 10 years of Impact

Stitching Together Chronic Kidney Disease Innovations to Transform Chronic Disease Management at Duke Health 64

Duke TeleHOPE Program Demonstrates Sustained Success in Reducing Readmissions and Costs 68

Summing up SymMon 70

Autism & Beyond 71

Voices of Duke Health Podcast 72

Evaluating and Scaling Innovation

Sepsis Watch 2.0 Expansion into Duke Regional Hospital 73

Diffusion of Capabilities: Launching a Practice Network 76

Pythion: New Models using Narratives 77

Maternal Early Warning System Update 79

Immune-Related Adverse Events Silent Evaluation 80

Duke Opioid Sedation Assessment Risk Score 82

Ongoing Support for Patients with Peripheral Artery Disease, Expanding to Diabetic Patients with Threatened Limbs 84

Implementing an Informatics-Driven Notification System for Patients with High-Risk Conditions Presenting with Fever In The Pediatric Emergency Department 87

Evaluation of a Digital Phenotype for the Early Recognition of Pediatric Sepsis 90

DIHI Interval Problem List Clean Up 2024 92

Scholar Experiences

Freya Gulamali 48

Henry Foote 65

Bruno Valen 67

Maryam Aziz 69

Kartik Pejavara 81

Noah Prizant 83

Samyukta Palle 91

Publications & Presentations 96

Team 94

Scholars 95

List of Collaborators 106



EXECUTIVE SUMMARY OF DIHI PROJECTS IN 2024

MEASURED OUTCOME IMPROVEMENTS

Monitoring Outpatient Parenteral Antimicrobial Therapy

Monitoring outpatient parenteral antimicrobial therapy (OPAT) facilitated early post-hospitalization follow-up for patients discharged from the hospital on OPAT. Early follow-up has increased from 26% to 40% of OPAT patients since July 1, 2023, which has been associated with a trend toward lower 30-day readmission rates.

Predicting Appropriateness for Telehealth in Rheumatology

We aimed to increase telehealth use by identifying patient-physician matches suitable for it. We redeveloped and prospectively tested a predictive model so that it automatically fits a workflow. If used as intended, we can expect to significantly and satisfactorily expand telehealth access from 3% currently scheduled to 20% over the next year.

Automating CMS Quality Measure Curation

This project created a dashboard to automate data collection for CMS SEP-1 bundle reporting. The dashboard went live in February 2024 and tracks compliance for 980 patients per month across the three hospitals. It is currently in use supporting quality improvement efforts at all three hospitals.

Duke Opioid Safety Assessment Risk Score

During a pilot implementation at Duke University Hospital, a 40% reduction in adverse events was observed in the high-risk group, demonstrating the score's effectiveness in improving opioid safety.

Implementing an Informatics-Driven Notification System for Patients with High-Risk Conditions Presenting with Fever in The Pediatric Emergency Department

A system was developed to immediately identify pediatric patients with high-risk conditions presenting with fever. Since the solution's go live in February 2024, ED arrival-to-antibiotics time has improved from 75.6 minutes to 42.0 minutes (44% decrease), and compliance for antibiotics within one hour of ED arrival has improved from 36.1% to 82.4%.

Sepsis Watch 2.0 Expansion into Duke Regional Hospital

Sepsis Watch 2.0 went live on June 24th, 2024, at Duke Regional Hospital. In the first 60 days, 614 pages were sent on 436 unique patient encounters. Since going live at Duke University Hospital in 2018, Sepsis Watch has increased SEP-1 bundle compliance by 89% and reduced observed-to-expected mortality by 27%.

PROCESS UPDATES

Evaluation of a Digital Phenotype for the Early Recognition of Pediatric Sepsis

This project developed a digital phenotype to identify pediatric patients at high risk of sepsis in real-time. It has a sensitivity of 76% and positive predictive value of 20% for sepsis ICD codes (i.e., a ground truth retrospective sepsis definition). The Duke Pediatric Sepsis Phenotype is now running in real time and is being temporally validated by pediatric sepsis expert clinicians.

Use of Machine Learning to Support Triage of Patients with Lower Back Pain

This project used data from over 137,000 patients to develop models that predict prolonged opioid use, pain reduction, and emergency room utilization post-intervention, per index provider type, for new lower back pain patients. The models will be piloted at the Spine Center to evaluate their ability to support scheduling with the right provider type for improved clinical outcomes and surgical consult-to-surgery rates.

Diffusion & Scaling: irAE Silent Evaluation

A machine learning model was developed to predict emergency department (ED) and hospital admissions due to immune-related adverse events (irAEs). The model demonstrated its potential to improve patient care by identifying high-risk patients and allowing for proactive interventions.

Cardiogenic Shock

A cardiogenic shock mortality model was developed to predict mortality risk at 12 and 24 hours after admission. The model incorporates data from various sources, including electrocardiograms and heart catheterizations, and has shown promising results in predicting mortality risk.

Pythion

Pythion, a graphical user interface that quickly predicts postoperative complication and mortality risk at the time of case posting, is being silently evaluated with the goal of reviewing over 100 cases.

EDUCATION, PATIENT CARE, RESEARCH, GROWTH

10 years catalyzing innovation • 92 RFA innovation projects • 10 staff average • Survey result of 50+ Duke Faculty:
 Best-in-class at operating efficiently. Excellent at commendably advancing satisfying ideation and solutions.
 Great at collaborative ideation, work, and follow-through aligned with Duke strategy.

536

Healthcare Co-Developers

- 17% Department of Medicine
- 8% Hospital Medicine
- 8% Nursing
- 13% Department of Surgery
- 25% Returned to do another project with DIHI
- 48% Clinical leads of RFA projects were women

87 Students

- 30 MEDICAL STUDENTS
- 54 DATA SCIENCE STUDENTS
- 3 MEDICAL FELLOWS

\$30M

Follow-on Research Sponsors include the National Institutes of Health, Gordon & Betty Moore Foundation, American Medical Association & General Services Administration

15+

NEW COMPANIES CREATED

98

Peer-Reviewed Publications

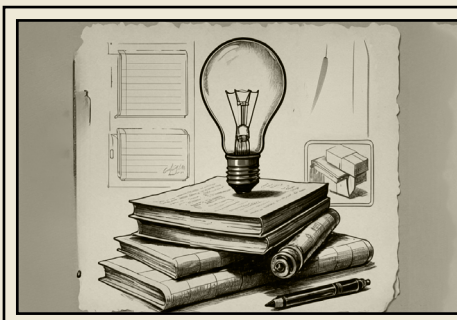
740

Innovative Ideas Surfaced to Leadership

>5,330 Hospital-Days Saved For Skilled Nursing Facility Patients

27%

REDUCED SEPSIS MORTALITY (OBSERVED-TO-EXPECTED)



693K

Active COVID-19 Check-Ins Through SymMon Monitoring
 KEEPING DUKE UNIVERSITY OPEN

59

Specialties using the e-Consult model

45+

INVENTION DISCLOSURES

1,919,477

CLINICAL DECISION COMPUTATIONS / DAY
211 MILLION DATA MODEL CALCULATIONS

- 3.14 M Unique Patients
- 207 M Patient Encounters
- 623 M Labs
- 195 M Notes
- 552 M Prescriptions
- 7.7 B Flowsheets
- 2190 Alerts 2021-2023





A Year of AI and Human Connection at DIHI



Over the past year, the Duke Institute for Health Innovation (DIHI) has forged ahead with several endeavors integrating artificial intelligence (AI) and machine learning solutions into healthcare operations, focused on working with our frontline clinicians. These technologies, while not as new as they might seem, have matured significantly, offering unprecedented opportunities for simultaneously enhancing patient care and operational efficiency.

This annual report showcases the successful outcomes of our efforts focused on leveraging AI and machine learning. These tools have enabled us to create more targeted interventions, streamline workflows, and ultimately, make healthcare more accessible and less burdensome. We also continue to enhance, scale, and complement our prior integrations, such as decompensation and sepsis in both adult and pediatric patients, reducing unplanned hospitalizations due to immune-related adverse events, improving outcomes for patients experiencing cardiogenic shock, supporting telehealth for rheumatology visits, and creating a groundbreaking tool for effective metadata management.

With a keen focus on engaging the next generation of clinical leaders, we have also intentionally grown our work with learners at Duke and beyond. This year, three MD students worked on various projects with us as part of their third-year curricular requirement; additionally, we worked with two Fellows—one in Pediatrics and one in Critical Care—to help drive innovations in each domain. Lastly, we have expanded our internship program to include local high school students and undergraduates from Duke, UNC, and UVA.

Recently, the world has seen a proliferation of AI-driven projects. Whether through AI-powered chatbots that can handle 24/7 customer support, using algorithms to personalize consumer preferences, or employing AI analytics to predict behaviors and optimize marketing and outreach efforts, evaluating risks, reducing costs, streamlining talent acquisition and employee engagement, ensuring better quality control or enhancing R&D, the impact of AI has been widely felt.

While the healthcare industry has started to use AI technologies such as machine learning and natural language processing, when fully realized, we believe that the implications for health and healthcare will be nothing short of revolutionary.

Over the past year, DIHI has worked with partners within and outside of Duke to empower our healthcare delivery colleagues to use AI to enhance efficiency and decision-making to serve patients better and improve outcomes. The report outlines our work in this area. Our RFA theme of using generative AI and large language models to improve operational efficiency and patient outcomes generated several impactful projects that are currently in the planning and implementation phases.

Additionally, we have successfully launched the Health AI Partnership, a collaboration enabled by the Gordon and Betty Moore Foundation and McGovern Foundation, for empowering healthcare delivery organization

leaders across the nation to responsibly adopt AI and to ensure the safe, equitable, and ethical use of AI, while also improving patient outcomes. As a part of this initiative, we have launched a Practice Network to provide access to one-on-one guidance, expert consultation, and a peer learning community for responsible integration of AI adoption. The focus is to support Federally Qualified Health Centers and community healthcare organizations that do not have ready access to AI expertise and resources.

While we have made significant strides in AI-driven technologies, in the new academic year, we reaffirm our commitment to human-centered care. We look forward to continuing to harness our talented and irreplaceable workforce that brings much-needed empathy, excellence, and compassion to patient care. Our goal is to integrate innovations in a way that enhances rather than replaces the human touch, ensuring that patient care remains at the heart of our work.

An Automated Post-Hospitalization Track for Patients Discharged on Outpatient Parenteral Antimicrobial Therapy

Problem

An estimated 1,400 patients are discharged each year from Duke University Hospital System on outpatient parenteral antimicrobial therapy (OPAT), which is managed by the Duke OPAT team (“our team”). These patients have complex infections requiring prolonged intravenous antimicrobials across multiple care settings, placing them at high risk for complications. OPAT patients are monitored by an interdisciplinary team of pharmacists, clinicians, and nurses in the Infectious Diseases (ID) clinic. Historically, the 30-day unplanned readmission rate in the Duke OPAT cohort was approximately 20%, a higher rate than OPAT programs at peer institutions. A previous pair-matched case-control study demonstrated that patients who had follow-up in an Infectious Diseases clinic within two weeks of hospital discharge had a significantly lower risk of readmission.¹ Quality improvement interventions, including an OPAT bundle (a comprehensive set of coordinated interventions designed to improve OPAT patient care),² have also been shown to improve readmission rates. Our team aimed to create a dashboard to track OPAT patients through their treatment course and improve ID provider follow-up within two weeks of discharge.

Solution

We developed and implemented OPATracker, an automated real-time dashboard to track OPAT patients. This dashboard visualizes a comprehensive list of OPAT patients, including admitted patients, upcoming follow-up appointments, anticipated stop dates of therapy, medication lists, and laboratory data. We have also implemented an automated daily email to notify our team of newly discharged OPAT patients so they can be enrolled in a transitions-of-care pathway. We concurrently opened a new telemedicine clinic in collaboration with the Duke Virtual Care Clinic to see these patients for early post-hospitalization clinic visits with an ID provider. These visits provide an opportunity to identify early challenges, facilitate additional follow-up care, and provide comprehensive patient and family education (Figure 1).

Impact

We utilized the OPATracker and automated notification system to facilitate early post-hospitalization follow-up for OPAT patients in the virtual care clinic. With these interventions, we have improved our rates of ID provider follow-up within two weeks of discharge from a process mean of 26% to 40% (Figure 2). The proportion of OPAT patients seen in telemedicine follow-up has increased since the virtual care clinic opened in September 2023 with a lower no-show rate than our conventional ID clinic (10.8% vs 19.5%). This positive trend in telemedicine follow-up is a promising sign for the future of our healthcare services. With improved access to early post-hospitalization ID care, we have seen an accompanying trend towards a decrease in readmission rates, with nine of the last ten or nine project phases showing readmission rates below our prior process mean (Figure 3).

Next Steps

In the coming year, our project team will work to refine our transitions of care pathway further. Specifically, our pharmacy team will discuss pending cases with our inpatient team prior to discharge to identify concerns, facilitate a transitions-of-care phone call with our OPAT nurse coordinator, and further improve our rates of early post-hospitalization ID clinic follow-up. We are committed to continuously improving our processes to ensure the sustainability and effectiveness of our initiative.

Academic output

We are writing a manuscript detailing the process improvements we have accomplished through these early interventions. We are also in the early stages of conducting a propensity-matched cohort study to evaluate the impact of early post-hospitalization follow-up for our patients.

References

1. Saini E, Ali M, Du P, Crook T, Zurlo J. Early Infectious Disease Outpatient Follow-up of Outpatient Parenteral Antimicrobial Therapy Patients Reduces 30-Day Readmission. *Clin Infect Dis* 2019; 69(5): 865-8.
2. Madaline T, Nori P, Mowrey W, et al. Bundle in the Bronx: Impact of a Transition-of-Care Outpatient Parenteral Antibiotic Therapy Bundle on All-Cause 30-Day Hospital Readmissions. *Open Forum Infect Dis* 2017; 4(2): ofx097.

PATIENT AND FAMILY EDUCATION
<ul style="list-style-type: none"> • Provide education on disease state • Review plan • Discuss prophylaxis • Return precautions • Shared decision making
ANTIBIOTIC ASSESSMENT
<ul style="list-style-type: none"> • Review regimen • Verify medication delivery and lab collection • Adjust for toxicity • Adjust for updated microbiology data • Adjust dosing schedules if needed
IV ACCESS
<ul style="list-style-type: none"> • Troubleshoot issues • Review plans for removal • Send orders to skilled nursing facility (SNF)
PRESCRIBE ORAL ANTIBIOTICS WHEN INDICATED
<ul style="list-style-type: none"> • Transition to complex oral regimens (CoPAT) • Oral suppression after initial induction course
OTHER FOLLOW-UP
<ul style="list-style-type: none"> • Order imaging • Coordinate with surgical teams • Schedule subsequent ID follow-up visit

TEAM

Molly Hillenbrand, MD
 Michael Yarrington, MD, MMCi
 Kristen Dicks, MD, MPH
 Jamie Daniel, BS
 Will Knechtle, MBA, MPH

IN BRIEF

Patients managed by the Duke University Hospital System outpatient parenteral antimicrobial therapy (OPAT) team have historically low rates of early post-hospitalization follow-up in infectious diseases (ID) clinics and high rates of unplanned readmission within 30 days. We built a dashboard to monitor OPAT patients and developed an automated system to notify our team when OPAT patients are discharged from the hospital. This system helped facilitate early post-hospitalization follow-up in a new virtual care clinic for OPAT patients. Since July 1, 2023, we've improved early post-hospitalization follow-up with an infectious diseases (ID) provider from 26% to 40% of OPAT patients, with an associated trend towards lower 30-day readmission rates in this cohort.

Figure 1. An overview of services provided in the OPAT follow-up virtual care clinic.

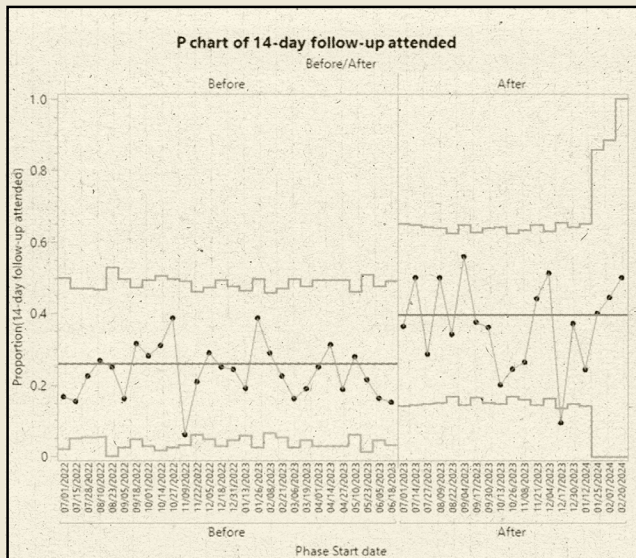


Figure 2. P-chart of early post-hospitalization follow-up for OPAT patients. Each point represents the proportion of patients discharged during a two-week time period that had ID follow-up within two weeks of discharge.

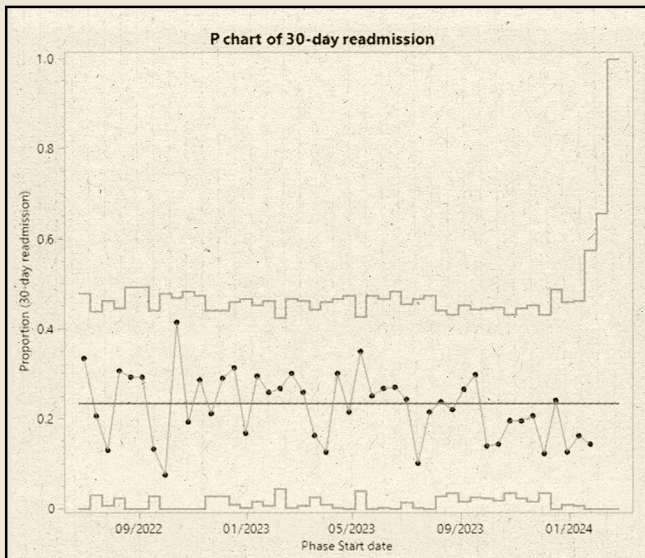


Figure 3. P-chart of 30-day readmission for OPAT patients. Each point represents the proportion of patients discharged during a two-week time period that were readmitted within 30 days of discharge.



RECENT RFA IMPACT

Making it EASY to Schedule Telehealth Visits in Rheumatology

David Leverenz, MD, MEd

Problem

Before this DIHI project began, David Leverenz, MD, MEd, and Jayanth Doss, MD, MPH, led a team to develop a novel scoring system called the Encounter Appropriateness Score for You (EASY), in which providers document their perception of the appropriateness of telehealth or in-person care after every encounter. An analysis of over 12,000 EASY scores in 2021 demonstrated that our providers think telehealth is acceptable for approximately 30% of all follow-up visits in rheumatology. However, Duke Rheumatology – which cares for over 13,000 patients in over 24,000 annual encounters - currently performs only about 15% of its visits by telehealth. As such, telehealth remains vastly underutilized at Duke Rheumatology in the post-pandemic era (Figure 1). If Duke rheumatologists can better identify patients who are appropriate for telehealth rather than in-person care, we can expand access to telehealth by over a thousand visits each year. The team used EASY scores to create a predictive model to identify future rheumatology visits appropriate for telehealth care. The initial model was a logistic regression algorithm based on data stored in the Duke Protected Analytics Computing Environment (PACE). Before collaborating with DIHI, the team could not fully implement this model or convert the model to a more actionable interface.

Solution

Our collaboration with DIHI was instrumental in putting our model to use. Their unique skills and understanding were crucial in designing the model for an impactful fit within our clinical workflow. DIHI are experts in shifting from a retrospective to a prospective mindset: model feature and operational design become dynamically different when you move from using a static dataset of the past to a continuously updating one that periodically re-accounts for the past, present, and future. Together, we have created a predictive model that can be accessed in real-time and implemented in clinical practice, providing healthcare professionals with a powerful tool for shared decision-making.

Impact

The reconfigured predictive model uses real-time data that can be accessed through a Tableau dashboard. DIHI’s updated design allows us to utilize this data in real-time for shared decision-making with patients. Additionally, periodic reviews of patient lists allow us to offer visit modality switches to patients identified by the model as appropriate for telehealth. The predictive model, with its ability to identify future visits appropriate for telehealth, has the potential to significantly improve patient care and practice efficiency.

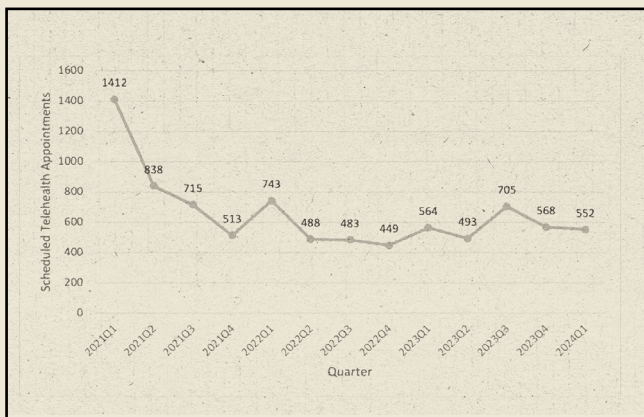


Figure 1: Quarterly Scheduled Telehealth Appointments (2021-2024)

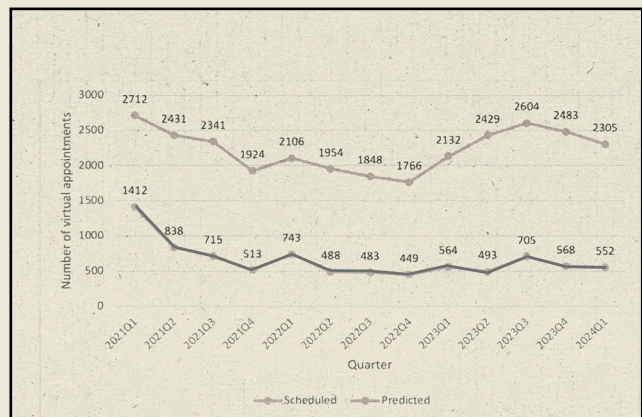


Figure 2: The actual and predicted virtual appointments per quarter (2021-2024)

To visualize the model’s performance, AUROC and PRC curves comparing the 2021 and 2023 models are presented. While AUROC and recall may show improvement, precision is expected to be lower in 2023 due to the drop in telehealth usage from 2021 to 2023. This reflects the current challenge of predicting telehealth-appropriate visits in an environment with low actual telehealth usage.

Next Steps

Future steps involve implementing the model within clinical care, measuring its impact, and disseminating it to other rheumatology practices. If our model was used in the last year, telehealth use would have been 18% rather than 5%. If used as intended, we can expect to significantly and satisfactorily expand telehealth access from 3% currently scheduled to 20% over the next year.

The DIHI team also contributed to other ongoing work on understanding the perspectives of rheumatology patients with regards to the appropriateness of telehealth visits. In particular, DIHI used our EASY data to help clinicians identify patients for qualitative interviews, which clinicians completed. Two manuscripts describing patient and provider perspectives of telehealth appropriateness are in preparation.

Year	AUROC	AP	Precision	Recall	F1 Score
2021	0.912284	0.714811	0.316433	0.855952	0.462052
2023	0.891032	0.597064	0.200352	0.829614	0.322758

Table 1. Compares the 2021 and 2023 model versions across key performance metrics such as AUROC, precision, recall, and F1 score.

Academic output

The initial model development (before this DIHI grant) resulted in several published manuscripts. Academic output from the novel version of the model is pending final implementation. Prior works include the following:

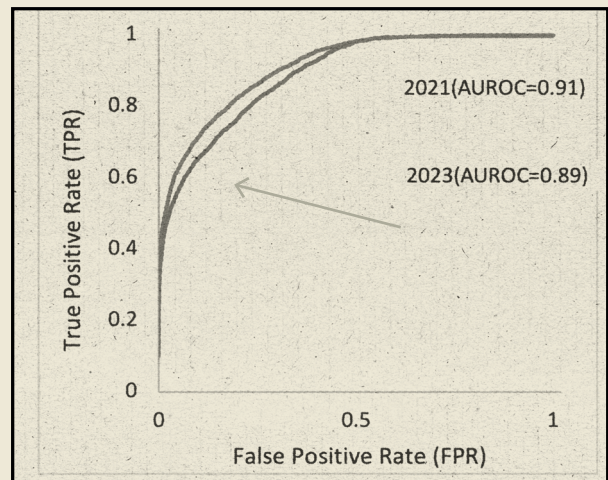
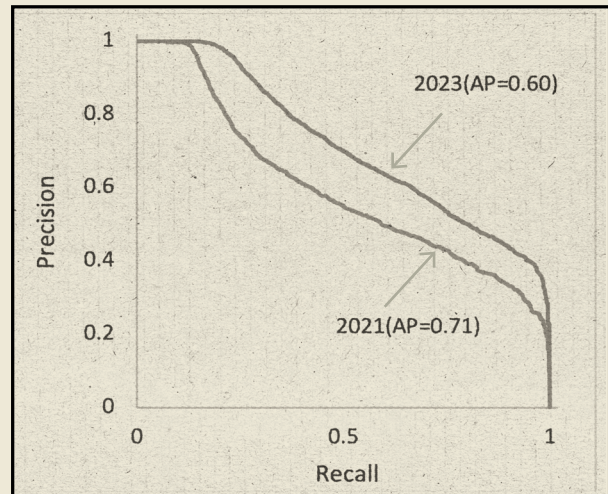
1. Smith ID, Coles TM, Howe C, Overton R, Economou-Zavlanos N, Solomon MJ, Zhao R, Adagarla B, Doss J, Henao R, Clowse MEB, Leverenz DL. Telehealth Made EASY: Understanding Provider Perceptions of Telehealth Appropriateness in Outpatient Rheumatology Encounters. *ACR Open Rheumatol.* 2022 Oct;4(10):845-852. doi: 10.1002/acr.2.11470. Epub 2022 Jul 19. PMID: 35855564; PMCID: PMC9555194.
2. Solomon M, Henao R, Economou-Zavlanos N, Smith I, Adagarla B, Overton AJ, Howe C, Doss J, Clowse M, Leverenz DL. Encounter Appropriateness Score for You Model: Development and Pilot Implementation of a Predictive Model to Identify Visits Appropriate for Telehealth in Rheumatology. *Arthritis Care Res (Hoboken).* 2024 Jan;76(1):63-71. doi: 10.1002/acr.25247. Epub 2023 Nov 13. PMID: 37781782.
3. Smith ID, Solomon MJ, Mulder H, Sims C, Coles TM, Overton R, Economou-Zavlanos N, Zhao R, Adagarla B, Doss J, Henao R, Clowse MEB, Bosworth H, Leverenz DL. Evaluating factors associated with telehealth appropriateness in outpatient rheumatoid arthritis encounters using the Encounter Appropriateness Score for You (EASY). *J Rheumatol.* 2024 May 15;jrheum.2024-0014. doi: 10.3899/jrheum.2024-0014. Epub ahead of print. PMID: 38749564.

TEAM

- David Leverenz, MD, MEd
- Jayanth Doss, MD, MPH
- Will Knechtle, MBA, MPH
- Bradley Hintze, PhD
- Samyukta Palle, MEng
- Michael Gao, MS
- Suresh Balu, MS, MBA

IN BRIEF

Telehealth in rheumatology remains vastly underutilized in the post-pandemic era. We created a predictive model to identify future rheumatology visits that could be safely performed via telehealth. Implementing this predictive model will improve patient access to appropriate, efficient, and effective care.



RECENT RFA IMPACT

Automating CMS Quality Measure Curation

Problem

Duke Health must review sepsis cases for SEP-1 public reporting to Centers for Medicare & Medicaid Services (CMS), which requires significant manual labor from clinicians. Unfortunately, due to current staffing constraints, individual sepsis case reviews are performed by an external practice management consultancy at a substantial cost to Duke Health. Every year, we estimate that SEP-1 case reviews cost Duke Health over \$75,000 when external consultant costs are combined with manual effort performed by Duke Health expert clinicians. Like other large health systems, Duke only reports a small subset (<20%) of its monthly qualifying cases for sepsis bundle treatment compliance. There is a two month delay in dissemination of the details of this subset of cases, due to the time requirement for manual review. In 2023, CMS announced that starting in 2025, Medicare will begin to reimburse health systems for sepsis cases based on SEP-1 Bundle Performance. Without comprehensive, salient reporting on all eligible bundle cases, Duke Health cannot effectively drive quality improvement on bundle performance. Automated data collection for reporting and quality improvement is essential to reduce reporting costs and sustain high performance on the SEP-1 Bundle compliance.

Solution

To address this gap, Duke University's Department of Medicine, the DUHS Quality Office, and the Duke Institute for Health Innovation (DIHI) formed a transdisciplinary team to automatically curate all relevant SEP-1 case reviews for external reporting and internal QI. Our goal is to reduce the manual effort and cost of reporting to CMS while also providing full visibility into all relevant data points for all eligible patients in order to improve and sustain each of Duke Health's three hospitals' SEP-1 Bundle compliance.

We used data collected from all inpatient encounters at three Duke University Health System (DUHS) hospitals from January 1, 2017, to yesterday (>500,000 total encounters). We excluded patients under 18 years old at the time of admission. On the day after discharge, an encounter and all associated data are added to the dashboard if the patient met at least one of three criteria during their hospitalization:

1. Met the real-time CMS criteria for severe sepsis
2. Had an ICD code for sepsis documented for the encounter
3. Met the Center for Disease Control and Prevention (CDC) retrospective Adult Sepsis Event (ASE) definition for sepsis

TEAM

Cara O'Brien, MD (PI)
 Will Ratliff, MBA
 Bradley Hintze, PhD
 Dustin Tart, BSN
 Catie Dunn, BSN
 Melanie Lockamy, BSN
 Cindy Corcoran
 Tom Ely, MHA
 Mark Sendak, MD, MPP
 Suresh Balu, MS, MBA

IN BRIEF

Duke Health's Severe Sepsis and Septic Shock Management (SEP-1) bundle compliance is reported to the Centers for Medicare & Medicaid Services (CMS) by an external consultancy using data collected via manual chart review by a Duke clinician. This is a significant yearly cost (~\$75K/year) to Duke Health and captures a small subset of randomly chosen SEP-1 bundle-eligible cases. Meanwhile, CMS has identified the SEP-1 bundle as an important quality metric determining financial reimbursement to Duke by Medicare for those cases. To help reduce the manual reporting effort and increase visibility for quality improvement, our team has automated the SEP-1 Bundle calculation for all eligible patients. We have implemented a Tableau dashboard solution that shows all compliance-related data for every Duke Health patient from 2014 to today and appends new patients daily as soon as they are discharged. We are completing the evaluation of the dashboard for automated reporting efforts and are applying the data to drive sepsis quality improvement (QI) across the health system.

The union of these three eligibility criteria is meant to provide the most comprehensive encounter cohort possible. The dashboard can then be used to filter as needed for various use cases. The complete list of data elements included in the dashboard is in Table 1.

Our goal is to reduce the manual effort and cost of reporting to CMS while also providing full visibility into all relevant data points for all eligible patients in order to improve and sustain each of Duke Health’s three hospitals’ SEP-1 Bundle compliance.

Impact

We went live with the Duke Health SEP-1 Bundle Dashboard in February 2024 (Figure 1). In June, as part of our Sepsis Watch go-live at Duke Regional Hospital (DRH), we began utilizing the sepsis bundle compliance data to drive quality improvement on the sepsis bundle by department, overseen by the DRH Sepsis Governance Committee.

Our next steps are to complete the validation of the data for automated reporting to CMS and the integration of the dashboard and its data into sepsis quality improvement work across Duke Health System. We aim to evaluate the solution’s potential to identify and improve bundle compliance overall, at the department level, and at the physician level. We also plan to measure the impact on subsequent clinical outcomes, such as length of stay, ICU requirement, in-hospital mortality, and readmission rate.

Next Steps

We are working with the Duke Quality Office to finalize the content for semi-automated reporting to CMS, with the goal of reducing the thousands spent per year in reporting costs.

DATA ELEMENT	EXAMPLE DATA	DESCRIPTION
encounter_id	123456	Contact Serial Number (CSN) for the hospital encounter
patient_id	ABC123	unique patient identifier
cms_sepsis_at	10/1/2024 14:25	the first time that the patient met CMS severe sepsis "time zero" (i.e., real-time definition of severe sepsis), if they met it during the hospitalization
septic_shock_at	10/1/2024 14:47	the first time that the patient met CMS septic shock (i.e., real-time definition of septic shock), if they met it during the hospitalization
first_blood_culture_ordered_at	10/1/2024 13:30	the blood culture order time for the first culture collected within the eligible time window for the CMS SEP-1 Bundle
first_blood_culture_collected_at	10/1/2024 13:56	the first eligible blood culture collection time within the eligible time window for the CMS SEP-1 Bundle
first_lactate_ordered_at	10/1/2024 11:18	the lactate order time for the first lactate collected within the eligible time window for the CMS SEP-1 Bundle
first_lactate_collected_at	10/1/2024 11:49	the first eligible lactate collection time within the eligible time window for the CMS SEP-1 Bundle
first_lactate_value	3.1	the resulting value for the first lactate collected within the eligible time window for the CMS SEP-1 Bundle

Table 1: Data elements included in the SEP-1 Bundle Compliance dashboard

NEXT PAGE



DATA ELEMENT	EXAMPLE DATA	DESCRIPTION
second_lactate_ordered_at	10/1/2024 13:20	the lactate order time for the second lactate collected within the eligible time window for the CMS SEP-1 Bundle
second_lactate_collected_at	10/1/2024 13:41	the second eligible lactate collection time within the eligible time window for the CMS SEP-1 Bundle
second_lactate_value	2.0	the resulting value for the second lactate collected within the eligible time window for the CMS SEP-1 Bundle
third_lactate_ordered_at	10/1/2024 14:20	the lactate order time for the third lactate collected within the eligible time window for the CMS SEP-1 Bundle
third_lactate_collected_at	10/1/2024 14:41	the third eligible lactate collection time within the eligible time window for the CMS SEP-1 Bundle
third_lactate_value	1.6	the resulting value for the third lactate collected within the eligible time window for the CMS SEP-1 Bundle
fourth_lactate_ordered_at	10/1/2024 15:20	the lactate order time for the fourth lactate collected within the eligible time window for the CMS SEP-1 Bundle
fourth_lactate_collected_at	10/1/2024 15:41	the fourth eligible lactate collection time within the eligible time window for the CMS SEP-1 Bundle
fourth_lactate_value	1.5	the resulting value for the fourth lactate collected within the eligible time window for the CMS SEP-1 Bundle
fifth_lactate_ordered_at	10/1/2024 16:20	the lactate order time for the fifth lactate collected within the eligible time window for the CMS SEP-1 Bundle
fifth_lactate_collected_at	10/1/2024 16:41	the fifth eligible lactate collection time within the eligible time window for the CMS SEP-1 Bundle
fifth_lactate_value	1.9	the resulting value for the fifth lactate collected within the eligible time window for the CMS SEP-1 Bundle
sixth_lactate_ordered_at	10/1/2024 17:20	the lactate order time for the sixth lactate collected within the eligible time window for the CMS SEP-1 Bundle
sixth_lactate_collected_at	10/1/2024 17:41	the sixth eligible lactate collection time within the eligible time window for the CMS SEP-1 Bundle
sixth_lactate_value	2.7	the resulting value for the sixth lactate collected within the eligible time window for the CMS SEP-1 Bundle
first_antibiotics_ordered_at	10/1/2024 12:59	the order time of the first qualifying antibiotic administered within the eligible time window for the CMS SEP-1 Bundle
first_antibiotics_given_at	10/1/2024 13:39	the first administration time of a qualifying antibiotic within the eligible time window for the CMS SEP-1 Bundle
first_iv_fluids_ordered_at	10/1/2024 12:59	the order time of the first IV fluids administered within the eligible time window for the CMS SEP-1 Bundle
first_iv_fluids_given_at	10/1/2024 14:40	the first administration time of IV fluids within the eligible time window for the CMS SEP-1 Bundle
iv_fluid_volume	400	the total volume of IV fluids ordered by the provider for IV fluids (mL) administered during the eligible time window for the CMS SEP-1 Bundle
iv_fluid_volume_sufficient	TRUE	Sufficient IV fluid volume compliance per CMS SEP-1 Bundle requirements: True (compliant) or False (not compliant)
null_iv_volume_dose_administrations_count	4	Count of IV fluids administered where a volume dose was specified by the provider



DATA ELEMENT	EXAMPLE DATA	DESCRIPTION
weight	79.5	the patient's weight (kg)
first_hypotension_at	10/1/2024 11:47	the first time that the patient met the CMS hypotension (i.e., real-time definition of hypotension), if they met it during the eligible time window related to the CMS severe sepsis "time zero"
first_hypotension_sbplt90_at	10/1/2024 11:47	the first time that the patient's systolic blood pressure was lower than 90, if it occurred during the eligible time window related to the CMS severe sepsis "time zero"
first_hypotension_sbplt90_at	10/1/2024 11:47	the first time that the patient's systolic blood pressure was lower than 90, if it occurred during the eligible time window related to the CMS severe sepsis "time zero"
first_hypotension_sbp40drop_at	10/1/2024 11:47	the first time that the patient's systolic blood pressure dropped by 40 or more (i.e., the second of two contributing timestamps), if it occurred during the eligible time window related to the CMS severe sepsis "time zero"
first_hypotension_maplt65_at	10/1/2024 11:47	the first time that the patient's mean arterial pressure (MAP) was lower than 65, if it occurred during the eligible time window related to the CMS severe sepsis "time zero"
first_hypotension_stabilizes_at	10/1/2024 15:15	the first time that the patient's hypotension stabilized (i.e., MAP >65) after a documented hypotension event, if it occurred during the eligible time window related to the CMS severe sepsis "time zero"
first_vasopressors_taken_at	10/1/2024 14:15	the first administration time of vasopressors within the eligible time window for the CMS SEP-1 Bundle, if the patient is hypotensive
first_vasopressors_ordered_at	10/1/2024 12:59	the order time of the first vasopressors administered within the eligible time window for the CMS SEP-1 Bundle, if the patient is hypotensive
bundle_met_three_hour	TRUE	3-hour bundle compliance status: True (compliant) or False (not compliant)
bundle_met_overall	FALSE	Overall bundle compliance status: True (compliant) or False (not compliant)
bundle_missed_antimicrobial	FALSE	Bundle failed due to antibiotics not administered during eligible time window: True (not compliant) or False (compliant)
bundle_missed_antimicrobial_after_culture	TRUE	Bundle failed due to antibiotics not administered after blood culture collected: True (not compliant) or False (compliant)
bundle_missed_blood_culture	FALSE	Bundle failed due to blood culture not collected during eligible time window: True (not compliant) or False (compliant)
bundle_missed_iv_fluids	FALSE	Bundle failed due to iv fluids not collected during eligible time window: True (not compliant) or False (compliant) or NA (not applicable- patient was not hypotensive)
bundle_missed_lactate	FALSE	Bundle failed due to lactate not collected during eligible time window: True (not compliant) or False (compliant)
bundle_missed_repeat_lactate	NA	Bundle failed due to repeat lactate not collected during eligible time window: True (not compliant) or False (compliant) or NA (patient did require repeat lactate)
bundle_missed_six_hour_shock_vasopressors	FALSE	Bundle failed due to vasopressors not administered during eligible time window: True (not compliant) or False (compliant) or NA (not applicable- patient was not hypotensive)

NEXT PAGE →



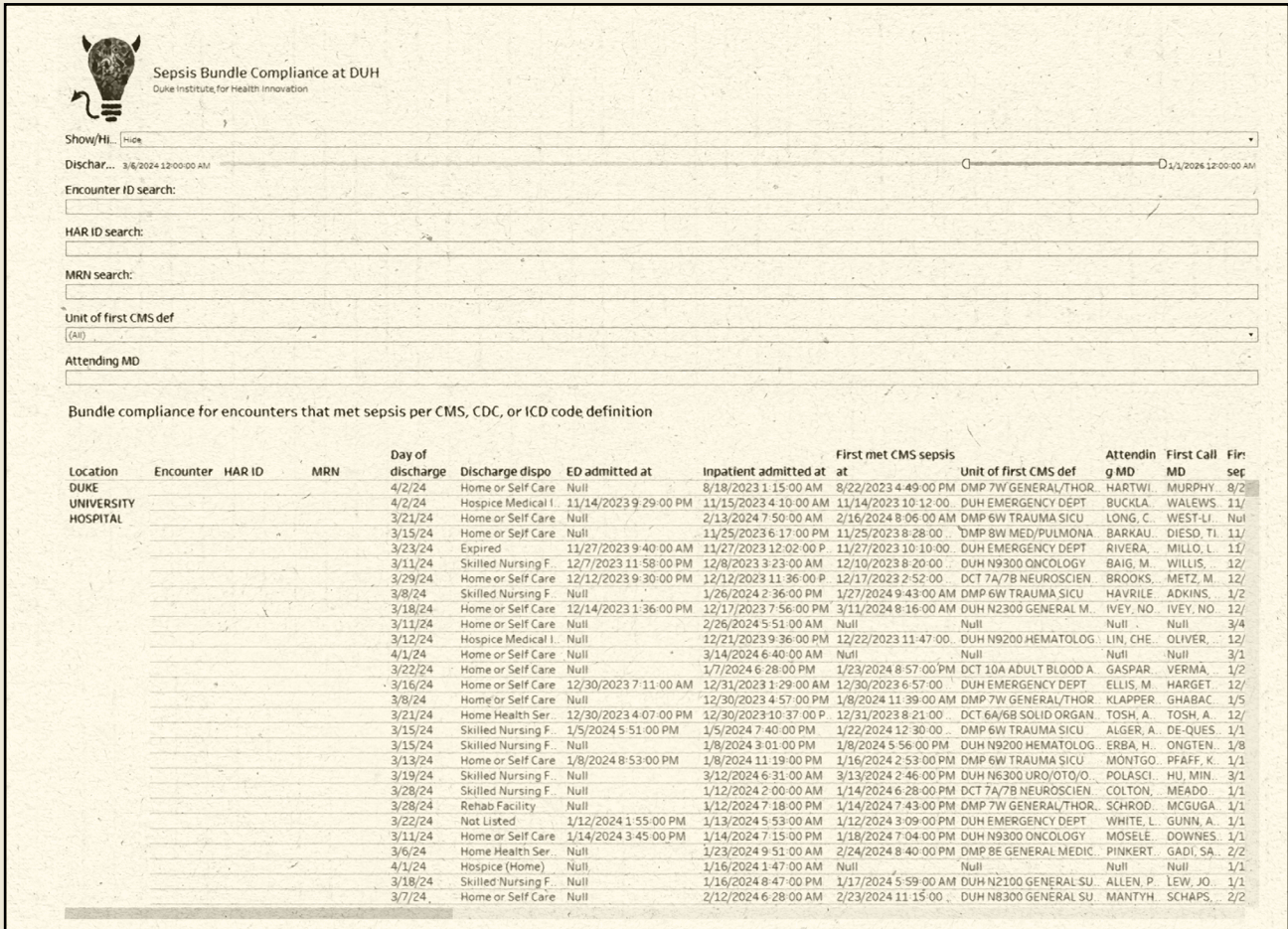
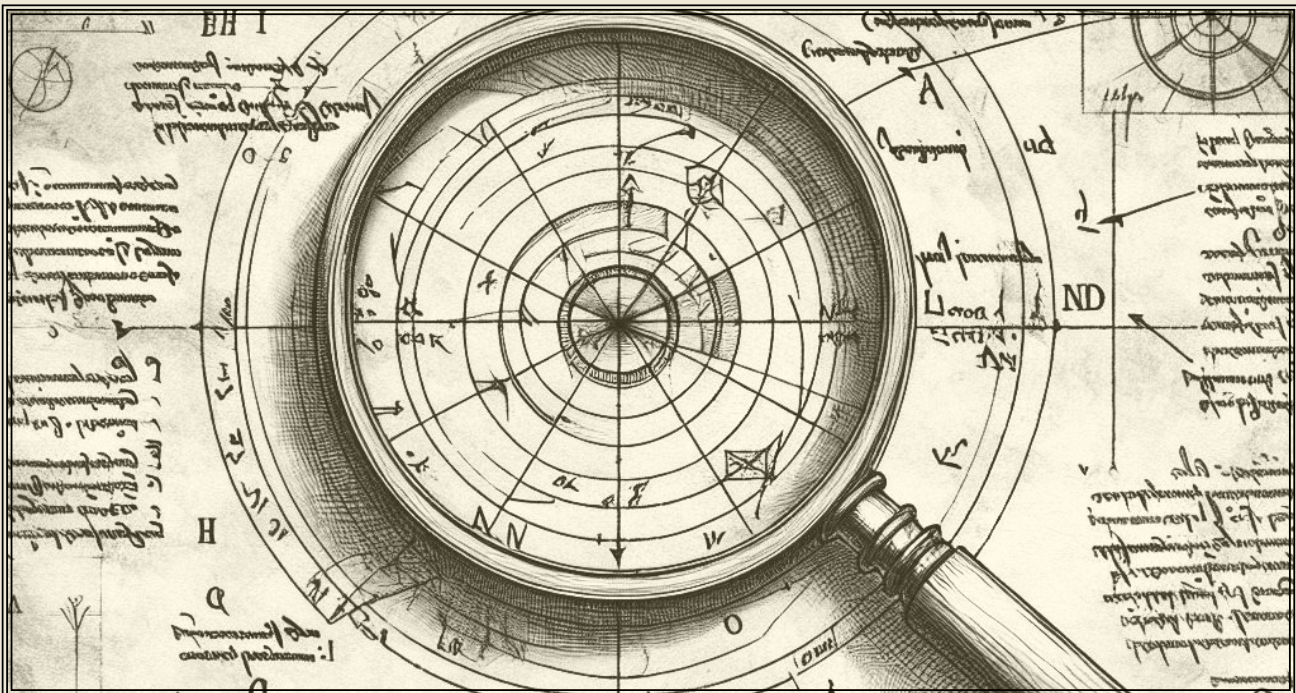


Figure 1: SEP-1 Bundle Compliance Dashboard



RECENT RFA IMPACT

Use of Machine Learning to Support Triage of Patients with Lower Back Pain

Problem

Lower back pain (LBP) remains the leading cause of years of life lived with a disability (YLLDS) worldwide and activity limitation in the world.¹ Despite innovations in therapeutic and diagnostic modalities, the impact of LBP on patients has continued to grow at an alarming rate.² The financial cost of LBP is estimated to be \$50-100 billion annually, with a significant proportion of the cost associated with indirect costs such as work absenteeism.^{3,4} Thus, the burden of LBP has been established as a significant burden on economic productivity and patient morbidity. This fiscal cost, as well as the burden of disability, make the efficient diagnosis and management of LBP a priority for both providers and healthcare systems worldwide.

Solution

To address this gap, Duke University's Departments of Neurosurgery and Orthopaedic Surgery, along with the Duke Institute for Health Innovation (DIHI), formed a interdisciplinary team to develop a machine learning model ensemble for the triage of new LBP patients using electronic health record (EHR) data. Our goal is to support a more standardized and effective intake process for LBP patients and ultimately achieve a higher rate of positive clinical outcomes for Duke patients. This project holds the promise of significantly improving the lives of patients with LBP and the efficiency of our healthcare system overall.

Our study cohort included data from 137,915 patients. The average age was 56.47 (SD: 27.21). 40% of the patients were male (n = 55,340), 63% were white (n = 87,556), and 91% of patients were not Hispanic (n = 125,074). The cohort was made up of 3,148 patients who received surgery within 90 days of their index encounter. The remaining patients were divided into sub-groups by their index intervention provider type: physical medicine and rehab ("PM&R": 8,239 patients), physical therapy ("PT": 4,889 patients), and primary care ("PCP": 52,886 patients). A total of 3 models were developed for each cohort predicting the use of opioids at 90, 180, and 365 days (3, 6, and 12 months), respectively.

TEAM

Muhammad Abd-El-Barr, MD, PhD
Christine Goertz, PhD
Will Ratliff, MBA
Bruno Valan, MEng
Andreas Seas
Abdullah Saleh
Michael Gao, MS
Suresh Balu, MS, MBA

IN BRIEF

Lower back pain (LBP) patients have a fragmented care pathway at Duke with multiple entry points and index interventions. Due to the variety of patient intake settings, presentations, and underlying etiologies, adherence to guideline-concordant care is challenging. Poor outcomes for LBP interventions, such as continued pain and opioid use, and a low conversion rate for surgical consult-to-surgery, present an opportunity for better prognostication of patients with LBP at time of triage. Our team has developed an ensemble of machine learning models to inform intake schedulers about the "right provider" for index intervention on patients with LBP. We are implementing the models to run in real-time. We will evaluate their performance with outcome metrics that include prolonged opioid use, reduction in pain, and urgent/emergent care visits due to complications.

NEXT PAGE →

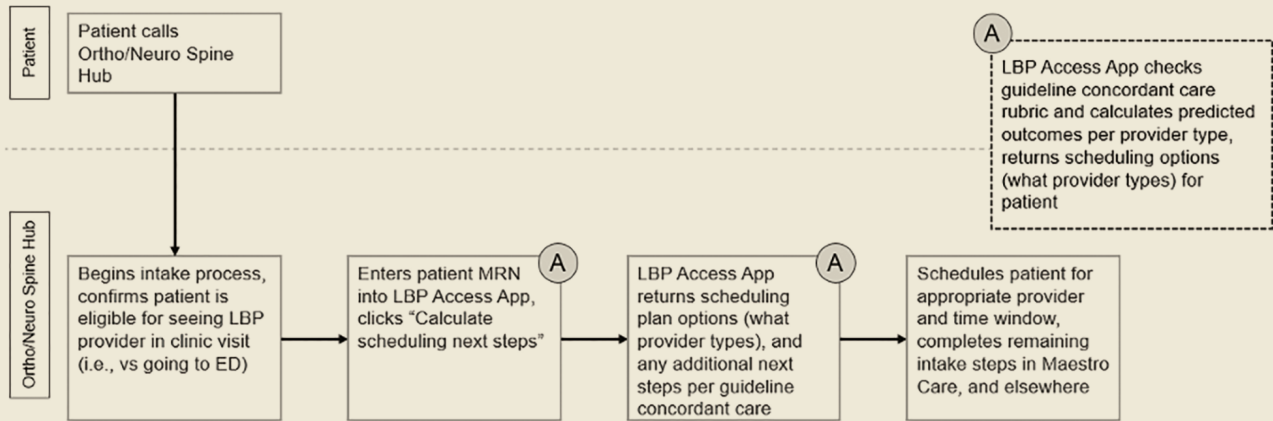


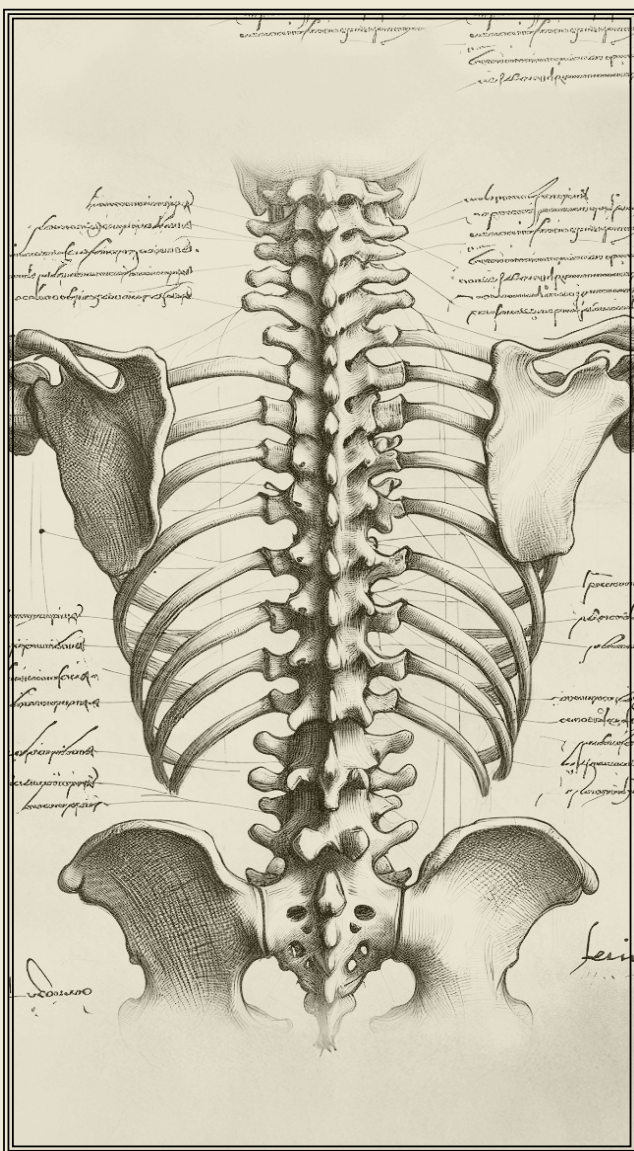
Figure 1: Future State Workflow

COHORT	OUTCOME	PERFORMANCE METRIC		OUTCOME PREVALENCE
		AUPRC	AUROC	
Early surgery (n=3,148)	Opioid Use @90 days	0.62	0.66	40%
	Opioid Use @180 days	0.47	0.67	29%
	Opioid Use @365 days	0.35	0.68	20%
PM&R (n=8,239 patients)	Opioid Use @90 days	0.58	0.71	32%
	Opioid Use @180 days	0.49	0.70	28%
	Opioid Use @365 days	0.40	0.70	20%
PT (n=4,889 patients)	Opioid Use @90 days	0.48	0.67	32%
	Opioid Use @180 days	0.40	0.68	26%
	Opioid Use @365 days	0.33	0.70	18%
PCP (n=52,886 patients)	Opioid Use @90 days	0.46	0.69	27%
	Opioid Use @180 days	0.43	0.69	23%
	Opioid Use @365 days	0.34	0.68	16%

Table 1: Model Performance



Our team has developed an ensemble of machine learning models to inform intake schedulers about the “right provider” for index intervention on patients with LBP. We are implementing the models to run in real-time.



Impact

For the cohort who received early surgery, the prevalence of opioid use was 40%, 29%, and 20% at 90, 180, and 365 days, respectively. The model’s performance in predicting opioid use showed an AUPRC of 0.62 and an AUROC of 0.66 at 90 days. At 180 days, the AUPRC was 0.47, and the AUROC was 0.67. By 365 days, the AUPRC had decreased to 0.35, while the AUROC improved to 0.68. Similar performance patterns were seen in the non-surgery groups. The results can be seen in Table 1.

Next Steps

We are expanding the model prediction outcomes to include pain reduction and emergent healthcare utilization, which will provide additional data points to support improved patient scheduling decision-making. We plan to pilot the LBP triage solution in the Duke University Health System (DUHS) Spine Center and select primary care sites to evaluate its performance. The clinical workflow utilizes a patient prediction calculation application while a patient is routed at the spine center. At the conclusion of the pilot period, we will evaluate the solution’s impact on the surgical consult-to-surgery rate and the clinical outcomes of prolonged opioid use, reduction in pain, and emergent hospitalizations.

References

1. GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Lond Engl.* 2016;388(10053):1545–1602. doi:10.1016/S0140-6736(16)31678-6
2. Vos T, Allen C, Arora M, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet.* 2016;388(10053):1545–1602. doi:10.1016/S0140-6736(16)31678-6
3. Hoy D, March L, Brooks P, et al. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis.* 2014;73(6):968–974. doi:10.1136/annrheumdis-2013-204428
4. Fatoye F, Gebrye T, Mbada CE, Useh U. Clinical and economic burden of low back pain in low- and middle-income countries: a systematic review. *BMJ Open.* 2023;13(4):e064119. doi:10.1136/bmjopen-2022-064119

RECENT RFA IMPACT

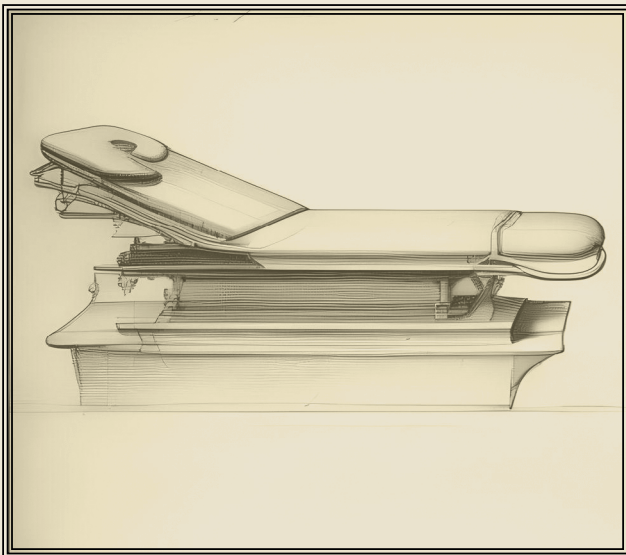
Duke Raleigh Hospital Automated Staff and Bed Assignment

Problem

Excess days in the hospital are driven by multiple factors. One significant factor affecting the length of stay is care team design. In 2021, patient care teams comprised of physicians, case managers, physical therapists, occupational therapists, and nurses were assigned through an ad hoc process. This method did not consider team members' familiarity with patients or optimize care team composition, leading to inefficiencies, poor communication, and extended hospital stays. Consequently, staff satisfaction decreased, and workload increased. Analysis revealed that prior patient assignment methods had wasted 4,700 hours yearly, costing more than \$300,000.

Solution

In 2022, Hospital Medicine leaders at Duke Raleigh, including lean transformation coaches, structured patient care teams and created a manual Excel-based process to document care teams, staff availability, and new floor patients. This approach aimed to support optimal staff-to-patient and staff-to-team matching. However, the manual process was labor-intensive and needed to be more scalable across the health system.



The Duke Institute for Health Innovation (DIHI) aimed to automate data entry for the Excel-based Scheduling Assistant Tool to streamline this process. DIHI used its Bed Watch solution to identify patient unit and bed locations in real time. By integrating real-time data pipelines, DIHI extracted physician and nurse scheduling data from QGenda and Symplr and treatment team relationships data from Epic APIs. This streamlined approach aimed to reduce time and errors in data entry and validation.

Impact

DIHI successfully integrated care team assignments into its data pipeline and created a dashboard to showcase this information. However, eliminating manual entry of patient names and locations, care member names, and their latest patient assignments provided negligible time savings. Unit-leading nurses using the spreadsheet already had timely knowledge of assignments, making the automated data extraction less valuable. As long as a unit leader formed the spreadsheet for the unit, its use was straightforward.

The project's focus shifted to curating incoming care teams and developing an algorithm for recommended provider-to-patient assignments and ideal care team mixes. Despite identifying incoming providers, challenges remained with linking them to specific floors, excluding travel and floating nurses, and balancing care assignments based on patient acuity and proximity.

While DIHI could estimate acuity by providing patients' Duke history of chemotherapy, restraint orders, or infection labs, the estimate would be incomplete compared to nurses' local undocumented knowledge of teammate proficiency in managing multiple patients with complex conditions or behaviors.

Given the incomplete information about incoming staff and the negligible time savings from other automation, the project was suspended. This decision was made after careful consideration of the project's goals and the challenges it faced. The team will revisit the automation approach once Duke University Health System and the Duke Office of Information Technology resolve the identified challenges.

TEAM

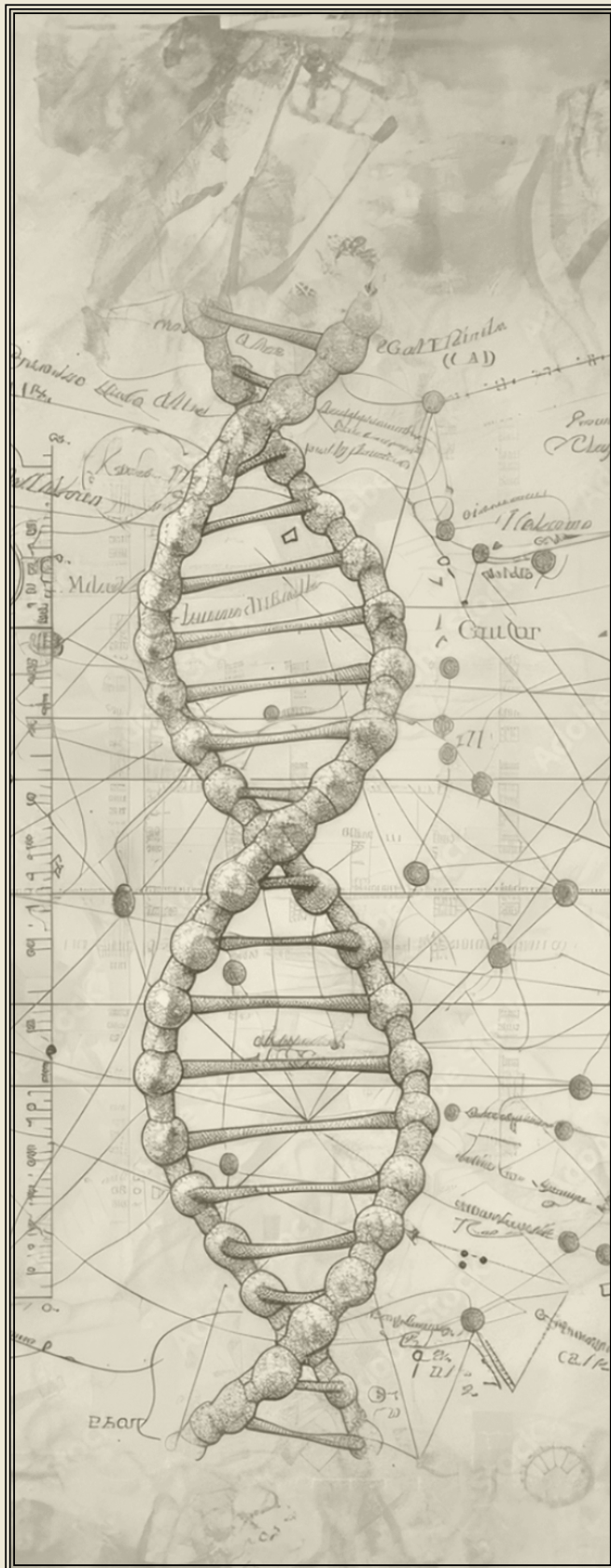
Zahra Mahbooba, MPH
Peter Berry, DNP
Mike Revoir, BS
Matt Gardner, BS
Will Knechtle, MBA, MPH
Eliseu Chuang, MD, MHA

IN BRIEF

A 2021 analysis at Duke Raleigh revealed inefficiencies in hospital care team design, leading to extended patient stays, extra work hours, and increased costs. In response, leaders introduced a well-formatted Excel process to optimize team assignments. DIHI explored automating several of its features using a real-time data pipeline. Despite improving data flow, the automation could not compete with nurses' specialized patient knowledge and incomplete information about incoming staff. The project was paused but was leveraged by DIHI to send alerts to a patient's current on-call provider.

Next Steps

Across a health system, identifying real-time care team members and matching them to their patients is crucial for clinical decision support. For example, early warnings of maternal sepsis or hemorrhage should be sent to the first-call provider and nurse best able to respond. Now, DIHI can automatically send those alerts quickly and specifically. DIHI filled this critical need by successfully developing a real-time pipeline for extracting physician and nurse-patient assignments, scheduling data from QGenda and Symplr, and treatment team relationships from Epic APIs. Today, DIHI can target information signals to the appropriate care team members.



CHAPTER II

EXPANDING THE PORTFOLIO



HEROIC-PrEP: Health Record Optimization to Identify Candidates for Pre-exposure Prophylaxis

Alifia Hassan, B.Pharm, MBA; Maryam Aziz, MS



Problem

Despite the availability of pre-exposure prophylaxis (PrEP) for human immunodeficiency virus (HIV) prevention, its adoption has been notably low, particularly in the South and among women. In 2020, Louisiana reported the fourth-highest HIV case rate in the nation (15.6 per 100,000 population), with its PrEP-to-need ratio ranked 46th nationally, indicating a significant gap in PrEP usage. Additionally, individuals with HIV in the South had on average, seven healthcare encounters prior to their diagnosis, showing missed opportunities for earlier identification and PrEP intervention. The existing HIV risk prediction tools based on Centers for Disease Control and Prevention (CDC) PrEP criteria consistently underperform in identifying at-risk populations. There is an urgent need to develop and implement feasible and effective strategies to identify PrEP candidates. The underperformance of existing HIV risk prediction tools based on CDC PrEP criteria exacerbates the challenge of low PrEP uptake in the fight against HIV.

Solution

As a potential solution to these challenges, electronic health record (EHR)-based models can serve as a useful tool to enhance population-level HIV prevention initiatives. In collaboration with frontline infectious disease care providers, the Duke Institute for Health Innovation (DIHI) developed an electronic health record (EHR) based HIV risk prediction model. This model was designed to identify individuals who would benefit from PrEP and was validated initially at DUHS. The model was developed to detect the risk of HIV in individuals seeking longitudinal care within Duke Health. The EHR data could also be used to provide important insight into the HIV risk of persons not engaging in longitudinal or primary care, including those receiving care in community-facing emergency departments (ED) and urgent care (UC) facilities.

State	2018 HIV Incidence Rate Per 100,000 (State Rank)	2018 HIV Incidence Rate Per 100,000 Females (State Rank)	2019 Prep to Need Ratio* (State Rank)	2019 Prep to Need Ratio, Females (State Rank)
Louisiana	25 (T-2 nd)	13 (1 st)	3.08 (46 th)	2.12 (30 th)
<i>Orleans Parish</i>	53	23	3.76	3.8
<i>E. Baton Rouge Parish</i>	48	31	2.34	1.72
Alabama	15 (T-9 th)	6 (T-8 th)	4.13 (39 th)	1.46 (42 nd)
Mississippi	19 (T-6 th)	8 (5 th)	2.03 (50 th)	1.15 (46 th)
New York	15 (T-9 th)	6 (T-8 th)	12.62 (10 th)	5.01 (9 th)
Washington	8 (T-28 th)	4 (T-17 th)	12.7 (9 th)	2.97 (23 rd)
United States	13.6	5.1	6	2.34

Table 1. HIV Incidence and PrEP Use in Selected States

This project has opened new avenues for the diffusion and scaling of the technologies built at DIHI. The EHR-based HIV risk prediction model will be implemented within OLOL Health system and LCMC Health

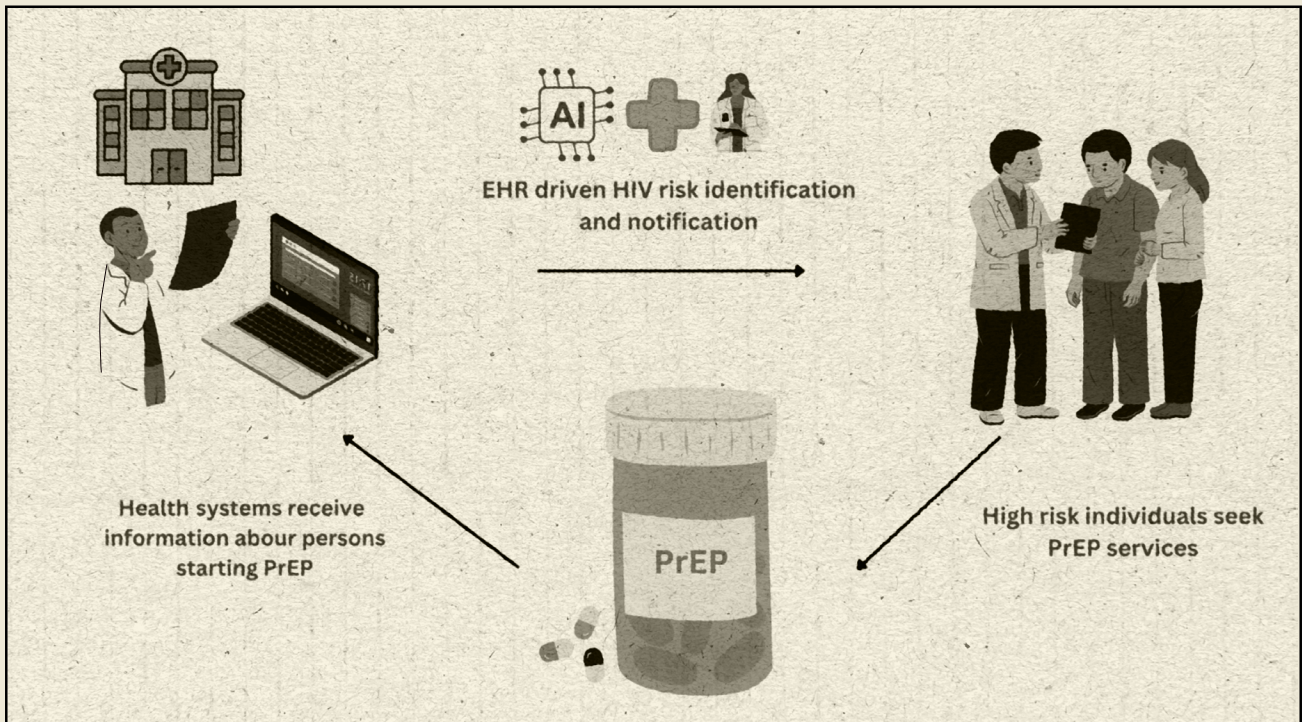
Our team at DIHI is building upon previous work on HIV risk prediction as part of the NIH-funded R01 project entitled “HEROIC-PrEP: Health Record Optimization to Identify Candidates for Pre-exposure Prophylaxis” led by Dr. Nwora Lance Okeke of Duke University, School of Medicine and Dr. Meredith Clement of Louisiana State University Health Sciences Center-New Orleans. This project also incorporates community-informed risk factors along with EHR-based data elements in the model development. The HIV risk prediction model will be built using clinical data from patients primarily receiving acute care in two large health systems in Southeastern Louisiana, Our Lady of the Lake (OLOL) Health System in Baton Rouge and LCMC Health in New Orleans.

Additionally, we will conduct a pilot trial to assess the feasibility and acceptability of two implementation strategies to expand the use of the EHR-model output beyond the health system and into the community, with input from stakeholders. The goals of this proposal are:

1. To create and validate an EHR-based HIV risk prediction model using clinical data from acute care encounters within two health systems in Southeastern Louisiana.
2. To develop implementation strategies, informed by stakeholders, to extend the reach of the EHR-based prediction model beyond the health system.
3. To assess the feasibility and acceptability of community-facing implementation approaches for deploying the EHR HIV risk prediction model.

Impact and Outcome

This project has opened new avenues for the diffusion and scaling of the technologies built at DIHI. The EHR-based HIV risk prediction model will be implemented within OLOL Health system and LCMC Health. This integration will require data infrastructure to be built in both institutions, and if successful, will demonstrate the scalability of the technology and its adaptability to diverse healthcare environments and systems. The most notable achievement of the project to date has been the successful engagement of these two major health systems to develop a robust data infrastructure.



NEXT PAGE →



This strategic partnership has facilitated conversations and solutions around how the data should be extracted, stored, cleaned, and harmonized across both institutions in similar geographies in a way that can inform the collaborative implementation of the EHR-based HIV risk prediction model within these systems, thereby enhancing their capacity for HIV prevention. At the time of this publication, over two million encounter records, spanning seven years of Electronic Health Record (EHR) data (2017-2023), are being processed to develop a robust HIV incidence model. Manual adjudication of positive HIV cases through chart reviews by healthcare professionals in the two institutions is currently underway to ensure data accuracy.

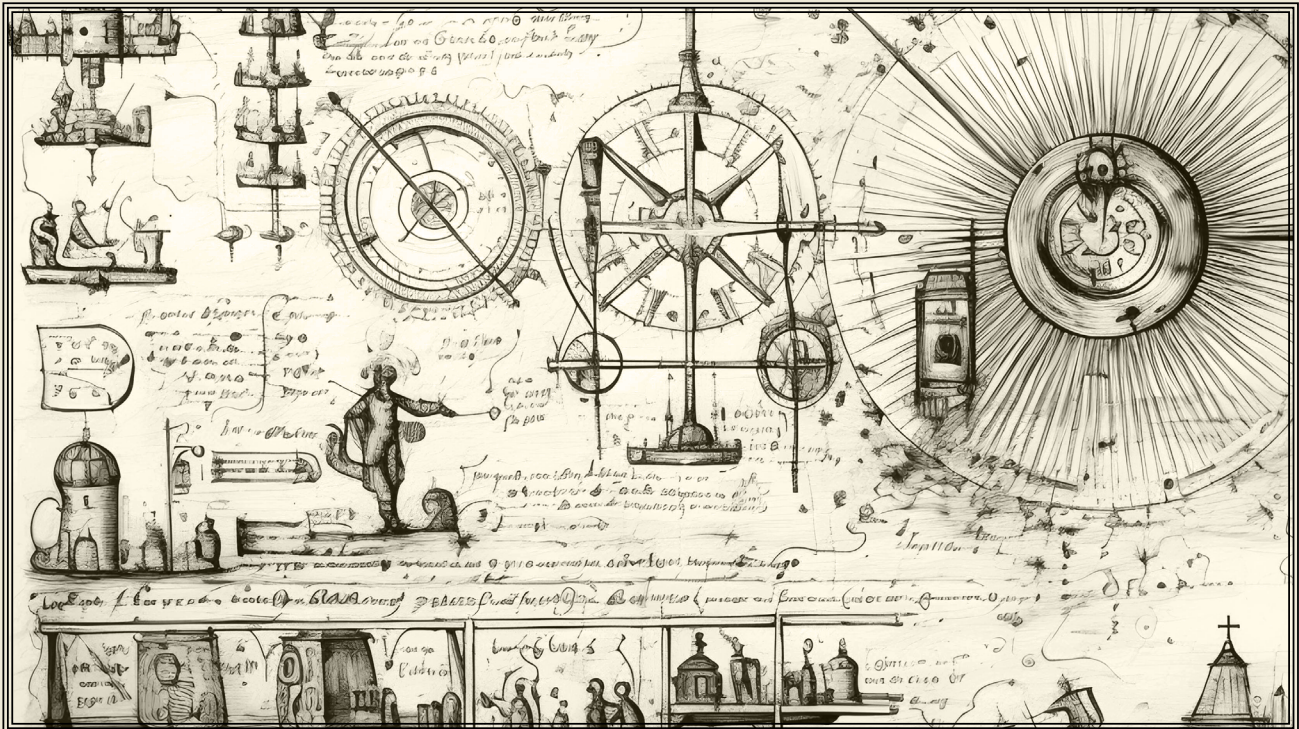
The significant innovation of this project is the use of community participation to inform model development based on knowledge of their local HIV epidemic. The next step of HEROIC-PrEP will determine if model features of HIV risk informed by community partner input leads to an improvement in model performance compared to the baseline iteration of the model. By combining comprehensive data with community insights, the teams are working towards a predictive model that will inform targeted interventions, helping reduce new infections and improve access to preventive care like PrEP in a state that suffers from a low PrEP-to-need ratio.

In light of recent advances in HIV prevention modalities, particularly the emergence of long-acting injectable PrEP, the DIHI/HEROIC-PrEP model is well-positioned to enable the efficient deployment of this high-impact public health intervention, a potentially significant step towards ending the HIV epidemic in Southeastern Louisiana.

Next Steps

The project will be a model for predictive AI technology dissemination and scaling at other institutions, ensuring local validation of these tools. This represents a significant step forward in using healthcare informatics and predictive analytics to improve population-level clinical outcomes relevant to a major public health threat. As this project progresses, we can surface technical and operational challenges in transferring and customizing technology-based solutions in other geographies and institutions.

As we move forward, we will continue to leverage our expertise in healthcare informatics and predictive analytics to develop innovative solutions for pressing healthcare challenges. We will also continue to work closely with our community partners to ensure that our efforts align with the needs of the communities in which the HIV risk prediction model will be implemented.



EXPANDING THE PORTFOLIO

Cardiogenic Shock

Henry Foote, MD



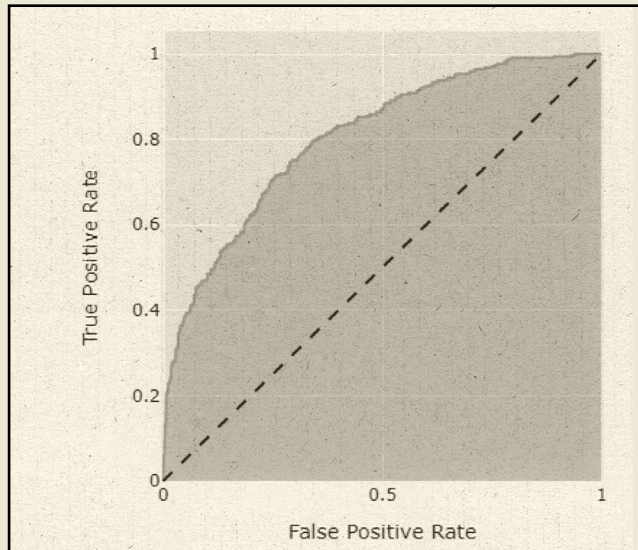
Problem

Cardiogenic shock is a physiologic state in which the heart can no longer adequately pump blood and oxygen to the body. Heart attacks and acute worsening of heart function in patients with underlying heart failure are the most common causes. Patients with cardiogenic shock are at high risk of kidney injury, respiratory compromise requiring mechanical ventilation, and death. Mechanical circulatory support such as intra-aortic balloon pumps, Impella heart pumps, and veno-arterial extracorporeal membrane oxygenation (VA-ECMO) can provide circulatory support for patients with cardiogenic shock.

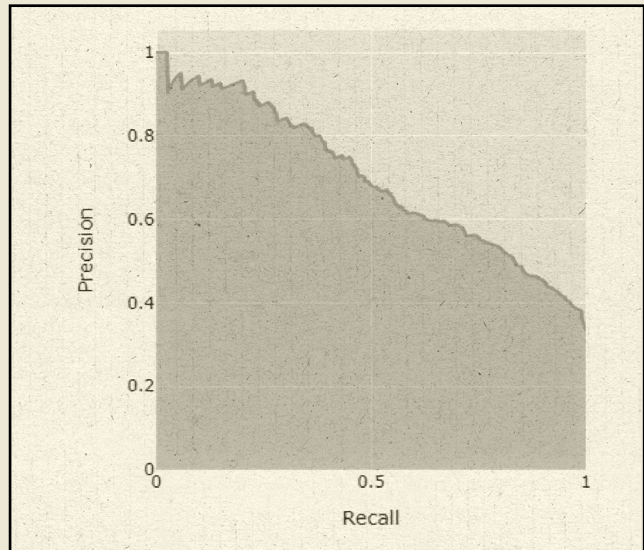
Yet, despite advances in medical technology and the coordinated efforts of medical teams, patients with cardiogenic shock have over 30% hospital mortality at Duke. This figure underscores the urgent need for a solution.

Solution

We aimed to develop a cardiogenic shock mortality model to improve outcomes for these patients. This model would leverage the infrastructure created by the Duke Institute for Health Innovation (DIHI) team for prior projects focused on sepsis and adult deterioration. Here, we trained a model to predict mortality risk at 12 hours and 24 hours into an admission. Similar to prior models, we incorporated the rich data available in the electronic health record, including demographics, comorbidities, vital signs, lab values, and nursing flowsheet documentation. Additionally, we added features reflective of patients' underlying cardiac physiology, including data from electrocardiograms, echocardiograms, and heart catheterizations. These inputs allow for a more complex and complete representation of patients' conditions than has been used in prior models developed by other institutions.



ROC Curve



Precision Recall Curve

The five most important features in predicting mortality were age at admission, mean serum albumin, mean Glasgow coma score, mean platelet value in the first six hours of admission, and total urine output.

NEXT PAGE ➔



Our model development also leveraged the AutoML pipeline for model training. This process, developed by the DIHI team, allows for rapid training of hundreds of models to build the best final model.

We aimed to develop a model that could be implemented in real-time to quickly and accurately identify the most at-risk patients, a development that holds the promise of significantly improving patient outcomes and supporting real-time clinical decision-making.

Progress

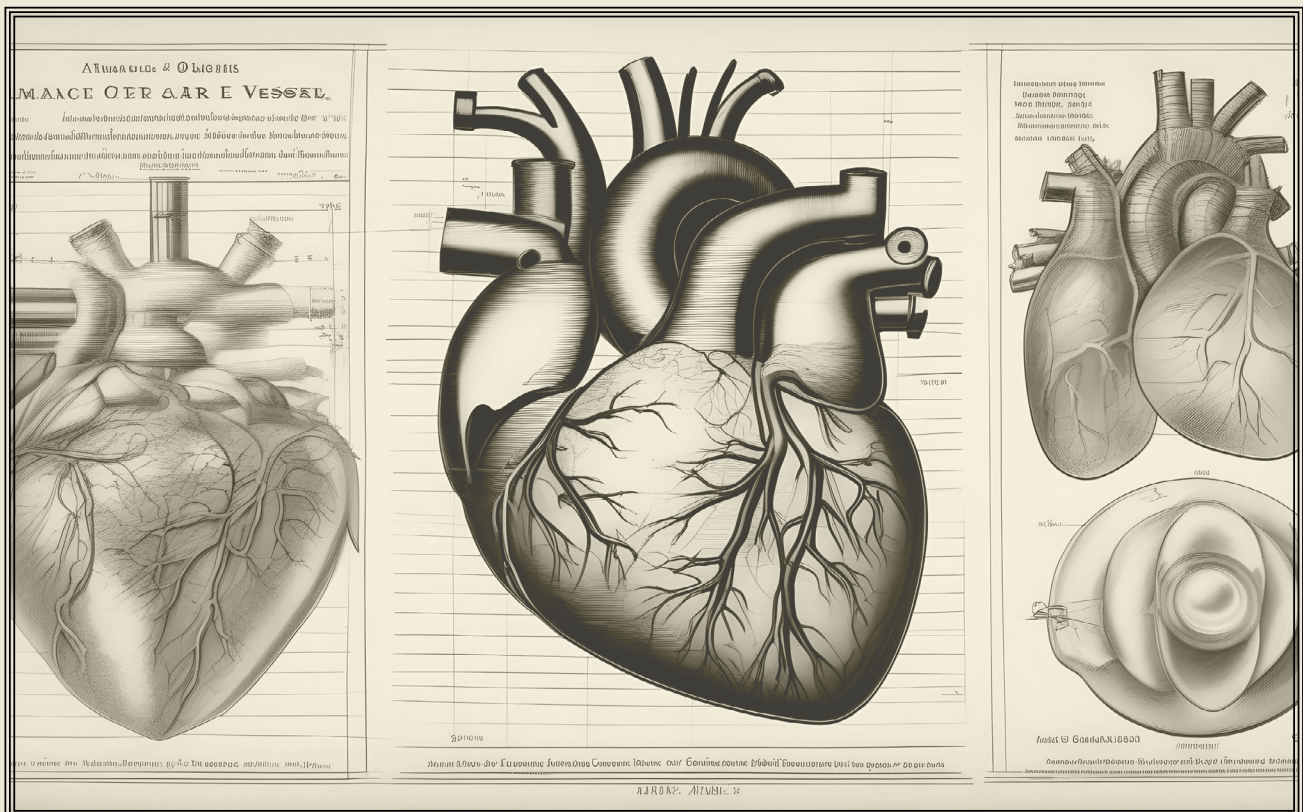
Our models at 12 and 24 hours after admission could reliably predict mortality risk, as evidenced by the area under the receiver operator characteristic curve (AUROC), the area under the precision-recall curve (AUPRC), and calibration. Notably, our model also outperformed a previously developed cardiogenic shock mortality prediction model, a testament to its technical robustness and reliability.

We are currently working on a ‘silent rollout’ of the model, which involves prospectively predicting mortality in patients to ensure its continued performance. Additionally, we are actively collaborating with other hospitals to test the model’s performance on their data, further validating its effectiveness and applicability in different clinical settings.

Next steps

Overall, it is encouraging that our model can accurately predict mortality in patients with cardiogenic shock. If our model continues performing well on prospective testing, we envision improving care in several ways. It will enable efficient identification of the most at-risk patients and rapid high-quality triage and care escalation. Interventions may include additional nursing support, consideration for additional therapies, or even transfer to a different hospital. The model will also allow for medical care to be tailored to patient and family goals by encouraging more informed discussions. We see this project as a collaborative effort between DIHI and the clinical teams at the patient bedside and the first step in improving outcomes in adults with cardiogenic shock. This project demonstrates DIHI’s commitment to leveraging data science and machine learning to create real-world solutions that directly improve the care of critically ill patients.

We look forward to continuing with our work.



EXPANDING THE PORTFOLIO

Federated Learning: Revolutionizing Healthcare AI with Owkin

Bradley Hintze, PhD



Machine learning (ML) has emerged as a powerful tool in healthcare, with algorithms analyzing large datasets to uncover hidden patterns and improve incidence recognition, treatment, and drug discovery. The utility of an ML-using tool is proportional to how well its results match reality. Meanwhile, the effectiveness of any ML model hinges on the size and quality of the data it's trained on. In healthcare, this presents a unique challenge, as data is often fragmented across different institutions, raising privacy concerns and limiting the scope of research.

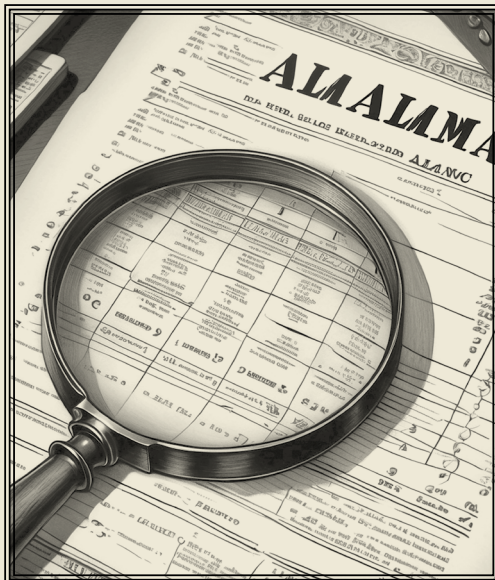
For example, when studying complicated or rare diseases, curating a dataset that meets the requisite size and quality parameters becomes very difficult. One challenge is that a hospital system usually only has access to its own patients' data. Furthermore, an ML model might be designed with a particular cohort – a subset of that data. This specificity will necessarily reduce the number of potential data points that can be fed into the model. Regarding quality, the data must be curated to be

formatted correctly and only include patients with all requisite data. These considerations can significantly reduce what one health system can do.

Owkin, an AI biotech company, offers a solution to this challenge through federated learning, an innovative approach to ML that allows models to be trained on decentralized datasets without compromising patient privacy. This means that instead of pooling sensitive patient data from multiple hospitals in a central location, the ML model is sent to where the data resides – on the individual hospitals' servers. Each hospital can then train the model on its own data, and only the learned model parameters are shared back with Owkin. This allows sensitive data to be kept safely on-premises while developing more robust and generalizable models from extensive datasets without compromising patient privacy.

Owkin works with numerous academic medical centers to curate many datasets that reside locally in the individual health center's servers. They employ a high-touch collaboration process to ensure the quality of the datasets. They collaborate with multiple health centers to get similar datasets to address size. Owkin leverages its expertise to effectively communicate specific dataset requirements and high parameter inclusion criteria.

The Duke Institute for Health Innovation (DIHI) has partnered with Owkin to explore the potential of federated learning in studying complex diseases like hypertrophic cardiomyopathy (HCM) and heart failure with preserved ejection fraction (HFpEF). DIHI, leveraging its expertise in clinical data management and curation, provides high-quality datasets to train Owkin's models. (Read the article about Poseidon Metadata Management on [page 29](#), which describes how DIHI has developed sophisticated systems



NEXT PAGE →

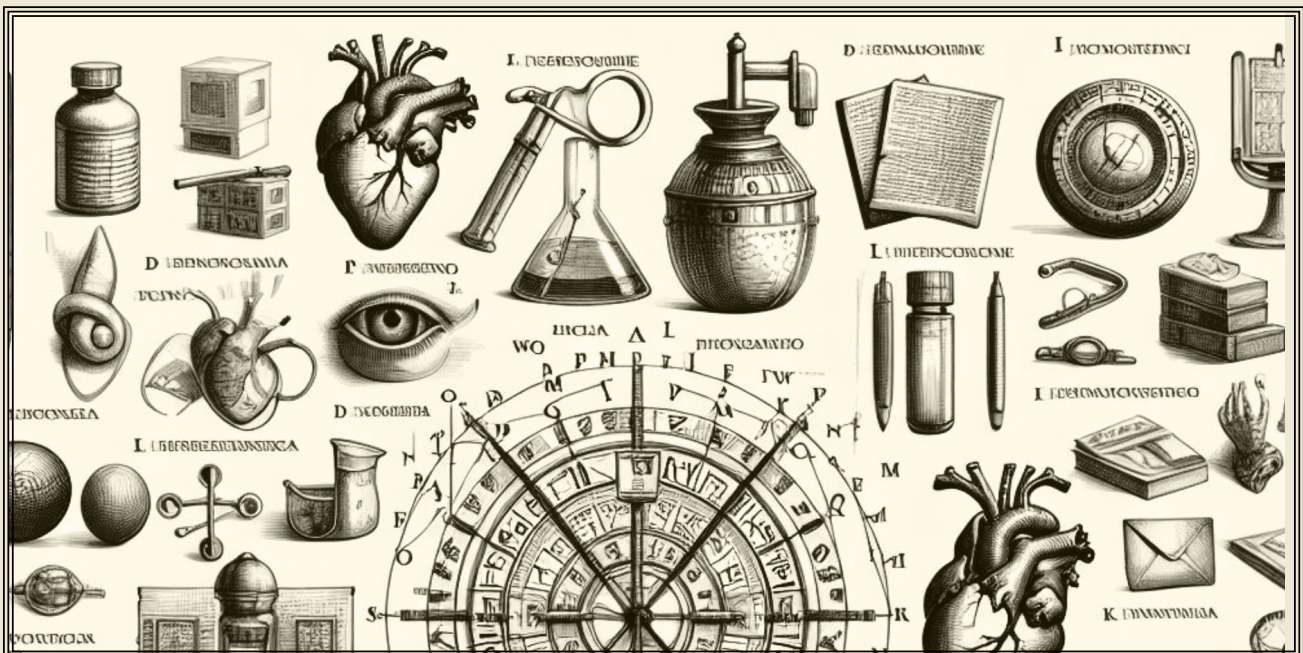
for managing clinical metadata, ensuring that data is accurately categorized and readily accessible for analysis.) High-quality data is critical because datasets of any size—federated or not—can become ineffective if the data is poorly formatted, incomplete, or irrelevant. Owkin’s emphasis on and need for data quality aligned with DIHI’s focus on robust data management practices, enabling both organizations to enhance the accuracy and reliability of their healthcare data analytics initiatives.

This collaboration has already yielded promising results, such as the preliminary finding that beta-blockers or non-dihydropyridine calcium channel blockers might not reduce the risk of progression in non-symptomatic HCM patients. This was found early in data analysis, with more results yet to be published. Despite Duke’s high activity in cardiac health, the need for comprehensive and stringent inclusion criteria underscores the need for federated learning to achieve dataset size.

The collaboration between Owkin and institutions like DIHI exemplifies the potential of federated learning in healthcare AI, addressing critical challenges of data privacy and fragmentation while enabling more comprehensive and accurate machine learning models. This innovative approach, as demonstrated by the preliminary findings from the HCM study, may accelerate medical research and

improve patient outcomes across a wide range of complex diseases. By enabling the analysis of diverse, decentralized datasets without compromising patient confidentiality, federated learning is poised to democratize access to large-scale, high-quality datasets, allowing smaller institutions to contribute to and benefit from cutting-edge research. The synergy between Owkin’s technology and DIHI’s expertise in data management illustrates how interdisciplinary partnerships can drive innovation in healthcare AI, fostering a more inclusive and comprehensive approach to medical research and discovery. As we navigate the complexities of healthcare in the digital age, federated learning offers hope for leverage data to improve human health and well-being.

High-quality data is critical because datasets of any size—federated or not—can become ineffective if the data is poorly formatted, incomplete, or irrelevant.



EXPANDING THE PORTFOLIO

Understanding and Managing Clinical Metadata with Poseidon

Bradley Hintze, PhD; Jamie Daniel, BS



Artificial intelligence (AI) has been transforming the healthcare landscape. Machine learning (ML) algorithms are now crunching through mountains of medical data, uncovering hidden patterns to aid in disease diagnosis, treatment planning, and drug discovery. From analyzing medical images for signs of cancer to personalizing treatment plans for chronic conditions, AI may accelerate healthcare, enhance accuracy, and even prevent disease from worsening. This technology holds immense promise for improving patient outcomes and revolutionizing medicine delivery.



As with any innovation, not all ML models have shown such promise. Some have demonstrated clear racial bias. Users of ML models released in the wild have been disappointed. The issue with such cases often has to do with the real-world input data. The model might be overfitting the training data, performing well on the pre-defined validation set but failing to generalize to real-world scenarios. One possible explanation is clinical metadata, an underappreciated component of ML models. Problems with metadata will degrade the quality of inputs and, thus, the model's performance. In the ever-evolving field of clinical research and healthcare, the management of clinical metadata is paramount.

What is clinical metadata? Metadata is data that describes and gives information about other data. Clinical metadata, which describes patient demographics, lab results, and treatment outcomes, is the backbone for various analytical processes. Effective metadata management ensures accurate, reliable, and timely insights that drive clinical decisions and research advancements. Tools designed to manage metadata have become indispensable, offering functionalities that streamline data organization, accessibility, and integration.

One of the key challenges in harnessing the power of clinical metadata lies in making data relatable and easily accessible. These tools must handle complex datasets, perform advanced analytics, and remain intuitive for users without extensive technical backgrounds. User interfaces become more accessible when designers prioritize simplicity and efficiency.

This dual focus on functionality and ease of use ensures that a wide range of healthcare professionals, from data scientists to clinicians, can effectively leverage these tools. By focusing on these aspects, developers ensure that their tools are technically capable and practically useful during day-to-day clinical and research activities.

The inherent complexities of clinical data itself compound this challenge. Within healthcare data, many clinical concepts have numerous names in electronic health records (EHRs). A few examples are blood pressure, glucose, and beta blockers. All of these have multiple identifiers.

- Blood pressure can be collected using many different methods, and each has a different label stored in the raw database.
- Healthcare professionals may test glucose on its own or as part of various panels, all of which have different names.
- Beta-blockers may have several identifiers, and the dose may be part of it.

All these different names represent one clinical concept: blood pressure, glucose, or beta blockers. ML models are generally interested in the underlying clinical concept rather than the specific names. They need to know your blood pressure at a given time or whether you've taken beta-blockers, regardless of the myriad names used to represent these concepts in the data.

NEXT PAGE →



A naïve approach would be to look in the EHR for all names that match the name of the concept, e.g., blood pressure or glucose. While one search term is likely to find results, this doesn't mean all data points related to the searched concept have been captured. This is due to data names not necessarily having the same name as the concept. An easy-to-understand example is vasopressors. Searching for 'vasopressor' in the name of the medication administered is inadequate to capture all vasopressors; you must search for epinephrine, norepinephrine, ephedrine, dopamine, and others.

Clinical metadata management offers a solution by categorizing clinical data into the essential categories that ML models require. Data scientists often refer to these categories as "groupers", each representing a single clinical concept and containing a list of names that match that concept.

The Duke Institute for Health Innovation (DIHI) has created a system that effectively manages clinical metadata using automation, a database, and a curation process. We have developed Poseidon, a groundbreaking tool that exemplifies the fusion of usability and power in metadata management. Poseidon excels in its ability to group related metadata from clinical data sources, binding them into a cohesive and meaningful base tool (It becomes a fundamental software component providing functionality needed for other solutions to build on or customize without starting from scratch). This allows statisticians, data scientists, and clinical researchers to access real-time, up-to-date metadata. Poseidon's design facilitates the grouping of related data points, providing a comprehensive view that enhances analytical capabilities and supports the development of robust clinical solutions. Its user-friendly features empower healthcare professionals by providing a clear and comprehensive view of related metadata, leading to more accurate and detailed analyses.

DIHI's robust system continuously examines labs, flowsheets, orders, and medications, updating the clinical metadata database and ensuring that groupers remain accurate. However, maintaining and updating these groupers presents an ongoing challenge. As healthcare practices evolve and new tests or treatments emerge, existing groupers might miss capturing this new data, leading to incomplete or inaccurate results. For example, if a new assay is introduced to measure bilirubin, older groupers might not include it, creating a discrepancy in the data. (See Box 1)

To address this, DIHI is further refining its automation and processes. Our plan involves using clinician-provided search terms to automatically identify potential new candidates for inclusion in groupers. When a new candidate is detected, the system flags it for project managers' and clinicians' review before it is published

into the DIHI Data Pipeline. This ensures that the groupers remain comprehensive, correct, and up-to-date.

Ultimately, clinical metadata management, while meticulous and complex, is paramount for realizing the full potential of AI and ML in healthcare. The quality of the metadata directly impacts the quality of the models, and degraded model results can erode trust in these innovative technologies. Managing clinical metadata through advanced tools like Poseidon represents a significant step forward in healthcare technology. The combination of user-friendly interfaces and powerful functionalities, achieved through automation, databases, and meticulous curation processes, ensures that a broad spectrum of users can effectively leverage these tools. DIHI's ongoing efforts in this domain underscore the critical role of metadata management in shaping the future of healthcare. While automation streamlines many aspects of metadata management, ongoing clinician input and project management remain essential to ensure accuracy and adapt to new clinical concepts and data sources. As the field continues to evolve, the innovation behind clinical metadata management will remain a critical factor in advancing healthcare and ensuring the success of ML innovation by improving clinical research and patient care.

The Duke Institute for Health Innovation (DIHI) has created a system that effectively manages clinical metadata using automation, a database, and a curation process.

Box 1.

Suppose Duke Health labs adopt a revolutionary new assay to measure bilirubin. In that case, bilirubin groupers created before the new assay will not capture this new result. If healthcare professionals suddenly use this new assay widely, any model running that requires bilirubin will be missing this data point; it will look as if patients are not getting bilirubin measured when they actually are. This degrades model results. Such omissions are challenging to identify.



AI GOVERNANCE & ETHICS



Framework for Assessing the Impact of AI on Health Equity: Health Equity Across the AI Lifecycle (HEAAL)

Jee Young Kim, PhD



Artificial intelligence (AI) use in clinical care has numerous benefits. For example, AI can support therapeutic and diagnostic decisions and promote operational efficiency by effectively prioritizing healthcare resources.¹⁻⁵ However, it also raises the potential for worsening health inequities and creating harm in patient care. In 2022 and 2023, government officials, including the White House, Health and Human Services (HHS) Office of Civil Rights, Office of the National Coordinator (ONC), and Office of the Attorney General in California, took regulatory actions to protect against such harms. Numerous academic papers have also discussed potential causes of bias in AI products and proposed mitigation strategies to address them.

Such efforts enhance awareness of the potential risk of AI products on health inequities. However, they still leave room for interpretation regarding how healthcare delivery

organizations can implement regulatory principles and mitigation strategies internally to mitigate the risk of AI products exacerbating health inequities. Health system leaders still need more practical guidance when adopting a new and uncertain AI product.

To close this gap, Health AI Partnership (HAIP) built an equity assessment framework grounded in real-world examples and contexts so health system leaders can adopt and use it. Our goal was to build a framework that answers the following question: “Our health system is considering adopting a new solution that uses AI. How do we assess the potential future impact on health inequities?” We recruited multi-disciplinary individuals and groups from health systems and ecosystem partners—including three innovation teams with experience in developing and implementing AI solutions in healthcare, six framework developers with

Participant		Role	Responsibilities
C	Case study presenters	3 innovation teams that develop and implement AI solutions in healthcare delivery organizations	Curated a case study, presented it at the workshop, and tested out the framework
W	Workshop participants	77 stakeholders from 10 healthcare delivery organizations and 4 ecosystem partners with clinical, technical, operational, regulatory, and AI ethics expertise	Contributed to developing the procedures of the framework
F	Framework developers	A clinician, a community representative, a computer scientist, a legal and regulatory expert, a project manager, and a sociotechnical scholar	Created a scaffolding of the framework and contributed to developing its procedures
H	HAIP leaders	A clinical data scientist, a community organizer, 2 computer scientists, 3 lawyers, and a program director	Evaluated the framework and provided feedback
D	Design researchers	A clinical data scientist, a project manager, and a qualitative research scientist	Facilitated the co-design process by collecting, iterating, and synthesizing data from all other participants

Figure 1. Roles and responsibilities of all participants.

NEXT PAGE →



various expertise, and 77 healthcare professionals—and engaged them at different phases of the framework development process (Figure 1). We developed the framework through two rounds of divergent and convergent design thinking (Figure 2). During the discover phase, we explored the problem extensively through a case study-based workshop using two real-world use cases. We discussed the question with healthcare professionals and documented their responses to gain a comprehensive understanding of the challenges related to health equity and potential mitigation strategies. During the define phase, we synthesized the collected responses. During the develop phase, we developed and tested framework prototypes. During the deliver phase, we refined the final prototype and prepared for dissemination of the framework.

The four phases of design research resulted in the creation of the Health Equity Across the AI Lifecycle (HEAAL) framework, with detailed step-by-step procedures organized around HAIP’s eight key decision points (Figure 3).⁶ The eight key decision points encompass decisions made in technical, strategic, operational, and clinical domains during the adoption of AI in healthcare. Using this framework in developing HEAAL helped evaluate both the technical aspects of the AI product and its work environment, including end users, clinical workflow, and business strategies concerning the health equity assessment. Some procedures assess data quality, population representativeness, and model performance across patient subgroups. Others focus on identifying disadvantaged subgroups, developing fairness metrics and equity objectives, and providing education and training to reduce user bias among end users. Across the entire AI lifecycle, HEAAL describes

what procedures are conducted under each key decision point, what health equity principles are assessed, who is engaged, and what data sources are examined.

The HEAAL framework is an initial attempt to address the potential risk of AI exacerbating existing health inequities through a comprehensive list of step-by-step procedures generated by the community of healthcare professionals. The HEAAL framework highlights three complex challenges that will require significant attention and investment by diverse stakeholders, specifically:

1. HEAAL is context-specific. An AI solution evaluated in one setting must still be reassessed when considered for use in a different context.
2. Successful implementation of HEAAL requires significant expertise, technology infrastructure, and personnel. Despite the publicly accessible framework and the consensus among healthcare leaders to mitigate bias in AI, healthcare delivery organizations will not be able to apply the entire framework to every AI solution without significant support. Therefore, healthcare delivery organizations that adopt HEAAL should be financially incentivized.
3. There is concern that HEAAL can serve as a rubber stamp for healthcare organizations, outwardly projecting a commitment to equity while minimizing changes to organizational practices. Independent registries may address these complex challenges by promoting transparency and traceability throughout HEAAL procedures to hold healthcare organizations accountable.

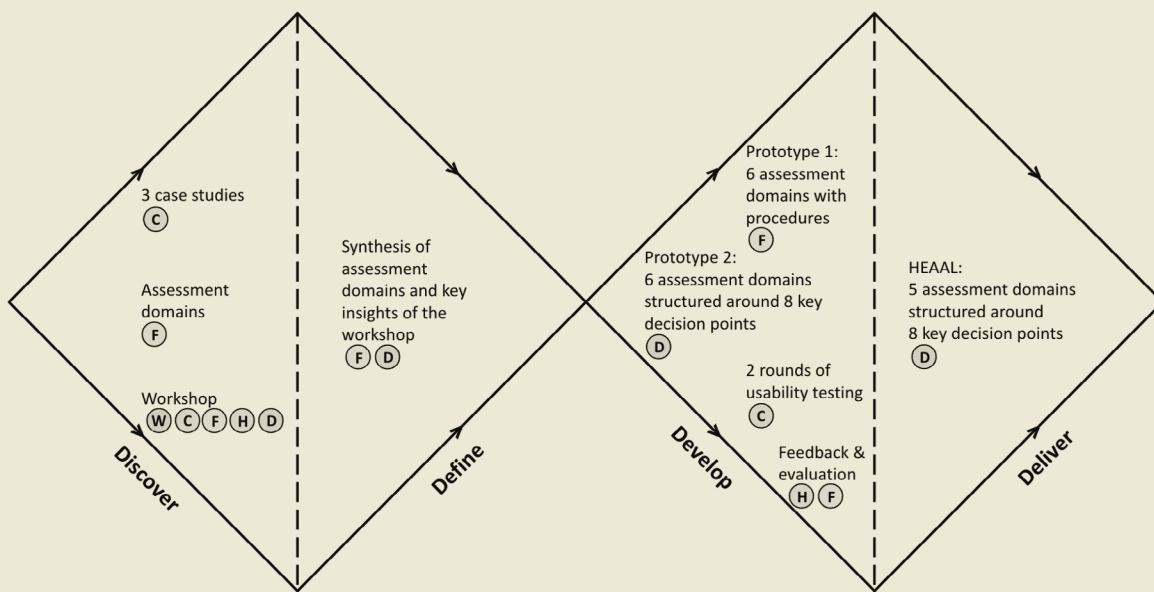


Figure 2. The process of the framework development and visual representation of participant engagement.

While HAIP has made significant strides in creating the framework, we recognize that our efforts to mitigate the potential risks of AI products on health inequities must continue. As best practices to address health inequities evolve, we plan to keep iterating on the HEAAL framework. We aim to engage more healthcare delivery organizations to adopt and test our framework, gather their feedback, and persist in addressing these organizations’ challenges. The HEAAL framework exemplifies Duke Institute for Health Innovation’s commitment to promoting healthcare equity by providing practical guidance for the responsible design, implementation, and monitoring of AI solutions. To learn more about the HEAAL, please visit healthpartnership.org or read its publication.⁷

References

1. Government Accountability Office. Artificial Intelligence in Health Care: Benefits and Challenges of Machine Learning Technologies for Medical Diagnostics With Content from the National Academy of Medicine. United States Government Accountability Office [Internet]. 2022 [Cited 2023 October 11]. Available from: <https://www.gao.gov/assets/gao-22-104629.pdf>
2. Raji ID, Smart A, White RN, Mitchell M, Gebru T, Hutchinson B, et al. Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. Proceedings of the 2020 conference on fairness, accountability, and transparency. 2020 Jan 27; 33–44.
3. Wu E, Wu K, Daneshjou R, Ouyang D, Ho DE, Zou J. How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. Nat Med. 2021;27(4):582–584 pmid:33820998
4. Agarwal R, Bjarnadottir M, Rhue L, Dugas M, Crowley K, Clark J, et al. Addressing algorithmic bias and the perpetuation of health inequities: an AI bias aware framework. Health Policy and Technology. 2023 Mar 1;12(1):100702.
5. The White House. Blueprint for an AI Bill of Rights. The White House [Internet]. 2022 [Cited 2023 October 11]. Available from: <https://www.whitehouse.gov/ostp/ai-bill-of-rights/>
6. Health and Human Services Office for Civil Rights. HHS Announces Proposed Rule to Strengthen Nondiscrimination in Health Care. Department of Health and Human Services [Internet]. 2022 [Cited 2023 October 11]. Available from: <https://www.hhs.gov/about/news/2022/07/25/hhs-announces-proposed-rule-to-strengthen-nondiscrimination-in-health-care.html>
7. Office of the National Coordinator for Health Information Technology, Department of Health and Human Services. Health Data, Technology, and Interoperability: Certification Program Updates, Algorithm Transparency, and Information Sharing. Federal Register [Internet]. 2023 [Cited 2023 October 11]. Available from: <https://www.federalregister.gov/documents/2023/04/18/2023-07229/health-data-technology-and-interoperability-certification-program-updates-algorithm-transparency-and>
8. Office of the Attorney General. Attorney General Bonta Launches Inquiry into Racial and Ethnic Bias in Healthcare Algorithms. State of California Department of Justice [Internet]. 2022 [Cited 2023 October 11]. Available from: <https://oag.ca.gov/news/press-releases/attorney-general-bonta-launches-inquiry-racial-and-ethnic-bias-healthcare>
9. Celi LA, Cellini J, Charpignon ML, Dee EC, Dernoncourt F, Eber R, et al. Sources of bias in artificial intelligence that perpetuate healthcare disparities: a global review. PLOS Digital Health. 2022 Mar 31;1(3):e0000022. pmid:36812532

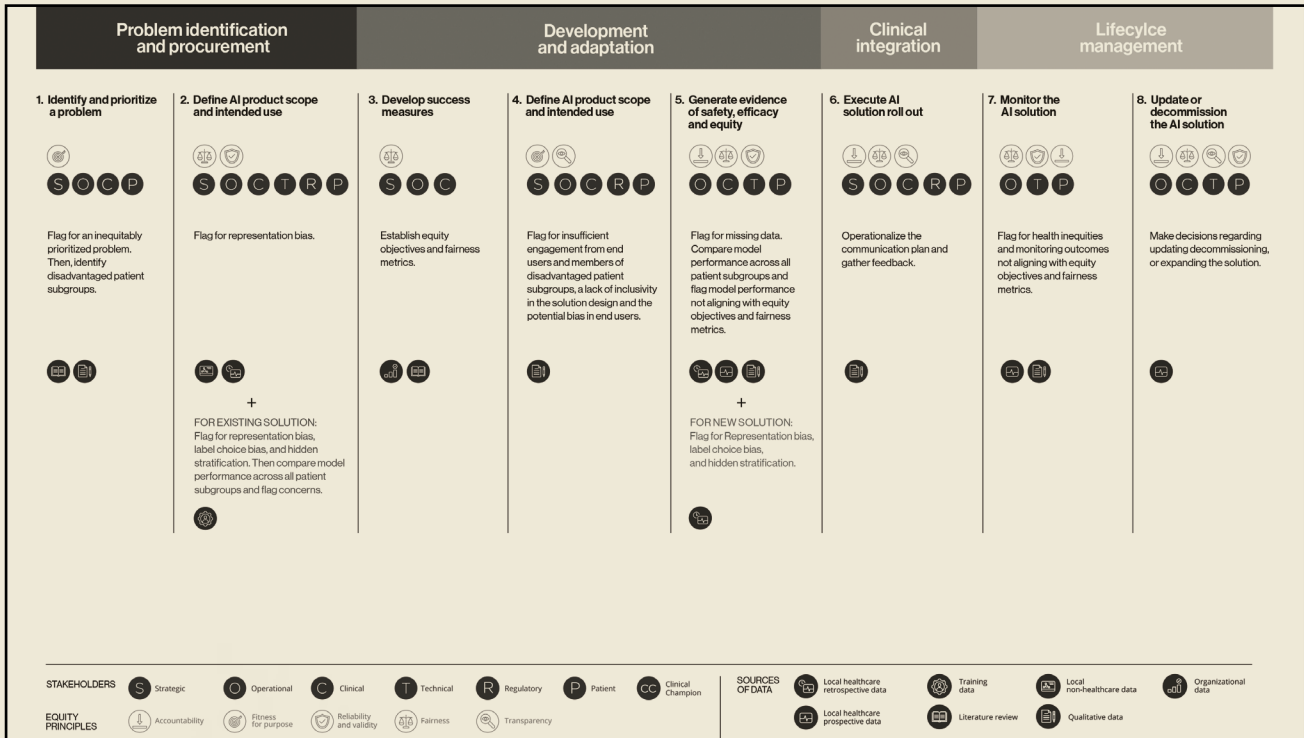


Figure 3. Overview of the Health Equity Across the AI Lifecycle (HEAAL).



AI GOVERNANCE & ETHICS

Creating an AI Ethics Training Program for Physicians in Collaboration with the American Medical Association

Jee Young Kim, PhD

As the use of artificial intelligence (AI) in clinical care has become more prevalent, physicians encounter novel concerns and challenges related to this emerging technology. Some of these concerns and challenges pertain to ethics in patient care and raise various questions. For instance, should patients be informed when healthcare professionals use AI in their care? Should they have a say in data ownership as data producers? Should they be asked to provide informed consent before individuals or organizations use their data to develop and commercialize future AI products?

To empower physicians in navigating these complex issues, in February 2024, we launched a 12-month-long collaboration with the American Medical Association (AMA) with the aim of developing an AI ethics training program and associated educational materials.

The program aims to help physicians gain a comprehensive understanding of the ethical challenges, dilemmas, and contexts associated with the use of AI within healthcare settings. It also seeks to develop physicians' abilities to effectively communicate the utilization and implications of AI technologies to patients, thereby promoting transparency and trust in healthcare delivery. Additionally, the program aims to enhance physicians' critical thinking skills, enabling them to analyze ethical considerations and apply their knowledge to real-world clinical scenarios, thereby ensuring ethically informed decision-making in practice.

To create the ethics training program, we developed a project plan composed of four phases: (1) discover, (2) design, (3) test, and (4) refine. In the "discover" phase, we aim to conduct interviews with physicians to identify key ethical questions and challenges they face regarding the use of AI in clinical care. In the "design" phase, we aim to curate case studies with real-world use cases based on the interview findings and to develop a training program along with educational materials derived from these case studies. We are employing a case study learning approach to promote active engagement and deep learning among physicians as well as to provide them with practical knowledge applicable in clinical settings. The "test" phase involves testing and evaluating the training program and educational materials with physicians to ensure they are comprehensive, educational, and

practical. In the last "refine" phase, the program will be refined based on feedback from the test phase and prepared for a broader dissemination to the AMA network as continuing medical education (CME). We have recruited a team of AI ethics experts to serve as advisors. They would provide further guidance on the development of the program.

We are employing a case study learning approach to promote active engagement and deep learning among physicians as well as to provide them with practical knowledge applicable in clinical settings.

The project is currently in the "design" phase after completing the "discover" phase. We have completed interviews with 14 physicians in the U.S. from diverse backgrounds, specializing in various clinical areas with different levels of experience. They serve inpatients and outpatients and are affiliated with either academic medical institutions or regional clinics in urban and rural areas. We have completed the analysis of these interviews and identified the most prevalent themes and subthemes in their responses. The findings highlighted health equity, transparency, and accountability as the most pressing ethical issues for physicians, and these principles will guide the focus of our training program and educational materials.

We are excited to be working on the next phases of this project and look forward to the potential impact that its outcomes will have on physicians, their knowledge, and the emerging landscape of AI technology in healthcare.

AI GOVERNANCE & ETHICS

Establishing an AI Governance System to Facilitate the Safe and Responsible Adoption of AI Software in a Canadian Context

Jee Young Kim, PhD

Trillium Health Partners (THP) is a hospital system that serves Mississauga and western Toronto in the Canadian province of Ontario. It is the largest community-based hospital affiliated with the University of Toronto. Since their electronic medical record (EMR) system went live in October 2020, THP has become increasingly interested in adopting artificial intelligence (AI) solutions in clinical practice. To gauge this interest, the AI Deployment and Evaluation (AIDE) Lab was created in 2021 to validate AI models for the safe, efficient, and equitable use of AI in healthcare. A few years later, senior executives and the AIDE Lab began planning to implement these AI solutions at THP. However, they soon recognized the need for organizational AI governance systems and federal regulations. They subsequently sought Duke Institute for Health Innovation (DIHI) assistance to identify best practices in governance to facilitate the safe and responsible adoption of AI in healthcare. At the end of this project, they wanted to draft organizational policies around AI governance using the best practices.

To identify governance best practices at THP, we conducted a series of activities with senior executives as well as diverse stakeholders at THP. This collaborative approach, which included interviews, journey mapping, surveys, and workshops (Table on the next [page](#)), ensured that all voices were heard and all perspectives were considered.

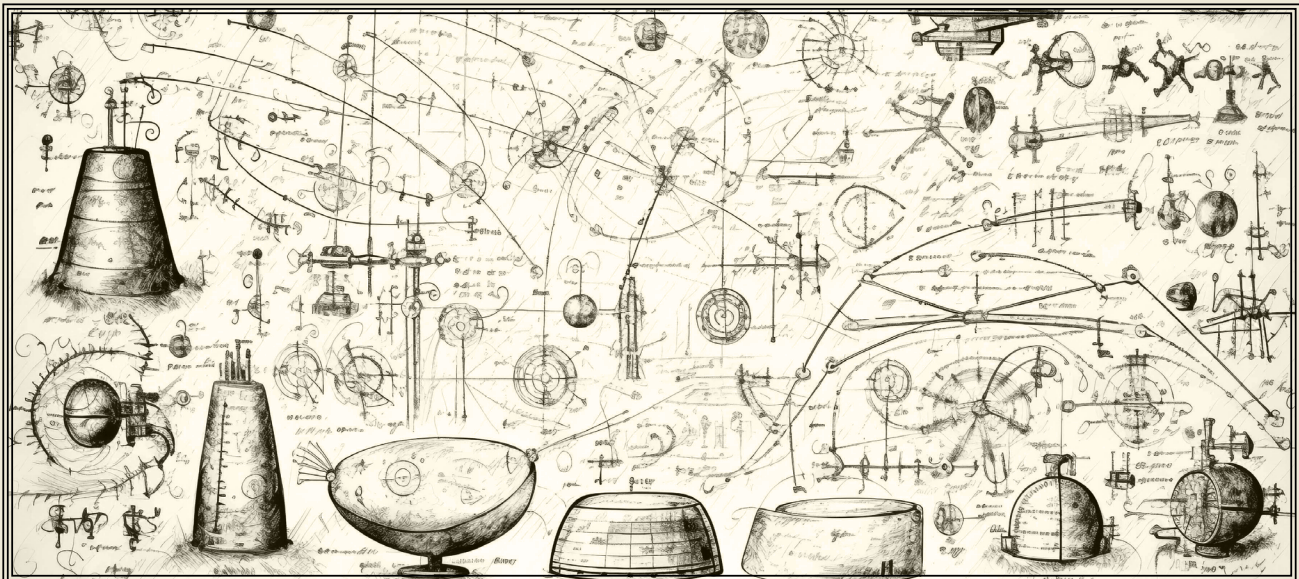
Upon the completion of these activities, we were able to develop a set of recommendations for implementing AI governance at THP. The recommendations spanned over the entire lifecycle of AI and presented best practices of AI governance in four different domains: People, Process, Technology, and Operation. This framework was adapted from the People, Process, Technology (PPT) framework that is widely used for building systems that effectively balance and coordinate critical factors in organizations.

As a result of this project, senior executives at THP enhanced their knowledge of AI governance and, more importantly, successfully formulated policies around AI governance aligned with our recommendations. They drafted and successfully gained approval for three main governance documents. The policy is now in effect as of October 2024. Establishing AI governance at THP marks a pivotal step towards ensuring the safe and responsible integration of AI technologies in healthcare. Through this effort, THP is poised to lead in responsible AI implementation. Our collaboration with THP not only facilitated the development of this critical governance framework but also underscored DIHI's commitment to advancing safe, effective, and equitable adoption of AI in healthcare, providing reassurance and confidence to all stakeholders.



NEXT PAGE →

ACTIVITY	DESCRIPTION	PARTICIPANTS	FINDINGS
Interviews	One-on-one interviews were conducted over Zoom for 30 minutes to an hour and later transcribed and coded for analysis.	12 stakeholders at Trillium Health Partners (THP) with expertise in operational, technical, and clinical domains	<ul style="list-style-type: none"> Organizational structure and key stakeholders involved in governing advanced technologies The current state of AI adoption at THP Strength and weakness in the existing governance process for advanced technologies Need for a centralized and standardized governance process The presence of robust resources and capabilities within THP
Journey mapping	A process map that outlines future AI governance was drafted through iterative processes.	A project team leader at Trillium Health Partners (THP)	<ul style="list-style-type: none"> Aspirational processes and key stakeholders involved in future AI governance Gaps between the existing governance system for advanced technologies and the future governance system for AI
Surveys	Online surveys were distributed to participants before and after the workshops	6 senior executives who serve on the existing governance committee for advanced technologies	<ul style="list-style-type: none"> Pre-workshop surveys: Initial set of recommendations across People, Process, Technology, and Operation domains Post-workshop surveys: Satisfaction on the workshops and the final set of recommendations
Workshops	Four workshops were conducted to refine the initial set of recommendations across People, Process, Technology, and Operation domains.	6 senior executives who serve on the existing governance committee for advanced technologies	<ul style="list-style-type: none"> Final set of recommendations across the People, Process, Technology, and Operation domains that are more tailored to THP



AI GOVERNANCE & ETHICS

Building a one-of-a-kind AI product evaluation rubric tool

Vaibhav Jha



Artificial intelligence (AI) products play an increasingly significant role in the healthcare industry as they are gradually adopted in clinical settings to improve patient outcomes and operational efficiency. As investors and purchasers search for promising clinical innovations, there has been a growing need for a systematic evaluation mechanism of AI products. Organizations are interested in various factors when choosing which products to buy, from clinical utility to a promising return on investment. With numerous vendors coming forward with their products, healthcare organizations have had to make investment decisions based on haphazard processes.

The Duke Institute for Health Innovation (DIHI) identified this problem through lengthy discussions with a biotech company focused on discovering and distributing new AI products. Interested in early-stage products, the organization aims to find value in products pre-commercialization. Investors and purchasers have found it difficult to choose between the numerous AI vendors that seek investment opportunities, resulting in a need to develop a standardized evaluation process for a diverse array of AI products. DIHI recognizes industry demands for an evaluation rubric tool that is specific enough to target the interests of a company and general enough to be applied across varied forms of AI products. We are developing a systematic tool to help organizations evaluate AI algorithms in the market.

Rubric construction began with numerous meetings to set the scope of the evaluation rubric. Here, representatives from both sides brought expertise in various aspects of AI algorithms, including but not limited to data science, business, and health equity perspectives. DIHI led a market scan through an intensive literature review to understand relevant data input considerations. Simultaneously, we set aside various positive and negative control cases at the onset of this project for later testing.

The team began developing the rubric with a two-day in-person workshop at DIHI. Both parties weighed different

framework options for the evaluation tool, ultimately agreeing on a two-part rubric. The first part of the rubric, called Stage Gate 1 (SG1) screening system, is a rapid quantitative assessment designed to filter out AI products that fail to meet the minimum requirements. SG1 evaluates products on ten mandatory criteria and assigns a weight to each criterion that reflects how much it contributes towards informing quality and commercial potential of the AI product. The outcome of SG1 is a list of products that meet the minimum threshold advance in the due diligence process.

The second part of the rubric, called Stage Gate 2 (SG2) screening system, is a comprehensive quantitative and qualitative assessment designed to prioritize AI products for commercial viability. SG2 evaluates products across seven domains: **(1) Model Characteristics and Clinical Integration, (2) Administrative and IT Integration, (3) Regulatory Compliance, (4) Intellectual Property, (5) Financial Assessment, (6) Ethical Development and Health Equity, and (7) Market Reputation and Impact.** A team of experts will review the AI product across different criteria under the seven domains and assign a weighted total score for product evaluation, level of potential risks, overall summary of product strengths, weaknesses, and recommendations for improvement. The outcome of SG2 provides a detailed and actionable assessment to inform commercialization decisions.

This AI product evaluation tool establishes a benchmark for assessing AI products across the healthcare industry. It is a comprehensive evaluation process that integrates quantitative and qualitative assessments for thorough, standardized AI product evaluation. It will also aid in informed strategic decision-making for identifying and prioritizing AI products for investment and partnership opportunities. It improves transparency in the industry by promoting clear criteria and processes, fostering trust and setting higher standards for product assessment. Lastly, it improves quality and reliability of AI products developed in healthcare.

Large Language Models

Leveraging a Large Language Model to deploy a tool for accurate care engagement classification among patients with HIV (PWH) within the Duke University Health System

ALIFIA HASAN, B PHARM, MBA; NASEEM ALAVIAN MD, MPH

Background

Southern states lead the nation in new human immunodeficiency virus (HIV) infections, accounting for 52% of new HIV infections.¹ Further, persons with HIV (PWH) have worse outcomes in the South as compared to other regions. Optimal retention in HIV Care and maintenance of life-saving antiretroviral medication is critical for reducing individual morbidity and mortality and reducing HIV transmission.^{2,3} The Centers for Disease Control and Prevention (CDC) estimates that 62% of HIV transmission occurs from PWH who are aware of their HIV status and are either not virally suppressed or not optimally engaged in HIV care.⁴

Despite the importance of regular HIV care engagement, national data suggests that at least 50% of PWH are not optimally engaged in HIV care.⁵ Our preliminary investigations of a 2932-person cohort of PWH who have received care in the Duke University Health System (DUHS) Infectious Disease HIV Clinic 2016-2021 suggested a large number of our patients (59%) had a period of HIV care disengagement (Figure 1).

A chart review of 110 patients, however, revealed that nearly 40% of the patients who appeared to be out of care based on health-system data had actually transferred care to another facility, often called a “silent transfer.” Misclassification of out-of-care patients and difficulty identifying “silent transfer” have been cited as creating major barriers in novel and effective re-engagement strategies.⁶⁻⁹

The data silos between significant healthcare organizations in the same jurisdiction represent a major structural barrier to fully harnessing the power of data science to address the gaps in the Southern HIV care continuum. Our preliminary work highlighted our inability to accurately identify patients with HIV in our region who are truly out of care, thus stunting the implementation of an effective HIV care re-engagement strategy for the health system.

We initiated a year-long project in January 2024 with the following aims:
AIM 1. Create an algorithm using structured electronic health record data to accurately identify PWH at DUHS who appear to have disengaged from care.

AIM 2. Leverage a Large Language Model (LLM) to identify the transfer of care documented in clinical notes to improve the precision of classifying patients’ care engagement status.

AIM 3. Deploy an HIV care engagement tool for real-time, accurate tracking of care engagement.

Approach and Current Progress

AIM 1. Create an algorithm using structured electronic health record data to accurately identify PWH at DUHS who seemed to have disengaged from care (defined as having no clinic visit for HIV care in ≥14 months). A revised algorithm was built within the DIHI data pipeline to correctly identify care engagement status among PWH ≥ 18 years old who have received care at the Duke Infectious Disease HIV clinic from January 2016 to April 2024. Care disengagement was defined as a period of ≥14 months with a clinical visit for HIV care in our health system.

Transfer to new HIV provider
Transfer to new provider
Transfer to new clinic
Transfer to different state/city/county
Transfer to a new health care system
Moved or moving to new state/city/county
PCP Providers HIV care
Other

Table 1. Transfer indication type identified by LLM found in clinical notes

Among 3431 PWH, we found that 52% appeared to have disengaged from HIV care at some point. Twenty-two percent of these individuals ultimately returned to care. However, 30% (1026 patients) appeared currently out of HIV care (Figure 2). A chart review of 200 patients who appeared out of care revealed that 53% transferred HIV care out of our health system, the algorithm misclassified 11%, and only 36% appeared truly out of care. Among those who transferred care, evidence of transfer was present within clinical notes two-thirds of the time.

AIM 2. Leverage a Large Language Model (LLM) to identify the transfer of care documented in clinical notes to improve the precision of classifying patients' care engagement status. Large Language Models (LLM), such as OpenAI and ChatGPT, are AI systems that leverage deep learning and incorporate prompts, such as reinforcement learning, to analyze and interpret extensive amounts of text-based data, like what is found in clinical notes. We hypothesized we could accurately identify "silent transfer" by incorporating text from clinical notes into our HIV Care Engagement Algorithm using an LLM.

We used selected clinical notes from patients who appeared to be disengaged to engineer a prompt to appropriately classify patients who appeared out of care as either transferred HIV care outside DUHS or with no evidence of care transfer (suggesting truly out-of-care status). The LLM further categorized notes to identify indications of transfer (Table 1). Preliminary investigation suggests we can identify "silent transfer" with high accuracy.

The combination of structured electronic health record data and unstructured data (clinical notes) will improve our classification of care engagement, allowing our HIV re-engagement team to focus on the highest priority patients who are truly disengaged from care. We are in the process of prompt iteration, which includes refining the prompt

and testing it on a more significant proportion of our retrospective cohort. Outcomes from this aim include model metrics, including area under the receiver operating characteristic (AUROC), sensitivity, and specificity of this model in accurately classifying care engagement status.

Further, through the application of this tool on a retrospective cohort of 3426 PWH who have received HIV care at Duke Infectious Disease Clinic, we can identify patterns of care engagement to determine which factors (clinical, demographic, social determinants of health, healthcare utilization) are associated with care retention, disengagement, and voluntary re-engagement after a period out-of-care (OOC).

AIM 3. Deploy an HIV Care Engagement tool for real-time, accurate tracking of HIV care engagement. We envision the application of this tool in a clinical setting with outcomes of care engagement presented in a clinical dashboard. The HIV re-engagement team, which consists of two physicians, two clinical social workers, and one HIV Nurse Coordinator, can receive care engagement reports regularly. Reports will highlight patients who have transferred care to another facility and those who have fallen out of care to conduct an outreach and target interventions to bring these patients back into care.

The product of this aim will inform the approach for the HIV Re-engagement Team and strengthen the Health System HIV re-engagement strategy, which includes emergency department (ED)-based care re-engagement intervention. Further, this work will enhance the implementation of a retention and re-engagement care model in collaboration with the Durham County Department of Public Health, providing unique opportunities for both entities (Figure 3).

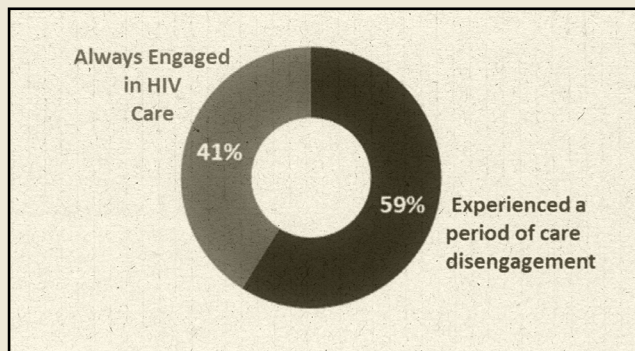


Figure 1. Care Engagement at Duke Infectious Disease Clinic 2016 - 2021 (n = 2963)

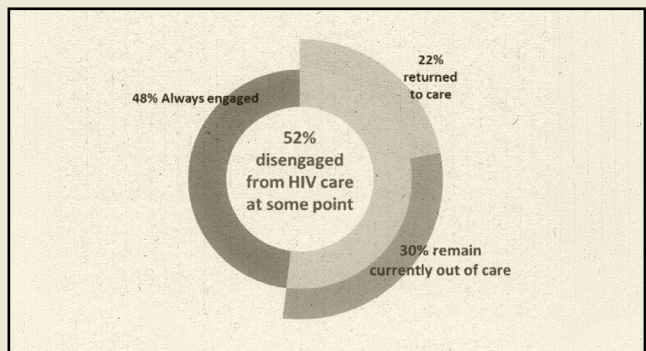


Figure 2. Refined Algorithm Care Engagement at Duke Infectious Disease Clinic 2016 - 4/2024 (n = 3431)

NEXT PAGE →



Impact

This project represents a meaningful step towards addressing high care disengagement rates among our growing HIV clinic population. The Duke Infectious Disease clinic sees over 2000 PWH annually, with 10% growth in the HIV clinic population yearly. Keeping PWH engaged in care improves clinical outcomes and can reduce costly ED visits or potential hospital admissions, particularly for underinsured or uninsured patients. If we can improve care engagement, we may also impact HIV transmission in our region, where Wake/Durham counties have seen a 16% increase in HIV transmission over the past five years.

Our use of machine learning/artificial intelligence to incorporate information from clinical notes to determine the disposition of patients who have fallen out of care has not yet been explored. It would overcome a well-described resource-intensive barrier to identifying out-of-care patients. The outcomes of this proposal hold the potential to further the innovative use of artificial intelligence addressing the HIV continuum of care.

References

1. HIV Surveillance Report: Diagnoses, Deaths, and Prevalence of HIV in the United States and 6 Territories and Freely Associated States, 2022. October 16, 2024. Accessed October 16, 2024. <https://stacks.cdc.gov/view/cdc/156509>
2. Giordano TP, Gifford AL, White AC, Jr., et al. Retention in care: a challenge to survival with HIV infection. *Clin Infect Dis.* Jun 1 2007;44(11):1493-9. doi:10.1086/516778
3. Mugavero MJ, Lin HY, Willig JH, et al. Missed visits and mortality among patients establishing initial outpatient HIV treatment. *Clin Infect Dis.* Jan 15 2009;48(2):248-56. doi:10.1086/595705
4. Li Z, Purcell DW, Sansom SL, Hayes D, Hall HI. Vital Signs: HIV Transmission Along the Continuum of Care - United States, 2016. *MMWR Morb Mortal Wkly Rep.* Mar 22 2019;68(11):267-272. doi:10.15585/mmwr.mm6811e1
5. US Department of Health and Human Services. Ending the HIV Epidemic: A Plan for America. . Accessed April 10, 2023. [https:// www.hrsa.gov/ending-hiv-epidemic](https://www.hrsa.gov/ending-hiv-epidemic)
6. Sack DE, Brantley M, Ratliff M, et al. Misclassification of Loss to Care Among Persons With Human Immunodeficiency Virus: Improved Capture of Silent Transfers Through Surveillance Linkage Using Statewide Mandatorily Reported Laboratory Measures. *Clin Infect Dis.* Jan 25 2024;78(1):118-121. doi:10.1093/cid/ciad461
7. Ridgway JP SJ, Almirol E, Millington M, Harding E, Pitrak D. Electronic data sharing between public health department and clinical providers improves accuracy of HIV retention data. *Open Forum Infect Dis.* Oct 4 2017;4(Suppl 1)(S421-2)doi:10.1093/ofid/ofx163.1059
8. Ridgway JP, Almirol E, Schmitt J, Wesley-Madgett L, Pitrak D. A Clinical Informatics Approach to Re-engagement in HIV Care in the Emergency Department. *J Public Health Manag Pract.* May/June 2019;25(3):270-273. doi:10.1097/PHH.0000000000000844
9. Gill MJ, Ody M, Lynch T, Jessiman-Perreault L, Krentz HB. Maintaining the continuity of HIV-care records for patients transferring care between centers: challenges, workloads, needs and risks. *AIDS Care.* Aug 2016;28(8):1073-8. doi:10.1080/09540121.2016.1139042

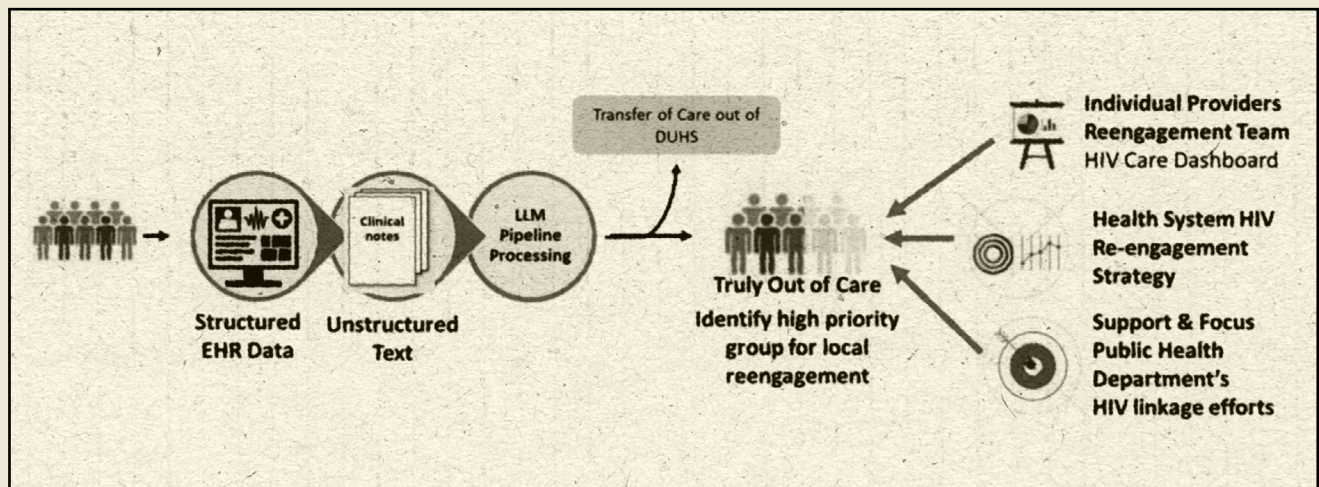


Figure 3. Application of Large Language Model to create tool to monitor HIV Care Engagement/Disengagement at DUHS

LARGE LANGUAGE MODELS

Inbasket Messaging with Large Language Models

KARTIK PEJAVARA, BS

Problem

Although adopting electronic health record (EHR) platforms has improved many aspects of patient care, it has also increased the burden on healthcare professionals. One of the largest contributors to this burden is patient portal messages. At Duke Health, providers spend considerable time in Epic addressing MyChart messages. To reduce the volume of messages that providers receive, Duke Endocrinology formed a triage team of nurses dedicated to sorting and addressing all messages sent by patients to the clinics. When a patient sends a message to anyone at Duke Endocrinology, the triage team first intercepts the message and a nurse will either respond to the message themselves or forward it to the appropriate person. This is beneficial for providers because messages that do not require medical input do not need their immediate response. For example, messages regarding prior authorization should go to the endocrine prior authorization office, but most patients do not know this and send them to their providers instead.

Similarly, patient messages about scheduling, financial assistance, and medical device usage should go to their designated offices. However, they are usually sent directly to physicians or advanced practice providers (APPs: predominantly nurse practitioners and physician assistants). APPs must then take the time to read these messages and forward them to their respective offices. With the triage team in place, messages are pre-filtered so that providers only receive messages regarding patient health and treatment plans.

Implementation of the triage team has significantly reduced the messaging burden on many healthcare professionals. However, the role of the triage nurse has evolved. It is now their responsibility to address many messages themselves, especially those concerning medication refill requests, lab orders, and specific symptoms that they must collect more information about. In combination with the hundreds of messages received daily, the expanding responsibilities of the triage team have led to the buildup of a message backlog. Not only do the triage nurses now face the same message burden originally placed on APPs, but patients must now wait days for their messages to be addressed.

A way to automatically triage low-importance messages to the correct team would significantly reduce the message load on the triage team and allow them to focus on addressing the messages relevant to them. Healthcare professional messaging needs a revolution similar to how telephone switchboards, invented in 1877-1878 and popularized during Duke University's first decade, improved communication by automating the connection between callers, ensuring they could reach the right destination quickly and efficiently. (Of note, Switchboard, MD has the same goal and branded themselves using this historical reference).

Our inbasket messaging project aims to use the latest in data science to streamline the flow of patient messages to healthcare professionals. A challenge we face is that we cannot simply train a model on previously classified messages. Duke Endocrine faces highly dynamic medical problems, which can result in entirely new message content that pre-trained classification models may have yet to see. In addition, workflows are constantly changing, limiting models trained for this task due to their inability to rapidly adjust to adding new categories and uncommon "edge-case" messages.



NEXT PAGE →

Solution

With their remarkable ability to comprehend natural language, Large Language Models (LLMs), such as ChatGPT, hold immense promise in improving healthcare operations. LLMs not only adapt their output to any given prompt, but they can also reason about questions beyond the scope of what we directly present to them. Research has demonstrated that prompting the model to reason about the question typically results in a more accurate answer. A recent study by Microsoft showcased how a general LLM, like GPT-4, could be used to solve domain-specific problems using a method called Chain-of-Thought (CoT) (Nori et al.). In this method, domain experts identify example questions, provide answers, and develop the reasoning for why that is the answer. Data scientists then encode this reasoning in an internal database accessible to the LLM. When presented with a new question through a prompt, the model retrieves some helpful question-answer-CoT examples from the internal database and inserts them into the prompt, providing the LLM with relevant examples and explanations before it provides its answer. This method significantly increases the model accuracy compared to simply asking the LLM to answer the question. This project report addresses the automatic triage problem by employing GPT-4 to classify patient portal messages using similar prompting techniques, offering a hopeful outlook for the future of healthcare operations.

The Process

Our collaborative project involves healthcare professionals' expertise and advanced technology capabilities. We started by extracting MyChart messages across Duke Endocrinology clinics. Working closely with the existing nurse triage teams at Duke Endocrinology, we developed a list of 11 message categories and their descriptions. This list served as the basis for classifying 81 messages and developing CoT for their classifications. These messages form our set of example (gold standard) messages. Using the OpenAI Embedding API, we embedded

these messages in an internal database. When presented with a new message, the model uses a similarity calculation to retrieve several gold standard messages similar to the new one. These annotations effectively encode the decision-making process.

The final solution involves prompting OpenAI's GPT-4 model when new messages are received. The model uses the message to retrieve several similar gold standard messages from its internal database. The new message and the gold standard messages, along with their corresponding classifications and CoT, are inserted into a prompt that contains the 11 category descriptions and asks the model to predict where the new message should go.

Testing

To ensure the robustness of our modeling framework, we conducted comprehensive testing. We compared a zero-shot approach (where the data scientists did not give the LLM examples but only descriptions of the categories) to our CoT approach. We also tested a setting in which the model was asked to return a generated reasoning for why the message should be returned. Lastly, we evaluated the monetary cost of running these models to illustrate the cost-to-performance trade-off of such an approach. To estimate relative cost, we used the current cost of the GPT-4 API, which assigns double the price to output tokens as input tokens. This thorough testing process instills confidence in the effectiveness of our proposed solution.

Results

The results show that the model's zero-shot performance results in roughly 16% of messages being misclassified. However, accuracy improved when our team added example messages to the prompt. Ultimately, the model showed the highest accuracy (88.9%) when asked to generate reasoning for its classification.

MODEL CONFIGURATION	ACCURACY (\pm SD)	MISCLASSIFICATION RATE	RELATIVE COST TO ZERO-SHOT
Zero-Shot + Answer Only	83.7 \pm 0.7%	47/289 (16.3%)	1x
Gold Standard + CoT + Answer Only Response	85.2% \pm 0.6%	43/289 (14.8%)	2.38x
KNN + CoT + Generated Response	88.9% \pm 0.5%	32/289 (11.1%)	2.60x

Table 1. Performances of the various models on classifying inbasket messages.

Anticipated Impact

This solution represents a flexible framework for automatically routing patient portal messages to an appropriate destination. DIHI’s collaboration with triage nurses and physicians has significantly advanced our ability to manage the ever-increasing communication load on healthcare providers. Not only will triage nurses be relieved of the impossible volume of messages, but they will also be able to address relevant messages more quickly; improving patient satisfaction.

What distinguishes this approach from other models is its flexibility. It allows for adjustments to the workflow by modifying the CoT examples to reflect the desired workflow changes. Moreover, we designed the system for continuous improvement. It can refine its classification accuracy over time by addressing messages that are difficult to classify or that healthcare professionals tell it that it has previously misclassified. This adaptability ensures the model remains relevant and effective even as healthcare practices and patient communication trends evolve.

Next Steps

At the time of this publication, we are discussing the best way to implement the model in Epic at Duke Endocrinology with the Duke Health Technology Solutions (DHTS) division.

Conclusion

The Inbasket Messaging project leverages the LLMs to offer a practical solution to a critical bottleneck in healthcare: patient message routing. By automating the classification and routing of these messages, the project aims to significantly reduce the workload on healthcare professionals, freeing up valuable time for patient care. This innovative approach, built on the analysis of extensive message data and collaboration with triage nurses, promises to enhance the efficiency of healthcare operations by improving response times and patient satisfaction. Furthermore, the project’s emphasis on standardization, data-driven decision-making, and measurable outcomes underscores its commitment to creating a tangible, scalable, and impactful solution in healthcare communication.

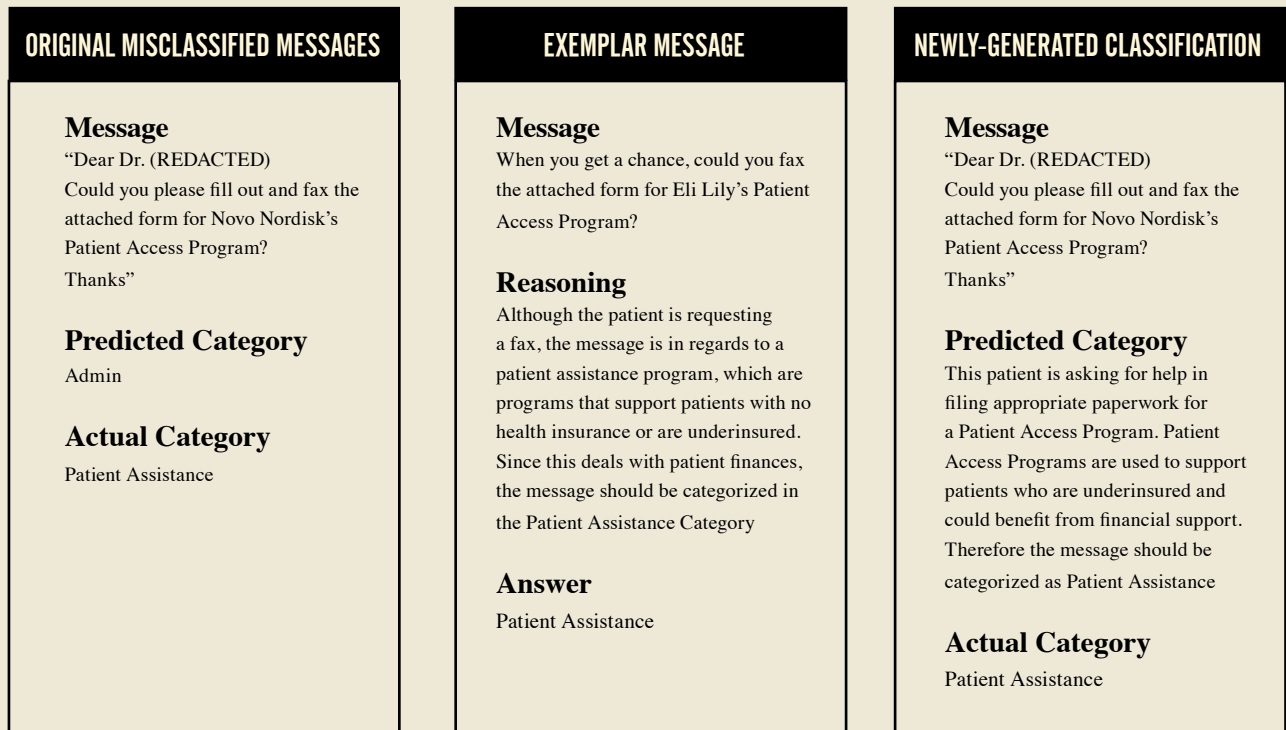


Figure 1. Improving model performance with “gold-standard” exemplar messages and reasoning.



LARGE LANGUAGE MODELS

Duke Health Policy Assistant – The RAG Journey

MATT GARDNER, BS

Problem

For Duke Health providers, timely access to over 4,000 policy documents in various formats, such as PDF and MS Word, poses a significant challenge. This not only makes efficient searching and comprehension difficult but also leads to care delays, compliance issues, staff disagreements, and ultimately, negative impacts on patient outcomes. The need for a solution is evident.

To address this, the Duke Institute for Health Information (DIHI) 2024 initiative titled “AI-Powered Knowledge Management System and Training Module” is focused on developing an artificial intelligence (AI)-powered assistant. This assistant aims to deliver accurate and timely responses to questions about Duke Health policies. In this article, we will share our progress in developing this AI assistant and outline the next steps in our journey.

The Intervention (Retrieval Augmented Generation)

Given the extensive and unstructured corpus of policy content at Duke Health, we have chosen to employ the AI technique known as Retrieval Augmented Generation (RAG). This technique, which utilizes mathematical text embeddings to compare a question posed to a large language model (LLM) with snippets of content from a reference corpus, is well-suited to address this challenge. It annotates questions asked of the LLM with semantically similar snippets, providing the knowledge needed for the LLM to generate an informative response.

The figure on the right illustrates the architecture DIHI has developed for the “DHE (Duke Health Enterprise) Policy Assistant”. It employs a standard RAG architecture tailored to meet our specific use case. (See Figure 1 – DHE Policy Assistant Architecture Diagram)

The Iterative Development Approach

While the technical aspects of configuring and deploying a RAG solution have become relatively straightforward, ensuring the accuracy of the generated responses remains a significant challenge. Next, we will outline our iterative approach aimed at continually improving accuracy, which is crucial for user trust and enterprise adoption.

To enhance the accuracy of our RAG solution, we are concentrating on the following key design components:

- Document Retrieval (Semantic Search)
- LLM token limits
- Prompt Engineering
- Continuous Expert Feedback

RAG Component: Document Retrieval (Semantic Search)

We leverage the HuggingFace Massive Text Embedding Benchmark (MTEB) Leaderboard, a comprehensive evaluation platform designed to assess and compare the performance of different text embedding models across a wide array of tasks. The MTEB gives us insight into the top-performing embedding models optimized for semantic search. To test search accuracy, we embedded the entire corpus of DHE policies. Policy content experts have created a bank of common questions along with the relevant policy documents they expect to be retrieved. We are translating these questions into automated test cases that can be run continuously. This approach ensures that iterative enhancements to document splitting strategies and the embedding model do not compromise search effectiveness and help identify areas where search parameters might need further optimization.

RAG Component: Context Awareness

Context Awareness refers to the system’s ability to comprehend, retain, and utilize the most pertinent content from Duke Health Policy documents in response to specific questions, such as: “How do I drape a patient for surgery?” We need to determine the optimal context scope to provide to the Large Language Model (LLM)—whether to use the most relevant chunks (e.g., sentences, paragraphs) from relevant documents or the entire documents.

This optimization challenge intensifies when RAG is enabled to support conversational interactions. To handle follow-up questions effectively, tracking prior questions and generated responses as part of the context is essential. For example, a follow-up question like “What about a craniotomy?” should be reformulated using the prior conversational history to something more context-aware, such as “How do I drape a patient for a craniotomy?”

The DIHI RAG framework keeps track of all prior context, including the document context retrieved via semantic search and the conversational history. This comprehensive tracking capability is instrumental in identifying areas where we need to optimize context awareness to improve conversational experiences.

RAG Component: Prompt Engineering

The challenge here is to craft prompts that guide the Large Language Model (LLM) to generate accurate, coherent, and contextually appropriate responses. There are two levels of prompting that we are currently focused on optimizing:

- Prompt for a new conversation topic: (system/LLM instructions + policy document context + user question)
- Prompt in the context of an existing conversation topic: (system/LLM instructions + policy document context + prior conversation history + user question).

We discovered several intricate considerations can significantly impact the quality, relevance, and coherence of the responses generated by the system. For example:

- Crafting clear and concise prompts helps avoid ambiguity and ensures the LLM can focus on the most pertinent details. For our use case, instructing the LLM to list the names of the policy documents referenced - and the Duke Entity they applied to - was critical.

- Structuring the prompt in a logical sequence that mirrors the conversation flow aids in maintaining coherence and relevance.
- Developing truncation strategies that retain the most significant information when the context and conversational history exceed LLM token limits.

The most important lesson we learned is that tracking prompt changes and iterations is important for understanding what modifications lead to improvements and replicating success. The DIHI RAG Framework records all prompts used so we can perform retrospective analysis on prompting strategies

RAG Component: Continuous Expert Feedback

The integration of continuous expert feedback is not just a feature, but a vital part of our ongoing enhancement of the RAG solution. We developed a Web application that allows end users to experiment with the “DHE Policy Assistant.” Conversational history with the AI assistant is captured in a database repository. This includes questions asked, document context referenced, responses generated, and the technical parameters used, such as the model used to embed documents/queries and the LLM used to generate responses. In addition, users can rate AI-generated responses and provide feedback. This continuous feedback loop ensures that our solution is not just effective but also user-centric and constantly improving, making our stakeholders involved and integral to the project.

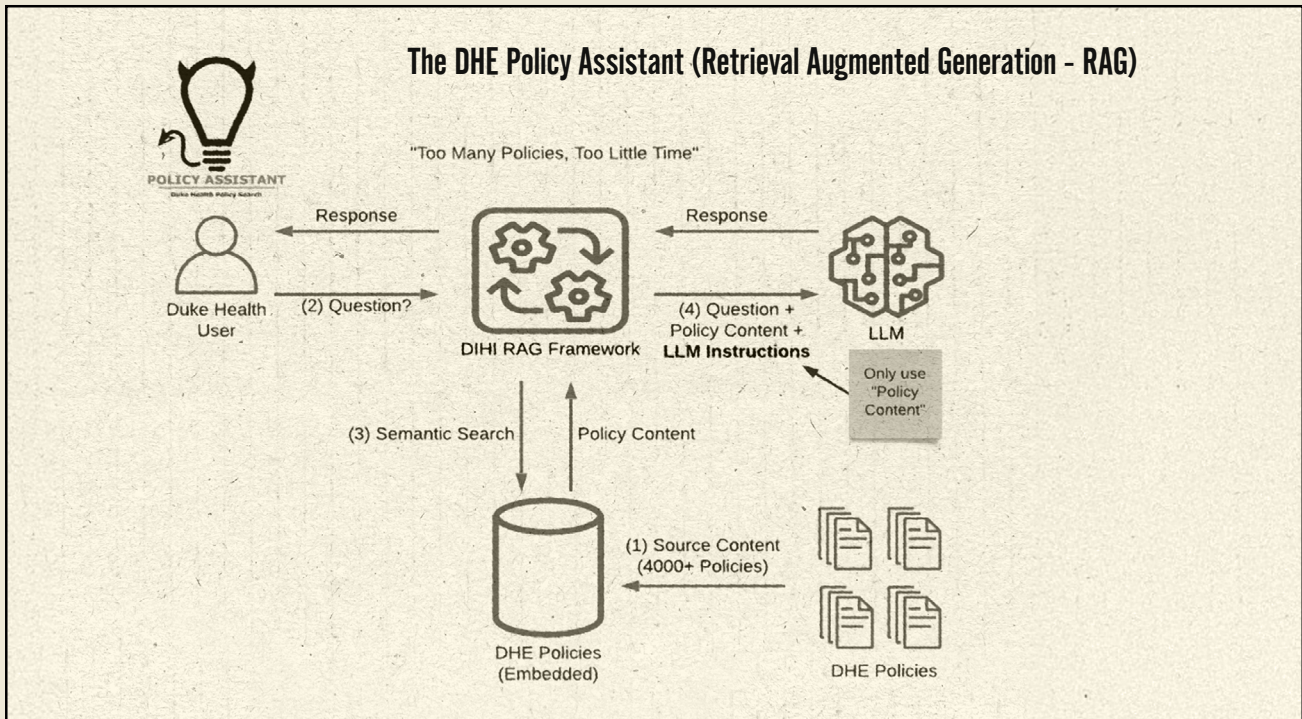


Figure 1 - DHE Policy Assistant Architecture Diagram

NEXT PAGE →



We also recognized that access to the source policy documents was necessary to facilitate user evaluation and feedback. This allows users to dig deeper to verify the accuracy of an LLM-generated response or to gain extended insight into discrepancies between generated responses and source documents. The Web application provides quick links to all referenced source documents in the context of a conversation. See Figure 2 – DHE Policy Assistant User Experience

Conclusion

The ongoing development of the “DHE Policy Assistant” through the DIHI 2024 RFA initiative underscores the transformative potential of AI-powered knowledge management systems. Using the sophisticated Retrieval-Augmented Generation (RAG) technique, we aim to address the pressing challenge of providing timely and accurate access to

Duke Health’s extensive policy documents. This project underscores DIHI’s commitment to addressing real-world healthcare challenges and improving clinical decision-making through efficient knowledge management and access to critical information.

We have begun a limited deployment of the system among a cohort of clinical domain experts. This phase will generate valuable data on usage patterns and user feedback, further informing our iterative improvements. Through continuous refinement and optimization, we aim to develop a robust AI assistant that significantly enhances access to policy information, ultimately improving decision-making and patient outcomes at Duke Health.

The screenshot displays the Duke Health Policies Assistant interface. At the top left is a lightbulb icon. The title "Duke Health Policies Assistant" is centered. Below the title is a text input field containing the query: "What is the mask policy when you are in the operating room?". To the right of the input field is a "Submit" button. Below the input field is a section titled "Query Response..." with a sub-header "Document Context...". The response text reads: "In the operating room, the mask policy is as follows:" followed by a numbered list of five items: 1. **Surgical Masks:** Properly secured surgical masks must be worn by all individuals entering the restricted areas of the Operating Room. The mask should cover the mouth and nose completely and be secured in a manner that prevents venting at the sides of the mask. Masks with ear loops are not approved within the restricted area as they may not provide a secure facial fit that prevents venting. 2. **Beard Covers:** Beard covers must be worn over facial hair that extends beyond the surgical mask. 3. **Mask Usage:** Surgical masks should not be worn hanging around the neck. The use of double masks is not recommended. A fresh surgical mask should be donned before the healthcare worker performs or assists with each procedure. The mask should be discarded and replaced whenever it becomes wet or soiled during the procedure and should be removed upon exiting the operating room suite. The surgical mask should be removed and discarded by handling only the mask ties. Hand hygiene should be performed after removal of masks. 4. **Special Masks:** Use a fit-tested N95 or high filtration mask for laser, TB, or other droplet/airway precautions procedures. 5. **Reusable Eye Protection:** Reusable eye protection devices should be cleaned according to the manufacturer's instructions before and after each use. Below the list is a section titled "Policies Referenced:" followed by two items: 1. DUHS Guidelines for Surgical Attire, Effective Date: 11/21/2022, Entities: AHS, DRAH, DRH, DUH 2. DRH DASC Operating Room Attire, Effective Date: 10/29/2020, Entities: DRH

Figure 2 - DHE Policy Assistant User Experience

LARGE LANGUAGE MODELS

Keep Large Language Models Local

MARSHALL NICHOLS, MS; MIKE REVOIR, BS

Generative AI (GenAI) and Large Language Model (LLM) technologies have the potential to significantly change healthcare delivery in many ways. For example, the Duke Institute for Health Innovation (DIHI) developed a natural language processing model to identify low-risk pulmonary embolisms in real time to facilitate safe outpatient management.¹ Our 2024 request for applications (RFA) was singularly focused on these types of AI solutions to improve staff and clinical efficiency, patient journey, and outcomes.

As expected, using these technologies is not without risk. While open-source LLMs (such as Llama 3.1) can be deployed within an enterprise (aka, on-premises or on-prem deployments), many popular LLMs (such as GPT-4 and o1) are not open-source and are only available in the cloud. Recent security breaches at AT&T, TicketMaster, Synnovis, MediSure, Kaiser Permanente, Cencora pharmaceutical,² UnitedHealth³ and Summit Pathology,⁴ have exposed millions of individuals' personal, financial, and health-related information. As healthcare professionals grapple with the complexities of patient data, regulatory compliance, and the need for securely efficient communication, the merits of on-prem solutions become increasingly apparent.

Enhanced Data Security and Compliance

One of the most pressing concerns in healthcare is data security. On-prem LLMs allow organizations to maintain complete control over sensitive patient information, ensuring compliance with regulations such as HIPAA (Health Insurance Portability and Accountability Act of 1996). Unlike closed-source models that operate in the cloud, thereby exposing data to potential breaches, on-prem solutions provide a fortified environment where data analysts, scientists, and engineers can process data locally. On-prem solutions minimize the risk of unauthorized access and align with the stringent data governance requirements that healthcare entities must adhere to.

Customization and Control

On-prem LLMs offer unparalleled customization capabilities. For instance, healthcare organizations can tailor these models to optimize for specific medical terminologies, integrate seamlessly with existing electronic health record (EHR) systems, or enhance functionalities like

clinical documentation and patient interaction. This customization level, often not feasible with open-source models, can significantly improve the efficiency and effectiveness of healthcare workflows.

Improved Performance and Reliability

Performance is critical in healthcare settings where timely and accurate information can significantly impact patient outcomes. On-prem LLMs can be optimized for speed and efficiency, reducing data retrieval and processing latency. Unlike cloud-based solutions that can suffer bandwidth limitations and downtime, on-prem systems provide a stable and reliable infrastructure, often at a lower cost. Lower costs can be a significant advantage for healthcare organizations, especially those operating on tight budgets.

Integration with Existing Systems

Healthcare organizations often rely on a multitude of existing systems and software. On-prem LLMs can be seamlessly integrated into these environments, allowing for a more cohesive workflow. This integration is crucial for tasks like clinical documentation, where LLMs can assist in summarizing patient notes and extracting pertinent information from EHRs. Such capabilities enhance productivity and improve the quality of care delivered to patients.

**Performance is critical
in healthcare settings where
timely and accurate information
can significantly impact
patient outcomes.**

NEXT PAGE →



Addressing Ethical and Legal Considerations

The deployment of LLMs in healthcare raises ethical and legal challenges, particularly concerning patient privacy and data usage. On-prem solutions allow organizations to implement ethical guidelines and governance frameworks, ensuring people use AI technologies responsibly. In contrast, the lack of control over data handling in closed-source models can lead to ethical dilemmas such as unauthorized data access and potential legal repercussions like data breaches and privacy violations.

Conclusion

Deploying GenAI and LLM solutions on-premises presents a compelling case for healthcare organizations seeking to leverage AI technology. By prioritizing data security, customization, performance, integration, and ethical governance, these solutions empower healthcare professionals to enhance patient care while navigating the complexities of modern healthcare environments. As the industry evolves, embracing on-prem LLMs may be a pivotal step toward achieving more efficient, effective, and secure healthcare delivery.

References

1. Amin KD, Weissler EH, Ratliff W, Sullivan AE, Holder TA, Bury C, Francis S, Theiling BJ, Hintze B, Gao M, Nichols M, Balu S, Jones WS, Sendak M. Development and Validation of a Natural Language Processing Model to Identify Low-Risk Pulmonary Embolism in Real Time to Facilitate Safe Outpatient Management. *Ann Emerg Med*. 2024 Aug;84(2):118-127. doi: 10.1016/j.annemergmed.2024.01.036. Epub 2024 Mar 2. PMID: 38441514.
2. <https://techcrunch.com/2024/08/12/2024-in-data-breaches-1-billion-stolen-records-and-rising/>
3. <https://www.reuters.com/technology/cybersecurity/hack-unitedhealths-tech-unit-impacted-100-mln-people-2024-10-24/>
4. <https://databreaches.net/2024/10/30/summit-pathology-laboratories-notified-1-8-million-patients-of-a-breach-less-than-48-hours-later-they-were-sued/>
5. <https://aisera.com/blog/large-language-models-healthcare/>
6. <https://www.allganize.ai/en/blog/the-rise-of-on-prem-llms-how-are-large-language-models-llm-changing-the-landscape-of-ai>
7. Abdul Basit et al. 'MedAide: Leveraging Large Language Models for On-Premise Medical Assistance on Edge Devices' <https://arxiv.org/html/2403.00830v1>

STUDENT EXPERIENCES WITH DIHI



**Freya
Gulamali**

I have been working with the Duke Institute for Health Innovation (DIHI) for the past two years. As a computer science student on the pre-med track, DIHI has been the perfect place for me to gain practical experience in applications of clinical data science and the importance of developing guidelines to ensure these applications are safe, effective, and equitable. I have had the opportunity to learn qualitative methods working with the Health AI Partnership, collaborate directly with clinicians on a project to identify venous thromboembolism (VTE) and stroke in the electronic health record (EHR) and understand the end-to-end process for developing an AI tool through Pythion, which assesses risks of post-surgical outcomes. I was privileged to have the opportunity to be the point person guiding the next steps for the VTE/stroke project. Through this experience, I learned how practicing medicine translates into the EHR system and how inconsistencies in this translation create barriers in data science projects. Additionally, I understood the importance of clinical-data science partnerships in making sense of the inconsistencies to create something that positively impacts patient care. The level of trust and support from my mentors at DIHI, enabling me to work on these real-world projects, has tremendously impacted my college experience.

With a team of people spanning all disciplines, DIHI embraces the importance of collaborative spirit in furthering the speed and quality of health innovation. Everyone is passionate about improving the way healthcare is delivered and is more than willing to share their knowledge and perspective on how they approach this shared mission. I have gained so many hard and soft skills required in this field and met so many people who have influenced and inspired my career aspirations. As I reflect on my experiences at DIHI, I feel the pioneering spirit Duke often aspires to. I am thrilled to continue working with DIHI over the coming year.

LARGE LANGUAGE MODELS

The Large Language Model (LLM) Frontier in 2024

MICHAEL GAO, MS

Large Language Models and generative AI continue to dominate the discourse across industries and the economy in 2024. Since the advent of ChatGPT, many other companies have contributed to the burgeoning landscape of LLMs and generative AI solutions. In the past two years, Anthropic, Meta, Mistral, and many other labs have contributed state-of-the-art models that have greatly improved the general performance of LLMs in what seems to be an arms race. In healthcare, generative AI solutions promise to automate burdensome processes, augment clinician workflows, and increase the overall efficiency of health systems, payors, and other healthcare players. Despite all of this activity, there is some emerging evidence that the investment into generative AI is still outpacing the return in terms of solutions, especially when it comes to investment in compute resources. On the other hand, there are exciting new developments in the world of large language models, especially when it comes to post-training, compute-heavy workflows. As I discussed last year, this is largely due to the last-mile problem of integrating generative AI into real-world workflows, which requires custom development and careful evaluation of solutions that go beyond demos that impress but are ultimately insufficient to transform the status quo.

The arms race

Training generative AI models takes a massive amount of compute resources. Today, those resources are predominantly Graphics Processing Units (GPUs). Large companies in the AI space have committed to building out immense GPU clusters. Estimates are that Meta has committed to spending around \$10 billion on GPUs from NVIDIA, with similar investments from other large players (Microsoft, Google, etc.). It is clear that generative AI has yet to generate a return when it comes to solutions that justify such a significant build-out of compute, but it is evident that the large AI players see a future where these clusters are necessary to keep their competitive advantage. What this means for the broader industry at large is that we can expect models to improve in their overall performance over time and that the cost of running these models will continue to be a race to the bottom. As of writing this article, models are significantly more performant and cheaper to run than they were at their inception. For specific numbers, in March of 2023, the cost of GPT-4 was \$60/million tokens, whereas

today, that number stands at \$5/million tokens. Although GPUs are used in both training the models as well as inference (actually running the trained models), there is an emerging class of hardware, such as those produced by Groq Inc., which are only used for the inference side of the equation which provide much faster inference than GPUs. However, the price per unit is prohibitively expensive for smaller companies. Nevertheless, the activity around compute has shown that it remains a primary axis by which companies are differentiated and how central compute is when it comes to generative AI.

The future of generative AI

Since the original advent of chat-based LLMs, there has been a flurry of research into extending these clearly powerful models into new workflows that will bring about the transformation that large companies are forecasting. One line of work involves making Large Language Models multi-modal, meaning that they can take in and generate not only text but also speech and images. This opens up a whole new world of opportunities where models seamlessly integrate world knowledge with these other modalities. The idea is that you can now speak to large language models and ask them to generate images or have the model generate speech given an image. Several demonstrations of these models have been shown, though their true commercial applications remain to be seen. Beyond these multi-modal models, there has been a whole class of solutions that rely on test-time compute, meaning scaling the performance and capabilities of models after they are trained by utilizing more compute resources to generate better answers and outputs.

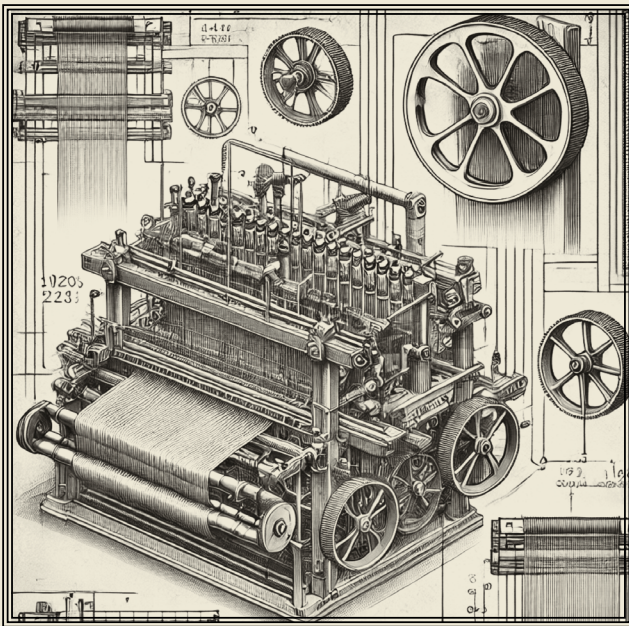
Generative AI solutions promise to automate burdensome processes, augment clinician workflows, and increase the overall efficiency of health systems, payors, and other healthcare players.

NEXT PAGE →



Test-time compute

The recent activity and focus on test-time compute is a natural progression of generative AI solutions. When foundation models like GPT-4, Llama, etc., take millions of dollars and months to train, the most meaningful way to make progress outside of scaling those processes has to take place after these models are already trained. One class of these is related to LLM agents. Agents are essentially LLMs that can utilize tools (generally software tools) to augment their ability to interact with complex systems. Whereas chat-only models can answer questions or perform text-based tasks, agents with tool usage can go beyond these types of workflows. Imagine that an LLM is equipped with the ability to query a database or to make phone calls. Combining these tools with LLM's understanding of language can open up a whole new surface for generative AI solutions. In addition, there is an increasing interest in LLM agents interacting with one another to produce and refine outputs from other LLMs. Imagine a group of LLMs, each with access to their own tools, working together to coordinate a complex task. At DIHI, we have already used such a system to automate literature reviews, by having certain LLMs perform search using terms and coming up with synonymous terms, one for culling down the results based on some relevant criteria, and another which is tasked with building out an annotated bibliography for a user to review. This allows for more flexible workflows, as these agents can be repurposed for other tasks that may utilize the same tools. We expect this frontier will continue to evolve and feature prominently in the next generation of generative AI products.



In addition to agents, there has also been some recent work on test-time reasoning, or essentially using more compute to give LLMs the ability to take more time to think. OpenAI's recently released o-1 preview and o-1 mini models are an excellent example of this. By taking several seconds to plan out a response, iteratively review the contents, and try to point out failure modes, the eventual outputs of these models often far surpass the outputs from previous models on certain reasoning tasks. One can imagine that this is somewhat like having a critiquing LLM, a planning LLM, and several others involved in the process of generating the best outputs. By utilizing more compute and layering LLM calls on one another, companies can leverage their compute resources to augment the performance of their already trained models to provide even better outputs. These reasoning tasks have already shown excellent performance on PhD-level tasks, which shows clear promise for this strategy moving forward. However, the amount of additional compute this takes is non-trivial. For example, OpenAI, as of the writing of this article, only allows for 50 submissions to their o1-preview API per week for their highest tier of paid users, whereas there is no limit for regular calls.

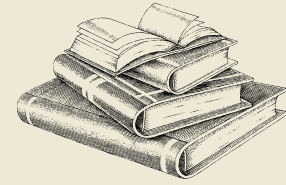
What we expect moving forward

Two emerging trends in 2024 are sure to continue into the future. Rapid investment into both train-time and test-time compute infrastructure from large corporations will continue to ramp up as the importance of compute resources continues to become more evident. Secondly, the advancement of utilizing those compute resources after the training process in order to create new technical workflows will continue to evolve and become more robust, leading to new possibilities for accomplishing complex and multi-modal workflows. It is not difficult to imagine a future in which an LLM could read a scanned document from the medical record, make inferences based on the results, search the EHR for more information about the patient, and then guess which team to consult about this patient and perhaps even make the phone call. Despite these promises, the last-mile problem still applies. Workflows tend to be complex, have outliers, and have exceptions to rules; which continue to be difficult to integrate into generative AI solutions. Careful work still needs to be done both from a technical standpoint and from a change management standpoint to ensure the effective and safe deployments of these solutions into critical workflows. Nevertheless, the promise of this AI revolution has not waned one bit, and we look forward to participating in its progress looking ahead.

OUR TEAM'S TAKE

Intrapreneurial Healthcare Strategy

Will Knechtle, MBA, MPH



This article explores how information fluidity, artificial intelligence (AI) in healthcare, and intrapreneurial spirit may be united to drive innovation within Duke Health. It emphasizes the importance of bolstering intellectual property, strategic AI adoption, and a resilient culture.

Healthcare innovation should be problem-driven, yet professionals should start by leveraging their assets.

Intrinsic motivation is key—and fortunately, healthcare is one of the most engaging academic fields. It is:

1. Ideal for the purposeful, its mission is to better human lives
2. Ideal for the curious, it faces vast unknowns and enormous ongoing challenges
3. Inspiring for the ambitious, it is one of the most difficult to master

For example, much is unknown about brain function, consciousness, and neurological disorders, yet we face ongoing challenges in cancer, aging, and emerging diseases. The medical field covers complex subjects such as physiology, pharmacology, epidemiology, and pathology, and its leaders should also be trained in policy and ethics. The healthcare field arguably has the most open frontier and highest demand for excellence, with providers needing to integrate vast amounts of information while performing intricate tasks in high-stakes situations.

A second asset is technological advancement, which accelerates our ability to master information and explore vast swaths of data. In October 2022, OpenAI unveiled ChatGPT, showcasing the capabilities of large language models (LLMs) on a grand scale for the first time. These technologies have demonstrated their ability to synthesize and summarize extensive amounts of data into concise, actionable information. Public access to foundational tools for digesting complex healthcare subjects has jumped forward. For example, public

access to Generative AI and LLMs is accelerating because they are becoming commoditized. The essential technical infrastructure to adapt them is also becoming increasingly open-sourced and generically applicable. Thanks to open-source code, software, GitLab, and GitOps, individuals with moderate technical expertise can learn methods, invent, test, and innovate.

Third, Duke leadership has reason to be confident in its people - its internal capabilities. Within Duke Health, the distinct value of our human expertise and personalized care remains irreplaceable and continues to grow in importance. In-house healthcare professionals can be the most effective healthcare innovators due to their close connection with patients and other care providers. For example, vendors visit us bringing technical infrastructure and materials Duke’s personnel already have access to while admitting they need our medical, business of care, and workflow expertise for their products to function. Therefore, Duke’s people already hold the keys to innovation. Despite its small size, the Duke Institute for Health Innovation continually turns research into innovation. There is great potential for Duke to leverage its extensive body of technical and healthcare experts further!

Generative AI and LLMs is accelerating because they are becoming commoditized. The essential technical infrastructure to adapt them is also becoming increasingly open-sourced and generically applicable.

NEXT PAGE →



Fourth, Duke has a 100-year legacy of bold and persistent intrapreneurial spirit. In 2017, President Vincent Price reminded us that “Duke has shown its boldness before in launching innovative interdisciplinary initiatives around – among others – the environment, the brain, and entrepreneurship. Few, if any, of our peers do it better.” The university’s legacy of fostering homegrown invention and innovation is a testament to its founders’ vision and the university’s commitment to nurturing its intellectual property (IP) and technological advancements. While Duke’s research is not unique among excellent universities, its gritty, resilient, scrappy confidence is.

Duke’s history is rich with examples of grit and innovation, particularly in medical care. Remember that Duke started with a cocksure ‘fake it till you make it’ attitude that Ivies may have considered laughable. Duke’s builders used stone quarried from local Hillsborough to create a presumptuous ivy-league imitation. Even so, “Duke, in fact, became a great university,” Tallman Trask suggested, “in part because it looked like one from the start.” Fans of Apple Inc. demos will respect the value of a ‘fake it till you make it’ attitude. This century, President Broadhead agreed that a “Let’s try that” spirit distinguishes Duke, and he’s described Duke as “the startup university.” Duke knew it was competing against legacy brands with much deeper pockets, and its employees and students scrapped and hustled until they had built a legacy. An intrapreneurial spirit is in our blue blood. We have the grit to innovatively develop intelligence and get information from noisy data as well as the humility to share it effectively.

We have a desperate need for information fluidity and local innovation.

The vast unknowns, ongoing challenges, and quantity of information to master have become too burdensome for healthcare professionals, patients, and our economy. For example, more three million articles are published annually in biomedical journals (Ghesemi, 2023). Population studies aside, the information required to assess one patient is tremendous. DIHI staff interviewed care providers and found that compiling a new patient’s information in preparation for an appointment can take up to two hours. Especially after COVID-19 and initiatives to promote access, virtualization has led to more patient messages and administrative Zoom meetings. Meanwhile, many healthcare professionals still rely on papers, faxes, scanned documents, and analog media to capture information. There are generous amounts of data, much of which is hard to access or complex to act upon.

Time is money, with both seeming increasingly challenging to find. In part because as the population grows, the sick percent remains the same (or has grown too), and the volume of healthcare providers hasn’t kept up, healthcare professionals must seamlessly access and share

extensive data to make informed decisions quickly. Information fluidity between healthcare professionals is not a convenience but a necessity—particularly crucial in high-stakes situations where every minute counts. Additionally, academic medical center economic margins have been suffering. (For an example, read FitchRatings’ December 2023 article about Duke University Health System). Consequently, the pressure to lower costs and optimize processes is relentless.

Duke has a 100-year legacy of bold and persistent intrapreneurial spirit.

Box 1. Article Definitions

DATA: Quantities, characters, or symbols. Simple facts or figures.

INFORMATION: Conveyance or signal from an arrangement or sequence of data.

FLUIDITY: easy movement, change, transferability.

TECHNOLOGY: Information applied for practical purposes in the real world.

AI: Computer science field focused on creating intelligent agents that can learn from data and experience. These agents process massive datasets and identifying patterns and relationships, which allow tasks such as diagnosis, decision-making, or creative problem-solving.

INVENTION: Creation of a new idea, product, or process.

INNOVATION: Invention applied for practical purposes in the world.

ENTREPRENEUR: Person who identifies a need in a market and works to fulfill it by starting a business.

INTRAPRENEUR: Person who identifies a need in a market and works to fulfill it by innovating within an existing business.

INTRAPRENEURIAL SPIRIT: Curious, Passionate, Optimistic, Determined, Resilient, Risk- Tolerant, Creative, Independent, Decisive.



Generative AI and LLMs are a tool, not a solution!

In response to urgent needs, Duke Health leaders requested that applicants for innovation funding in October 2023 submit ideas using generative AI and LLMs specifically. This was warranted because the problem of information abundance and viscosity is where Generative AI and LLMs can make a significant impact. Long patient narratives can be distilled into sentences that healthcare workers need at critical moments. Whether it is making a prior authorization decision, formulating a surgical policy, managing a shift change handoff, or matching a donor organ to a recipient, these tools can provide essential information at the minute needed. The healthcare sector has struggled to keep up with its burgeoning information demands, but with these advanced tools, it might finally move towards managing it efficiently.

By making information more accessible, Generative AI and Large Language Models are likely to improve staff and clinician efficiency, patient journey, and outcomes. While ChatGPT and image generators make things look magical in just a few clicks. “Realizing their full potential demands meticulous research, development, and a dedicated, multidisciplinary team to ensure that solutions are efficient and secure and uphold quality standards. The mere existence of this technology does not imply that Duke will emerge as a leader in their use or that the technology will resolve all of our challenges unless there is a genuine commitment in time, energy, and resources toward the development and rollout of LLM-centric solutions.” -- Michael Gao, Impact Vol 24. In short, the material, personnel, and expertise to catalyze LLM/Generative AI adoption are in high demand due to low information fluidity, high healthcare professional burnout, and low margins. These catalysts and their integration into daily workflow are necessary to transform Generative AI and LLM inventions into innovations. How should Duke leadership solve this problem?

Healthcare leaders should focus on complete solutions while valuing their organization’s human touch.

1. Empowering transformative ideas to strengthen operations
 One solution could be to amass and equip quality improvement engineers to optimize efficiency by strengthening existing processes. However, process perfectionism can stifle creative thinking and flexible experimentation. While keeping track of metrics is crucial for improvement, not all solutions yield immediate, measurable results. We must balance both data-driven optimization of existing workflows and investment in transformative ideas. We need an intrapreneurial culture to foster the development and implementation of new technologies and solutions, ensuring that healthcare systems can rapidly adapt to increased information demands while reducing costs and improving care quality.

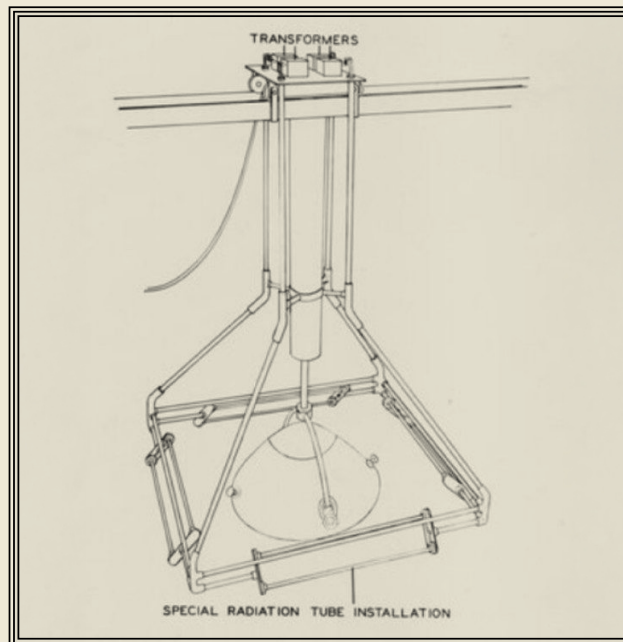
2. Valuing Human Influence

Information fluidity is a lubricant to lower care costs because it can catalyze operational efficiency improvements and improve patient access. Sharing basic medical knowledge or foundational data science tools can accelerate patient care improvements, creating a ripple effect of improving health and the economy. However, it is crucial to recognize that most medical knowledge is not a commodity and should not be treated as such. Freely sharing all information is rarely sustainable and can undervalue the time, expertise, and innovation behind it (Box 2).

A more viable approach is to strategically capitalize on healthcare knowledge. Healthcare professionals and institutions can share their expertise as paid consultants, industry employees, or entrepreneurs. When they combine their expertise with advanced software that improves healthcare practice, they create valuable IP. Patenting and licensing valuable IP generates revenue streams that can be reinvested in further research and development. This approach sustains innovation and motivates growth while protecting the value of unique medical insights.

3. Balancing external solutions with internal growth

Considering the urgent demand for material, personnel, and expertise, reaching out to external vendors seems sensible. Vendors could help Duke speed to market, reduce startup costs, increase competitive advantage, and earn high early revenues. Nevertheless, weighing the potential drawbacks of buying external tech support is essential.



NEXT PAGE →



Four costs of external technology are:

- **Purchasing cost:** The upfront cost may exceed its benefits if it doesn't fluidly integrate with existing systems or comply with healthcare regulations, leading to additional customization, security, and privacy expenses. The cost is sunk if the product fails to meet its intended purpose, is cumbersome, or does not advance healthcare effectively (read [page 47](#), Keep LLMs Local).
- **Switching cost:** Without the necessary technical expertise to maximize and maintain the product, an organization may become dependent on a vendor, resulting in higher long-term costs and reduced influence. This dependency can make maintaining, modifying, or replacing the product more challenging. High switching costs weaken Duke Health's negotiating position, especially in a rapidly changing tech environment where innovations quickly disrupt existing products. (See Box 3)
- **Opportunity cost:** A missed opportunity for lower-cost, similar-value alternatives. Reduced long-term differentiation. Reduced staff and faculty development and a reduced ability to adapt or innovate the next time it's required.
- **IP leakage:** This hands companies the knowledge they capitalize on for free and even with payment. In other words, there's a risk of underselling our expertise, knowledge, and experience. IP leakage may reduce financial incentives for internal researchers, hinder innovation, reduce ability to attract top talent, and limit revenue generation. It often slows how we innovate.

The overwhelming influx of technology and marketing can lead purchasers to undervalue internal expertise and resources. Simultaneously, we could value internal knowledge and resources so much that we just 'put our heads down' and lean into refining current processes without making impactful changes. Ultimately, we miss the opportunity to increase Duke's IP (future value) and inspire our passionate people.

A solution: catalyze internal innovation and build a community of intrapreneurs

Duke University and Duke Health stand at a pivotal crossroads, where they must invest in technology, materials, and people to foster innovation. The Duke community should ask itself:

- Would Duke Health advance its vision or "boldly contribute" if it pays someone else to accelerate discovery and translation?
- Are we advancing the frontiers of knowledge or creating transformative education if we pay others to develop solutions outside and bring them in?
- If a university is not investing in its internal capabilities, making information more accessible, or training people to solve real-world problems, can it call itself an institution of learning?

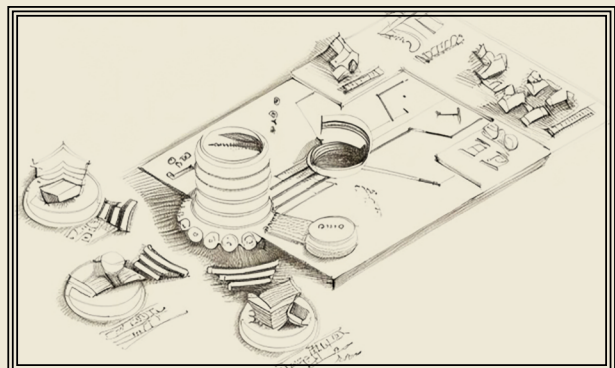
Box 2. Case Example

During the COVID-19 pandemic, the open inter-organizational sharing of research and clinical data worldwide led to the swift development of vaccines and treatment protocols, demonstrating how free information exchange can profoundly impact public health and the economy. Commercial collaboration allowed Moderna to leverage MIT's research to create viable products, showing how academic research can be translated into practical applications that benefit society and generate economic growth—furthermore, sharing internal information, such as who was COVID-19 positive and where mechanical ventilators were, enabled efficient treatments and lowered wasted materials.

However, today, pharmaceutical companies are heavily focused on developing new methods to recover financially from their generosity.

Box 3. Case Example

A few decades ago, the United States (US) outsourced chip manufacturing to Taiwan Semiconductor Manufacturing Company Limited (TSMC). This decision created a significant disadvantage in the era of AI, increasing our risks in defense and cybersecurity. Today, it forces the US to invest heavily to bring those capabilities back home to reduce risk and stay ahead in the global AI race.



As long as Duke University and Duke Health share educational values, Duke should prioritize its internal expertise and foster an intrapreneurial spirit to maintain its leadership in medicine and influence over its contributions.

Initial steps to foster innovation are building confidence, increasing internal information fluidity, and encouraging resilience (See Box 4). Hence, we should be more conscious of holding our keys to the future. Duke must invest in and leverage our internal expertise, ensuring we remain at the forefront of medical innovation while valuing our unique offerings.

As Duke navigates the rapidly evolving landscape of higher education and AI, it must carefully balance external acquisitions with internal development. While buying technological solutions can offer short-term benefits, it risks long-term dependencies and challenges. The temptation to leap-frog development by acquiring ready-made solutions is understandable. Still, it must account for the time required to localize innovation within hospitals and clinics to create value. Information fluidity is critical in this process, catalyzing operational efficiency and strategic decision-making and enabling Duke to harness its internal expertise effectively. By prioritizing, developing, and commercializing its IP, Duke can ensure its innovations remain rooted in its mission to nurture talent, expand humanity’s knowledge base, and advance patient healthcare. This approach will enable Duke to continue leading in its field, leveraging its unique strengths and maintaining its influence over its contributions to society. DIHI is committed to homegrown innovation as Duke navigates a rapidly evolving landscape of healthcare, higher education, and AI.

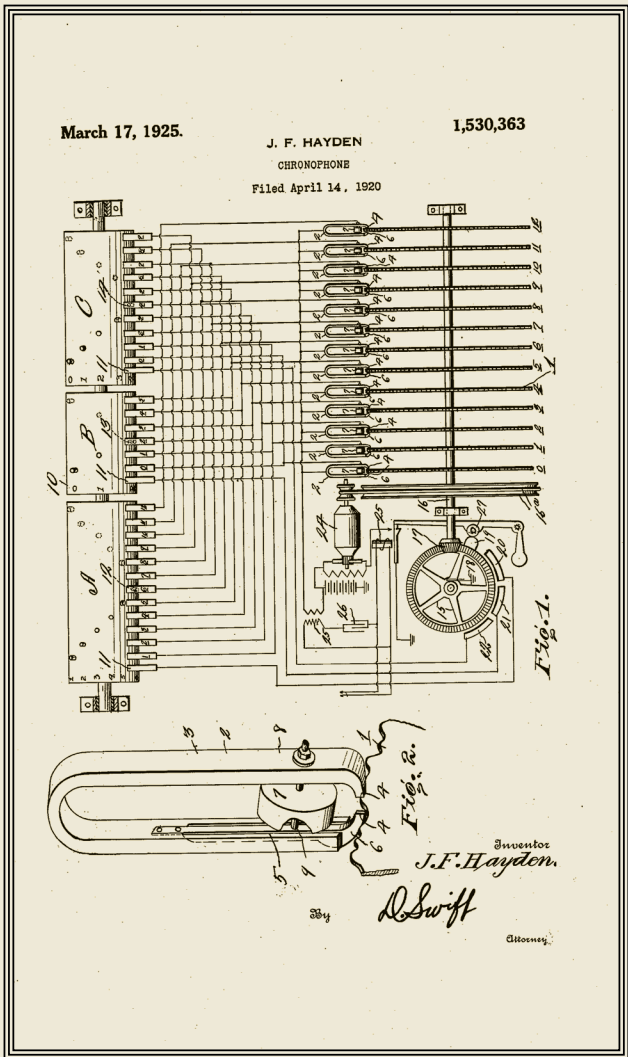
References

Huler, Scott; Hill, Corbie. (2024, March 26). ‘A look back, a look ahead: We celebrate Duke’s centennial by honoring what came before and imagining what the future holds.’ DukeMag. <https://dukemag.duke.edu/stories/look-back-look-ahead>.

Beauvais, B., Richter, J., & Fulton, L. (2016). Moving the Needle: Evaluating the Impact of New Care Delivery Models on Hospital Profitability. <https://core.ac.uk/download/216389182.pdf>

Box 4. Here are a few suggestions for how to foster innovation:

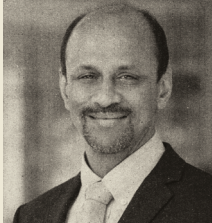
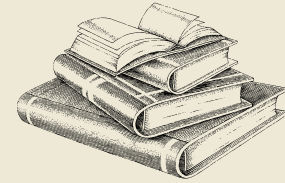
- Seek inspiration from Duke founders’ era of innovation
- Seek solutions from staff, faculty, & students
- Support git and secure on-premise computing operations
- Support intellectual property protection
- Provide resources and time for invention
- Provide freedom for piloting and rapid failure
- Train and hire for curiosity and humility
- Visit Cameron Indoor Stadium



OUR TEAM'S TAKE

Commercializing Innovation: An Interview with Suresh Balu

Will Knechtle, MBA, MPH; Fedor Kossakovski



FEDOR K:

A new health innovation itself is great, but how do you get it out there to the masses?

SURESH BALU:

Most innovations arise from addressing a specific problem. At DIHI, the focus is on solving existing problems rather than finding applications for pre-existing technology or solutions. The process begins by gaining a deep understanding of the problem, then creating a proof of concept that demonstrates value by effectively solving the problem.



WILL K:

When innovators develop new solutions, are they even thinking about the problem or just the solution? How do they bridge problems to solutions to applications?

SURESH BALU:

Innovators often start with a solution in mind. At times, innovators are enamored with new technology, such as Generative AI or ChatGPT, and attempt to find a use for it. This approach is risky. The DIHI model is to start with a foundation of deep understanding of the problem, how it affects the various stakeholders, and then determine the best approach to developing the solution. While ideating and prototyping, we iterate on designs to ensure usability and effectiveness. Even after launch of a solution, we remain committed to monitoring performance and making ongoing improvements to ensure continued efficacy, sustainability and scalability.

WILL K:

Are you saying that if you solve a big problem, the masses will come? After that, is the work simply to scale to meet the market with good business operations? What does it take to scale the solution?

SURESH BALU:

People are willing to adopt solutions that provide real solutions to problems, especially in healthcare, where the need for improvement is constant.

While solving a problem is a starting point, scaling is influenced by many factors like regulatory compliance, reimbursement, market considerations and economics. It's not just about generating a return on investment (ROI), but rather about understanding the problem and ensuring that the solution effectively addresses all other impact measures. The key is to demonstrate that the solution works consistently in real-world settings.

WILL K:

How do you get from a small working solution to mass mobilization?

SURESH BALU:

Scaling a small working solution to mass mobilization requires a strategic approach and one that involves several key steps:

- Understand the problem and the customer.
- Develop a solution.
- Validate its clinical and economic utility for the user / customer.
- Assess its economic impact. This includes evaluating potential cost reductions, burden reduction, outcome improvements, and efficiency gains.
- Leverage expertise in translating research into market-ready products and operating technology development to determine the best scaling approach.
- Collaborate with the Duke University Office for Translation & Commercialization (OTC) to develop a strategic approach to scaling (e.g., licensing the solution, developing strategic partnerships and collaborations).
- Work with the licensee to revalidate outcomes (both financial and economic) and adapt the solution to their specific market needs.

Scaling in healthcare differs from other industries because healthcare solutions often require customization to fit unique local workflows, support complex value chains and meet regulatory requirements.

WILL K:

Would you work with one customer at a time (like from one health system to another)? Would you do small batches? Would you do the same it the way if we were a much larger team?

SURESH BALU:

A larger team could scale solutions in parallel, aggregating learnings. Segmenting the market is essential to customize solutions. For example, academic medical centers (AMCs) may require a different approach than low-resource community hospitals. Segmentation allows for customization without resorting to a purely individualized (“N of 1”) approach.

Healthcare innovations often require a segmented approach because workflows vary significantly.

FEDOR K:

What are the unique challenges of commercializing healthcare innovations?

SURESH BALU:

Commercializing healthcare innovations is a complex and risky endeavor. A big challenge is the complex payment model in healthcare. Unlike other industries, healthcare innovation is not simply about selling a product. It’s crucial to align the innovation with existing business models and understand how value flows through the system.

- Who pays for the innovation? And, what is the reimbursement model for the solution?
- Who benefits from it?
- How are funds allocated?
- What is the right channel for selling the product? (One option is a group purchasing organization)
- What evidence is required prior to commercialization (clinical trial, regulatory review and approval)?
- What are the ethical and equity considerations?

These questions must be carefully considered to ensure successful commercialization.

Commercializing healthcare innovation is difficult because the payment, purchasing, and procurement processes are often unclear. Even established companies and startups struggle with this complexity, highlighting the need for specialized talent. Another complexity of commercialization in healthcare is the workflow. There is no other industry sector where the actual value created depends so much on the workflow. Everything depends on it, and even the best innovations will have limited value if we fail to integrate with the workflow of healthcare providers and patients.

Another significant challenge is the constantly changing regulatory landscape, especially in AI and digital health space. As regulations evolve, products that were not previously considered medical devices may now require Food and Drug Administration (FDA) approval. This shift creates uncertainty for innovators and increases the complexity and cost of bringing a product to market. Different states within the US and regions like the EU may also have varying AI regulations, adding further complexity to the process.

FEDOR K:

How does DIHI interact with OTC?

SURESH BALU:

DIHI and OTC collaborate in several ways. First, OTC assists us with invention disclosure filings (IDFs), patent applications, and intellectual property (IP) protection. OTC provides guidance on navigating the licensing process, both for out-licensing DIHI innovations and in-licensing technologies from other organizations. Second, OTC is always engaged in DIHI’s important innovation sourcing events such as “Innovation Jam.” In addition, OTC has been instrumental in helping the teams in developing the terms for licensing and commercialization. Just through our Innovation Jam events, OTC has supported the creation of more than ten companies.

Furthermore, OTC helps DIHI explore alternative business models for commercialization and partnerships (with industry and other academic partners). We receive strategic advice on the best pathways for protecting our IP, commercialization, and value capture for specific companies spun out of DIHI innovations. An example is CohereMed, a licensee of DIHI’s several clinical decision support models, for which OTC specifically advised us to explore STTR as a way to secure non-dilutive funding to scale the innovation into other healthcare delivery organizations. The annual Invented at Duke event by OTC has significantly helped expand our collaborations and form new partnerships.

WILL K:

Tell us more about Sepsis Watch and CohereMed.

SURESH BALU:

We developed Sepsis Watch at Duke and demonstrated its value here. We completed an IDF, had many publications, and presented the impact at conferences. We met CohereMed leaders there, who were experienced in healthcare product management, and they had heard of us. Sepsis Watch was licensed to CohereMed. OTC played a crucial role in facilitating this partnership, setting up terms and conditions, and representing DIHI and Duke throughout the process.

NEXT PAGE →



One unique challenge was that Sepsis Watch involved licensing an AI model, a first in the industry. This required extensive discussions with OTC to determine what exactly was being licensed (model coefficients, code, etc.).

DIHI provided ongoing support to CohereMed, helping with code implementation, testing, and validation. Commercializing Sepsis Watch was challenging, requiring retrospective, prospective, and silent validation to demonstrate value. OTC also assisted CohereMed in securing a Small Business Technology Transfer (STTR) grant, providing non-dilutive funding.

After a lengthy process, CohereMed is on the verge of securing its first sale to a healthcare system, marking a significant milestone in the commercialization journey.

FEDOR K:
What are some lessons learned in DIHI's commercialization efforts?

- SURESH BALU:**
 DIHI has gained valuable insights through its commercialization efforts:
- Submit an IDF as early as possible. This helps crystallize the project's scope, identify key personnel, and outline the steps required for commercialization.
 - Clearly demonstrate value creation upfront. Healthcare systems are more likely to adopt solutions that have a proven impact.
 - Publish research findings promptly. Disseminating research findings through publications increases visibility and credibility of the product and solutions from Duke.
 - Provide ongoing support to the licensee. This includes technical assistance, knowledge transfer, and guidance on adapting the solution to different settings.
 - Embrace continuous learning. Commercialization is a dynamic process, and DIHI recognizes the importance of adapting and improving its approach.

FEDOR K:
What does the future of DIHI project commercialization look like?

SURESH BALU:
 DIHI aims to expand its commercialization efforts by solving more tough problems with more products, offering services, particularly innovation consulting services, to maximize value creation for partners, attracting larger investments from partners, and staying nimble and adapting to rapid technological advancements. We also look forward to continuing to focus on solving problems, while recognizing that commercialization is a byproduct of successful problem-solving. Lastly, we would like to build partnerships with regulators to better align commercialization with reimbursement and regulatory pathways across the world.

WILL K:
What closing statements would you like to share?

SURESH BALU:
 DIHI's success stems from its focus on solving real-world problems within our healthcare system for our patients, community and our care providers. We have an exceptional team that, in addition to the technology and management skills, brings a huge amount of curiosity to learn and passion for problem-solving. Our focus will continue to be on catalyzing and implementing pragmatic innovations that address real-world challenges. We look forward to continued investments in this area that will enable Duke Health to reach even higher levels of excellence and impact.

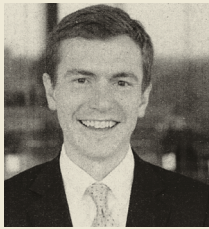
Fedor Kossakovski is the Director of Marketing & Communications for Duke's Office for Translation and Commercialization (OTC).



OUR TEAM'S TAKE

Utilizing Statistical Process Control to Track the Impact of Our Health Innovations: Lessons from a Pediatrics Project

Will Ratliff, MBA



“Will this intervention have a positive impact on my patient?” is the core question of healthcare. It transcends the segments of our industry and is the most critical metric of impact across health innovation, quality improvement, and research. At the start of each clinical decision support project (i.e., most of our work at DIHI), we identify and measure impact metrics once the solution is live at the bedside. So, the question, “What is the impact, and how can we measure it?” is always on our mind and drives us to improve and expand our methodologies with each go-live. The burden to answer this question effectively can be enormous, and typically there are not enough resources to do so. However, by applying our existing data assets and project knowledge, we can use basic statistical methods to sustainably track impact at scale.

Here is a case study from our High-Risk Condition (HRC) + Fever project (see [page 87](#) for project summary) to illustrate this concept:

In Spring 2024, we went live with a solution to page Emergency Department (ED) frontline clinicians as soon as a high-risk pediatric patient presents in the emergency department with a fever and has not yet received antibiotics. With the guidance of our project PI, Dr. Emily Sterrett, we applied Statistical Process Control (SPC) methodology to our standard automated data monitoring and reporting approach to yield time-to-antibiotic run charts in near real-time for our cohort of high-risk patients (Figures 1a, 1b). These charts visually represent the impact of our real-time page interventions over time on our key outcome variable, time-to-antibiotics. The charts support our clinical leads, Emily Sterrett and Emily Greenwald, to identify variation by case and track overall improvement for reporting to their ED colleagues.

The Value of Run Charts in the HRC Fever Project

A comparison of the run charts to a tabular representation of our overall reportable metric, median time from ED arrival-to-antibiotics (Tables 1a, 1b) reveals the following findings.

Increased visibility into time-to-antibiotics metric

Whereas the overall median time-to-antibiotics showed some improvement (75.6 minutes pre-go-live compared to 63.0 minutes post-go-live), the run chart shows evidence of six measurements below the latest process sample average (a “centerline”) in September and October. Even with a small data shift, seeing multiple points consistently below the centerline indicates a statistically significant improvement because it is evidence of a trend toward a lower average value. The run chart also enabled us to observe high variability among cases (i.e., the spikes above 240 minutes), compelling us to investigate those events for root-cause analysis and quality improvement takeaways.

The continuous use of SPC tools like run charts in the HRC Fever Project is an achievable methodology to link clinical action and patient outcomes.

NEXT PAGE



Variation by sub-group for workflow-driven quality improvement:

During our investigation into the high variability of certain cases, we discovered that a specific sub-cohort, transplant list patients, represents about one-third of our cohort but two-thirds of our high variability cases. When we exclude outlying cases from the analysis, we see much less variability in time from ED arrival-to-antibiotics and a more substantial impact. Additionally, this chart helps hone our focus on transplant patient workflow to reduce unwarranted variation.

The promising trends in September and October coincide with the start of monthly dissemination of our impact tracking data (i.e., tables and figures below) directly to the frontline ED care teams. We will continue to track these metrics to understand whether the real-time pages plus dissemination of metrics are having a sustained, statistically significant impact.

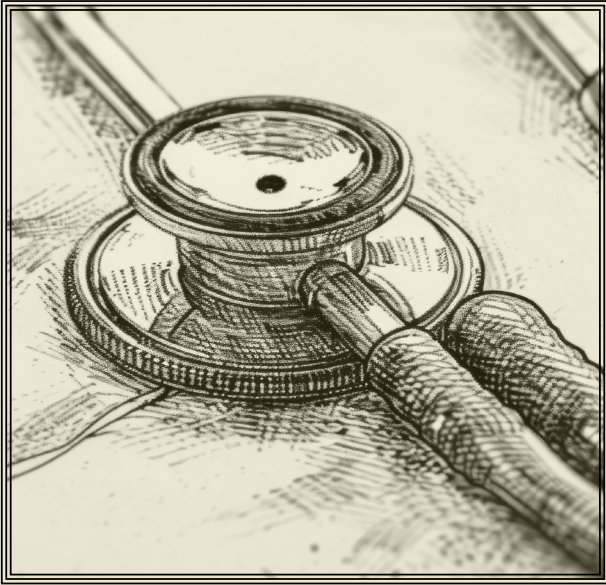
Conclusion

Leveraging Data for Continuous Improvement

The continuous use of SPC tools like run charts in the HRC Fever Project is an achievable methodology to link clinical action and patient outcomes. By understanding trends and real-time shifts in critical metrics, we can quickly iterate to optimize workflows and understand the impact of each iteration. We can then apply what works to other projects to assess generalizable efficacy.

As healthcare innovation leaders, we must leverage our data assets and established methods like SPC to support better patient outcomes and advance the quality of care provided across our health system. By closely monitoring these key metrics and implementing timely changes, we ensure that our interventions not only address immediate clinical needs but also contribute to the broader goal of long-term healthcare improvement.

Editorial Comment:
 When we use SPC tools well, we merge knowledge from the best of the 1920s – the decade Trinity University became Duke – with the applications of today’s newest technologies. SPC tools were pioneered by a self-taught statistician working for Bell Telephone Laboratories, Walter A. Shewhart, in 1924. By leveraging Shewhart’s foundational work on statistical process control and applying it to modern data analytics and machine learning, we can improve levels of quality control and efficiency in service industries, such as the healthcare provider space.



Outcome	Silent trial	Live in ED
	(1/1/24-2/25/24)	(2/26/24-10/31/24)
Patients meeting HRC + Fever phenotype, n	40	161
Reminder pages sent, n	n/a	74
Actionable initial page, n (%)	36 (90.0%)	138 (85.7%)
Median ED arrival to antibiotic administration time, minutes (SD)	75.6 (50.1)	63.0 (87.3)
Arrival-Abx compliance (<1 hour ED arrival to Abx time), %	36.1%	45.9%

Table 1a: Overall metrics for HRC+Fever, comparing pre vs post go live



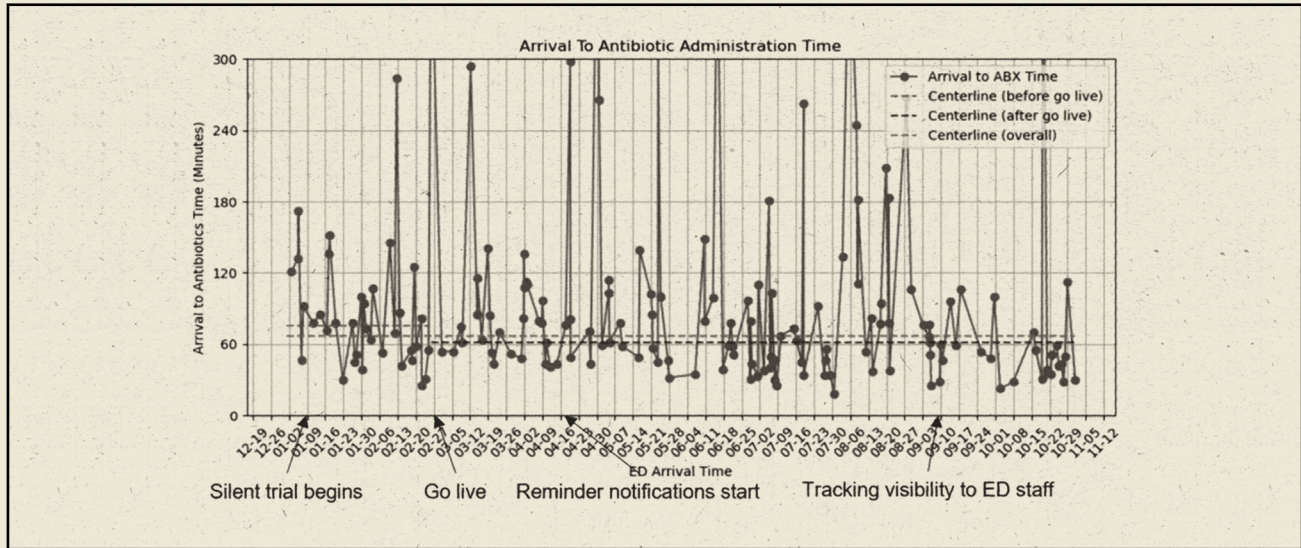


Figure 1a: Time to antibiotics, HRC+Fever Patient Cohort

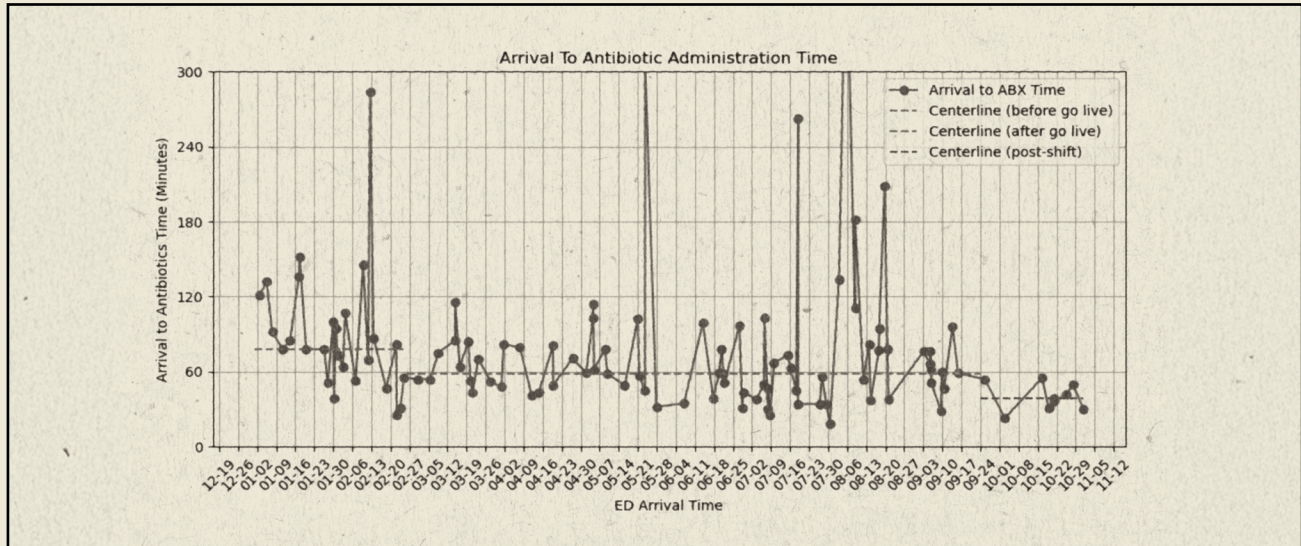


Figure 1b: Time to antibiotics, HRC+Fever Patient Sub-Cohort: non-transplant patients

Outcome (excludes transplant patients)	Silent trial	Live in ED
	(1/1/24-2/25/24)	(2/26/24-10/31/24)
Patients meeting HRC + Fever phenotype, n	28	100
Reminder pages sent, n	n/a	34
Actionable initial page, n (%)	27 (96.4%)	89 (89.0%)
Median ED arrival to antibiotic administration time, minutes (SD)	78.0 (51.2)	58.2 (59.4)
Arrival-Abx compliance (<1 hour ED arrival to Abx time), %	25.9%	55.2%

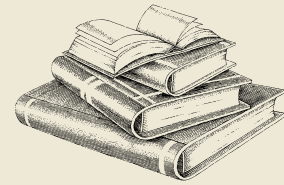
Table 1b: Overall metrics for HRC+Fever non-transplant sub-chort, comparing pre vs post go live



OUR TEAM'S TAKE

Perspectivation in Healthcare: Driving Innovation Through Diverse Perspectives

Jamie Daniel, BS



The healthcare industry, faced with the monumental task of catering to the diverse needs of global populations, is increasingly turning towards a new concept known as “Perspectivation.” This approach, which merges “perspectives” and “innovation,” is a strategy for problem-solving and innovation, particularly in healthcare, that emphasizes the integration of diverse perspectives to achieve more comprehensive and effective solutions. Perspectivation is proving to be a transformative strategy in tackling some of the most complex challenges in healthcare today.

Understanding Perspectivation in Healthcare

Perspectivation combines the diverse insights and experiences of individuals from various backgrounds to fuel innovative healthcare solutions. It underscores the crucial role of each individual, from patients to practitioners and policymakers to researchers, in creating more effective, inclusive, and sustainable healthcare solutions. Your perspective matters.

The Need for Perspectivation

Healthcare systems worldwide are often criticized for their one-size-fits-all solutions that may not adequately address the needs of diverse populations. Challenges such as health disparities, accessibility, and personalized medicine require a more nuanced approach. For example, digital health solutions designed for the average patient often overlook the specific requirements of those with more severe conditions, lower health literacy, or limited access to technology. Additionally, standardized treatment guidelines may not consider the unique circumstances of underrepresented populations, such as women, older adults, and racial minorities, who are often underrepresented in clinical trials. Perspectivation offers a way forward, ensuring that multiple perspectives inform healthcare innovation, leading to more equitable and effective care delivery.

Implementing Perspectivation: Key Strategies

1. MULTIDISCIPLINARY TEAMS

By forming teams that include healthcare professionals, data scientists, ethicists, patient advocates, and others, hospitals and research institutions can ensure that a variety of perspectives are considered in the development of new treatments, technologies, and systems.

2. PATIENT-CENTRIC INNOVATION

Engaging with patients from diverse demographic and socio-economic backgrounds can provide insights that drive more personalized and effective care. This includes understanding cultural sensitivities, language barriers, and accessibility needs that might otherwise be overlooked.

3. POLICY INTEGRATION

Policymakers play a crucial role in enabling an environment where Perspectivation can thrive. They can do this by crafting regulations that encourage collaboration across different sectors, emphasizing ethical standards, and providing funding for innovative projects. This can facilitate innovations that are not only effective but also equitable.

Benefits of Perspectivation in Healthcare

ENHANCED PATIENT OUTCOMES

By considering a wide array of patient experiences and needs, healthcare providers can develop more targeted and effective interventions. This improves patient outcomes and enhances patient satisfaction and trust in healthcare systems. Considering a wide array of patient experiences allows healthcare providers to identify specific subgroups or individual characteristics that may require tailored approaches, enabling them to move from broad, generalized care to more personalized and precise interventions.

REDUCED DISPARITIES

Perspectivation helps to address health inequities by bringing attention to the specific needs of underrepresented groups. This leads to more accurate diagnoses, better patient education, and tailored healthcare services. For instance, by analyzing diverse patient data, providers can recognize patterns in treatment responses among different demographic groups, leading to the development of targeted therapies or adjusted care protocols that address the unique needs of specific patient populations.

Increased Innovation

Cross-pollinating ideas from different fields and perspectives leads to breakthrough innovations. For example, leveraging insights from behavioral science can enhance the effectiveness of digital health tools, making them more user-friendly and widely adopted. Toyota implemented the “Toyota Production System” (TPS) by borrowing ideas from American supermarkets and applying them to its manufacturing process, resulting in significant efficiency gains and cost reductions. At the dawn of Duke University, Henry Ford was inspired by flour mills and meat packing plants, revolutionizing German automotive ideas to make cars more affordable for the average American. Recently, collaboration between the automotive and technology industries led to the development of electric and self-driving cars, as well as innovative infotainment systems.

Challenges to Overcome

Implementing Perspectivation is not without its challenges. It requires significant changes in organizational culture, including training to enhance empathy and understanding across diverse groups and systems to manage and integrate varied inputs effectively. Additionally, ongoing funding is needed to support collaborative research and development efforts. Our recommendations are to:

- Explore diverse funding sources, including government grants, private foundations, and industry partnerships (Read article by Vaibhav Jha on [page 37](#))
- Encourage leaders to model inclusive behaviors and actively seek out diverse viewpoints in decision making processes
- Establish a centralized system for collecting and organizing ideas and feedback from all levels in the organization (Read about the Duke Institute for Health Innovation (DIHI) and its Request for Application (RFA) process)
- Create cross-functional teams for key projects to ensure diverse perspectives are represented from the outset (Read about the DIHI RFA)
- Implement a structured process for evaluating and prioritizing ideas based on their potential impact and feasibility (Read about the DIHI RFA)

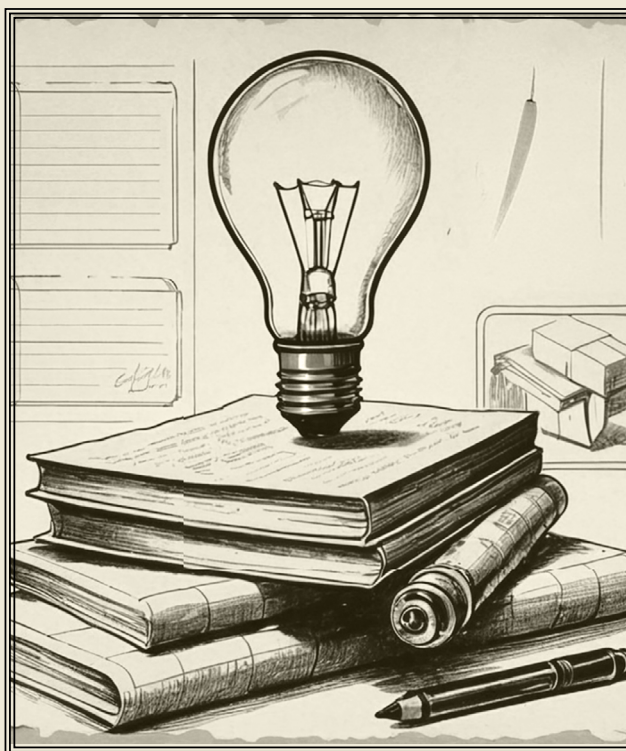
- Conduct regular diversity and inclusion workshops to foster awareness and sensitivity towards different perspectives (Read about Health AI Partnership (HAIP))
- Implement cross-departmental job shadowing programs to help employees understand challenges faced by colleagues in different roles
- Regularly solicit feedback, evaluate processes, and measure performance indicators

Conclusion

Perspectivation is setting a new standard for innovation in healthcare. It fosters a more holistic approach to solving health issues, which is particularly crucial in a field as inherently personal and diverse as healthcare. As this approach continues to gain traction, it promises to reshape the future of health services, making them more inclusive, effective, and responsive to the needs of all patients. The journey is complex and fraught with challenges. Still, the potential to transform healthcare for the better makes it a path worth pursuing.

References

Carl Friedrich Graumann and Werner Kallmeyer. Perspective and Perspectivation in Discourse. November 2002. e-Book ISBN: 9789027296931. DOI: <https://doi.org/10.1075/hcp.9>



LEGACY INNOVATIONS

Reflecting on 10 years of Impact



Stitching Together Chronic Kidney Disease Innovations to Transform Chronic Disease Management at Duke Health

Mark Sendak, MD, MPP

In the summer of 2014, Dr. Blake Cameron was a nephrology and informatics fellow with bold ideas about leveraging data and technology to improve healthcare. He sought out universal problems hiding in plain sight.

The idea at that time was to leverage existing patient data to detect and intervene on progressing chronic kidney disease (CKD) to prevent dialysis “crash starts” - when patients are put on dialysis emergently. These patients, at higher risk of poor outcomes, were frequently unaware of their progressing CKD and not receiving adequate treatment.

Individuals from DIHI, Duke Connected Care, the Department of Statistics, and the Department of Medicine, came together to prepare a proposal for the 2015 DIHI Request for Applications (RFA). Starting with this RFA, Blake and his colleagues from Duke Primary Care (DPC) and Duke Population Health Management Office (PHMO) would transform chronic disease management at Duke.

Scaling innovation within the CKD vertical

Between 2015 and 2020, DPC and PHMO completed a series of three DIHI RFA projects to diffuse chronic disease management support to primary care providers (PCPs) and patients.

POPULATION HEALTH ROUNDS

The first project, awarded in 2015, created risk-based population health rounds. A predictive model ran monthly on patient data to identify patients at high risk of progression to end-stage renal disease (ESRD). An interdisciplinary clinical team consisting of a specialist, PCP, pharmacist, and care manager would meet weekly to review the high-risk patients for CKD concerns and broader social drivers of care. Their recommendations were sent to PCPs.

Blake emphasized that “if we were able to prevent one adverse outcome of a dialysis crash start, we would cover the cost of the program for a full year.”

Population health rounds began in 2015 with an eligible population of about 50,000 Medicare patients enrolled in Duke’s Medicare Shared Savings Program (MSSP). By 2024, the program expanded to cover all patients participating in value-based care insurance programs, which now includes about 400,000 patients with Medicare, Medicaid, and private insurance.

Dr. Cameron appreciates how the innovation pilot allowed us to demonstrate that not only is this a good idea, but it’s a better way to provide care and tackle many of the structural problems that our health system faced in terms of access and quality.



E-CONSULTS

The second innovation project was e-Consults, which started in 2016, during Blake’s first year on faculty at Duke. Blake worked closely with Daniel Costello, Chief Operating Officer of the PHMO, to develop the e-Consult program, which allowed PCPs to request remote assistance from specialists.

Before e-Consults, PCPs often referred patients to specialists as soon as they felt uncomfortable managing a complex condition independently. Patients would see the specialist once, and the specialist would send the patient back to the PCP with specific treatment recommendations. The back-and-forth between PCPs and specialists created frustration through delays in care and lack of coordination on both sides.

Former DPC Chief Medical Officer, Dr. John Anderson, sent Blake the first ever e-Consult for a patient who came in for pre-travel vaccinations and lab tests. When the patient’s results showed reduced kidney function and elevated urine protein, John placed an e-Consult. Hours later, Blake pulled up the patient’s medical record. Seeing the patient’s complaint of lower back pain during a recent visit, Blake called John and walked through potential causes for the lab abnormalities and back pain, including multiple myeloma, a form of blood cancer.

John and Blake agreed on a set of follow-up tests to order immediately. Within two days, lab tests were consistent with multiple myeloma, and the patient canceled his trip. Three days later, multiple myeloma was confirmed via a bone marrow biopsy, and the patient was started on chemotherapy a few days later. Within one week, a primary care doctor and specialist worked together to confirm a cancer diagnosis and initiate treatment. Without an e-Consult, the traditional referral process could have taken months.

In 2016, Blake completed about 40 nephrology e-Consults per month. By 2024, specialists completed over 2,000 e-Consults per month. Since more than half of patients who received an e-Consult did not need to see a specialist in person, Dr. Ed Cooner, Chief Patient Safety Officer for Duke Primary Care explained, e-Consults “allowed us to free up access for patients who did need to be seen.”

Starting in September 2024, Duke Health began billing for e-Consults, enabled by payment policy reforms during the pandemic. The team expects to bring in several hundred thousand dollars annually through direct billing and incremental downstream appointments. Ed said, “The bigger the program got, the more work it put on specialists. Without some model for reimbursement and billing, we could not [continue to] grow the way we have been.”

STUDENT EXPERIENCES WITH DIHI



Henry Foote, MD

I work clinically as a pediatric cardiology fellow and will complete an additional fellowship in pediatric critical care to work clinically in the pediatric cardiac intensive care unit. Here, I will care for critically ill children with heart disease and hope to improve their care with electronic health record (EHR)-based models.

My two-year tenure at the Duke Institute for Health Innovation (DIHI) has been a journey of continuous learning and growth. In the past year, as an innovation scholar, I have had the privilege of contributing to a diverse range of projects, from predicting mortality in adults with cardiogenic shock to identifying risk factors for deterioration in children with single ventricle heart disease. This role has been both challenging and rewarding.

My time at DIHI has exposed me to the intricate and gritty teamwork involved in developing and implementing models using EHR data. With the guidance and support of the DIHI team, I have gained hands-on experience in extracting, cleaning, feature-selecting, and engineering EHR data for model development. The automated machine learning (autoML) tool allows for the rapid cycling of models. For example, we can train multiple models predicting mortality in cardiogenic shock by predicting mortality at different time points with various patient populations.

My work at DIHI not only involved model development but also evaluation and policy. I have been exposed to various tools and am particularly excited about the potential of LLM-based solutions.

NEXT PAGE →



DIRECT-TO-PATIENT ENGAGEMENT

The third innovation was direct communication and engagement with patients. Blake describes this as “the next frontier.” Instead of requiring a predictive model to flag patients for a specialist to review or rely on a PCP to identify a kidney-related concern, why not automatically calculate CKD risk and communicate it directly with patients?

Today, CKD risk scores and education material are shared directly with patients. Patients at high risk of disease progression are directly notified and are sent educational material with information about how to reduce the risk of disease progression.

Scaling to Other Chronic Disease Verticals

Oftentimes, PHMO and DPC start by piloting care delivery innovations for a single specialty to then scale across specialties. Blake explained that the three CKD programs (population health rounds, e-Consults, and direct-to-patient engagement) “provided a framework for similar programs to be developed and scaled across disease conditions.”

In 2015, there was one specialty rounds for nephrology. By 2024, there were eight specialty rounds for nephrology, endocrinology, cardiology, behavioral health, internal medicine, peripheral artery disease, geriatric complexity, and pediatrics. In 2016, there was one e-Consult program for nephrology. By 2024, there were e-Consults across 59 specialties, sub-specialties, and disease conditions. Demand from clinicians across specialties to participate in care delivery model innovations led e-Consults to become a pillar for how Duke provides specialty care.

Strengthening the Duke Health System

Despite the narrow scope of the three separate CKD innovation pilot projects, DPC and PHMO leaders explained how the entire Duke Health system has benefitted.

First, the programs fill gaps between in-person visits. Daniel, the PHMO COO, explained, “For CKD, we are there all the time, every day. We are there monitoring lab indicators, separating signals from the noise to identify issues we need the health system to act upon.”

Second, the programs fill gaps in resource availability and awareness. Pam Cohen, Program Coordinator for PHMO Population Health Rounds, prepares the rounding list for population health rounds every week and completes follow-up referrals assigned during rounds. She knows “there are so many resources available” that patients and PCPs don’t know about. Without the added layer of support, she explains how too often “the patient may have given up. But no, we want to help you get what you need to be as healthy as you can be.” Third, the programs streamline complex care pathways at Duke Health that confuse patients and even PCPs. Ed walked through an example of a patient receiving

a new cancer diagnosis. He explained, “When a PCP gets an imaging report back that shows a new large mass, it is daunting to figure out what to do next.” They need to determine how to get a biopsy, if surgery is required, etc. With e-Consults, patients are connected to the right oncologists. Ed has heard repeatedly that this “relieved so much of the stress from new diagnoses.”

Fourth, the programs broadly improve primary care. The sustained investment in specialty support has built trust among PCPs and specialists providing remote support. Blake attests to this, “You’re much more likely to take advice from a specialist like me when I reach out to you if I’ve been helpful to you for eight years when you’ve asked for help.” PHMO also provides PCPs with actionable data about their patients. Daniel often hears from PCPs that “the more I can know about a patient’s health trajectory, the better I can serve a patient.”

Perhaps most surprisingly, PCPs who participate in the programs improve their care for entire populations. Ed explained how e-Consults build capacity for chronic disease management. The first time a PCP “used an e-Consult, they would get guidance on the appropriate steps and tests to start the evaluation and management. They would learn from that. The next time they had a patient with CKD, before they would put in an e-Consult or refer, they would do A, B, and C.”

Ed, a PCP himself, explained, “PCPs want to feel empowered to treat their patients.” The innovations have done just that.

The Value of Innovation

There are three main lessons about how innovation can have an outsized impact.

The first lesson is that even big organizations with exceptional leaders need to invest in small innovation projects.

Blake emphasized the need for investing in small projects to “prove it could be done.” He explained how quickly ideas can be dismissed within large organizations. In early conversations about e-Consults, he said, “There was a lot of skepticism about the clinical value. Don’t patients need to be seen [by the specialist]? There was skepticism around the economic value. We’re fee for service; do we really want to deflect this [revenue from the visit]?... It was suggested this could not be done inside of Epic. We were able to overcome all of those.”

After demonstrating the potential for e-Consults through a pilot project, the care model became obvious. Blake appreciates how the innovation pilot “allowed us to demonstrate that not only is this a good idea, but it’s a better way to provide care and tackle many of the structural problems that our health system faced in terms of access and quality.”

The second lesson is to be more confident about the big impact of small projects.

Suresh Balu, the DIHI team director, often describes the DIHI approach as ‘slingshots’ rather than ‘moonshots.’ The US spent 3% of its national GDP to send astronauts to the moon. Most health systems in the US don’t operate with a greater than 3% margin; for Duke Health specifically, 3% of total revenue is \$90M. While moonshots are out of reach, small, targeted investments of \$30K-\$75K are feasible.

Blake understood how to make the most of small projects, explaining that “while the DIHI RFAs are individually narrow in scope, when you stitch these things together in a cohesive, programmatic way, you accomplish a lot more.” He encourages other aspiring innovators “to think cohesively about technologies...How we can link together all these things ranging from predictive analytics to professional oversight to e-Consults to self-scheduling to virtual care to patient-facing education.”

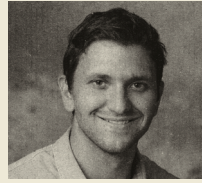
The third lesson is that to fully realize the benefits of innovation, you must invest in people.

Beyond validating and de-risking investments, DIHI projects also facilitate investments in people who become future leaders. Blake is confident that training at a place like Duke set him up for a successful career in nephrology. But medical training alone couldn’t have set him up to have such a broad impact through his PHMO roles. “I don’t know what my path would have looked like because it was creating roles that didn’t exist,” said Blake. Thankfully, Blake continues to receive support to create brand-new ways of delivering care for the better.

With such strong leadership, some may question the need for a team like DIHI to conduct short-term innovation projects. Why doesn’t Duke Health just encourage business units to innovate themselves?

**“Could we have done it alone?”
asked Daniel, COO of PHMO.
“Not at the speed, breadth,
depth, or quality with what
DIHI brought to the table.”**

STUDENT EXPERIENCES WITH DIHI



**Bruno
Valan, BS**

My time as a scholar at the Duke Institute for Health Innovation (DIHI) has been an enriching experience. On the lower back pain project, I worked with clinicians, data scientists, and clinical researchers to build an algorithm that predicted the correct treatment for different patients with chronic lower back pain—whether it be physical therapy, surgery, or physiatry—based on their baseline health characteristics. This endeavor allowed me to apply my background in machine learning, which had primarily focused on generic commercial applications, to the healthcare sector. It was a unique and challenging experience that provided a valuable bridge between my clinical background as a medical student and my expertise in machine learning.

One of the main takeaways from my time at DIHI was the crucial importance of understanding the clinical context in which a model is deployed. Regardless of how well a model performs, its impact on patient care and outcomes is ultimately determined by how healthcare professionals use its results. This insight deepened my appreciation for the intersection of data science and healthcare, emphasizing that innovators must couple technological advancements with practical, clinically-informed implementation strategies. DIHI’s collaborative approach, bringing together diverse perspectives from medicine, data science, and technology, is crucial for developing and implementing AI solutions that can be seamlessly integrated into clinical practice and improve patient outcomes.

DIHI is a remarkable organization with a diverse team that combines entrepreneurship, clinical expertise, and engineering. The collaborative environment fostered here is truly one of a kind, and it has been an honor to be a part of this innovative team for a year. The experiences and skills I have gained at DIHI will undoubtedly have a lasting impact on my career, equipping me to continue bridging the gap between machine learning and healthcare to improve patient outcomes.



LEGACY INNOVATIONS

Duke TeleHOPE Program Demonstrates Sustained Success in Reducing Readmissions and Costs

Aubrey Jolly Graham, MD; Rachel Hughes, MD; Geraldine Kanne, NP; Cindy Roberson, Pharm D;
Will Knechtle, MBA, MPH; Will Ratliff, MBA; Julia Bellantoni, MD

The Duke TeleHOPE program, a multidisciplinary initiative to improve care transitions between hospitals and skilled nursing facilities (SNFs), has successfully reduced readmission rates and reduced SNF costs over five years. This innovative program leverages weekly teleconferences to facilitate communication and collaboration between hospital and SNF care teams.

Problem

The transition from hospital to SNF is a critical period for patients, often marked by complexity and a high risk of readmissions. In 2018, 20% of Medicare beneficiaries were discharged to SNFs, incurring a cost of \$28.5 billion for Medicare. The national 30-day rehospitalization rate in 2015 was 19%, with studies indicating that 31% to 67% of these readmissions were potentially preventable. Recognizing the need to address this challenge, the Duke University Health System (DUHS), which discharges approximately 5,200 patients to SNFs annually, launched the TeleHOPE program in 2019.

Solution

The Duke TeleHOPE program was initially funded by the Duke Institute for Health Innovation (DIHI) as a pilot project. Following the pilot's success, the program received sustained funding and became integral to the Duke Population Health Management Office (PHMO). The program focuses on patients discharged from Duke University Hospital (DUH) and Duke Regional Hospital (DRH) to four partner SNFs. The core components of the program include:

- **Weekly Multidisciplinary Teleconferences:** These conferences bring together key hospital and SNF stakeholders to discuss patient care transitions. Duke's TeleHOPE team participants include a hospitalist, a geriatric nurse practitioner, a pharmacist, and learners from various healthcare disciplines. SNF representatives often include a nurse, a social worker, or an advanced practice provider.

- **Pre-Conference Patient Review:** The Duke TeleHOPE team reviews each patient's case before the conference.
- **Collaborative Care Planning:** During the teleconference, the team discusses the patient's needs, identifies potential safety gaps in the transition process, and develops recommendations for ongoing care at the SNF.
- **Documentation, Communication, and Analysis:** Clinical recommendations and medication reconciliation notes are documented in the patient's medical record and transmitted to the SNF care team. The program tracks patient outcomes, including readmission rates, length of stay at the SNF, and associated costs. The data is compared to similar non-participating SNFs to assess the program's impact.

Patients in the TeleHOPE SNFs had a significantly shorter average length of stay (26.45 days) than their counterparts in the non-TeleHOPE cohort (30.18 days).



Impact and Outcomes

Since its inception, the Duke TeleHOPE program has facilitated 1,815 weekly teleconferences, discussing the care transitions of 1,429 patients. The majority of patients were discharged from DUH (57%), with 30.3% identified as high-risk for readmission. Conferences typically occurred an average of 6.87 days after hospital discharge. Importantly, patients in the TeleHOPE partner SNFs were demographically similar to those in the comparison SNF cohort. Over the five-year evaluation period, the program demonstrated the following key outcomes:

- **Sustained Reduction in Readmission Rates:** Patients discharged to TeleHOPE partner SNFs consistently experienced lower readmission rates than those in the non-participating SNF cohort. This positive trend was observed across each 12-month interval throughout the study period.
- **Shorter Length of Stay in SNFs:** Patients in the TeleHOPE SNFs had a significantly shorter average length of stay (26.45 days) than their counterparts in the non-TeleHOPE cohort (30.18 days). This finding suggests that the program's interventions may contribute to more efficient care transitions and potentially quicker patient recovery.
- **Lower SNF Costs:** The program reduced mean SNF payments for patients in the TeleHOPE facilities (\$12,694) compared to the non-TeleHOPE group (\$13,178). This cost reduction highlights the potential financial benefits of improved care coordination and reduced readmissions.

Conclusion

The Duke TeleHOPE program provides a compelling model for improving the quality and efficiency of care transitions between hospitals and SNFs. By fostering collaboration between care teams and proactively addressing potential care gaps, the program has successfully reduced readmissions, shortened SNF lengths of stay, and lowered costs. The program's impact underscores the value of multidisciplinary teamwork and communication in optimizing patient care during this vulnerable transition period. The program's ability to identify systemic issues and initiate quality improvement efforts further strengthens its contribution to enhancing patient care. The financial sustainability of the program, supported by cost savings and alignment with Medicare's shared savings goals, ensures its continued impact on patient outcomes.

STUDENT EXPERIENCES WITH DIHI



Maryam Aziz, MS

I joined the Duke Institute for Health Innovation (DIHI) as an Artificial Intelligence Research Intern in November 2023, during the first year of my Ph.D. I enrolled in my Ph.D. program to learn about and contribute to machine learning projects in healthcare, and I'm grateful that DIHI has provided the perfect opportunity for me to do that.

DIHI involved me in two exciting projects. The first was developing an HIV risk prediction model using clinical data from patients in two large health systems in Southeastern Louisiana. My responsibilities included automating the preprocessing and filtering of data, as well as data cleaning and exploratory analysis of patient-level data. The second project was equally exciting. It focused on using prompt engineering and large language models (LLMs) to identify HIV patients who have fallen out of care through patient notes and messages. Joining early in the model development process has allowed me to understand the foundational steps in building a model. My tasks included preprocessing data and creating prompts to complete tasks effectively. While maintaining an intrapreneurial spirit, DIHI taught me to pay attention to the grit within datasets and grease the cogs of data processing.

My time at DIHI has been a rich learning experience, providing me with valuable insights into the application and implementation of machine learning in healthcare. These insights have not only broadened my understanding but also equipped me with the knowledge to make a meaningful impact in the field. It has also been a platform for me to connect with the brilliant minds at DIHI. The skills I've gained and the connections I've made here have been invaluable, and I'm excited about the prospect of continuing my work with DIHI.

LEGACY INNOVATIONS

Summing up SymMon

Mike Revoir, BS

The Duke COVID-19 Symptom Monitoring App, known as SymMon, was a cornerstone of Duke University’s pandemic response. The app’s development, spearheaded by the Duke Institute for Health Innovation (DIHI) and the Duke Office of Information Technology (OIT), played a pivotal role in the successful reopening of campus. The app was launched for iOS and Android phones in July 2020 to facilitate the safe return of student-athletes to campus, roughly three weeks after the project team first met to conceive the solution. It quickly ramped up to over 16,000 users in the next two weeks as faculty, students, and staff returned to campus for the Fall 2020 semester. The app included several key features that made it highly effective in managing the health and safety of the Duke community during the pandemic:

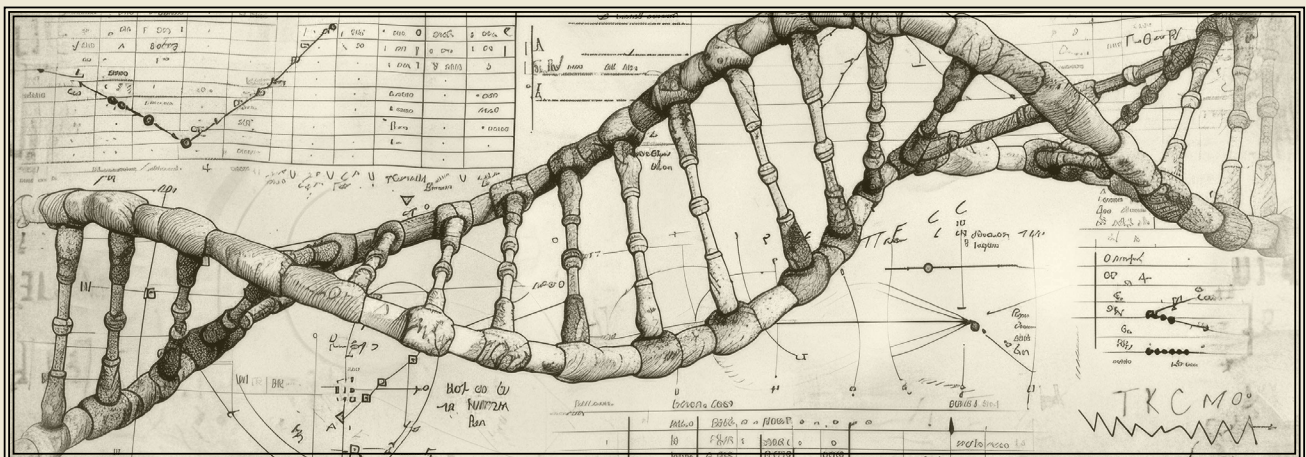
1. Health Checks: Users could complete daily health checks, which included symptom reporting and temperature logging. This feature helped in early detection of potential COVID-19 cases. Usage scaled from 11,000+ health checks in July 2020 as student-athletes returned to 195,700+ in August and 291,000+ in September as the Fall 2020 Semester began.

2. Campus Check-Ins: The app facilitated campus check-ins, allowing the University to monitor and control access to various campus facilities. This was crucial in maintaining social distancing and managing crowd control.

3. COVID-19 Test Kit Scans: Users could scan their COVID-19 test kits directly through the app, streamlining the sample collection process and reporting test results, and ensuring timely updates to the health monitoring system. Including this feature was instrumental in making it possible for Duke to test well north of 12,000 students, faculty, and staff on a weekly basis from the Fall 2020 Semester through the Spring 2022 Semester.

4. User Engagement: The app saw significant user engagement, with over 43,600 unique users, 4 million health checks, 693,000 campus check-ins completed, and 993,000 COVID-19 test kits scanned. This high level of participation was essential for the University’s success in tracking and managing the spread of COVID-19 on campus from July 2020 through April 2022.

It is clear that the success of SymMon was not just in its technical achievements but also in the collaborative spirit that drove its development. Despite the challenging circumstances, the dedication and teamwork of everyone involved made this one of DIHI’s most successful projects. The goal of promoting the health of Duke students, staff, and faculty as well as our community at large continues to drive our efforts today, and collaboration between healthcare professionals and technology will be a critical component of addressing the challenges that lie ahead.



LEGACY INNOVATIONS

Autism & Beyond

Jamie Daniel, BS

'Autism & Beyond' was a study of young children's mental health that sought to better understand and identify risks for development. The study aimed to test video technology to analyze a child's emotions and behavior. This technology was adopted to screen young children in their homes for autism and mental health challenges, such as anxiety or tantrums. Early intervention is crucial for improving outcomes in children with autism.

The app, which leveraged Apple's ResearchKit, used the front-facing camera on an iPhone to record a child's reactions to stimulus videos and used facial recognition software to automatically locate key features of the child's face so that we could measure emotions and head position. Most of our innovation came in the video portion of the app. As the child sat on the parent's lap watching four short video clips, the device's user-facing 'selfie' camera recorded the child's reactions to the videos, which were designed to make them smile, laugh and be surprised. Analytical software in the app automatically marked key landmarks on the child's face and assessed their emotional responses. For example, in a short video of bubbles floating across the screen, the video algorithm looked for movements of the face that would indicate joy. We made an accessible tool that combined this video analysis with patient-reported questionnaires. It is important to note that all data collected was anonymized and stored securely, adhering to the highest standards of privacy and data security.

Beginning in October 2015, 'Autism & Beyond' was offered from the Apple App Store for use on iOS devices. The development of the app was made possible with the generous support of the National Center for Advancing Translational Sciences (NCATS) of the National Institutes of Health.

As a developer on the Autism & Beyond project, I was part of an interdisciplinary team from across the Duke community that had expertise in pediatrics, psychiatry, engineering, and data science. As someone with a family that autism has touched, I knew how important this application was. As a developer, I wanted to ensure this was easy to use and robust enough that parents and providers alike would find it beneficial. The validation of our project through the gathering of large amounts of data from across the country reassured us of its credibility and the need for such a tool in the field of autism research.

I found this project to be a fantastic experience. Not only did we increase the tools needed to advance the diagnosis and treatment of autism, but we also advanced our knowledge of how to collaborate across a large organization and leverage experience and expertise from groups that may not have had the opportunity to work together previously. Winning the Duke President's Award for our efforts was a testament to the impact and success of our project. It showed that a solution was possible, collaboration is possible, and as always in the Duke community, when we come together, we create incredible things that benefit the world.



LEGACY INNOVATIONS

Voices of Duke Health Podcast

Freya Gulamali

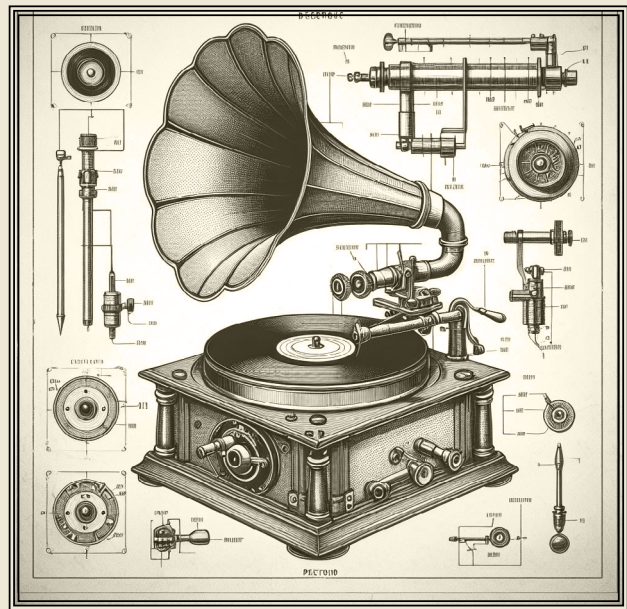
The Department of Medicine and the Duke Health Office for Patient Safety and Clinical Quality developed the Voices of Duke Health initiative with support from DIHI's 2018 Innovation Grant. We designed this podcast as a "listening booth" with the mission of giving Duke Health providers, students, and patients a new platform to be vulnerable and share their stories. This mission to bridge understanding amongst our listeners provided us with a different way to explore how we could better support each other. In 2019, Voices of Duke Health won the ABIM Foundation Trust Practice Challenge, an initiative to build or rebuild trust in healthcare systems – a testament to the success of the project. It was also recognized by the AAMC Group on Institutional Advancement (GIA) during the Best Practices in Digital Storytelling from Around the GIA session.

The Voices of Duke Health group set up a recording booth within the Duke Medicine Pavilion Patient Resource Center waiting room and ventured into various departments with a mobile cart to meet people where they were. Within the first year of the project launch, the Voices of Duke Health recorded 150 voices for 15 podcast episodes. The total number of listens before April 25, 2019 surpassed seven thousand. One of the most impactful episodes in the first season was "Episode 10: The Gauntlet". In this episode, Dr. Anthony Galanos spoke about the death of his son, Nicholas, and coping with grief at home and work.

Dr. Galanos discussed his thoughts about how to support someone who has lost a family member by being present and "do[ing] what's good for the person as opposed to what's good for you." Beyond the impact that this podcast has had on listeners, participants in the listening booth reported feeling more energized, centered, or empowered even when they were nervous before entering the listening booth.

This project was, in part, a response to the 2016 Duke Health Advancing Health Together strategy, which aimed to work towards Duke being a "place where everyone thrives and is valued." When annual workforce surveys indicated that Duke Health employees were facing higher levels of stress and burnout, we learned from the past and incorporated storytelling.

Historically, the emphasis on storytelling in medicine has roots dating back to significant developments in the early 20th century, when Duke University was founded. The Flexner Report of 1910 laid the groundwork for reforming medical education by highlighting the need for understanding patients beyond their diseases. During the 1920s and '30s, William Carlos Williams began integrating his medical experiences into his literary works. This culminated in his publication of *The Doctor Stories* in 1934, which explored doctor-patient relationships through autobiographical essays. His contributions continued through the 1940s and '50s, influencing later movements in narrative medicine that gained formal recognition in the late 1990s, with advocates like Rita Charon emphasizing storytelling's importance in improving patient care.



As Voices of Duke Health exhibits, storytelling remains a vital component of healthcare practices today, echoing sentiments from past pioneers who recognized its significance in fostering empathy, understanding, and peace within medical contexts.

Evaluating and Scaling Innovation

Sepsis Watch 2.0 Expansion into Duke Regional Hospital

Will Ratliff, MBA

What is the solution?

Every year, roughly 1.7 million American adults develop sepsis, with 270,000 (16%) dying as a result of the disease. To continue addressing this challenge at Duke, we collaborated with clinical, technical, and operational colleagues to bring Sepsis Watch to the bedside beginning in November 2018 at Duke University Hospital and then in June 2019 at Duke Raleigh Hospital. Fast forward to 2024: Sepsis Watch 2.0 combines a real-time sepsis detection phenotype (i.e., the Centers for Medicare & Medicaid Services [CMS] real-time sepsis definition), an “imminent sepsis” prediction machine learning model, a custom-built web application and/or paging notification system (depending on workflow needs), with sepsis-treatment-bundle tracking functionality (real-time and retrospective bundle tracking support).

How was it implemented?

There are two main delivery mechanisms for Sepsis Watch: (1) the Sepsis Watch 2.0 Web Application, which features four columns of patient “cards” that can be re-arranged easily depending on status (Triage/Screened/Monitoring/Sepsis Bundle) and live bundle status tracking, and (2) the Sepsis Watch 2.0 secure paging system, which sends a “high risk” or “met sepsis” page (including how the phenotype was met) to a Spok or pager number. Then, as soon as any patient with sepsis is discharged, we automatically compile and post their sepsis definition and bundle compliance details to our new set of hospital-specific sepsis-compliance-tracking Tableau dashboards. We will use these Tableau dashboards to understand performance and drive improvement by unit, provider, bundle component, etc. In total, this represents the Sepsis Watch 2.0 program.

When and where did it go live?

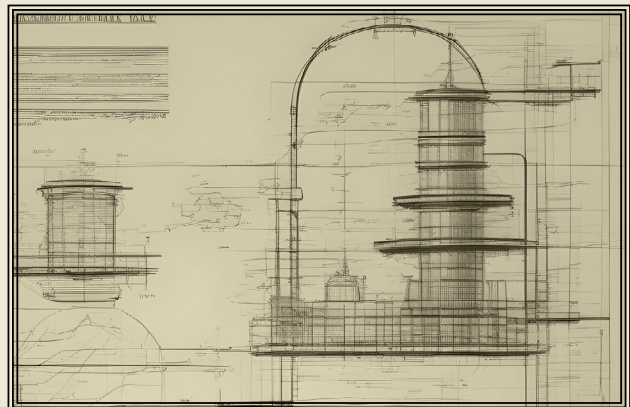
Sepsis Watch went live on June 24th, 2024, at Duke Regional Hospital’s (DRH) Emergency Department (ED) and four intermediate/step-down units for General Medicine, Telemetry, and Neurology/Oncology specialties. See Figures 1a and 1b for the workflows currently used in

the ED and inpatient settings. We established a Sepsis Governance Committee to oversee the go-live and ongoing implementation effort. We regularly meet with this Committee to report progress and respond to collected feedback.

So what is the impact?

In the first 60 days since going live, there were 614 total pages sent on 436 unique patient encounters, with 387 (63%) of these pages occurring in the ED. In the ED, clinical teams responded to 6.5 pages per day (5.2 high risk, 1.3 met sepsis) compared to 3.8 pages per day (1.3 high risk, 2.5 met sepsis) in the inpatient units.

Among the patients who triggered a page in the ED, 305 (81%) were admitted and 30 (8%) died in the hospital. 10 of the 72 patients (13.9%) who triggered a page in the inpatient units died in the hospital. Tables 1a and 1b show the diagnoses associated with these encounters. We are currently assessing the pre vs. post-go-live impact on bundle compliance, length of stay, ICU requirement, ICU length of stay, and in-hospital mortality. With the guidance of our clinical leads at DRH, we will continue to optimize the workflow and study its effectiveness in improving patient care.



NEXT PAGE →

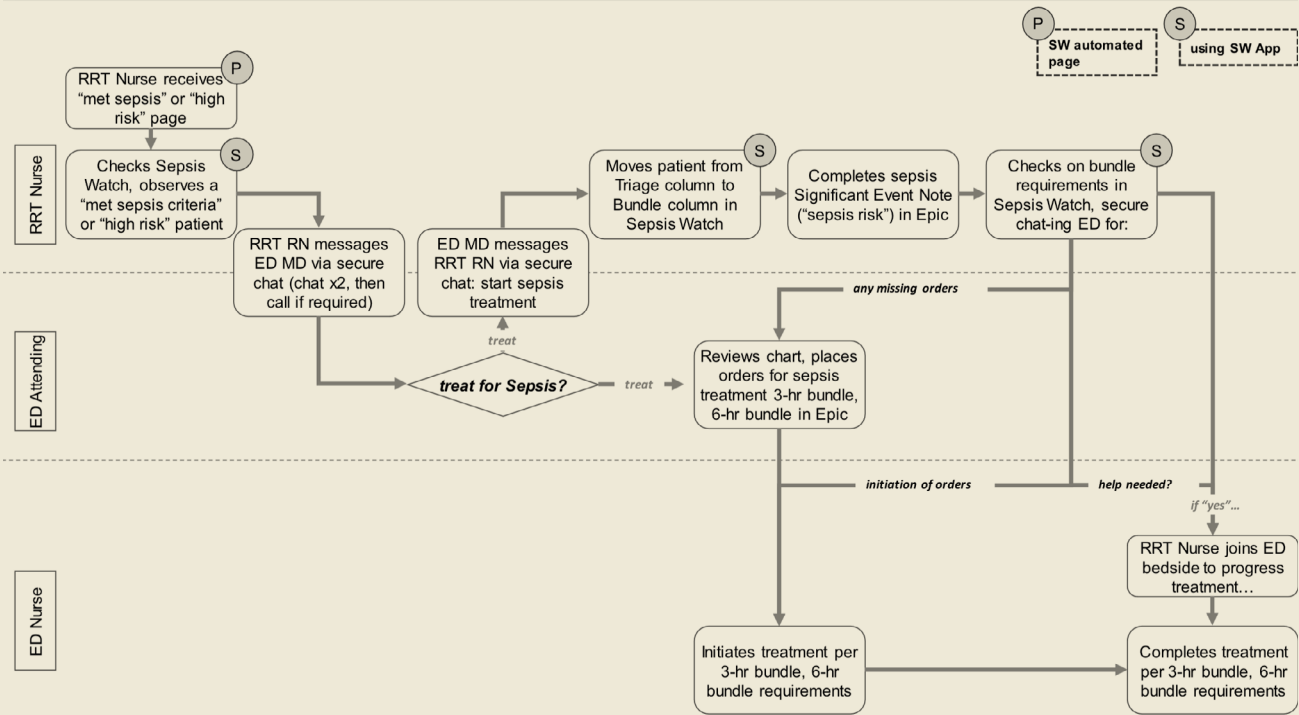


Figure 1a. Emergency workflow

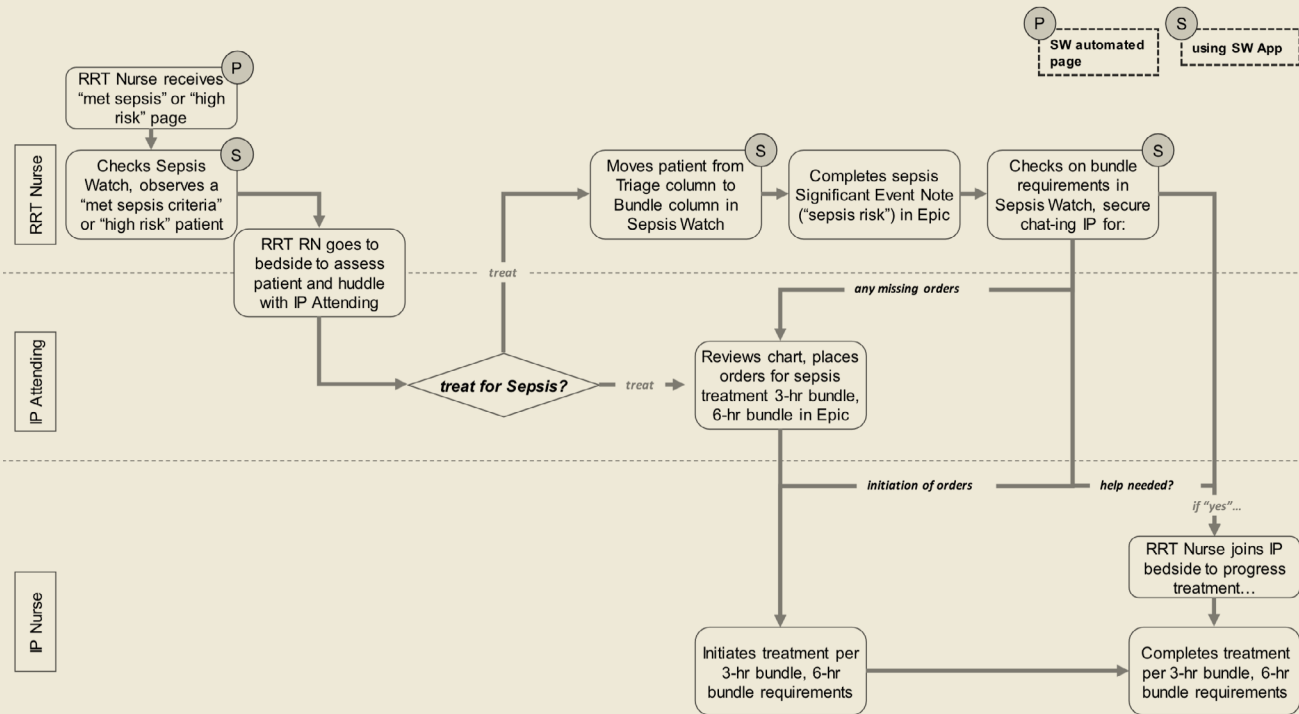


Figure 1b. Inpatient units workflow

RANK	CCSR CATEGORY	UNIQUE PATIENTS WITH CCSR	DELTA OF NEW PATIENTS ADDED BY CCSR	CUMULATIVE % OF 376 PATIENTS
1	Fluid and electrolyte disorders	182	182	48.53
2	Disorders of lipid metabolism	139	49	61.6
3	Septicemia	131	27	68.8
4	Abnormal findings without diagnosis	118	24	75.2
5	Respiratory signs and symptoms	117	25	81.87
6	Essential hypertension	113	5	83.2
7	Other specified and unspecified nutritional and metabolic disorders	111	0	83.2
8	Diabetes mellitus with complication	102	2	83.73
9	Acute and unspecified renal failure	102	3	84.53
10	Abdominal pain and other digestive/abdomen signs and symptoms	97	17	89.07
11	Chronic kidney disease	95	0	89.07
12	Circulatory signs and symptoms	79	9	91.47
13	Urinary tract infections	79	5	92.8
14	Heart failure	77	0	92.8
15	Fever	76	5	94.13
16	Nervous system signs and symptoms	75	5	95.47
17	Cardiac dysrhythmias	71	0	95.47
18	Esophageal disorders	70	0	95.47
19	Respiratory failure; insufficiency; arrest	69	0	95.47
20	Pneumonia (except that caused by tuberculosis)	63	1	95.73

Table 1a. Diagnoses of Emergency Department patients who triggered a page

RANK	CCSR CATEGORY	UNIQUE PATIENTS WITH CCSR	DELTA OF NEW PATIENTS ADDED BY CCSR	CUMULATIVE % OF 72 PATIENTS
1	Fluid and electrolyte disorders	42	42	58.33
2	Acute and unspecified renal failure	31	13	76.39
3	Respiratory signs and symptoms	30	3	80.56
4	Septicemia	29	1	81.94
5	Other specified and unspecified nutritional and metabolic disorders	28	2	84.72
6	Abdominal pain and other digestive/abdomen signs and symptoms	28	4	90.28
7	Abnormal findings without diagnosis	26	0	90.28
8	Disorders of lipid metabolism	25	0	90.28
9	Urinary tract infections	24	1	91.67
10	Essential hypertension	21	0	91.67
11	Aplastic anemia	20	0	91.67
12	Chronic kidney disease	20	1	93.06
13	Fever	20	2	95.83
14	Heart failure	19	0	95.83
15	Nervous system signs and symptoms	19	0	95.83
16	Malnutrition	18	0	95.83
17	Circulatory signs and symptoms	18	0	95.83
18	Cardiac dysrhythmias	17	1	97.22
19	Other nervous system disorders (neither hereditary nor degenerative)	17	0	97.22
20	Esophageal disorders	16	0	97.22

Table 1b. Diagnoses of Inpatient unit patients who triggered a page



EVALUATING AND SCALING INNOVATION

Diffusion of Capabilities: Launching a Practice Network

Alifia Hasan, B Pharm, MBA

The Health AI Partnership (HAIP), a collaborative effort representing all healthcare organizations that the Duke Institute of Health Innovation coordinates, has recently launched a Practice Network. This initiative aims to empower healthcare professionals to use AI effectively, safely, and equitably through community-informed, up-to-date standards. The Practice Network is a testament to HAIP’s commitment to support both high- and low-resource environments and to bridge the gap between AI technology and healthcare, especially for smaller and rural hospitals.

The Practice Network is a network of healthcare organizations interested in adopting AI solutions through implementing AI best practices in collaboration with the HAIP Corps Sites and Coordinating Center. The Practice Network Sites will contribute to the HAIP community by implementing HAIP best practices as they implement AI-based tools in their settings and by providing feedback and highlighting gaps. They will also share learnings and knowledge, advocate for sustainability, and potentially become leaders in AI adoption.

HAIP recently selected five sites to be the inaugural cohort of the Practice Network program (Figure 1). The chosen sites bring a variety of proposed use cases; will explore the use of large language models,

clinical decision support, and imaging AI solutions; and represent both rural and urban populations, including FQHCs and larger hospital systems, and range in volume from ~45,000 to 200,000+ outpatient visits per year.

Over the next year (August 2024 to August 2025), we will work with selected sites to build organizational capabilities for the safe, effective, and equitable implementation of AI. The program is free to participate in, and participating sites will get immediate access to a network of AI experts and healthcare leaders through one-on-one guidance and expert consultations. They will also share insights and learn from each other as a peer learning community and receive educational materials and other relevant resources to stay up-to-date with real-time learning about emerging policies, research, and industry trends.

The Practice Network is not only focused on closing the resource gap between large and small healthcare providers but also on establishing best practices for testing and monitoring AI in local settings. This initiative is a significant step towards making AI in medicine accessible to smaller and more rural hospitals, and it holds the promise of transforming healthcare delivery across the United States.

	COMMUNITY SERVED	USE CASE PROPOSED	SOLUTION PROPOSED	ORGANIZATION TYPE
Community-University Health Care Center	49% prefer care in another language, 54% have Medicaid/CHIP, and 91% live at or less than 200% of the Federal Poverty	Reduce patient no-show rate	EPIC’s Risk of Patient	Federally Qualified Health Center
Health Center Of Southeast Texas	Clinics located in underserved areas primarily cater to low-income individuals and those without insurance	Reduce documentation load and time spent by providers on admin	Sunoh AI	Federally Qualified Health Center
North Country Healthcare	7+ specialties available to 55,000 unique patients; 91% of communities served are Medically Underserved Areas	Reduce clinical documentation time for providers	Abridge, Clinically Speaking, ChartNote, etc.	Federally Qualified Health Center
San Ysidro Health Center	Majority of patient population faces high rates of poverty, is uninsured/underinsured, and has low educational attainment	Improve compliance for annual diabetic retinopathy screenings	EyeArt AI	Federally Qualified Health Center
WakeMed	Close to 100 clinics serving ~2 million residents in the Research Triangle areas	Reduce sepsis mortality rates by increasing early detectio	Epic’s Cognitive Computing Sepsis Version 2	Hospital System

Figure 1. Sites in the Inaugural Practice Network Cohort



EVALUATING AND SCALING INNOVATION

Pythion: New Models Using Narratives

Freya Gulamali, Will Knechtle, MBA, MPH

Introduction

The Pythion initiative aims to enhance preoperative planning and patient outcomes through advanced predictive models. “Pythion” is a graphical user interface that, when a user enters any patient identifier, presents the patient’s postoperative complication and mortality risk using their available electronic health record data. It is unusual among postoperative prediction products because it uses only features typically available when a surgeon posts a case. Therefore, it aims to provide an honest accounting of risk-knowledge weeks before surgery. Notably, the only user-entry requirement is a medical record number and planned procedure. In contrast, its most-used competitor, the American College of Surgery’s National Surgical Quality Improvement Program (NSQIP) model, requires more than five manual entries.

The Pythion initiative aims to enhance preoperative planning and patient outcomes through advanced predictive models.

“Healthcare remains largely a reactive practice, and the current predictive tools are limited in time horizon, ungeneralizable cohorts, and the ability to integrate into clinical workflows. Pythion breaks the glass ceiling on these issues by providing unique upstream predictive capabilities that allow for preoperative optimization, resource allocation, and planning to impact measurements that clinically matter the most. Critically, the model framework allows for the integration of novel features that can scale to a host of additional outcomes important to both patients and healthcare systems.” – Dr. Michael Kent, Associate Vice Chair, Perioperative Informatics, Dept. of Anesthesiology. Pythion’s newest predictive models have scaled to support critical outcome areas such as venous thromboembolism (VTE) prevention, postoperative intensive care unit (ICU) needs, and discharge planning. By leveraging narrative data over traditional ICD-10 codes, Pythion is setting new clinical validity data standards.

New Models

Venous Thromboembolism (VTE) Prediction

One of the most significant advancements in the Pythion project is the development of a new model for predicting VTE using narrative data from ultrasound and CT reports. Traditional models rely heavily on ICD-10 codes, which have shown variable sensitivity and specificity. Internal chart reviews have also indicated significant delays between the occurrence of VTE and the assignment of diagnosis codes. The new narrative-based model, however, offers significantly higher sensitivity and specificity. This model provides a timelier and clinically valid outcome by identifying VTE from the same sources that inform clinical judgment. The performance metrics of the narrative-based model are comparable to those of the ICD-10 model, but the former’s clinical applicability makes it a game-changer in VTE prediction.

Postoperative ICU Needs

Another critical area where Pythion is making strides is in predicting postoperative ICU needs. Early recognition of patients requiring acute care after surgery can significantly improve patient outcomes. By combining this predictive capability with other outcomes, the model offers a more in-depth understanding of a patient’s needs, enabling healthcare providers to allocate resources more efficiently and effectively.

Discharge Planning

Effective discharge planning is crucial for reducing hospital stays and minimizing post-discharge complications. Pythion’s model for early discharge planning ensures that transportation, caregivers, family support, and other resources are available during discharge. This proactive approach reduces hospital costs and lowers the likelihood of post-discharge complications, improving overall patient care.

Oversight And Evaluation

The Pythion project is subject to rigorous oversight and evaluation to ensure its robustness and fairness. Multiple subgroups have studied the models, and we presented a comprehensive bias and mitigation plan. The Duke Institute for Health Innovation (DIHI) has recruited surgeons

NEXT PAGE 



and anesthesiologists to practice using the application and validate the models. This evaluation is “silent”; providers test the model in real-time without taking clinical action based on the model’s predictions. The project is progressing towards reviewing over 100 cases, ensuring the models are statistically sound and practically applicable.

VTE poses a significant risk, particularly post-operatively, due to immobility. The risk after major orthopedic surgery is 4.3% , with additional hospital-acquired VTE costs estimated at \$17,367 per patient. Traditional methods of determining VTE rates using ICD-10 codes have shown variable sensitivity and specificity, with significant delays in diagnosis.

The Python project addresses these issues by developing a language-based definition for VTE using ultrasound and CT reports. This method aims to create a more accurate ground truth for VTE, increasing the model’s clinical applicability and utility.

The model was trained on a cohort of 364,528 surgical cases from 238,058 patients, using a LightGBM ensemble of decision trees. The final language-based definition achieved a positive predictive value (PPV) of 95% and a negative predictive value (NPV) of 99% for acute VTE. The model’s performance metrics, including AUROC and AUPRC, were comparable to those of the ICD-10 model but with a slight increase in identified VTE, indicating its superior detection capabilities.

Upcoming Interests

DIHI is not stopping at VTE prediction. Future interests include developing models for acute kidney injury (AKI) and using the VTE phenotype to monitor hospital-acquired VTE separately from Python.

In conclusion, DIHI’s Python project is setting new benchmarks in predictive healthcare. By leveraging narrative data and focusing on clinically valid outcomes, “Python” is poised to transform postoperative care, ICU needs, and discharge planning, ultimately improving patient outcomes and reducing healthcare costs.

References:

1. Liu, B., Hadzi-Tosev, M., Eisa, K., Liu, Y., Lucier, K. J., Garg, A., Li, S., Xu, E., Mithoowani, S., Ikesaka, R., Heddle, N. M., Rochweg, B., & Ning, S. (2024a). Accuracy of venous thromboembolism ICD-10 codes: A systematic review and meta-analysis. *Thrombosis Update*, 14, 100154. <https://doi.org/10.1016/j.tru.2023.100154>
2. Falck-Ytter Y, Francis CW, Johanson NA, Curley C, Dahl OE, Schulman S, Ortel TL, Pauker SG, Colwell CW Jr. Prevention of VTE in orthopedic surgery patients: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest*. 2012 Feb;141(2 Suppl):e278S-e325S. doi: 10.1378/chest.11-2404. PMID: 22315265; PMCID: PMC3278063.
3. Estimating the additional hospital inpatient cost and mortality associated with hospital-acquired conditions. Agency for Healthcare Research and Quality. (2017, November). <https://www.ahrq.gov/sites/default/files/publications2/files/hac-cost-report2017.pdf>

LABEL	LANGUAGE
Positive	“deep”, “pulmonary”, “dvt”, “pe”, “clot”, “thrombus”
Negative	“no”, “negative”! “without evidence” “chronic”, “superficial”! “unchanged”

Table 1. Example Language Indicative of a Positive or Negative VTE in Ultrasound/CT Reports

VTE GROUND TRUTH	CASES	VTE PREVALENCE	AUROC	AURP
ICD-10 Model	364,528	0.7%	0.87 (0.85, 0.88)	0.20 (0.17,0.23)
Language-Based Model	364,528	0.8%	0.87 (0.85, 0.88)	0.17 (0.14,0.20)

Table 2. Performance of VTE Models Trained on ICD-10 Codes and Ultrasound/CT Reports to Determine Positive VTE

EVALUATING AND SCALING INNOVATION

Maternal Early Warning System (MEWS) Update

Will Knechtle, MBA, MPH

In 2023, we developed time-based machine learning models to predict the likelihood of sepsis within four hours, calculated every hour between admission to and discharge from a maternal and fetal medicine floor. Additionally, we trained models to predict postpartum hemorrhage at critical stages, including inpatient admission, labor, and delivery. Our commitment to integrating these machine learning models into clinical workflows led us to initiate a rigorous validation process in December 2023. This process involved data analysis, prediction accuracy assessment, user interface design, and workflow validation with active participation from Maternal and Fetal Medicine (MFM) physicians and leaders. Our close collaboration with MFM physicians was critical in refining MEWS, enabling it to consistently align with clinical judgment. This partnership aimed to build trust in the model and lead to faster, more reliable interventions, directly improving patient care outcomes.

These efforts ensured that our MEWS solution effectively presented the right patients with the appropriate data at the right time. The validation process included a silent evaluation of patient risk by physicians, followed by a comparative analysis of the model's predicted risk results against the physicians' assessments. This collaboration fostered discussions among engineers, managers, and physicians, building trust in the model as a decision-making tool due to its consistent alignment with clinical judgments and its enhanced decision-making speed.

To communicate these warnings effectively, we employed the 'Analytic and Logic Driven Intimation System' (ALDIS), as described in DIHI Impact Volume 22 on [page 20](#). We conducted interviews with clinicians to identify, test, and refine best practices for integrating alerts and monitoring within hospital operations.

In 2024, we began a prospective real-time review of MEWS alerts. We received valuable feedback from leaders within and outside the Labor and Delivery (L&D) floors, leading to improvements in our definitions and measurements of hemorrhage and sepsis at Duke Health. In alignment with DIHI's emphasis on robust clinical metadata management, as exemplified by the Poseidon tool, the MEWS development team rigorously curated and validated data inputs, ensuring the accuracy and reliability of the system's

predictions. Committees also validated model performance across various dimensions, including time, units, race, ethnicity, and payer type. Furthermore, we developed a plan for addressing potential biases or errors.

This summer, we responded to feedback from MFM colleagues by condensing the alert format. We eliminated "Medium Risk" alerts based on the consensus that High-Risk alerts provide more significant value. Additionally, we renamed "Sepsis [MET]" to "Sepsis Critical Risk," clarifying that the previous label did not accurately reflect actual sepsis, which is better defined by a positive blood culture. Our Sepsis-related phenotype aligns with a quick Sequential Organ Failure Assessment (SOFA), which can be utilized at the bedside for suspected sepsis cases. We learned that, due to delayed documentation, hourly alerts tend to be accurate but untimely, so we will further limit alerts and "MET" updates so that they only are visible in a "Grease Board" at the relatively precise times of inpatient admission and labor onset.

We also revisited triage, labor and delivery, and postpartum units this summer to observe how alerts functioned in real-time clinical workflows. We shadowed nurses, nurse managers, anesthesiologists, and obstetricians while conducting interviews lasting 30 to 60 minutes each. Mapping the overlap between clinical and technical processes allowed us to refine the alert system further. These detailed visits and process maps will enable us to fine-tune the alert mechanism effectively, thereby reducing alert burden. The shadowing experiences also facilitated knowledge transfer among MFM members regarding MEWS, expanding our reviewing team's capacity.

References:

1. Agarwal R, Goyal P, Mohta M, Kar R. Comparison of Sequential Organ Failure Assessment (SOFA) and Sepsis in Obstetrics Score (SOS) in Women with Pregnancy-Associated Sepsis with Respect to Critical Care Admission and Mortality: A Prospective Observational Study. *J Obstet Gynaecol India*. 2021 Feb;71(1):45-51. doi: 10.1007/s13224-020-01375-9. Epub 2020 Sep 25. PMID: 33814798; PMCID: PMC7960863.
2. Singer M, Deutschman CS, Seymour CW, et al. The third international consensus definitions for sepsis and septic shock (sepsis-3) *JAMA*. 2016;315:801-810. doi: 10.1001/jama.2016.0287.



EVALUATING AND SCALING INNOVATION

irAE Silent Evaluation

Noah Prizant, BA

As part of the 2021 RFA cycle, the Duke Institute for Health Innovation (DIHI) developed a machine learning model to predict emergency department (ED) and hospital admissions due to immune-related adverse events (irAEs). These are potentially dangerous side effects that occur in up to 76% of patients undergoing immune checkpoint inhibitor (ICI) therapy for various types of cancer.¹ irAEs may increase the risk of ED and hospital admissions, significant morbidity, and even death among cancer patients.²⁻⁴ Our goal was to use this model to identify high-risk patients, enabling healthcare professionals to intervene earlier and decrease the number of preventable acute care visits associated with irAEs.

Our model performed well upon retrospective validation, yielding an AUROC (Area Under the Curve) of 0.76 and a AUPRC (Precision-Recall Curve) of 0.27. These values indicate the model's ability to distinguish between high and low-risk patients and its precision in identifying high-risk patients. We used the model results to divide the patient cohort according to risk thresholds. The high-risk threshold (recall 0.04, precision 0.59) yielded a PPV of 0.59 with a sensitivity of 0.04, and the medium-risk threshold (recall 0.49, precision 0.27) yielded a PPV of 0.27 with a sensitivity of 0.49. The next step was to perform a silent validation of a real-time clinical tool using this model.

The model underwent prospective silent validation starting on February 16, 2023, during which we assessed performance measures and clinical actionability. (Silent validation refers to testing the model in real-time without healthcare providers taking clinical action based on the model's predictions.) The cohort included all patients receiving at least one ICI dose within the prospective validation period. Once per week, our clinical tool identified new adult infusion patients, pulled new data on previously identified patients from the EHR, and updated patients' probability of admission within the next three weeks of the latest ICI infusion. We instructed the model to categorize as high, medium, or low risk for admission according to established cutoffs. Advanced practice providers (APPs) reviewed each patient identified as a high or medium risk to assess for agreement with risk designation and documented if action (e.g., call patient or contact provider) was warranted.

Thus far, 9982 infusion visits occurred during the prospective validation period. During this time, 608 (6.1%) patients had an ICI infusion followed by re-admission to the hospital or ED within 21 days. A total of 687 total encounters were "alerted" for meeting medium or high-risk category criteria. The relatively low rate of ED visits and hospital admissions during this period likely contributed to a slightly decreased performance in the model compared to the retrospective validation. The performance showed lower sensitivity (24.7%) than expected, likely because the model missed many admitted patients. However, the model was very specific (94.3%) and had a relatively high PPV (21.5%) for admissions in the silent trial cohort.

In the clinical review portion, a total of 134 patients were included in clinical adjudications by APPs. Patients identified by the model were almost exclusively (98.5%) in the medium-risk category, and reviewers felt that only 61.4% of these patients were higher than low-risk. These results suggest that reconsidering the threshold for the medium risk category may be warranted before implementation to reduce false positives. The high-risk category (50% agreement) is challenging to assess due to the low number of patients during the silent validation period. Overall, 14.1% of patient cases in the medium and high-risk categories were deemed actionable, which supports the model's utility despite the lower-than-expected relevance of cases.

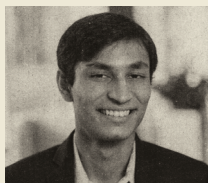
**We must determine whether
the model identifies high-
risk patients that healthcare
providers would not otherwise
quickly identify.**

Overall, silent validation on a prospective cohort supported the real-world utility of our model. The model’s performance, despite the challenges, demonstrated its potential to improve patient care by identifying high-risk patients and allowing for proactive interventions. Future implementation efforts must evaluate clinical impact and aim to ease provider burden. For instance, we must determine whether the model identifies high-risk patients that healthcare providers would not otherwise quickly identify. The project thus far represents a significant step towards improving patient care in oncology and a shift from reactive to proactive approaches in treating patients with immunotherapy.

References:

1. Ramos-Casals, M., et al. Immune-related adverse events of checkpoint inhibitors. *Nature Reviews Disease Primers* 6(2020).
2. Song, P., Zhang, D., Cui, X. & Zhang, L. Meta-analysis of immune-related adverse events of immune checkpoint inhibitor therapy in cancer patients. *Thoracic Cancer* 11, 2406-2430 (2020).
3. El Majzoub, I., et al. Adverse Effects of Immune Checkpoint Therapy in Cancer Patients Visiting the Emergency Department of a Comprehensive Cancer Center. *Annals of Emergency Medicine* 73, 79-87 (2019).
4. Wang, D.Y., et al. Fatal Toxic Effects Associated With Immune Checkpoint Inhibitors. *JAMA Oncology* 4, 1721 (2018).

STUDENT EXPERIENCES WITH DIHI



Kartik Pejavara, BS

Interning at the Duke Institute for Health Innovation (DIHI) has been one of the most incredible experiences of my undergraduate time at Duke. With the advent of large language models (LLMs) in healthcare occurring close to my start, I conducted much exploratory work on LLMs. Specifically, I researched how Duke Health clinics could use LLMs to accurately and automatically triage MyChart messages from patients to the appropriate provider or office. Right from the beginning, I was part of a team, testing various ideas and models. I learned a lot about artificial intelligence and machine learning (AI/ML) and how to work with large amounts of data.

As the MyChart Message project evolved, I got the chance to learn about the intricacies of the routing process directly from our nurses. For example, it may seem natural that when you have a question about your medication, treatment plan, or symptoms, you message your doctor for help. However, at Duke Endocrine, several offices offer distinct support and advice to patients. In fact, many treatment-related messages, such as trouble with blood-glucose monitors, medication refills, or prior authorizations, go to offices that help patients with their specific problems. We analyzed

three years of message data and shadowed triage nurses for hours to make the tool as accurate as possible at distinguishing between the offices. Nurse expertise and workflow intellect were indispensable.

In addition to collaborative sweat, LLM data science innovation is hard work. At first, working with the dataset was tricky — MyChart messages are highly unstructured, and there was no “ground-truth label” provided for each message. Pulling together various pieces of information, including previous routing information, employee IDs, and job titles, we constructed our model training and testing sets. The learning process was engaging and interesting, and although challenging, it has helped me build confidence in taking on future projects. In addition to this technical experience, I was fortunate to sit in on various meetings about implementing the solution. Seeing the forethought, resilience, and teamwork required to turn inventions into impactful innovations has been amazing. I have genuinely felt like a part of something bigger, gaining skills that will have a lasting impact on my career.



EVALUATING AND SCALING INNOVATION

Duke Opioid Sedation Assessment Risk Score

Padma Gulur, MD; Mary Cooter Wright, MS; Eugenia R McPeck Hinz, MD, MS;
Marshall Nichols, MS; Matt Gardner, BS; Suresh Balu, MS, MBA

The Duke Opioid Safety Assessment (DOSA) Risk Score was developed to address the challenges of opioid-induced respiratory depression (OIRD) in hospitalized patients—a serious complication that can lead to significant morbidity and mortality. We integrated the DOSA risk score in the electronic health record (EHR) and completed a limited go-live pilot at Duke University Hospital this past year. During the initial implementation phase, we observed a reduction of adverse events in the high-risk group by over 40% compared to the silent evaluation period (3.4% vs. 6.2%, $p=0.036$).

Problem

Opioids have a narrow window of therapeutic benefit yet remain one of healthcare's primary tools to address acute pain and anesthesia in a hospital setting. Opioid-induced respiratory depression (OIRD) is a common and often underdiagnosed cause of postoperative respiratory depression. Unrecognized or untreated episodes can lead to significant morbidity and potential mortality events. Other concurrent patient factors such as residual anesthesia, muscle paralysis, concurrent use of other sedatives, splinting from inadequate pain control, or obstructive sleep apnea can increase the risk of a respiratory event necessitating clinical rescue with naloxone.

Solution

Detection of a patient's declining respiratory status before progression to respiratory depression can help avert transfers to intensive care units (ICUs) and adverse outcomes. In 2020, Dr. Padma Gulur recognized the complexity of factors contributing to OIRD incidents and set out to develop a risk score model to alert the clinical care team to the patient's risk. This would allow standard-of-care actions such as continuous pulse ox monitoring. Effective monitoring and tailored pain management strategies can mitigate the risks of opioids and ensure patient safety.

The Duke Opioid Safety Assessment (DOSA) Risk Score categorizes patients as low, moderate, and high risk for opioid-related adverse events based on clinical factors such as recent surgeries, existing comorbidities, polypharmacy, and opioid tolerance. The DOSA Score is calculated hourly and subsequently integrated back into the

electronic health record (EHR) as a flowsheet score. For moderate or high-risk scores (~15% of patients), this result is translated into clinical decision support in the form of Best Practice Advisory (BPA) alerts. These alerts prompt the clinical team to implement appropriate additional monitoring measures, such as continuous pulse oximetry.

The BPA was made clinically visible for a limited pilot at Duke University Hospital on February 13, 2024 after approval from local oversight committees. Based on feedback from clinical users, location criteria were narrowed to eliminate ICU locations.

Outcome

During the initial implementation phase, from March 1 to June 1, 2024, we observed a reduction of adverse events in the high-risk group by over 40% compared to the silent evaluation period (3.4% vs. 6.2%, $p=0.036$). However, no similar benefit was found for the moderate-risk group. A secondary analysis of adherence to the BPA recommendations found inconsistent actions, with 54% of alerts not followed by the recommended actions.

A total of 7145 qualifying encounters with 55 events were identified, resulting in a rate of 7.7 events per 1000 patients. The first 5515 encounters from the Go-Live period were matched one-to-one with encounters from the prior silent evaluation window to assess efficacy. The event rate among patients with moderate risk was no different between the live-efficacy period 1.7% (10/602) vs. 1.0% (4/420) for the matched silent evaluation period (risk difference of 0.7%, chi-square $p=0.34$). When comparing the event rate among patients flagged as high risk, the Go-Live efficacy rate was 3.4% (19/568) vs. 6.2% (24/386) in the matched silent evaluation period (absolute risk difference of 2.9% (2.9 events per 100 patients), chi-square $p=0.036$).

Fifteen percent of the 54% who did not follow the recommended action, took no action at all, even after acknowledging the alert. This was a missed opportunity for intervention.

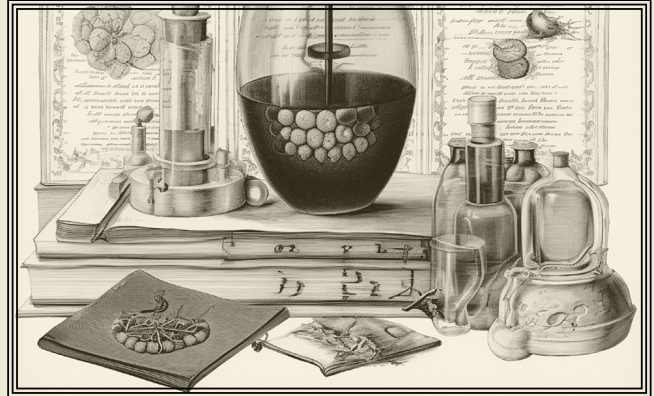
Overall, the DOSA Risk Score tool supported our highest-risk individuals more effectively, translating to improved opioid safety in the inpatient setting at Duke Health.

Next Steps

The Go-Live limited pilot evaluation has been shared with stakeholders and awaits final approval to expand to Duke Regional and Duke Raleigh Hospitals.

Academic Output

Publication in process.



STUDENT EXPERIENCES WITH DIHI



**Noah
Prizant, BA**

Deciding to spend my third year as a Duke Institute for Health Innovation (DIHI) Scholar was one of the best decisions I've made as a medical student. I've enjoyed venturing outside the traditional medical research realm and learning to push the boundaries of what is possible in healthcare. I have learned healthcare strategy and enjoyed meeting leaders in the field, learned to code in Python, and, most importantly, had the opportunity to contribute to informatics and machine learning (ML) projects that DIHI is actively implementing across the health system.

My primary project was "Developing a Digital Phenotype for the Early Recognition of Pediatric Sepsis." This project was originally launched in the RFA 2021 cycle. However, I joined at an exciting time when we were launching a "High-Risk + Fever" tool to be used by the pediatric emergency room team. I had the chance to be an integral part of its silent evaluation and deployment in March 2023, working closely with the DIHI team, clinical leads, and hospital stakeholders to ensure successful adoption. Additionally, DIHI teammates involved me in evaluating and working toward silent validation of their complete informatics-based pediatric sepsis phenotype. I had the opportunity to present these projects at Duke internal research events and at the Pediatric Academic Society (PAS) international conference in Toronto.

I have also been fortunate to contribute to other efforts at DIHI, including the "Predicting Hospital Admissions and ER Visits from Immune-related Adverse Events" project. Its leaders involved me in the silent validation phase, publication writing, and stakeholder interviews. Furthermore, Health AI Partnership (HAIP) leaders invited me to be involved in their cutting-edge policy work; I contributed to a manuscript about key decision points in AI healthcare best practices, "Aligning AI Principles and Healthcare Delivery Organization Best Practices: Navigating the Shifting Regulatory Landscape."

My time at DIHI has been overwhelmingly positive – so much so that I have decided to take another year to continue my research at DIHI and be involved in the next cycle of innovation projects. DIHI has enabled me to have the quintessential Duke experience: here, people can weave together interdisciplinary interests and launch commercial and nonprofit companies based on inventions, ideas, and solutions to social challenges. These experiences will surely shape who I will become as a physician and professional and will provide a lifelong community. I hope to incorporate what I've learned here in my future career and am excited for my second year working with DIHI.

EVALUATING AND SCALING INNOVATION

Ongoing Support for Patients with Peripheral Artery Disease, Expanding to Diabetic Patients with Threatened Limbs

Dennis Narcisse, MD; Hope Weissler, MD; Will Ratliff, MBA; Bradley Hintze, PhD; Pamela Cohen, BA; Daniel Costello, MPA; E. Holly Alvarado, PharmD, CCP; Manesh Patel, MD; Jill Engel, DNP, ACNP; Mary Schilder, RN; S. Benjamin Smith, PhD; Steven Lippmann PhD; Tara Kinard, RN, MSN, MBA; Faye Farber, MD; Dawn Coleman, MD; W. Schuyler Jones, MD

Peripheral artery disease (PAD) remains underdiagnosed and undertreated, which puts patients at risk for cardiovascular and neurovascular complications. We incorporated a predictive model for PAD into a virtual rounding program and have intervened on over 350 patients since our go-live in January 2022. Now, we are measuring the downstream impact of these interventions for our patients with PAD. We are also expanding our scope to proactively intervene on diabetic patients with threatened limbs.

Problem

Peripheral artery disease (PAD) is often underdiagnosed and undertreated, leading to significant health risks. Patients with PAD face higher chances of heart attack, stroke, amputation, and death, with groups/communities with minority or lower socioeconomic status being disproportionately affected. At Duke, a previous study showed patients with untreated PAD experience a 49% hospitalization rate, 8% amputation rate, and 15% mortality rate, placing their outcomes among the worst nationally.^{1,2} These adverse clinical outcomes not only affect patient health but also contribute to increased costs for both patients and our health system. We have started to expand care for this disease through population health monitoring of PAD. Meanwhile, we have an opportunity to expand our proactive interventions to Duke diabetic patients with threatened limbs to mitigate the risk of lower extremity amputation (LEA) and other poor outcomes. Studies show that a high rate of patients in North Carolina receive LEA even though preventive treatments exist to prevent this outcome.^{3,4}

Solution

Our team is improving the care of Duke patients with PAD through proactive identification and intervention to address gaps in their care. We designed and implemented a PAD virtual rounding program that features a PAD predictive model and generates intervention recommendations to patients' primary care physicians (PCPs) ahead of scheduled visits.³ The model runs on all adult patients with a Duke

University Health System (DUHS) clinical encounter that included at least one of 108 PAD-related diagnosis codes. A risk score is generated for each patient in the model cohort, indicating the likelihood that the patient has PAD. Each week prior to the rounds, a PAD specialist reviews the charts of "high risk for PAD" patients who have a primary care provider (PCP) appointment in the upcoming week to generate actionable recommendations to the PCP at a time when the PCP is thinking about the patient. Final recommendations are discussed at the weekly rounds by a multidisciplinary team that includes the PAD specialist, a pharmacist, and a Duke population health lead and are then sent to the patient's PCP. Figure 1 illustrates the workflow of these ongoing rounds.

Our team is improving the care of Duke patients with PAD through proactive identification and intervention to address gaps in their care.

For our diabetic patients with threatened limbs, we created a threatened limb patient registry, which went live on a Tableau Dashboard in December 2023. The registry focuses on two cohorts: (1) patients who had an encounter that included a diagnosis of diabetes mellitus and that co-occurred within one year of an encounter that included a diagnosis of lower extremity wound, and (2) patients who show uncontrolled diabetes mellitus via a hemoglobin A1C result above nine percent (9%) that co-occurred within one year of an encounter with either a foot deformity or peripheral artery disease diagnosis. We scan Duke's

EHR data once per week to assess all patients for whether they meet the inclusion criteria for either cohort. We upload newly identified patients to the registry with relevant data on recent and upcoming care. Figure 2 shows the Threatened Limb Patient Registry.

Impact

Since going live with the PAD virtual rounds workflow in January 2022, the PAD rounding team has intervened on over 350 patients with a variety of patient-tailored intervention suggestions to the patients’ PCPs. Table 1 shows the distribution of interventions sent. We are currently assessing the intervention cohort for subsequent positive outcomes in care.

Concurrently, for our diabetic patients with threatened limbs, we identified opportunities to close gaps in care and are working with health leadership to increase Duke Health resources to accomplish this. Through the Threatened Limb Dashboard, we observe opportunities to increase clinic follow-up for our 3,065 diabetic patients with a lower extremity wound and our 982 patients with uncontrolled diabetes mellitus and a foot deformity/PAD. Table 2 shows these opportunities.

Statin prescription	169
Smoking cessation referral	123
PCSK9 inhibitor prescription	43
PAD exercise therapy referral	20
Cardiology referral for inclisiran	14
Anti-platelet prescription	9
Endocrine lipid clinic referral	6
Cardiac rehab therapy referral	6
Ankle Brachial Index measurement needed	6
Ezetimibe prescription	6
SglT2 inhibitor prescription	3
Dukewell Pharmacy Tech outreach	3
Pharmacy referral (A1C and LDL management)	3
Lipid panel required	3
Cardiology referral	3
Endocrine/cardiometabolic clinic referral	3

Table 1. Intervention sent to PCPs for 353 patients

Next steps

We will evaluate the impact of our pilot on clinical and cost outcomes for our patients compared with the Duke patient population pre-implementation. We will maintain the PAD virtual rounding process, seeking feedback for optimization from key stakeholders, especially our PCPs. For our Threatened Limb Initiative, we will optimize the Threatened Limb Patient Registry Dashboard, finalize a distinct rounding workflow for these patients, and plan to go live with interventions and impact measurement tracking in early 2025.

References:

1. Kalbaugh CA, Loehr L, Wruck L, et al. Frequency of care and mortality following an incident diagnosis of peripheral artery disease in the inpatient or outpatient setting: the ARIC (Atherosclerosis Risk in Communities) study. *J Am Heart Assoc* 2018;7(8):e007332. <https://doi.org/10.1097/PCC.0000000000000254>
2. Weissler EH, Ford CB, Narcisse DI, Lippmann SJ, Smerek MM, Greiner MA, Hardy NC, O’Brien B, Sullivan RC, Brock AJ, Long C, Curtis LH, Patel MR, Jones WS. Clinician Specialty, Access to Care, and Outcomes Among Patients with Peripheral Artery Disease. *Am J Med.* 2022 Feb;135(2):219-227. doi: 10.1016/j.amjmed.2021.08.025. Epub 2021 Oct 7. PMID: 34627781; PMCID: PMC8840959. <https://pubmed.ncbi.nlm.nih.gov/34627781/>
3. Weissler EH, Lippmann SJ, Smerek MM, Ward RA, Kansal A, Brock A, Sullivan RC, Long C, Patel MR, Greiner MA, Hardy NC, Curtis LH, Jones WS. Model-Based Algorithms for Detecting Peripheral Artery Disease Using Administrative Data From an Electronic Health Record Data System: Algorithm Development Study. *JMIR Med Inform.* 2020 Aug 19;8(8):e18542. DOI: 10.2196/18542. PMID: 32663152; PMCID: PMC7468640. <https://pubmed.ncbi.nlm.nih.gov/32663152/>
4. Badjatiya A, Merrill P, Buse JB, Goodman SG, Katona B, Iqbal N, Pagidipati NJ, Sattar N, Holman RR, Hernandez AF, Mentz RJ, Patel MR, Jones WS. Clinical Outcomes in Patients With Type 2 Diabetes Mellitus and Peripheral Artery Disease. *Circ Cardiovasc Interv.* 2020 May;13(5):e008018
5. Margolis DJ, Hoffstad O, Nafash J, Leonard CE, Freeman CP, Hennessy S, Wiebe DJ. Location, location, location: geographic clustering of lower-extremity amputation among Medicare beneficiaries with diabetes. *Diabetes Care.* 2011 Nov;34(11):2363-7. doi: 10.2337/dc11-0807. Epub 2011 Sep 20. PMID: 21933906; PMCID: PMC3198303.



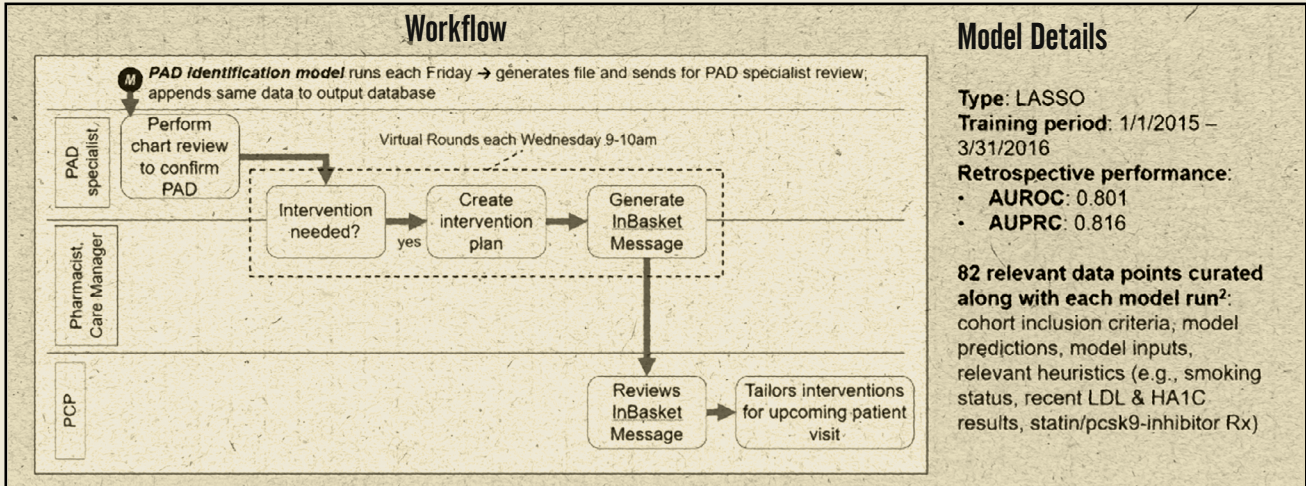


Figure 1. Peripheral artery disease population surveillance workflow

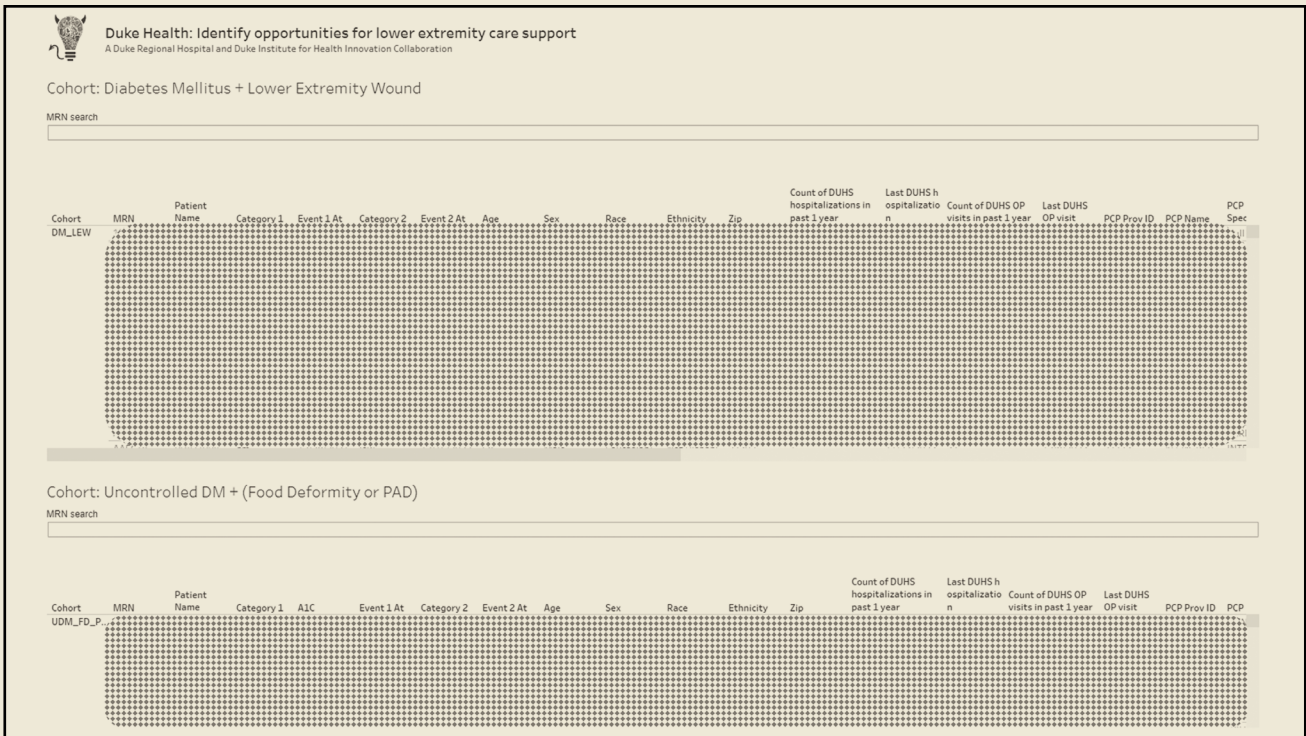


Figure 2. Threatened Limb Initiative: new patients automatically added daily to registry dashboard

COHORT	SERVICE	Patients with 1+ visit within past 1 year	Patients with visit scheduled in next 90 days
Diabetes Mellitus + Lower Extremity Wound (n=3065)	PCP	1233 (40.23%)	655 (21.37%)
	Vascular Surgery	542 (17.68%)	150 (4.89%)
	Podiatry	615 (20.07%)	241 (7.86%)
	Endocrine	741 (24.18%)	410 (13.38%)
Uncontrolled DM + (Foot Deformity or Peripheral Artery Disease) (n=982)	PCP	478 (48.68%)	293 (29.84%)
	Vascular Surgery	225 (22.91%)	73 (7.43%)
	Podiatry	83 (8.45%)	18 (1.83%)
	Endocrine	359 (36.56%)	209 (21.28%)

Table 2. Identified gaps in care for Threatened Limb Population

EVALUATING AND SCALING INNOVATION

Implementing an Informatics-Driven Notification System for Patients with High-Risk Conditions Presenting with Fever In The Pediatric Emergency Department

Noah Prizant; Shems Saleh; Will Ratliff, MBA; Marshall Nichols, MS; Mike Revoir, BS; Matt Gardner, BS; Michael Gao, MS; Suresh Balu, MS, MBA; Crystal Crider; Jenny Mando, Pharm D; Emily Greenwald, MD; Emily Sterrett, MD

Problem

Patients with high-risk conditions (HRC) who present to the pediatric emergency department (ED) with fever are at a significantly increased risk of developing systemic infections or sepsis. These conditions include the patient recently receiving chemotherapy, being on the transplant list, or having sickle cell disease. Timely evaluation and administration of antibiotics, ideally within one hour of presentation, are critical for these patients. However, achieving this goal has been challenging. From 2016 to 2023, only 42% of patients admitted to Duke Children's Hospital who met the HRC-plus-Fever criteria received antibiotics within the recommended one-hour window. Delays in antibiotic administration have been associated with poorer outcomes, making it imperative to find a solution that ensures timely treatment for these vulnerable patients.

Solution

To address this issue, we developed an informatics-driven system to immediately identify patients with high-risk conditions presenting with fever in the Duke Pediatric ED. The system aims to reduce the time-to-antibiotics (TTA) by leveraging real-time electronic health record (EHR) data. High-risk conditions were defined as active chemotherapy, being on the solid organ transplant list, or having sickle cell disease. Fever was defined as a temperature of $>38^{\circ}\text{C}$ measured in the ED or by chief complaint. Live data was obtained every 15 minutes from the electronic health record (EHR).

For each patient meeting the HRC-plus-Fever criteria, an automated notification is sent to the ED charge nurse and ED pharmacy if antibiotics still needed to be administered. This system was initially validated silently, starting on January 1, 2024, generating notifications without sending them to clinicians. The notifications went live on February 26, 2024, and reminder pages were added on April 13, 2024 to prompt action if antibiotics had not been administered within 45 minutes. In early September, clinical leads on our project team disseminated the project vision and monthly tracking metrics on time-to-antibiotics.

Impact

From February 26- October 31, 161 patients were flagged by the system. An additional 74 reminder pages were triggered. 86% of initial alerts were actionable, meaning that antibiotics were administered after the care team received the alert. Overall, median ED arrival-to-antibiotics time decreased by 12 minutes between the silent trial and go-live period (Table 1). However, run charts detailing TTA for all alerted patients showed special cause variation (extreme outliers) in cases of significantly delayed antibiotic administration (Figure 1). This finding led us to investigate these cases further, where we found instances where consult teams (primarily transplant) were deferred to for antibiotic decision-making. We determined that the transplant patient group should be a key focus of workflow improvements to decrease overall TTA in the ED. Excluding transplant patients, median TTA decreased by 20 minutes and we observed a centerline shift in October, suggesting that HRC-plus-Fever alerts are starting to impact other high-risk populations (Table 2, Figure 2).

We began developing an informatics-driven system to immediately identify patients with high-risk conditions The system aimed to reduce the time-to-antibiotics (TTA) by leveraging real-time electronic health record (EHR) data.

NEXT PAGE 



Next Steps

While the initial implementation of the informatics-driven notification system has shown promising results, there is still room for improvement. The next steps involve continuing to track impact metrics to ensure sustained effectiveness and identifying additional workflow enhancements. We will also focus on workflow improvements for instances where antibiotic administration was significantly delayed due to decision-making by consult teams, such as transplant or infectious disease specialists.

Future efforts include:

- Expanding the notification criteria to include additional patient populations
- Integrating a predictive model
- Exploring complementary notification delivery mechanisms for increasing awareness in real-time in the ED

We will continue iterating on this informatics-driven approach to identify interventions that improve our critical metric to reduce time-to-antibiotics. We aim to ensure that all high-risk pediatric patients presenting with fever receive timely interventions to improve clinical outcomes.

Outcome	Silent trial (1/1/24-2/25/24)	Live in ED (2/26/24-10/31/24)	Last Month (10/1/24-10/31/24)
Patients meeting HRC + Fever phenotype, n	40	161	18
Reminder pages sent, n (%*)	n/a	74 (58.7%*)	9 (50.0%)
Actionable initial page, n (%)	36 (90.0%)	138 (85.7%)	17 (94.4%)
Median ED arrival to antibiotic administration time, minutes (SD)	75.6 (50.1)	63.0 (87.3)	42.0 (106.3)
Median alert to antibiotic administration time, minutes (SD)	56.4 (46.0)	42.0 (80.9)	41.0 (108.1)
Arrival-Abx compliance (<1 hour ED arrival to Abx time), %	36.1%	45.9%	82.4%
Alert-Abx compliance (<1 hour alert to Abx time), %	61.1%	62.2%	88.2%

*denominator is all notifications after reminder go live, 4/13/2024

Table 1. HRC + Fever (non-transplant patients) Impact Metrics Pre vs Pre Go Live

Outcome (excludes transplant patients)	Silent trial (1/1/24-2/25/24)	Live in ED (2/26/24-10/31/24)	Last Month (10/1/24-10/31/24)
Patients meeting HRC + Fever phenotype, n	28	100	8
Reminder pages sent, n (%*)	n/a	34	1 (12.5)
Actionable initial page, n (%)	27 (96.4)	89 (89.0)	8 (100.0)
Median ED arrival to antibiotic administration time, minutes (SD)	78.0 (51.2)	58.2 (59.4)	37.5 (10.7)
Median alert to antibiotic administration time, minutes (SD)	58.2 (46.0)	37.2 (42.7)	17.4 (8.4)
Arrival-Abx compliance (<1 hour ED arrival to Abx time), %	25.9%	55.2%	100.0%
Alert-Abx compliance (<1 hour alert to Abx time), %	55.6%	71.3%	100.0%

*denominator is all notifications after reminder go live, 4/13/2024

Table 2. HRC + Fever (non-transplant patients) Impact Metrics Pre vs Post Go Live

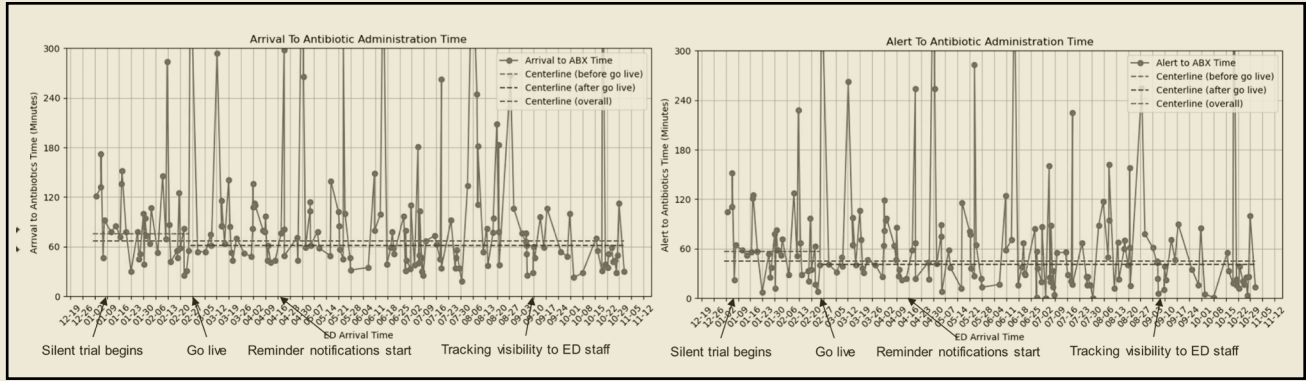


Figure 1. HRC + Fever Run Charts

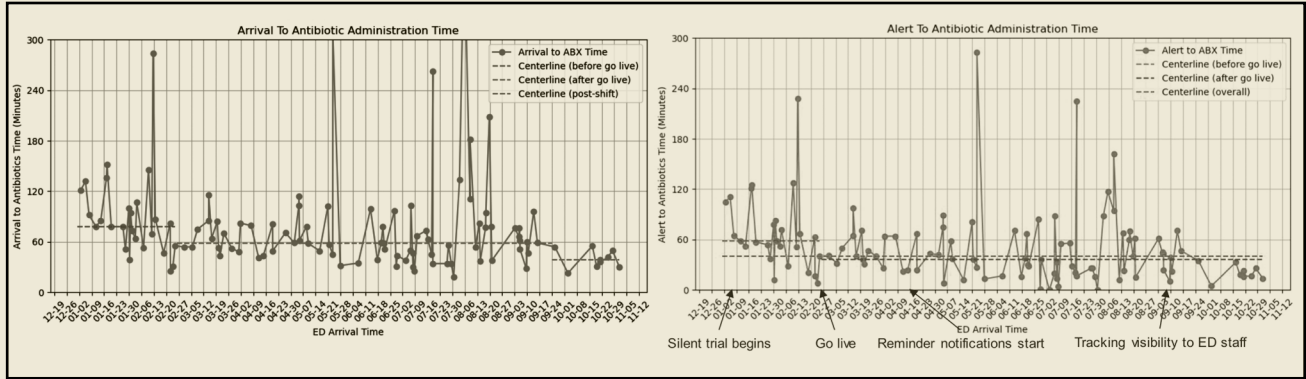


Figure 2. HRC + Fever (non-transplant patients) Run Charts



EVALUATING AND SCALING INNOVATION

Evaluation of a Digital Phenotype for the Early Recognition of Pediatric Sepsis

Noah Prizant; Shems Saleh; Will Ratliff, MBA; Marshall Nichols, MS; Mike Revoir, BS, Michael Gao, MS; Mark Sendak, MD, MPP; Suresh Balu, MS, MBA; Emily Greenwald, MD; Emily Sterrett, MD

Problem

Sepsis, a life-threatening response to infection, is a significant concern in pediatric care. In the United States, it accounts for an estimated 3.1% of all pediatric hospitalizations, with a mortality rate of 8.2%. Despite international guidelines emphasizing timely recognition and treatment, sepsis can be challenging to diagnose in children due to its varied presentations and nonspecific symptoms. At Duke Children’s Hospital, 38% of children receive timely recognition, and 28% receive prompt treatment.

Solution

The Duke Pediatric Sepsis Phenotype (DPSP)

Our project team has developed a digital phenotype, the DPSP, to identify pediatric patients at high risk of sepsis in real-time. See Figure 1 for a visualization of our pediatric patient cohort and the overlap of our sepsis definitions within that cohort. By combining two existing definitions – the Real-Time Weiss (RTW) and the Duke Children’s Trigger Tool (TT) – the DPSP can accurately identify sepsis in real-time, with a sensitivity of 0.79 and positive predictive value (PPV) of 0.18 for ICD codes. See Table 1 for details on DPSP sensitivity and positive predictive value for ground truth definitions of sepsis, as time constraints are placed on RTW and TT being met within n hours of each other. DPSP is now running in real-time, and the temporal validation of the phenotype kicked off on October 1st, 2024. We are currently calculating the Phoenix criteria for pediatric sepsis to add to retrospective definition comparisons in addition to refining workflow with the project team and key stakeholders.

Time constraint	ICD		Weiss	
	Sensitivity	PPV	Sensitivity	PPV
None	0.79	0.18	0.95	0.30
48 hours	0.78	0.19	0.93	0.30
24 hours	0.78	0.19	0.93	0.30
12 hours	0.77	0.19	0.91	0.31
8 hours	0.76	0.19	0.91	0.31
4 hours	0.76	0.20	0.90	0.32

Table 1. DSPTS Encounter Level, Time-Constrained Performance Analysis

Impact

Using the DPSP as a real-time pediatric sepsis definition, along with a predictive model, can potentially improve patient outcomes in pediatric care. By identifying patients at high risk of sepsis earlier, healthcare professionals can respond more quickly with timely recognition and treatment. Timely intervention can lead to reduced mortality rates, shorter hospital stays, and improved overall health outcomes for children. The tool can also be audited for equity, specifically fairness of treatment across racial subgroups.

When evaluating the phenotype, we found that patients meeting the DPSP criteria had significantly longer mean hospital length-of-stay (23.48 vs. 6.72 days) and mortality rates (5.67% vs. 0.74%) compared to the full retrospective cohort. These findings highlight the importance of early detection and treatment in improving patient outcomes.

Next Steps

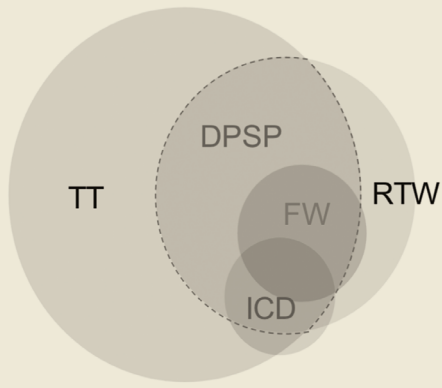
While the DPSP has shown promise, further temporal validation is necessary to confirm its clinical relevance. The next steps include:

- Prospective validation.
- Clinical adjudications.
- Development of a predictive model.
- Deployment at the bedside – accounting for workflow considerations.

References

1. Balamuth F, Weiss SL, Neuman MI, et al. Pediatric severe sepsis in U.S. children’s hospitals. *Pediatr Crit Care Med.* 2014;15(9):798–805.
2. Weiss SL, Peters MJ, Alhazzani W, et al. Surviving sepsis campaign international guidelines for the management of septic shock and sepsis-associated organ dysfunction in children. *Pediatr Crit Care Med.* 2020; 21(2):e52–e106.





SUB GROUP	ENCOUNTERS
Trigger Tool (TT)	7,314 (25.8%)
Real-Time Weiss (RTW)	3,953 (13.9%)
Duke Pediatric Sepsis Phenotype (DPSP)	3,104 (10.9%)
"Full" Weiss definition of sepsis (FW)	978 (3.4%)
ICD codes for sepsis (ICD)	718 (2.5%)
Population	28,399

Figure 1. Cohort characterization and Venn diagram visualization of the DPSP (to scale)

STUDENT EXPERIENCES WITH DIHI



**Samyukta
Pale, MEng**

In the spring of my first year of Masters in Artificial Intelligence studies at Duke, I joined the Duke Institute for Health Innovation (DIHI) as a data science intern. Coming from India, I was initially unfamiliar with the US healthcare system but eager to apply my skills to real-world healthcare challenges. DIHI quickly broadened my understanding of clinical processes, medical data, and health policies, bridging the gap between AI and the medical field.

During my opportune time at DIHI, I worked on a wide range of projects, each with its unique set of challenges and learning opportunities. These projects included developing predictive models for rheumatology telehealth visits, postoperative complications, and primary hyperparathyroidism (PHPT) surgical candidates. My role varied with each project, from evaluation and data analysis to modeling using natural language processing (NLP) techniques for PHPT. These experiences provided me with a comprehensive understanding of the AI workflow, from data analysis to deployment and clinical integration.

One of the most significant insights I gained from my time at DIHI was the importance of prioritizing the end-user experience. This was particularly evident in our discussions about sepsis predictions, where we learned about the critical balance required in alert frequency for patients at risk of sepsis. It was crucial to ensure that the alerts were effective in their purpose but not

overwhelming for the healthcare professionals. This respect for and alliance with care providers and patients has shaped my approach to AI in healthcare.

I still recall my first conversation with the Innovation Program Managers at DIHI, who asked me what innovation and teamwork meant to me. This sympathetic conversation sparked my excitement for DIHI's culture, and my appreciation for it has only grown as I've experienced the team's energy, expertise, and genuine mentorship. Working with a group that combines diverse expertise—spanning medicine, technology, management, and policy—with on-the-ground insights was incredibly rewarding. The value of teamwork in achieving our goals at DIHI cannot be overstated. It's through our shared successes that we build trust between developers and clinicians, making the inventive technology we build truly innovative, effective, and reliable.

As I look forward to the future, I am excited to continue my journey with DIHI. I am eager to build on the momentum we have created and to contribute to the impactful work we will accomplish in the coming year. I am confident that my experiences and learnings at DIHI will continue to shape my approach to AI in healthcare and will enable me to make a meaningful contribution to data science.



EVALUATING AND SCALING INNOVATION

DIHI Interval Problem List Clean Up 2024

Eugenia R McPeek Hinz, MD, MS; Rashaud Senor, MD; Tim Tsai, DO; Lisa Nadler, MD; Tres Brown;
Matt Gardner, BS; Will Ratliff, MBA, MBA; Suresh Balu, MS, MBA

Problem

The problem list in the Electronic Health Record (EHR) needs to be frequently reconciled. Duplicated, outdated diagnoses, and conflicting information make it challenging for clinicians to identify relevant patient problems. This disorganization impacts not only direct patient care but also affects downstream processes such as billing, predictive modeling, and research.

Without a standardized process for maintaining the problem list, diagnoses accumulate, leading to issues such as conflicting diagnoses (e.g., different classifications of asthma) or lapsed conditions (e.g., acute myocardial infarction after 30 days). Insufficient ownership and accountability for maintaining the problem list means clean-up efforts are typically secondary to direct patient care. This disorganization can create significant challenges that only worsen over time as more and more diagnoses accumulate in the patient's problem list.

Solutions

We approached this project in two phases:

PHASE ONE

Standardizing diagnosis organization using SNOMED CT groupers. To improve the organization and management of diagnoses, we developed and implemented 21 SNOMED CT groupers to organize the problem list by system in 2022. Building off of Epic grouper functionality, the SNOMED CT groupers were optimized utilizing Boolean logic to create comprehensive diagnosis groupers linked to ICD-10 codes. These groupers were subsequently shared with Epic and incorporated into the Foundation System for organization in 2023.

PHASE TWO

Automating clean-up of well-defined problem list findings, such as duplicates and outdated diagnoses. In February 2024, the Duke Institute for Health Innovation (DIHI), in partnership with Duke Health Technology Solutions (DHTS), implemented an automated clean-up of the problem list for exact duplicate diagnoses across the entire problem list. The initial run identified 15,322 diagnoses across 12,360 patients,

of which 14,522 were successful. This was consistent with the predicted extrapolation of duplicate diagnoses. Subsequently, we worked with Cardiology leadership to implement the automated resolution of lapsed cardiovascular diagnoses to the history of prepended with ICD-10 modification. Since then, these processes run weekly to address duplicate and lapsed diagnoses.

Outcomes

These automated clean-ups reduced the cognitive and time burdens on clinicians for clean-up, improving the usability of the problem list and facilitating more accurate patient care.

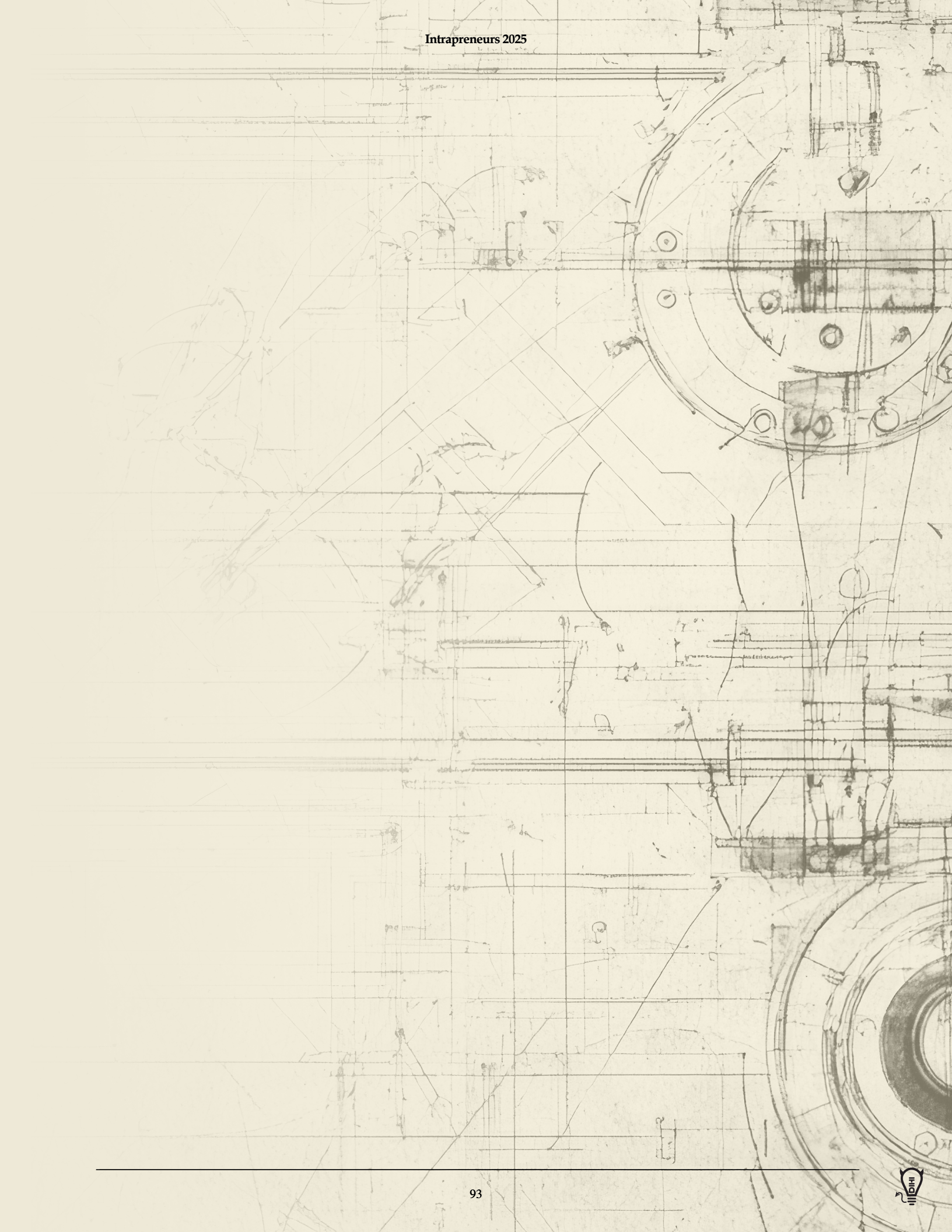
In September 2024, we published "Evaluation of SNOMED CT Grouper Accuracy and Coverage in Organizing the Electronic Health Record Problem List by System: Observational Study" in *Journal of Medical Internet Research Medical Informatics*. This study described the accuracy and efficiency of SNOMED CT groupers for patient diagnoses organization. We showed improved organization for clinical use of the new format. The results indicate that this approach could enhance the clinical utility of problem lists by reducing duplication and improving relevance for care coordination.

Summary

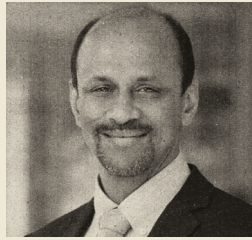
The Duke problem list clean-up project, using SNOMED CT groupers, has successfully improved the organization and accuracy of the problem list in the EHR. By addressing duplicated, outdated, and conflicting diagnoses, Duke has enhanced both patient care and secondary uses of clinical data. The automation of this process ensures continued clean-up for duplicated and lapsed diagnoses.

Reference

Senior R, Tsai T, Ratliff W, Nadler L, Balu S, Malcolm E, McPeek Hinz E, Evaluation of SNOMED CT Grouper Accuracy and Coverage in Organizing the Electronic Health Record Problem List by Clinical System: Observational Study, *JMIR Med Inform* 2024;12:e51274 doi: 10.2196/51274



DIHI Innovation Staff



Suresh Balu, MS, MBA

Director for DIHI
Associate Dean,
Innovation and Partnership,
School of Medicine



Jeffrey Ferranti, MD, MS

Executive Director, DIHI,
Senior Vice President,
Chief Digital Officer
Duke University Health System



Jamie Daniel, BS
Solutions Architect



Michael Gao, MS
Data Scientist



Alifia Hasan, B.Pharm, MBA
Innovation Program Manager



Bradley Hintze, PhD
Data Engineer



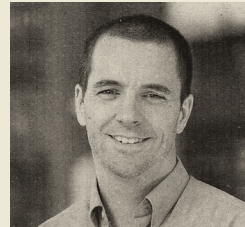
Jee Young Kim, PhD
Data Scientist



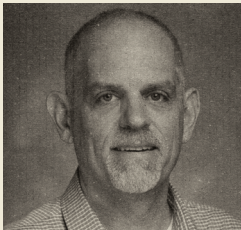
Will Knechtle, MBA, MPH
Innovation Program Manager



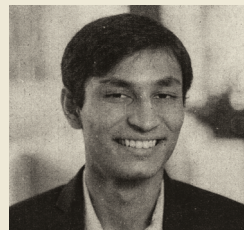
Sena Kpodzro, MPH
Research & Product Manager



Marshall Nichols, MS
Data Engineer



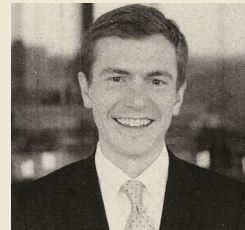
Matt Gardner, BS
Solutions Architect



Kartik Pejavara, BS
Data Scientist



Anusha Prakash, MSc
Data Scientist



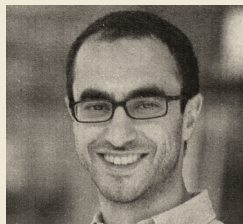
Will Ratliff, MBA
Innovation Program Manager



Mike Revoir, BS
Solutions Architect



Shems Saleh, MSc
Data Scientist



Mark Sendak, MD, MPP
Clinical Data Scientist



Ciera Thomas, MPH
Research & Product Manager

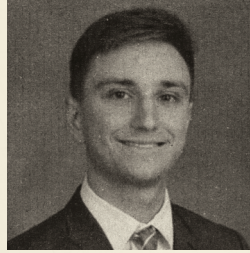
DIHI Scholars



Mahmoud Alwakeel, MD



Seyi Atoyebi, BChD



Ian Bracken



Henry Foote, MD



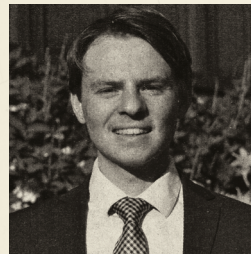
Freya Gulamali



Vaibhav Jha



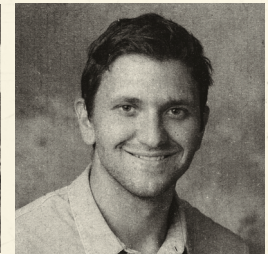
Samyukta Palle, MEng



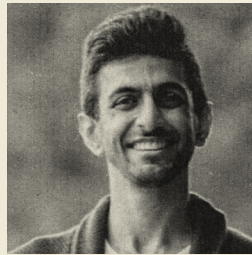
Noah Prizant, BA



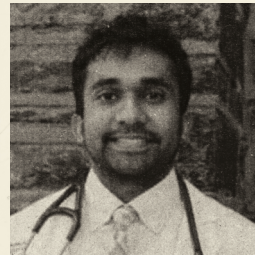
Islam Tayeb



Bruno Valen, BS



Srijan Bhasin, BS



Jay Swayambunathan, BS

Publications

1. Amin, K. D., Weissler, E. H., Ratliff, W., Sullivan, A. E., Holder, T. A., Bury, C., Francis, S., Theiling, B. J., Hintze, B., Gao, M., Nichols, M., Balu, S., Jones, W. S., & Sendak, M. (2024). Development and Validation of a Natural Language Processing Model to Identify Low-Risk Pulmonary Embolism in Real Time to Facilitate Safe Outpatient Management. *Annals of Emergency Medicine*, 84(2), 118–127. <https://doi.org/10.1016/j.annemergmed.2024.01.036>
2. Boag, W., Hasan, A., Kim, J. Y., Revoir, M., Nichols, M., Ratliff, W., Gao, M., Zilberstein, S., Samad, Z., Hoodbhoy, Z., Ali, M., Khan, N. S., Patel, M., Balu, S., & Sendak, M. (2024). The algorithm journey map: a tangible approach to implementing AI solutions in healthcare. *Npj Digital Medicine*, 7(1). <https://doi.org/10.1038/s41746-024-01061-4>
3. Economou-Zavlanos, N. J., Bessias, S., Cary, M. P., Bedoya, A. D., Goldstein, B. A., Jelovsek, J. E., O'Brien, C. L., Walden, N., Elmore, M., Parrish, A. B., Elengold, S., Lytle, K. S., Balu, S., Lipkin, M. E., Shariff, A. I., Gao, M., Leverenz, D., Henao, R., Ming, D. Y., ... Poon, E. G. (2024). Translating ethical and quality principles for the effective, safe and fair development, deployment and use of artificial intelligence technologies in healthcare. *Journal of the American Medical Informatics Association*, 31(3), 705–713. <https://doi.org/10.1093/jamia/ocad221>
4. Foote, H. P., Shaikh, Z., Witt, D., Shen, T., Ratliff, W., Shi, H., Gao, M., Nichols, M., Sendak, M., Balu, S., Osborne, K., Kumar, K. R., Jackson, K., McCrary, A. W., & Li, J. S. (2024). Development and Temporal Validation of a Machine Learning Model to Predict Clinical Deterioration. *Hospital Pediatrics*, 14(1), 11–20. <https://doi.org/10.1542/hpeds.2023-007308>
5. Kim, J. Y., Hasan, A., Kellogg, K. C., Ratliff, W., Murray, S. G., Suresh, H., Valladares, A., Shaw, K., Tobey, D., Vidal, D. E., Lifson, M. A., Patel, M., Raji, I. D., Gao, M., Knechtle, W., Tang, L., Balu, S., & Sendak, M. P. (2024). Development and preliminary testing of Health Equity Across the AI Lifecycle (HEAAL): A framework for healthcare delivery organizations to mitigate the risk of AI solutions worsening health inequities. *PLOS Digital Health*, 3(5), e0000390. <https://doi.org/10.1371/journal.pdig.0000390>
6. Khan, S. D., Hoodbhoy, Z., Raja, M. H. R., Kim, Jee Young, Hogg, H. D. J., Manji, A. A. A., Gulamali, Freya, Hasan, Alifia, Shaikh, A., Tajuddin, S., Khan, N. S., Patel, M. R., Balu, S., Samad, Z., & Sendak, M. P. (2024). Frameworks for procurement, integration, monitoring, and evaluation of artificial intelligence tools in clinical settings: A systematic review. *PLOS Digital Health*, 3(5), e0000514. <https://doi.org/10.1371/journal.pdig.0000514>
7. Mandall, M., Halstater, B., Arguello, M., & Viera, A. (2024). The eProvider Model: A Novel Approach to the Modern Problem of Portal Message Overload. *Family Practice Management*, 31(2), 13–18. <https://www.aafp.org/pubs/fpm/issues/2024/0300/eprovider-model.html>
8. Nixon, M. P., Momotaz, F., Smith, C., Smith, J. S., Sendak, M., Polage, C., & Silverman, J. D. (2024). From pre-test and post-test probabilities to medical decision making. *BMC Medical Informatics and Decision Making*, 24(1). <https://doi.org/10.1186/s12911-024-02610-3>
9. Patel, M. N., Mara, A., Acker, Y., Gollon, J., Setji, N., Walter, J., Wolf, S., Zafar, S. Y., Balu, S., Gao, M., Sendak, M., Casarett, D., LeBlanc, T. W., & Ma, J. (2024). Machine learning for targeted advance care planning in cancer patients: a quality improvement study. *Journal of Pain and Symptom Management*. <https://doi.org/10.1016/j.jpainsymman.2024.08.036>
10. Rice, H. E., Owen, M., Johnson, A., Swandell, J., Phifer Nicholson, C., Provencher, S., Horne, E., Solomon, C., Ratliff, W., Knechtle, W., Campbell, D., Smith, R., & Graves, L. (2024). Lived experiences of people impacted by gun violence: qualitative analysis of the prescriptions for repair project. *Trauma Surgery & Acute Care Open*, 9(1), e001503. <https://doi.org/10.1136/tsaco-2024-001503>
11. Sendak, Mark P, Liu, V. X., Beecy, A., Vidal, D. E., Shaw, K., Lifson, M. A., Tobey, D., Valladares, A., Loufek, B., Mogri, M., & Balu, S. (2024). Strengthening the use of artificial intelligence within healthcare delivery organizations: balancing regulatory compliance and patient safety. *Journal of the American Medical Informatics Association*, 31(7), 1622–1627. <https://doi.org/10.1093/jamia/ocae119>
12. Sendak, Mark P., Kim, J. Y., Hasan, A., Ratliff, W., Lifson, M. A., Patel, M., Raji, I. D., Sehgal, A., Shaw, K., Tobey, D., Valladares, A., Vidal, D. E., & Balu, S. (2024). Empowering US healthcare delivery organizations: Cultivating a community of practice to harness AI and advance health equity. *PLOS Digital Health*, 3(6), e0000513. <https://doi.org/10.1371/journal.pdig.0000513>
13. Senior, R., Tsai, T., Ratliff, W., Nadler, L., Balu, S., Malcolm, E., & McPeck Hinz, E. (2024). Evaluation of SNOMED CT Grouper Accuracy and Coverage in Organizing the Electronic Health Record Problem List by Clinical System: Observational Study. *JMIR Medical Informatics*, 12, e51274–e51274. <https://doi.org/10.2196/51274>
14. Gao, Michael, Pejavara, Kartik, Balu, S., & Henao, R. (2024, November 11). Development of a Flexible Chain of Thought Framework for Automated Routing of Patient Portal Messages. *AMIA Annual Symposium Proceedings*. AMIA 2024 Annual Symposium, San Francisco, CA
15. Kim, J. Y. (n.d.). AI Governance in Healthcare in a Learning Health System: Case Study of a Canadian Hospital System. (Poster Presentation presented at the Machine Learning for Healthcare Conference, Toronto, Canada, Aug 16, 2024)
16. Prizant, Noah et al., “Evaluation of a Digital Phenotype for the Early Recognition of Pediatric Sepsis” (Poster Presentation presented at the Pediatric Academic Society, Toronto, Canada, May 3, 2024)

Publications

17. Prizant, Noah et al., “Evaluation of a Digital Phenotype for the Early Recognition of Pediatric Sepsis” (Machine Learning for Healthcare Conference, Toronto, Canada, Aug 16, 2024)
18. Prizant, N., Saleh, S., Ratliff, W., Nichols, M., Revoir, M., Gardner, M., Gao, M., Sendak, M., Balu, S., Greenwald, E., & Sterrett, E., “Implementing a Notification System for Patients with High-Risk Conditions Presenting with Fever in the Pediatric Emergency Department” (Poster Presentation presented at the Duke Pediatric Research Retreat, Durham, NC, April 10, 2024)
19. Prizant, N., Saleh, S., Ratliff, W., Nichols, M., Revoir, M., Gardner, M., Gao, M., Sendak, M., Balu, S., Greenwald, E., & Sterrett, E., “Implementing a Notification System for Patients with High-Risk Conditions Presenting with Fever in the Pediatric Emergency Department” (Poster Presentation at the DukeMed Research Symposium, Aug 1, 2024)
20. Prizant, N., Saleh, S., Ratliff, W., Nichols, M., Revoir, M., Gardner, M., Gao, M., Sendak, M., Balu, S., Greenwald, E., & Sterrett, E., “Implementing a Notification System for Patients with High-Risk Conditions Presenting with Fever in the Pediatric Emergency Department” (Machine Learning for Healthcare Conference, Toronto, Canada, Aug 16, 2024)
21. Brown, Britney E., and Uzuri I. Holder. “The Proliferation of Hospital-Based Violence-Intervention Programs in North Carolina.” *North Carolina Medical Journal* 84, no. 4 (July 5, 2023). <https://doi.org/10.18043/001c.81266>
22. Burns, Charles M., Leland Pung, Daniel Witt, Michael Gao, Mark Sendak, Suresh Balu, Douglas Krakower, Julia L. Marcus, Nwora Lance Okeke, and Meredith E. Clement. “Development of a Human Immunodeficiency Virus Risk Prediction Model Using Electronic Health Record Data From an Academic Health System in the Southern United States.” *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 76, no. 2 (January 13, 2023): 299–306. <https://doi.org/10.1093/cid/ciac775>
23. Davis SE, Matheny ME, Balu S, Sendak MP. A framework for understanding label leakage in machine learning for health care. *J Am Méd Inform Assoc.* 2023
24. Foote, Henry P., Grace S. Lee, Carla Dominguez Gonzalez, Zohaib Shaik, William Ratliff, Michael Gao, Bradley Hintze, et al. “Risk of In-Hospital Deterioration for Children with Single Ventricle Physiology.” *Pediatric Cardiology* 44, no. 6 (2023): 1293–1301. <https://doi.org/10.1007/s00246-023-03191-0>
25. Khan Sarim Dawar, et al. “Frameworks for Procurement, Integration, Monitoring, and Evaluation of Artificial Intelligence Tools in Clinical Settings: A Systematic Review.” *PLOS Digital Health*, vol. 3, no. 5, May 2024, p. e0000514, <https://doi.org/10.1371/journal.pdig.0000514>
26. Kim, J. Y., Boag W, Gulamali F, Hasan A, Hogg HDJ, Lifson M, Mulligan D, Patel M, Raji ID, Sehhal A, Shaw K, Tobey D, Valladares A, Vidal D, Balu S, Sendak M. Organizational Governance of Emerging Technologies: AI Adoption in Healthcare. In 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23), June 12–15, 2023, Chicago, IL, USA. ACM, Chicago, IL, USA, 23 pages. <https://doi.org/10.1145/3593013.3594089>
27. Ming DY, Wong W, Jones KA, Antonelli RC, Gujral N, Gonzales S, Rogers U, Ratliff W, Shah N, King HA. Feasibility of Implementation of a Mobile Digital Personal Health Record to Coordinate Care for Children and Youth With Special Health Care Needs in Primary Care: Protocol for a Mixed Methods Study. *JMIR Res Protoc* 2023;12:e46847. URL: <https://www.researchprotocols.org/2023/1/e46847>. DOI: [10.2196/46847](https://doi.org/10.2196/46847)
28. Movva, Naimisha, Mina Suh, Heidi Reichert, Bradley Hintze, Mark P Sendak, Zachary Wolf, Shannon Carr, et al. “Respiratory Syncytial Virus During the COVID-19 Pandemic Compared to Historic Levels: A Retrospective Cohort Study of a Health System.” *The Journal of Infectious Diseases* 226, no. Supplement_2 (August 15, 2022): S175–83. <https://doi.org/10.1093/infdis/jiac220>
29. Price WN, Sendak M, Balu S, Singh K. Enabling collaborative governance of medical AI. *Nat Mach Intell.* 2023;1–3
30. Sandhu, Sahil, Mark Sendak, William Ratliff, William Knechtle, William J. Fulkerson, and Suresh Balu. “Accelerating Health System Innovation: Principles and Practices from the Duke Institute for Health Innovation.” *Patterns* 4 (April 14, 2023). <https://doi.org/10.1016/j.patter.2023.100710>
31. Sendak, Mark P, David Vidal, Sylvia Trujillo, Karandeep Singh, Xiaoxuan Liu, and Suresh Balu. “Editorial: Surfacing Best Practices for AI Software Development and Integration in Healthcare.” *Frontiers in Digital Health* 5 (2023): 1150875. <https://doi.org/10.3389/fgdth.2023.1150875>
32. Sendak, M., Balu, S., & Hernandez, A. F. (2023). Proactive Algorithm Monitoring to Ensure Health Equity. *JAMA Network Open*, 6(12), e2345022. <https://doi.org/10.1001/jamanetworkopen.2023.45022>
33. Shah, Nishant P., Courtney Page, Cynthia L. Green, Michael Gao, Joanna Cavalier, Robert W. McGarrah, Tracy A. DeWald, Devdutta Sangvai, Manesh R. Patel, and Neha J. Pagidipati. “Bending the Cardiovascular Event Curve by Evaluating the Potential Impact of Achieving Low-Density Lipoprotein Cholesterol Goal Across a Large Health System Among Secondary Prevention Patients.” *The American Journal of Cardiology* 186 (January 1, 2023): 91–99. <https://doi.org/10.1016/j.amjcard.2022.10.03>

NEXT PAGE →



Publications / continued

34. Wang S, Hogg H, Sangvai D, Patel M, Weissler E, Kellogg K, Ratliff W, Balu S, Sendak M. Development and Integration of Machine Learning Algorithm to Identify Peripheral Arterial Disease: Multistakeholder Qualitative Study. *JMIR Form Res* 2023;7:e43963. URL: <https://formative.jmir.org/2023/1/e43963>. DOI: 10.2196/43963
35. Wong, Samantha C., William Ratliff, Meng Xia, Christine Park, Mark Sendak, Suresh Balu, Ricardo Henao, Lawrence Carin, and Meenal K. Kheterpal. "Use of Convolutional Neural Networks in Skin Lesion Analysis Using Real World Image and Non-Image Data." *Frontiers in Medicine* 9 (October 19, 2022): 946937. <https://doi.org/10.3389/fmed.2022.946937>
36. Zanolli, Nicole C, Stephanie Lim, William Knechtle, Kelvin Feng, Tracy Truong, Laura J Havrileskey, and Brittany A Davidson. "Implementation of a Validated Post-Operative Opioid Nomogram into Clinical Gynecologic Surgery Practice: A Quality Improvement Initiative." *Gynecologic Oncology Reports* 49 (October 1, 2023). <https://doi.org/10.1016/j.gore.2023.101260>
37. Boag, W., Knechtle, W., Gao, M., Sendak, M., Richards, M., Kansal, N., Mayya, V., Talley, H., Habib, A., Allen, T., Bauer, M., Mitchell C.; Gilner, J.; Hughes, B. (August 12, 2023) "Phenotype Development and Validation for a Maternal Early Warning System." Poster presented at the Machine Learning in Healthcare Conference (MLHC), Columbia University
38. O'Brien, C.; Ratliff, W.; Gao, M., Hintze, B., Sendak, M., Saleh, S., Balu, S., Nichols, M., Gardner, M., Revoir, M., Tang, L., Kester, K., Dunn, C., Tart, D., Shirey, K. (2023, May 2). Predicting Behavioral Emergencies in the Hospital [Slides and demonstration presented]. GSA Applied AI Healthcare Challenge – competition winner [\$25,000 award]
39. Ochoa, T., Tang, L., Knechtle, W., Deshpande, K., Sendak, M., Gao, M., Balu, S., Klapper, J., Bottiger, B., Snyder, L., & Hartwig, M. (2023, April 26). Development of a Machine Learning Model to Predict Textbook Outcomes in Lung Transplant Patients. Surgical Research Day, Durham, NC
40. Ratliff, W., Gao, M., Hintze, B., Sendak, M., Saleh, S., Balu, S., Nichols, M., Gardner, M., Revoir, M., Tang, L., Kester, K., Dunn, C., Tart, D., Shirey, K., O'Brien, C. (2023, August 12). Predicting Behavioral Emergencies in the Hospital [Poster presented]. Machine Learning in Healthcare Conference/ New York, NY, USA
41. Tang, L., Gao, M., Ratliff, W., Saleh, S., Balu, S., Nichols, M., Revoir, M., Sterrett, E., Sendak, M. (2023, August 12). A Case Study on Deep Learning Label Leakage Identified during a Silent Trial [Poster presented]. Machine Learning in Healthcare Conference/New York, NY, USA
42. Bellantoni, Julia B. MD, et al. "Implementation of a Telehealth Videoconference to Improve Hospital-to Skilled Nursing Care Transitions: Preliminary Data." *Journal of the American Geriatrics Society*, Mar. 2022, pp. 1–10, <https://doi.org/10.1111/jgs.17751>
43. Chan Norine W., et al. "Social Determinants of Health Data in Solid Organ Transplantation: National Data Sources and Future Directions." *American Journal of Transplantation*, vol. 22, no. 10, 2022, pp. 2293–301, <https://doi.org/10.1111/ajt.17096>
44. Corey, Kristin M, et al. "Exposure and Outcome in Practice: A Retrospective Cohort Study between Fibrinolytic Suppression and Hypercoagulability, the Severity of Hypoxemia, and Mortality in COVID-19 Patients." *Anesthesiology*, Apr. 2022, p. 10.1097/ALN.0000000000004239, <https://doi.org/10.1097/ALN.0000000000004239>
45. Fenn, Alexander, et al. "Development and Validation of Machine Learning Models to Predict Admission From Emergency Department to Inpatient and Intensive Care Units." *Annals of Emergency Medicine*, vol. 78, no. 2, 2021, pp. 290–302, <https://doi.org/10.1016/j.annemergmed.2021.02.029>
46. Honeycutt Christopher Cole, et al. "Assessment of Practices Affecting Racial and Ethnic COVID-19 Vaccination Equity in 10 Large US Cities." *Journal of Public Health Management and Practice*, vol. 28, no. 6, 2022, pp. E778–88, <https://doi.org/10.1097/PHH.0000000000001610>
47. McClain, Micah T., et al. "A Blood-Based Host Gene Expression Assay for Early Detection of Respiratory Viral Infection: An Index-Cluster Prospective Cohort Study." *The Lancet Infectious Diseases*, vol. 21, no. 3, 2021, pp. 396–404, [https://doi.org/10.1016/S1473-3099\(20\)30486-2](https://doi.org/10.1016/S1473-3099(20)30486-2)
48. Sendak, Mark P, et al. Development and Validation of ML-DQA -- a Machine Learning Data Quality Assurance Framework for Healthcare. 2022, <https://doi.org/10.48550/ARXIV.2208.02670>
49. Sendak, Mark P, Hulamali F, Balu S. Overcoming the Activation Energy Required to Unlock the Value of AI in Healthcare. NBER [Internet]. 2022 Sep 22; Available from: <https://www.nber.org/books-and-chapters/economics-artificial-intelligence-health-care-challenges/overcoming-activation-energy-required-unlockvalue-ai-healthcare-sendak>
50. Xia, Meng, Meenal K. Kheterpal, Samantha C. Wong, Christine Park, William Ratliff, Lawrence Carin, and Ricardo Henao. "Lesion Identification and Malignancy Prediction from Clinical Dermatological Images." *Scientific Reports* 12, no. 1 (September 23, 2022): 15836. <https://doi.org/10.1038/s41598-022-20168-w>

Publications / continued

51. Xiu, Zidi, et al. "Variational Disentanglement for Rare Event Modeling." Proceedings of the... AAAI Conference on Artificial Intelligence. AAAI Conference on Artificial Intelligence, vol. 35, no. 12, May 2021, pp. 10469–77. <https://ojs.aaai.org/index.php/AAAI/article/view/17253/17060>
52. Yang Zhou, et al. "Advancing Primary Care with Artificial Intelligence and Machine Learning." Healthcare, vol. 10, no. 1, 2022, p. 100594, Journal of Healthcare, Elsevier. December 2021. <https://doi.org/10.1016/j.hjdsi.2021.100594>
53. Zullig, Leah L, et al. "Low-Touch, Team-Based Care for Co-Morbidity Management in Cancer Patients: The ONE TEAM Randomized Controlled Trial." BMC Family Practice, vol. 22, no. 1, Nov. 2021, p. 234, <https://doi.org/10.1186/s12875-021-01569-8>
54. Chan, N. W. et al. "Social Determinants of Health Data Capture Within National and Health System Data Sources", is accepted for an oral presentation in the Scientific Forum at Clinical Congress 2022, taking place October 16-20 in San Diego, CA
55. Mark Sendak, Gaurav Sirdeshmukh, Timothy Ochoa, Hayley Premo, Linda Tang, Kira Niederhoffer, Sarah Reed, Kaivalya Deshpande, Emily Sterrett, Melissa Bauer, Laurie Snyder, Afreen Shariff, David Whellan, Jeffrey Riggio, David Gaieski, Kristin Corey, Megan Richards, Michael Gao, Marshall Nichols, Bradley Heintze, William Knechtle, William Ratliff, Suresh Balu. Development and Validation of ML-DQA -- a Machine Learning Data Quality Assurance Framework for Healthcare. Presented at 2022 Machine Learning in Health Care Conference
56. Niederhoffer K, Knechtle W, Uronis H, Shariff A, et al. (2022). A Machine Learning Model to Predict Hospital Admissions and Emergency Department Use in Patients' Immune Checkpoint Inhibitors. Machine Learning Healthcare Conference (MLHC) 2022 poster. Durham, NC, 2022
57. Ochoa T, Knechtle W, Sendak M. Development of a Machine Learning Model for Prediction of Mortality in Lung Transplant Patients. Poster presentation at Machine Learning for Healthcare (MLHC), August 2022, Durham, NC
58. Premo H, Shi H, Knechtle W, Kazaura H, et al. (2022). A Geriatric-Specific Morbidity and Mortality Perioperative Risk Stratification Tool. Machine Learning for Healthcare (MLHC) 2022. Poster presentation. Durham, NC, 2022
59. Shen R, Weisse EH, Ratliff W, Nichols M, Hintze B, Gao M, Sendak M, Balu S, Jones S, et al. Improving Equity and Value of Peripheral Artery Disease Care at a Population Level. Machine Learning for Healthcare 2022 – Clinical Abstract, Software, and Demo Track. August 5-6, 2022
60. Shen R; Ratliff W; Sendak M; Kapadia N; Burrows B, et al. Algorithm Development for Duke Emergency Pre-hospital Capacity Management. Machine Learning for Healthcare 2022 – Clinical Abstract, Software, and Demo Track. August 5-6, 2022
61. Tang, L, Ratliff, W, Sendak, M, Gao, M., Nichols, M, Revoir, M, Yashar, F, Yao, J, Balu, S, Subramanian, N, Uhl, T, Denis, L, Sterrett, E. Identifying Sepsis in real-time for Duke University Hospital Pediatric Patients [Poster presented]. Machine Learning for Healthcare (MLHC) Conference. August 2022. Durham, NC, USA
62. Zanolli N, Knechtle W, Havrilesky L, Davidson B. Integration of a Post-operative Opioid Calculator into an Academic Gynecologic Surgery Practice. North Carolina Obstetrical and Gynecologic Society (NCOGS) 2022 Annual Meeting. Kiawah, SC, 2022
63. Zanolli N, Knechtle W, Sendak M, Havrilesky L, Davidson B, Integration of a Post-operative Opioid Calculator into an Academic Gynecologic Surgery Practice. Machine Learning Conference for Healthcare. Durham, NC, 2022
64. Dohlman, A. B., Arguijo Mendoza, D., Ding, S., Gao, M., Dressman, H., Iliev, I. D., Lipkin, S. M., & Shen, X. (2021). The cancer microbiome atlas: a pan-cancer comparative analysis to distinguish tissue-resident microbiota from contaminants. Cell Host & Microbe, 29(2), 281-298.e5. <https://doi.org/10.1016/j.chom.2020.12.001>
65. Fenn Alexander, et al. "Development and Validation of Machine Learning Models to Predict Admission From Emergency Department to Inpatient and Intensive Care Units." Annals of Emergency Medicine, vol. 78, no. 2, 2021, pp. 290–302, <https://doi.org/10.1016/j.annemergmed.2021.02.029>
66. Fenn, A., Davis, C., Buckland, D. M., Kapadia, N., Nichols, M., Gao, M., Knechtle, W., Balu, S., Sendak, M., & Theiling, B. J. (2021). Development and Validation of Machine Learning Models to Predict Admission From Emergency Department to Inpatient and Intensive Care Units. Annals of Emergency Medicine, 78(2), 290–302. <https://doi.org/10.1016/j.annemergmed.2021.02.029>
67. Honeycutt, C. C., Bueno, K. M., Tran, T., Gao, M., Balu, S., & Sendak, M. (2021). Assessment of Spanish Translation of Websites at Top-Ranked US Hospitals. JAMA Network Open, 4(2), e2037196. <https://doi.org/10.1001/jamanetworkopen.2020.37196>
68. Kansal, A., Gao, M., Balu, S., Nichols, M., Corey, K., Kashyap, S., & Sendak, M. (2021). Impact of diagnosis code grouping method on clinical prediction model performance: A multi-site retrospective observational study. International Journal of Medical Informatics, 151, 104466. <https://doi.org/10.1016/j.ijmedinf.2021.104466>

NEXT PAGE 



Publications / continued

69. Kansal, A., Green, C. L., Peterson, E. D., Newby, L. K., Wang, T.Y., Sendak, M., Balu, S., Patel, M. R., & Fanaroff, A. C. (2021). Electronic Health Record Integration of Predictive Analytics to Select High-Risk Stable Patients With Non–ST-Segment–Elevation Myocardial Infarction for Intensive Care Unit Admission. *Circulation: Cardiovascular Quality and Outcomes*, 14(4), e007602. <https://doi.org/10.1161/CIRCOUTCOMES.120.007602>
70. Sendak, Mark P, Gao, M., Ratliff, W., Nichols, M., Bedoya, A., O'Brien, C., & Balu, S. (2021). Looking for clinician involvement under the wrong lamp post: The need for collaboration measures. *Journal of the American Medical Informatics Association*, ocab129. <https://doi.org/10.1093/jamia/ocab129>
71. Sendak, Mark P, Gao, M., Ratliff, W., Whalen, K., Nichols, M., Futoma, J., & Balu, S. (2021). Preliminary results of a clinical research and innovation scholarship to prepare medical students to lead innovations in health care. *Healthcare (Amsterdam, Netherlands)*, 9(3), 100555. <https://doi.org/10.1016/j.hjdsi.2021.100555>
72. Sendak, Mark P, Roth, C., & Balu, S. (2021). AI Transparency: Why We Should Label Algorithms Like Food Products. *Data Science Institute; American College of Radiology*, 20 Oct. 2021, <https://www.acrdsi.org/DSIBlog/2021/10/20/AI-Transparency-Why-We-Should-Label-Algorithms-Like-Food-Products>
73. Xia, M., Kheterpal, M., Wong, S., Park, C., Ratliff, W., Carin, L., & Henao, R. (2021). Malignancy Prediction and Lesion Identification from Clinical Dermatological Images. *ArXiv*, abs/2104.02652
74. Amin, K., Weissler, E. H., Ratliff, W., Sullivan, A., Gao, M., Sendak, M., Balu, S., Francis, S., Bedoya, A., Washington, L., Black, M., Dahhan, T., & Jones, W. S. Natural Language Processing for Early Identification of Pulmonary Embolism [Poster presented]. *Machine Learning for Healthcare (MLHC) Conference*. (2021, August). Award: GlaxoSmithKline (GSK) James B.D. Palmer, MD Memorial Scholarship - Awarded \$10,000 on August 6, 2021 for outstanding research in pulmonary medicine
75. Patel, A.D., Knechtel, W., Smith, S., Gao, M., Nichols, M., Balu, S., Spiritos, M., Sangvai, D., & Aroumougame, V. (2021, August 06). Acute hospital care at home: A retrospective financial analysis for hospital systems [Poster presented]. *Duke University Medical Student Research Day/Durham, NC, USA*
76. Patel, A.D., Knechtel, W., Smith, S., Gao, M., Nichols, M., Balu, S., Spiritos, M., Sangvai, D., & Aroumougame, V. (2021, October 28). Acute Hospital Care at Home: A Retrospective Financial Analysis from One Hospital System [Poster presented]. *Hospital at Home Users Group Annual Meeting/ Virtual, NC, USA*
77. Wong, S.C., Xia, M., Ratliff, W., Park, C., Balu, S., Sendak, M., Gao, M., Nichols, M., Hintze, B., Shah, K., Henao, R., Kheterpal, M. (2021, July 31). Use of Convolutional Neural Networks in Skin Lesion Analysis Using Real World Image and Non-Image Data [Poster presented]. *Machine Learning in Healthcare Conference/Virtual, NC, USA*
78. Aggarwal, N., Ahmed, M., Basu, S., Curtin, J. J., Evans, B. J., Matheny, M. E., Nundy, S., Sendak, M. P., Shachar, C., Shah, R. U., & Thadaney-Israni, S. (2020). Advancing Artificial Intelligence in Health Settings Outside the Hospital and Clinic. *National Academy of Medicine Perspectives*. <https://doi.org/https://doi.org/10.31478/202011f>
79. Bedoya, A., Futoma, J., Clement, M. E., Corey, K., Brajer, N., Lin, A., Simons, M. G., Gao, M., Nichols, M., Balu, S., Heller, K., Sendak, M., & O'Brien, C. (2020). Machine Learning for Early Detection of Sepsis: an Internal and Temporal Validation Study. *Journal of the American Medical Informatics Association*, 3(2), 252–260. <https://doi.org/https://doi.org/10.1093/jamiaopen/ooaa006>
80. Bradley, K. E., Cook, C., Reinke, E. K., Mather, R. C., 3rd, Riboh, J., Lassiter, T., & Wittstein, J. R. (2020). Comparison of the Accuracy of Telehealth Examination versus Clinical Examination in the Detection of Shoulder Pathology. *Journal of shoulder and elbow surgery*, S1058-2746(20)30689-3. Advance online publication. <https://doi.org/10.1016/j.jse.2020.08.016>
81. Brajer, N., Cozzi, B., Gao, M., Nichols, M., Revoir, M., Balu, S., Futoma, J., Bae, J., Setji, N., Hernandez, A., & Sendak, M. (2020). Prospective and External Evaluation of a Machine Learning Model to Predict In-Hospital Mortality of Adults at Time of Admission. *JAMA Network Open*, 3(2), e1920733. <https://doi.org/10.1001/jamanetworkopen.2019.20733>
82. Chen, X. J., LaPorte, E. T., Collins, L. M., Patel, P., Karra, R., & Mainsah, B. O. (2020). Feasibility of Heart Sound Analysis in Individuals Supported with Left Ventricular Assist Devices. *ArXiv:2002.12305 [Eess]*. <http://arxiv.org/abs/2002.12305>
83. Clement, M. E., Nicchitta, B., Ying, S., Ellis, A., Chakraborty, H., McGee, K., Eagle, C., Frye, J., Taylor, D., Okeke, N. L., Johnston, B., Sena, A. C., & McKellar, M. (2020). PrEP Outcomes in an Urban Community in North Carolina: Discontinuation of Care and Sexually Transmitted Infections. *Sexually Transmitted Diseases*. <https://doi.org/10.1097/OLQ.0000000000001288>
84. Corey, K. M., Helmkamp, J., Simons, M., Curtin, L., Marsolo, K., Balu, S., Gao, M., Nichols, M., Watson, J., Mureebe, L., Kirk, A. D., & Sendak, M. (2020). Assessing Quality of Surgical Real-World Data from an Automated Electronic Health Record Pipeline. *Journal of the American College of Surgeons*, S1072751520300612. <https://doi.org/10.1016/j.jamcollsurg.2019.12.005>

Publications / continued

85. Denny, T. N., Andrews, L., Bonsignori, M., Cavanaugh, K., Datto, M. B., Deckard, A., DeMarco, C. T., DeNaeyer, N., Epling, C. A., Gurley, T., Haase, S. B., Hallberg, C., Harer, J., Kneifel, C. L., Lee, M. J., Louzao, R., Moody, M.A., Moore, Z., Polage, C. R., ... Wolfe, C. R. (2020). Implementation of a Pooled Surveillance Testing Program for Asymptomatic SARS-CoV-2 Infections on a College Campus — Duke University, Durham, North Carolina, August 2–October 11, 2020. *MMWR. Morbidity and Mortality Weekly Report*, 69(46). 1743-1747. <https://doi.org/10.15585/mmwr.mm6946e1>
86. Elish, M. C., & Watkins, E. A. (2020). Repairing Innovation: A Study of Integrating AI in Clinical Care. *Data & Society*. www.datasociety.net
87. Sandhu S, King Z, Wong M, Bissell S, Sperling J, Gray M, Ratliff W, Herring K, LeBlanc T. (2020). Implementation of Electronic Patient-Reported Outcomes in Routine Cancer Care at an Academic Center: Identifying Opportunities and Challenges. *JCO Oncology Practice* 16:11, e1255-e1263. <https://ascopubs.org/doi/abs/10.1200/OP.20.00357>
88. Sandhu, S., Lin, A. L., Brajer, N., Sperling, J., Ratliff, W., Bedoya, A. D., Balu, S., O'Brien, C., & Sendak, M. P. (2020). Integrating a Machine Learning System Into Clinical Workflows: Qualitative Study. *Journal of Medical Internet Research*, 22(11), e22421. <https://doi.org/10.2196/22421>
89. Sendak, Mark P, Gao, M., Brajer, N., & Balu, S. (2020). Presenting machine learning model information to clinical end users with model facts labels. *Npj Digital Medicine*, 3(1), 41. <https://doi.org/10.1038/s41746-020-0253-3>
90. Sendak, Mark P, Ratliff, W., Sarro, D., Alderton, E., Futoma, J., Gao, M., Nichols, M., Revoir, M., Yashar, F., Miller, C., Kester, K., Sandhu, S., Corey, K., Brajer, N., Tan, C., Lin, A., Brown, T., Engelbosch, S., Anstrom, K., ... O'Brien, C. (2020). Sepsis Watch: A Real-World Integration of Deep Learning into Routine Clinical Care. *Journal of Medical Internet Research: Medical Informatics*. <https://doi.org/10.2196/15182>
91. Alexander Fenn, Connor Davis, Neel Kapadia, Daniel Buckland, Marshall Nichols, Michael Gao, William Knechtle, Suresh Balu, Mark Sendak, B. Jason Theiling. Development of Machine Learning Models to Predict Admission from ED to Inpatient and Intensive Units. Presented at 2020 Society of Academic Emergency Medicine (SAEM) Southeastern Regional Conference, Greenville, SC. February 22, 2020
92. Clark E, Jolly Graham A, Bellantoni J, Malone D, Knechtle W, White H, Pavon J. Uncovering Errors in Transitions from Hospital to Nursing Home: A Video Telehealth Transitions Conference. NC ACP Meeting 2020 QI Category Winner. February 24, 2020; The Society for Post-Acute and Long-Term Care Medicine 2020 Annual Conference, Chicago, IL; American Geriatrics Society 2020 Annual Meeting, Long Beach, CA
93. Cortina G, Zhong S, Nichols M, Gao M, Ratliff W, Knechtle W, Balu S, Kester K, Lindsay M, Engle J, Bhatta A, Schroder J, Henao R, Sendak M, Podgoreanu M. Development and Validation of a Machine Learning Model to Predict ICU Readmission or Mortality After Discharge From the Cardiothoracic ICU. Presented at the Annual Meetings of the International Anesthesia Research Society, the Association of University Anesthesiologists, the Society of Critical Care Anesthesiologists, and 2020 Machine Learning for Healthcare conference. August 8, 2020; formerly Durham, NC (virtual)
94. Fenn, A. (2020, February 22). Development of Machine Learning Models to Predict Admission from ED to Inpatient and Intensive Units [Poster Presentation]. 2020 Society of Academic Emergency Medicine (SAEM) Southeastern Regional Conference, Greenville, SC. <https://ghsme.ethosce.com/courses/2020SAEM> <https://www.youtube.com/watch?v=QImENo3BbA>. Best Student Poster February 2020
95. Fenn, A. (2020, May 12). Development of Machine Learning Models to Predict Admission from ED to Inpatient and Intensive Units [Oral Abstract]. 2020 Society of Academic Emergency Medicine (SAEM) National Conference, Denver, CO [Cancelled due to COVID-19]. <https://www.youtube.com/>
96. Jedrek Wosik, Mike Revoir, Patrik Doshi, Will Knechtle, Suresh Balu, Mark Sendak, Will Ratliff, James Teheng. Chest Pain Assessment Tool (CPAT): A Real-time Clinical Decision Support Aid for Evaluation of the ED Patient with Chest Pain. Presented at AMIA 2020 Virtual Annual Symposium: HL7 FHIR Applications Competition. November 17, 2020
97. Ratliff W, Wegermann Z, Shi H, Gao M, Sendak M, Kansal A, O'Brien C, Skove S, Zhao S, Kashyap S, Nichols M, Jones S, Patel C, Katz J, Balu S, Kochar A, Patel, M. Early Identification of High Risk Cardiac Decompensation Phenotypes via Real-time Electronic Health Record Data. Poster presented at: American Heart Association Scientific Sessions 2020. November 13, 2020; formerly Dallas, TX (virtual)
98. Sendak, Mark P, et al. "The Human Body Is a Black Box": Supporting Clinical Decision-Making with Deep Learning." Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, ACM, 2020, pp. 99–109, <https://doi.org/10.1145/3351095.3372827>
99. Shaikh Z, Witt D, Shen T, Ratliff W, Shi H, Gao M, Nichols M, Sendak M, Balu S, Osborne K, Kumar K, Jackson K, McCrary A, Li J. Development of Machine Learning Models for Early Prediction of Clinical Deterioration in Pediatric Inpatients. Poster presented at: Machine Learning for Healthcare 2020. August 8, 2020; formerly Durham, NC (virtual)

NEXT PAGE →



Publications / continued

100. Shi H, Ratliff W, Sendak M, Gao M, Nichols M, Revoir M, Balu S, Zhao S, Pencina M, Kester K, Jones WS, Patel C, Katz J, Kansal A, Kochar A, Wegermann Z, Patel M. Predicting Cardiac Decompensation and Cardiogenic Shock Phenotypes for Duke University Hospital Patients. Poster presented at: Machine Learning for Healthcare 2020. August 8, 2020; formerly Durham, NC (virtual) <https://doi.org/10.1016/j.jamcollurg.2019.08.203> Kansal, A., Huang, Z., Rockhold, F. W., Baumgartner, I., Berger, J. S., Blomster, J. I., et al. (2019). Impact of Procedural Bleeding in Peripheral Artery Disease. *Circulation: Cardiovascular Interventions*, 12(10), 1–9. <http://doi.org/10.1161/CIRCINTERVENTIONS.119.008069>
101. Si, S.; Wang, R., Dov, D., Wosik, J., Henao, R., Carin, L. Students Need More Attention: BERT-based Attention Model for Small Data with Application to Automatic Patient Message Triage. Poster presented at: Machine Learning for Healthcare 2020. August 8, 2020; formerly Durham, NC (virtual). Tan, C. (2019, March 24). Characterizing Sepsis Encounters Across Community and Quaternary Hospitals Within Academic Health System. *Hospital Medicine* 2019, National Harbor, Md. <https://www.shmabstracts.com/abstract/characterizing-sepsis-encounters-across-community-and-quarternary-hospitals-within-academic-health-system/>
102. Skove S, Shi H, Shen Z, Gao M, Cui M, Nichols M, Balu S, Bedoya A, Tart D, Goldstein B, Ratliff W, Sendak M, O'Brien C. Development of Machine Learning Model to Predict Risk of Inpatient Deterioration. Poster presented at: Machine Learning for Healthcare 2020. August 8, 2020; formerly Durham, NC (virtual)
103. William Ratliff, Zachary Wegermann, Harvey Shi, Michael Gao, Mark Sendak, Aman Kansal, Cara O'Brien, Stephanie Skove, S Zhao, Sehj Kashyap, Chetan Patel, Ajar Kochar. Early Identification of High Risk Cardiac Decompensation Phenotypes via Real-time Electronic Health Record Data. Presented at American Heart Association Scientific Sessions 2020, formerly Dallas, Tx (virtual). 24th of November 2020
104. Wosik, J., Revoir, M., Doshi, P., Knechtel, W., Balu, S., Sendak, M., Ratliff, W., & Tchong, J. (2020, November 17). Chest Pain Assessment Tool (CPAT): A Real-time Clinical Decision Support Aid for Evaluation of the ED Patient with Chest Pain. AMIA 2020 Virtual Annual Symposium: HL7 FHIR Applications Competition
105. Wosik, J., Si, S., Henao, R., Sendak, M., Ratliff, W., Balu, S., Poon, E., Carin, L., Patel, M. Topic Modeling of Patient Portal and Telephone Encounter Messages: Insights from a Cardiology Practice. Poster presented at: Machine Learning for Healthcare 2020. August 8, 2020; formerly Durham, NC (virtual)
106. Brajer, N., Cozzi, B., Gao, M., Revoir, M., Nichols, M., Futoma, J., Bae, J., Setji, N., Balu, S., Hernandez, A., & Sendak, M. (2019). Prospective and External Evaluation of a Machine Learning Model to Predict In-Hospital Mortality. *MedRxiv*, 19000133. <https://doi.org/10.1101/19000133>
107. Corey, K., Helmkamp, J., Kirk, A. D., Balu, S., Thompson, D., Mureebe, L., Watson, J., Marsolo, K., Curtis, L., & Sendak, M. (2019). Assessing Quality of Real-World Data Supplied by an Automated Surgical Data Pipeline. *Journal of the American College of Surgeons*, 229(4), S89 <https://doi.org/10.1016/j.jamcollurg.2019.08.203> Kansal, A., Huang, Z., Rockhold, F. W., Baumgartner, I., Berger, J. S., Blomster, J. I., et al. (2019). Impact of Procedural Bleeding in Peripheral Artery Disease. *Circulation: Cardiovascular Interventions*, 12(10), 1–9. <http://doi.org/10.1161/CIRCINTERVENTIONS.119.008069>
108. Lewinsky, A. A., Drake, C., Shaw, R., Jackson, G. L., Bosworth, H. B., Oakes, M., Gonzales, S., Jelesoff, N. E., & Crowley, M. J. (2019). Bridging the Integration Gap between Patient-Generated Blood Glucose Data and Electronic Health Records. *Journal of the American Medical Informatics Association*, 26(7), 667–672. <https://doi.org/10.1093/jamia/ocz039>
109. Lin, A., Sendak, M., Bedoya, A. D., Clement, M. E., Brajer, N., Futoma, J., Bosworth, H. B., Heller, K. A., & O'Brien, C. L. "Evaluating Sepsis Definitions for Clinical Decision Support against a Definition for Epidemiological Disease Surveillance." *BioRxiv*, no. 648907, May 2019, <https://doi.org/https://doi.org/10.1101/648907>
110. Sendak, Mark P, Elish, M., Gao, M., Futoma, J., Ratliff, W., Nichols, M., Bedoya, A., Balu, S., & O'Brien, C. (2019). "The Human Body is a Black Box": Supporting Clinical Decision-Making with Deep Learning. *ArXiv:1911.08089 [Cs]*. <http://arxiv.org/abs/1911.08089>
111. Sendak, Mark P, Gao, M., Nichols, M., Lin, A., & Balu, S. (2019). Machine Learning in Health Care: A Critical Appraisal of Challenges and Opportunities. *EGEMS (Washington, DC)*, 7(1), 1. <https://doi.org/10.5334/egems.287>
112. Simons, M. G., Futoma, J. D., Gao, M., Corey, K., Sendak, M., Whalen, K. B., Doshi-Velez, F., McGee, A. S., & Setji, T. (2019). 1185-P: Predictive Model for Hyperglycemic Events after High Dose Corticosteroid Administration. *Diabetes*, 68(Supplement 1), 1185-P. <https://doi.org/10.2337/db19-1185-P>
113. Vasavda, C., Kothari, R., Malla, A. P., Tokhunts, R., Lin, A., Ji, M., et al. (2019). Bilirubin Links Heme Metabolism to Neuroprotection by Scavenging Superoxide. *Cell Chemical Biology*, 26(10), 1450–1460.e7. <http://doi.org/10.1016/j.chembiol.2019.07.006>
114. Wiens, Jenna, Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., Jung, K., Heller, K., Kale, D., Saeed, M., Ossorio, P. N., Thadaney-Israni, S., & Goldenberg, A. (2019). Do no harm: a roadmap for responsible machine learning for health care. *Nature Medicine*, 25(9), 1337–1340. <https://doi.org/10.1038/s41591-019-0548-6>
115. Wosik, J., Shijing, S., Henao, R., Carin, L., & Patel, M. R. (2019). Artificial Intelligence to Identify Commonly Asked Questions via an Electronic Patient Portal: Lessons from a Cardiology Department within a Large Health System. *Journal of the American Heart Association*, 140(Suppl_1). Pre-2019

Publications / continued

116. Christelle Tan, Kristen Corey, Mark Sendak, Michael Gao, Marshall Nichols, Mike Revoir, Armando Bedoya, Suresh Balu, Cara O'Brien. Characterizing Sepsis Encounters Across Community and Quaternary Hospitals Within Academic Health System. Poster presented at Hospital Medicine meeting, National Harbor, Md. March 24, 2019
117. Fenn A, Davis C, Kapadia N, Buckland D, Nichols M, Gao M, Knechtle W, Balu S, Sendak M, Theiling Jason B. Using Machine Learning in Emergency Department Patient Flow. 2019 Duke AI Health Data Science Showcase. November 25, 2019; Durham, NC
118. Kansal A, Kashyap S, Ratliff W, Sendak M, Sriram K, Simons M, Corey K, Nichols M, Gao M, Futoma J, Revoir M, Balu S, Pencina M, Kester K, Miller C, Wegermann Z, Granger C, Schroeder J, Milano C, Patel M, Jones S, Patel C, Kochar A. Using Predictive Mortality and Cardiogenic Shock Identification Tools to Support Team Based Treatments. Poster presented at: Machine Learning for Healthcare 2019. August 10, 2019; Ann Arbor, MI
119. Tan, Christelle. Characterizing Sepsis Encounters Across Community and Quaternary Hospitals Within Academic Health System. <https://www.shmabstracts.com/abstract/characterizing-sepsis-encounters-across-community-and-quaternary-hospitals-within-academic-health-system/>. Hospital Medicine 2019, National Harbor, Md
120. Berkowitz, C. M., Allen, D. H., Tenhover, J., Zullig, L. L., Fischer, J. E., Pollak, K. I., Hicks, M. R., Hillson, J. V., & Koontz, B. F. (2018). Head and Neck Cancer Survivors: Specific Needs and Their Implications for Survivorship Care Planning. *Clinical Journal of Oncology Nursing*, 22(5), 523–528. <https://doi.org/10.1188/18.CJON.523-528>
121. Corey, K. M., Kashyap, S., Lorenzi, E., Lagoo-Deenadayalan, S. A., Heller, K., Whalen, K., Balu, S., Heflin, M. T., McDonald, S. R., Swaminathan, M., & Sendak, M. (2018). Development and validation of machine learning models to identify high-risk surgical patients using automatically curated electronic health record data (Pythia): A retrospective, single-site study. *PLoS Medicine*, 15(11), e1002701. <https://doi.org/10.1371/journal.pmed.1002701>
122. Egger, H. L., Dawson, G., Hashemi, J., Carpenter, K. L. H., Espinosa, S., Campbell, K., Brotkin, S., Schaich-Borg, J., Qiu, Q., Tepper, M., Baker, J. P., Bloomfield, R. A., & Sapiro, G. (2018). Automatic emotion and attention analysis of young children at home: a ResearchKit autism feasibility study. *Npj Digital Medicine*, 1(1), 20. <https://doi.org/10.1038/s41746-018-0024-6>
123. Elish, M. C. (2018). The Stakes of Uncertainty: Developing and Integrating Machine Learning in Clinical Care. Social Science Research Network, 2018 Epic Proceedings. <https://ssrn.com/abstract=3324571>
124. Helmkamp, J. K., Bullock, G. S., Amilo, N. R., Guerrero, E. M., Ledbetter, L. S., Sell, T. C., & Garrigues, G. E. (2018). The clinical and radiographic impact of center of rotation lateralization in reverse shoulder arthroplasty: a systematic review. *Journal of Shoulder and Elbow Surgery*, 27(11), 2099–2107. <http://doi.org/10.1016/j.jse.2018.07.007>
125. Kansal, A., Long, C. A., Patel, M. R., & Jones, W. S. (2018). Endovascular treatment of femoro-popliteal lesions. *Clinical Cardiology*, 42(1), 175–183. <http://doi.org/10.1002/clc.23098>
126. O'Brien, S. M., Feng, L., He, X., Xian, Y., Jacobs, J. P., Badhwar, V., Kurlansky, P. A., Furnary, A. P., Cleveland, J. C., Lobdell, K. W., Vassileva, C., Wyler von Ballmoos, M. C., Thourani, V. H., Rankin, J. S., Edgerton, J. R., D'Agostino, R. S., Desai, N. D., Edwards, F. H., & Shahian, D. M. (2018). The Society of Thoracic Surgeons 2018 Adult Cardiac Surgery Risk Models: Part 2-Statistical Methods and Results. *The Annals of Thoracic Surgery*, 105(5), 1419–1428. <https://doi.org/10.1016/j.athoracsur.2018.03.003>
127. Patel, M. P., Schettini, P., O'Leary, C. P., Bosworth, H. B., Anderson, J. B., & Shah, K. P. (2018). Closing the Referral Loop: an Analysis of Primary Care Referrals to Specialists in a Large Health System. *Journal of General Internal Medicine*, 33(5), 715–721. <https://doi.org/10.1007/s11606-018-4392-z>
128. Ryan, S., Visgauss, J., Kerr, D., Helmkamp, J., Said, N., Vinson, E., et al. (2018). The Value of MRI in Distinguishing Subtypes of Lipomatous Extremity Tumors Needs Reassessment in the Era of MDM2 and CDK4 Testing. *Sarcoma*, 2018, 1–7. <http://doi.org/10.1155/2018/1901896>
129. Shahian, D. M., Jacobs, J. P., Badhwar, V., Kurlansky, P. A., Furnary, A. P., Cleveland, J. C., Lobdell, K. W., Vassileva, C., Wyler von Ballmoos, M. C., Thourani, V. H., Rankin, J. S., Edgerton, J. R., D'Agostino, R. S., Desai, N. D., Feng, L., He, X., & O'Brien, S. M. (2018). The Society of Thoracic Surgeons 2018 Adult Cardiac Surgery Risk Models: Part 1-Background, Design Considerations, and Model Development. *The Annals of Thoracic Surgery*, 105(5), 1411–1418. <https://doi.org/10.1016/j.athoracsur.2018.03.002>
130. Zullig, L. L., Ramos, K., Berkowitz, C., Miller, J. J., Dolor, R. J., Koontz, B. F., et al. (2018). Assessing Key Stakeholders' Knowledge, Needs, and Preferences for Head and Neck Cancer Survivorship Care Plans. *Journal of Cancer Education: the Official Journal of the American Association for Cancer Education*, 34(3), 584–591. <http://doi.org/10.1007/s13187-018-1345-5>
131. Corey, Kristin M. PYTHIA: Automated Surgical Outcomes Data Pipeline and Prediction Engine. Presented at Machine Learning for Healthcare. Duke Clinical Research Day. May 17, 2018

NEXT PAGE →



Publications / continued

-
132. Lin, Anthony. Investigating the Clinical and Operational Impact of Sepsis Phenotype Selection. August 2018
133. Livingston, Jack. One Thing Straight: A Posture Coach for Individuals with Parkinson's Disease. 2018 LiveWell Student App Challenge
134. Berkowitz, C. M., Zullig, L. L., Koontz, B. F., & Smith, S. K. (2017). Prescribing an App? Oncology Providers' Views on Mobile Health Apps for Cancer Care. *JCO Clinical Cancer Informatics*, 1, 1–7. <https://doi.org/10.1200/CCI.17.00107>
135. Berkowitz, C., Allen, D. H., Tenhover, J., Zullig, L. L., Ragsdale, J., Fischer, J. E., Pollak, K. I., & Koontz, B. F. (2017). Knowledge and Preferences of Primary Care Providers in Delivering Head and Neck Cancer Survivorship Care. *Journal of Cancer Education: The Official Journal of the American Association for Cancer Education*, 33(6), 1323–1327. <https://doi.org/10.1007/s13187-017-1250-3>
136. Clement, M. E., Seidelman, J., Wu, J., Alexis, K., McGee, K., Okeke, N. L., Samsa, G., & McKellar, M. (2017). An educational initiative in response to identified PrEP prescribing needs among PCPs in the Southern U.S. *AIDS Care*, 30(5), 650–655. <https://doi.org/10.1080/09540121.2017.1384534>
137. Clement, M., Okeke, N. L., Munn, T., Hunter, M., Alexis, K., Corneli, A., Sena, A., McGee, K., & McKellar, M. (2017). Partnerships between a University-Affiliated Clinic and Community Based Organizations to Reach Black Men who have Sex with Men for PrEP Care. *Open Forum Infectious Diseases*, 4(suppl_1), S438–S438. <https://doi.org/10.1093/ofid/ofx163.1110>
138. Futoma, J., Hariharan, S., Sendak, M., Brajer, N., Clement, M., Bedoya, A., O'Brien, C., & Heller, K. (2017). An Improved Multi-Output Gaussian Process RNN with Real-Time Validation for Early Sepsis Detection. ArXiv: 1708.05894 [Stat]. <http://arxiv.org/abs/1708.05894>
139. Patel, M. P., Schulman, A., Shah, K. P., Anderson, J. B., & Polascik, T. J. (2017). Engaging the primary care community to encourage appropriate prostate cancer screening. *Therapeutic Advances in Urology*, 10(1), 11–16. <https://doi.org/10.1177/1756287217735799>
140. Schettini, P., Shah, K. P., O'Leary, C. P., Patel, M. P., Anderson, J. B., Cho, A. H., Long, A. L., Bosworth, H. B., & Cameron, C. B. (2017). Keeping care connected: e-Consultation program improves access to nephrology care. *Journal of Telemedicine and Telecare*, 25(3), 142–150. <https://doi.org/10.1177/1357633X17748350>
141. Sendak, Mark P, Balu, S., & Schulman, K. A. (2017). Barriers to Achieving Economies of Scale in Analysis of EHR Data. A Cautionary Tale. *Applied Clinical Informatics*, 8(3), 826–831. <https://doi.org/10.4338/ACI2017-03-CR-0046>
142. Aubrey Kumm, Kathleen Cambell, Sam Marsan, Jordan Hashemi, Steven Espinosa Rick, Ricky Bloomfield, Dawson Geraldine, Guillermo Sapiro, Petrus J De Vries. Feasibility of a smartphone application to identify young children at risk for Autism Spectrum Disorder in a low-income, community setting in South Africa. Presented May 2017
143. Futoma, J., Hariharan, S., & Heller, K. (2017). Learning to Detect Sepsis with a Multitask Gaussian Process RNN Classifier. *Proceedings of Machine Learning Research* 70. The 34th International Conference on Machine Learning, Sydney, Australia
144. Malhar, Priscille. Closing the Specialty Referral Loop: An Imperative for a Resilient Primary Care Network. Presented at 2017 Society of General Internal Medicine (SGIM) Annual Meeting. April 2017
145. American Medical Informatics Association (AMIA): Guiding Principles for the Duke Connected Care Predictive Modeling Pilot
146. American Medical Informatics Association (AMIA): Rationale and Design for the Duke Connected Care Predictive Modeling Pilot with a Medicare Shared Savings Program Population
147. Annual Duke Health Patient Safety and Quality Conference: E-Consultation Program Improves Access to Nephrology Care (poster). Annual Scientific Session of the American College of Physicians (ACP): Closing the Specialty Referral Loop: Are Referrals from Primary Care to Specialists Actually Completed? (poster)
148. Bayesian Young Statisticians Meeting: Scalable Joint Modeling of Longitudinal and Point Process Data for Disease Trajectory Prediction and Improving Management of Chronic Kidney Disease
149. 12th Annual Duke Health Patient Safety and Quality Conference: What Happens to a Referral?: Life Cycle Analysis of Referrals from Duke Primary Care to Duke Specialty (poster) 2017 Annual Scientific Session of the North Carolina Chapter of American College of Physicians (NCACP): Improving Access to Nephrology Care through an E-Consultation Program (poster) 2017 Annual Scientific Session of the North Carolina Chapter of American College of Physicians (NC ACP): Closing the Specialty Referral Loop: Are Referrals from Primary Care to Specialists Actually Completed? (poster)
150. Machine Learning in Health @ NIPS: Transfer Learning via Sparse Latent Factor Analysis to Improve Prediction of Surgical Outcomes
151. Machine Learning in Health Care: CALYPSO: Precision Medicine at Point-Of-Care Management of Surgical Complications
-

Publications / continued

152. Machine Learning in Health Care: Scalable Modeling of Multivariate Longitudinal Data for Prediction of Chronic Kidney Disease Progression
153. Society for General Internal Medicine (SGIM) 2016: Developing a Data-Driven Workflow for Population Health Rounding
154. Society of General Internal Medicine (SGIM) Annual Meeting: Closing the Specialty Referral Loop: An Imperative for a Resilient Primary Care Network (oral)
155. The Institute for Operations Research and the Management Sciences, 2017 Optimizing Operations and Outcomes Conference (Informs Healthcare 2017): Accurate and Timely Real-time Prediction of Sepsis Using an End-to-end Multitask Gaussian Process Rnn Classifier
156. UAI 2016 Workshop on Bayesian Applications: Scalable Joint Modeling of Longitudinal and Point Process Data for Disease Trajectory Prediction and Improving Management of Chronic Kidney Disease
157. Uncertainty and Artificial Intelligence: Scalable Joint Modeling of Longitudinal and Point Process Data for Disease Trajectory Prediction and Improving Management of Chronic Kidney Disease
158. Schulman, K. A., Balu, S., & Reed, S. D. (2015). Specialty Pharmaceuticals for Hyperlipidemia--Impact on Insurance Premiums. *The New England Journal of Medicine*, 373(17), 1591–1593. <https://doi.org/10.1056/NEJMp1509863>
159. Hirsch, B. R., Balu, S., & Schulman, K. A. (2014). The impact of specialty pharmaceuticals as drivers of health care costs. *Health Affairs (Project Hope)*, 33(10), 1714–1720. <https://doi.org/10.1377/>

DIHI Collaborators

Aatif Husain, MD	Becky Smith, MD	David Ming, MD
Abhi Kapuria, MD	Blake Cameron, MD, MBI	David Thompson, MD
Aditee Narayan, MD	Bradi Granger, MSN, PhD	Dawn Dolby, RN
Adrian Hernandez, MD	Brandi Bottiger, MD, MBI	Deborah Engle, EdD, MS
Afreen Shariff, MD	Brenna Hughes, MD, MSc	Deborah Kaye, MD, MS
Aimee K. Zaas, MD, MHS	Brian Burrows, MD	Debra L. Sudan, MD
Ajar Kochar, MD, MHS	Brian Griffith, MD, MMCi	Dennis Narcisse, MD
Alexander Borun, MD	Brian Halstater, MD	Derek Robinson, MBA
Alice Cooper, OGNP	Brittany Davidson, MD	Devdutta Sangvai, MD, MBA
Alison Toth, MD	Cara O'Brien, MD	Devon Check, PhD
Allan Kirk, MD, PhD	Carey Unger, MHA	Dustin Tart, RN
Amanda Bisset, MD	Carly Kelley, MD	E. Hope Weissler, MD
Amy Podraza, MBA	Catherine Kuhn, MD	Elisabeth Kidd, PA
Andrew Barbas, MD	Catie Dunn, BSN	Emily Norboge
Andrew Berchuck, MD	Charles Gerardo, MD	Emily Sterrett, MD
Andrew Godfrey, MD	Charlie Kneiffel, PhD	Eric Poon, MD, MPH
Andrew Muir, MD, MHS	Chet Patel, MD	Erik Carvalho, PT, DPT, MHA
Andrii Kuraksa, MBA	Christine Goertz, DC, PhD	Erin Hamtil, RN
Angela Lowenstern, MD, MHS	Christopher R Polage, MD, MAS	Eugenia McPeck Hinz, MD, MS
Angelo Milazzo, MD, MBA	Colette Allen, NP	Gail Shulby, RN, MA, CPPS
Ann McGee, PharmD	Colleen Shannon, JD	Gary Maslow, MD
Ann Reed, MD	Cooper Linton, MHA, MBA	Genie Komives, MD
Anthony Viera, MD, MPH	Cory Miller, RN	Geraldine Dawson, PhD
Aparna Kamath, MD	Crystal Crider, RN	Hadiza Kazaure, MD
Aparna Swaminathan, MD	Dan Bruno, MS	Hayden Bosworth, PhD
Ashraf Habib, MBBCh, MSc	Dan Dardani, CLP	Heather Marstiller, MBA
Aubrey Joly Graham, MD	Daniel Costello, MPA	Heather Talley, MSN
Azalea Kim, MD, MBA, MPA	David Attarian, MD	Heidi White, MD, MHS, MEd, CMD
Azmen Johnson, BA	David Casarett, MD, MA	Helen Egger, MD, MPH
Barbara Griffith, MD	David Leverenz, MD	Henry Rice, MD

DIHI Collaborators

Holly Alvarado, PharmD, CCP	Kasia Konopka, MFA	Lisa Parnell, MD, MPH
Hope Uronis, MD, MHS	Kate Ulrich, MS, BSN, RN	Lisa Pickett, MD
James Tcheng, MD	Kathleen Cooney, MD	Liset Denis, RN
Jared Lowe, MD	Kathryn Pollak, PhD	Loretta Matters, MSN, RN
Jason Theiling, MD	Kathy Brink, RN, BSN	Lynn Bowlby, MD
Jedrek Wosik, MD	Katia Ferguson, MSN	Madhav Swaminathan, MBBS, MD, MMCi
Jeffrey Clarke, MD	Katie Flanagan, MSW, LCSW, ACM-SW	Manesh Patel, MD
Jennifer Gilner, MD, PhD	Katy Clark, RN	Marcia Owen
Jennifer Kennedy-Stovall	Kay Lytle, DNP	Marcus Strickland, MPH
Jennifer Li, MD	Keith Stover, MAcc	Maribeth Kuntz, PA-C
Jessica Seidelman, MD	Kelley Kester, DNP, MSN	Marvin Swartz, MD
Jill Engel, DNP	Kevin Oeffinger, MD	Mary Klotman, MD
Joanna Kipnes, MD, MS	Kevin Shah, MD, MBA	Mary Lindsay, MSN
Jocelyn Wittstein, MD	Kitty O'Hare, MD	Mary Martin, MPA
Joe Futoma, PhD	Kristen Dicks, MD, MPH	Mary Schilder, RN
Joel Boggan, MD, MPH	Kristen Lakis, Mdiv, LCSW	Matt Roman, MHA, MMCi
John Anderson, MD, MPH	Kristen Lott, MSN, BSN	Matt Rougeux, MHA
John Bonini	Kristen Shirey, MD	Matthew Barber, MD, MHS
John Haws, PhD	Lalit Verma, MD	Matthew Ellis, MD
John Herr, BFA	Lance Okeke, MD	Matthew Engelhard, MD, PhD
John Sampson, MD, PhD, MBA, MHSc	Larry Carin, PhD	Matthew Hartwig, MD
Jon Bae, MD	Laura J. Havrilesky, MD, MHSc	Meenal Kheterpal, MD
Jonathan Fischer, MD	Laura Roe, MMCi	Megan Arguello, RN
Joshua Watson, MD	Lauren Siewny, MD	Melissa Bauer, DO
Judy Prewitt, RN	Laurie Snyder, MD, MHS	Meredith Clement, MD
Juliessa Pavon, MD	Lawrence Carin, PhD	Merhri McKellar, MD
Kaivalya Deshpande, MD	Lawrence Greenblatt, MD	Michael Kent, MD
Kara McGee, PA-C, DMS, MSPH	Leah Zulig, PhD, MPH	Michael Lipkin, MD, MBA
Kara Penna, ANP	Lisa Ferrari, MD	Michael Pencina, PhD
Karen Osborne, RN, BSN	Lisa Nadler, MD	Michael Spiritos, MD

NEXT PAGE 



DIHI Collaborators

Michael Yarrington, MD, MCCi

Mihai Podgoreanu, MD

Mike Datto, MD, PhD

Mike Haglund, MD, PhD, MEd

Mo Shahsahebi, MD, MBA

Moira Rynn, MD

Molly Hillenbrand, MD

Momen Wahidi, MD, MBA

Muhammad Abd-El-Barr, MD, PhD

Muhammad Zafar, MD

Mya Sendak, MD

Neel Kapadia, MD

Neel Subramanian, MD

Nelson Chao, MD, MBA

Nirmish Shah, MD

Oren Gottfried, MD

Padma Gulur, MD

Patrick Codd, MD

Paul R. Newman, MHA

Paula McKenzie, RN

Priscilla Ramseur, DNP, RN, CNOR, NEA-BC

Raj Gupta, MD

Raquel R. Bartz, MD, MHA, MMCi

Ricardo Henao, PhD

Richard Chung, MD

Rob Mentz, MD

Rob Odom, MA

Robert Qi, MD

Robin Rasor, MS

Rohit Tejwani, MD

Ryan Smith, MPP, MTS

S. Benjamin Smith, PhD

Sachin Mehta, MD

Sam Fiske, BSN

Sandhya Lagoo-Deenadayalan, MD, PhD

Sara Johnson, MBA

Sarah Tallent, DNP

Sarah Tucker, BSN, exp MHA

Saumil Chudgar, MD, MEd

Sharah Collier, RN

Shelley McDonald, DO, PhD

Shilpa Shelton, MHA

Song Betancur

Sophia Smith, PhD

Steven Lippmann, PhD

Stuart J Knechtle, MD

Stuart Smith, MHA

Suresh Kumar Agarwal, MD

Susan Spratt, MD

Susanna Naggie, MD, MPH

Tammy Uhl, RN

Tara Kinard, RN, MSN, MBA

Teri-Lynne Sennett

Terrence Allen, MBBS

Terry McDonnell, DNP

Theodore Pappas, MD

Thomas LeBlanc, MD, MA

Timothy Dunn, PhD

Tingrui Zhao, MD

Ton Wang, MD, MS

Tres Brown III, BS

W. Schuyler Jones, MD

Walt Hutchins

Walter Kwiatek

Walter Wiggins, MD, PhD

Warren Kibbe, PhD

Wendy Webster, MBA

Yuval Patel, MD, MHS

Zac Ginsberg, MD, MPP

Zach Wegermann, MD

Zachary Cockerham, RN



dihio.org | [@DukelInnovate](https://twitter.com/DukelInnovate)

[linkedin.com/company/DukelInnovate](https://www.linkedin.com/company/DukelInnovate)