

Proceedings of the 13th Conference on Health IT & Analytics



Edited by

Ritu Agarwal PhD¹, Gabriel Brat MD², Guodong (Gordon) Gao PhD¹, Jeffrey McCullough, PhD³

¹Center for Digital Health and Artificial Intelligence (CDHAI), Johns Hopkins University Carey Business School.

²Beth Israel Deaconess Medical Center, Harvard Medical School (Secondary)

³School of Public Health, University of Michigan, Ann Arbor, MI

This project was supported by grant number 1R13HSO28541 from the Agency for Healthcare Research and Quality. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality.

Table of Contents

Proceedings Abstract	5
Current Trends and Future Agenda for Health Information Technology and Health Analytics Research	6
Regular Session 1A: Equity & Fairness (Speed Presentations and Panel Discussion).....	7
Reevaluating Human Values for Patient Care in the Age of Artificial Intelligence.....	7
Racial Differences in Medical Technology Diffusion	8
Counterfactual Fairness Analysis in Treatment of Substance Use Disorder Patients	8
Modeling and Inferring Dynamic Moral Judgments towards Allocations in High-stakes Domains	9
The Health Equity Discourse of Indian Digital Health Interventions: An Application of Large Language Models	10
Regular Session 1B: Telemedicine 2 (Speed Presentations and Panel Discussion)	11
Developing a Framework for Remote Patient Monitoring	11
Does Telehealth Reduce Rural-Urban Care-Access Disparities? Evidence from COVID-19 Telehealth Expansion.....	12
Tele-Follow-Up and Outpatient Care	12
Novel population health management systems: Insights into the evolution process of a telemonitoring system.....	13
Caring for Rural Health: Socially-oriented Organisations Catalysing Healthcare Provision through Technology-enabled Social Innovations	14
When Less is More: An Analytical Analysis and Empirical Evidence of a Telemedicine Platform.....	14
Regular Session 2A: Equity & Fairness	15
When Systemic Biases Taint Algorithms: A Path to More Equitable Access in Healthcare	15
People Talking and AI Listening: How Stigmatizing Language in EHR Notes Affects AI Performance and Fairness	16
FAIR-DSP: Fair Dynamic Survival Prediction on Longitudinal Electronic Health Record	16
Regular Session 2B: Novel Techniques	17
Federated Causal Inference in Heterogeneous Observational Data.....	17
Machine Learning-Based Precision Nudging: Evidence from a Randomized Field Experiment on a Mobile Healthcare Platform	17
Telemedicine and its impact on Atrial Fibrillation Patients	19
Regular Session 3A: Social Media	20
Influence of Twitter on Hydroxychloroquine Medication Prescriptions for COVID-19 Patients.....	20
The Vanishing Voice: Disappearing Local Newspapers and Public Health Outbreaks	21
Who is Your Doctor and Why? The Role of Information Signaling in Physician Selection.....	21
Regular Session 3B: Health Services	22
AI Adoption and Access to Healthcare Resources: An Empirical Analysis of Mental Health Therapies	22
Do You Remember What Your Doctor Instructed? Impacts of Length, Readability and Diagnosticity of Discharge Instructions on Out of Hospital Mortality.....	23
Do Patient Navigators Improve Health Outcomes?	24

Regular Session 4A: Social Media	26
Combating health misinformation: A multi-level analysis of discussion ratio, temporal, and sentiment effects of health news stories on Twitter	26
Gender Bias in Physician Communities of Practice: The Case of Pathologists on Twitter	26
Effect of Warm Glow Nudge and Privacy Control on Online Physician Reviews: Evidence from a Field Experiment.....	27
Communication Accommodation in Online Health Communities: Implications for Patients with Opioid Use Disorder	28
Regular Session 4B: HIT & Providers.....	29
The Effects of Physician Documentation Support.....	29
Informal Hierarchies - Ramifications For Nursing and Interprofessional Communication In German Hospitals	30
Documenting Complexity: A Longitudinal Textual Analysis of Electronic Medical Records	31
Theory-driven Evaluation of Usefulness of Explanations in CDSS.....	31
Can Information Sharing Reduce Diagnostic Bias? Evidence from a Health Information Exchange....	32
Regular Session 5A: Machine Learning /Analytical Methods	33
A novel visualization framework for explaining healthcare AI decisions.....	33
Detecting signs of depression from social media content: A deep learning model enhanced by depressive emotion recognition	34
Predicting Medical Device Recalls: An ML-Based Framework Leveraging Regulatory Submission Characteristics	35
Estimating Heterogeneous Survival Treatment Effects Among Veterans with Metastatic Castrate-Resistant Prostate Cancer: An Application of Causal Inference Using Machine Learning	36
Using Early Engagement Data from a Digital Health Solution to Predict Future Engagement Patterns	36
Regular Session 5B: IT Value/Outcomes	37
Assessing the Impact of Health Information Exchange on Hospital Data Breach Risk	37
Value of Healthcare IT and Advances in Medical Sciences: The Role of AI Research	38
Balancing Technology and Humanity in Palliative Care: The Limits of Digitalization.....	39
Health Information Technology and Patient Outcomes: Evidence from Hacker Attacks	40
Spillover Effects of a Trauma I Unit on Hospital Operations.....	41
Adapting with Health Information Technologies in Joyfully Productive Ways	41
Regular Session 6A: Deep Learning	43
What Symptoms and How Long? An Interpretable AI Approach for Depression Detection in Social Media	43
Diagnosis of Brain Tumors through MRI Images: A Novel Approach using Local Binary Patterns and Convolutional Neural Networks.....	43
Did we personalize? Assessing personalization by an online reinforcement learning algorithm using resampling	44
Detecting the Human Behaviors Associated with Depression via Sensor Signal Analysis: An Agreement-based Self-Attention Deep Learning Approach	45

Unsupervised discovery of biomedical associations from clinical and auxiliary data	45
Fast Multi-Dimensional Subset Scan With Sparse Tensor Decomposition For Pattern Detection	46
Regular Session 6B: HIT & Patient Outcomes	47
Impact of telehealth on appointment adherence in ambulatory care	47
What Factors Affect the Digital Technology Adoption by the Elderly for Daily Life and Health Management - a UTAUT Analysis.....	48
Balancing Patient Convenience and Healthcare Costs: A personalized recommendation approach using reinforcement learning.....	49
Evaluation and Revision of a Postpartum Depression Risk Prediction Model in Preparation for Implementation	50
Hybrid Artificial intelligence and Inverse Learning for Diet Recommendation.....	52
Regular Session 7A: AI in Action	53
Will Machines Take Over? Algorithms for Human-Machine Collaborative Decision Making in Healthcare.....	53
Artificial Intelligence on Call: The Physician's Decision of Whether to Use AI in Clinical Practice.....	53
Love and Robots: An Empirical Investigation of Warehouse Automation and Worker Injury	54
Regular Session 7B: Telemedicine 1	55
Does Telemedicine Affect Physician Decisions? Evidence from Antibiotic Prescriptions.....	55
How Does the Adoption of Telehealth and Telepsychiatry Services Level in US counties Affect Community Mental Health Outcomes: Results from a Quasi-Natural Experiment	56
Does Telehealth Mitigate or Exacerbate Privacy Concerns in Mental Healthcare Delivery?.....	57
Regular Session 8A: M-health	58
A Fully Automated and Culturally Adapted mHealth Intervention for Diabetes Self-Management Among Vietnamese Adults	58
Covid-19 Apps - Population Health Crisis Management and Impact around the world	60
Regular Session 8B: HIT & Industry Structure	61
Are FDA Enforcement Letters the Cure for Misleading Pharma Ads? An Empirical Investigation of Prescription Claims Data	61
Let it Ride! An Empirical Investigation of Problem Gambling and the Implications of Legalized Sports Betting on Societal Outcomes.....	62

Proceedings Abstract

The Conference on Health IT and Analytics (CHITA) is an annual health information technology and analytics research summit, including a doctoral consortium that each year gathers prominent scholars from more than 40 research institutes, and leading policy and practitioner attendees in a vibrant setting to discuss opportunities and challenges in the design, implementation and management of health information technologies and advanced analytics including artificial intelligence and machine learning systems. CHITA's goal is to deepen our understanding of strategy, policy and systems fostering health IT and analytics effective use, to stimulate new ideas with both policy and business implications, and to support the development of a health IT and healthcare advanced analytics research agenda. These Proceedings of the 13th Conference on Health IT & Analytics (CHITA 2023) represent the work of 67 papers and 135 authors working on the digitally-enabled future of healthcare.

Current Trends and Future Agenda in Health Information Technology and Advanced Analytics Research

Ritu Agarwal PhD¹, Gabriel Brat MD², Guodong (Gordon) Gao PhD¹, Jeffrey McCullough, PhD³

¹Center for Digital Health and Artificial Intelligence (CDHAI), Johns Hopkins University Carey Business School.

²Beth Israel Deaconess Medical Center, Harvard Medical School (Secondary)

³School of Public Health, University of Michigan, Ann Arbor, MI

Research Objective: To identify emerging trends, knowledge gaps, and research priorities in health information technology (HIT), analytics and artificial intelligence (AI) from the perspective of multiple stakeholders.

Study Design: The research presented and discussions conducted at the 13th Annual Conference on Health IT and Analytics (CHITA) held on May 4-6, 2023 as a hybrid event were analyzed. A primary aim for the AHRQ-supported CHITA is to promote the development of a research agenda through interactions between stakeholders, including academics, policymakers, and practitioners. Initial program committee discussions were crafted into the call for submissions. Clinician leaders from the health systems and practitioners industry shared their strategic research priorities. Audience questions and comments were documented. Senior researchers led an open discussion session to synthesize a research agenda, and participants were surveyed for priority future research areas. This information was distilled to identify common themes, priority areas and gaps.

Population Studied: Researchers, practice leaders, and policymakers. CHITA 2023 participants included 148 attendees from over 73 organizations representing a diverse mix of academic programs, including schools of public health, medicine, business, policy, and economics as primary disciplines, and policy and practitioner informants.

Regular Session 1A: Equity & Fairness (Speed Presentations and Panel Discussion)

Reevaluating Human Values for Patient Care in the Age of Artificial Intelligence

Monica Lopez¹

1. Cognitive Insights for Artificial Intelligence

Abstract

Artificial intelligence (AI) is rapidly revolutionizing healthcare to diagnose, treat and manage diseases. Innovation to maximize the opportunities of AI in healthcare has, however, created a gap between utility and responsible AI practices. Several challenges that have arisen and require determined efforts to solve include data diversification and non-biased models; model explainability and algorithmic transparency; provider and patient education on the current limitations of the technology; and overall human-centered AI considerations of ethics, fairness and human well-being. As AI-enabled use cases proliferate and the integrated use, for example, of clinical decision support systems blurs with commercial mobile technology to monitor patient health outside of the hospital, responsible AI practices become fundamental to not just improving the healthcare outcomes of all, but protecting patients from possible violation of their rights and any resulting harms. To address these concerns, emerging soft law guidance and regulatory regimes both globally and in the U.S. offer a starting point. As examples, the World Health Organization has established guidelines for the ethical use of AI in healthcare; they include recommendations for transparency, accountability, and the protection of patient privacy and confidentiality. For medical devices, the International Medical Device Regulators Forum has developed guidelines for the regulation of AI-enabled medical devices. In the U.S., agencies such as the FDA has established a regulatory framework for software as a medical device; the Office for Civil Rights, within the Department of Health and Human Services, enforces the Health Insurance Portability and Accountability Act which sets standards for the protection of patient health information including regulating the use of AI algorithms that involve the processing of patient data. Additionally, the National Institute of Standards and Technology plays a role in providing guidance on AI, particularly through its AI Risk Management Framework. In this paper, I argue that building responsible AI in healthcare necessitates an end-to-end integration of clearly delineated human values across the AI product development lifecycle from design and development to deployment and real-time monitoring and evaluation. Moreover, it calls for a human-centered approach in which humans remain at the center of the AI lifecycle to ensure that AI technologies are ethical, fair and enhance the human condition. The four guiding principles of fairness, integrity, resilience and explainability (FIRE) become paramount to providing the flexibility required for continuous measurement and reevaluation of outcomes that an AI governance framework ultimately needs to succeed in the long-term. I provide examples of how FIRE can be integrated from the questions we ask, experiments we run and data we collect to the model we determine, implementation we adopt, evaluation methods we use and patient education strategy we promote.

Racial Differences in Medical Technology Diffusion

Kartik Ganju¹, Seth Carnahan², Brad Greenwood³ and Jeffrey McCullough⁴

1. University of Minnesota, 2. Washington University in St Louis, 3. Temple University, 4. University of Michigan

Abstract

We examine racial differences in the diffusion of surgical robots. Although black patients are generally less likely to obtain the robot, we find that in the first year of adoption, black patients are more likely to be operated on with surgical robots. We are able to rule out the hypothesis that surgeons are training on black patients. Instead, it appears that surgeries are offered to white suburban patients after initial adoption. Expansion at the extensive margin of surgical utilization appears to drive some racial disparities in healthcare.

Counterfactual Fairness Analysis in Treatment of Substance Use Disorder Patients

Motahareh Pourbehzadi¹, Giti Javidi¹, Kimberly Johnson¹ and Ehsan Sheybani¹

1. University of South Florida

Abstract

Substance use disorder (SUD) is a condition that affects an individual's behavior and leads to their inability to control their substance intake such as legal or illegal drugs, alcohol or medications. Despite significant advances in SUD treatment, the efficacy of treatment interventions is still limited and not all individuals benefit equally from available treatments. The treatment process of SUD patients is subjective and can vary based on the cooperation and determination of the patient, as well as their race, ethnicity and gender, which highlights the need for a fair and equitable approach to SUD treatment. Patients in SUD treatment programs go over different stages of screening, comprehensive assessment, stabilization, treatment, and continuing care and recovery support. We can argue that SUD treatment is a patient-centered care. The service offering patterns in SUD treatment are indicators of different types of disparities in the system. These services are affected by environmental factors (e.g. state-level policies), organizational characteristics (clinic attributes), managerial factors and the composition of patients and staff. In other words, similar to any other healthcare outcome, this care path is subject to disparities and potential sources of bias. Counterfactual fairness analysis (CFA) is a recently developed method based on counterfactual reasoning about the outcomes that would have occurred if a different decision had been made. In the SUD treatment context, CFA is used to compare the outcomes of different treatment interventions for different groups of patients. In this study, we will use the secondary data provided by the Florida Department of Family and Children, where we will compare the outcomes of different medication-assisted treatments (MAT) for individuals. Next, we use statistical methods to estimate the counterfactual outcomes of each treatment intervention for each group of patients, which includes the probability of each patient receiving a specific treatment as well as the probability of different outcomes. We can then compare the estimated outcomes of each treatment intervention for each patient group to assess the interventions' fairness. Finally, we will use CFA to help identify the causal mechanisms that

underline disparities in treatment outcomes and their effects on different groups of patients, which can lead to evidence-based care for the treatment of SUD patients.

Modeling and Inferring Dynamic Moral Judgments towards Allocations in High-stakes Domains

Violet Chen¹, Joshua Williams², Derek Leben² and Hoda Heidari²

1. Stevens Institute of Technology, 2. Carnegie Mellon University

Abstract

Practical resource allocation in high-stakes domains, such as public health, education, and more, is fundamental to promoting the society towards long-term justice and equity. For policymakers and social planners, it is important to understand people's moral judgments, which influence their beliefs on who and how allocation policies should prioritize and benefit. In recent years, as AI and ML tools are increasingly applied to inform or automate high-stakes allocation decisions, an overwhelming body of evidence has established the need for AI and ML to reflect people's collective moral values, such as justice and fairness. However, translating these principles into computationally tractable and verifiable forms has proven challenging.

A growing body of work has called on the AI-ethics community to bring stakeholders' judgments into the process of formulating values, such as fairness, for AI. Following this human-centric view, we first posit that moral judgments vary across situations and contexts. Moreover, allocation policies are often sequential in nature, and their deployment leads to changes in moral judgments through shifting key contextual determinants. So, an ethically minded social planner needs to understand and potentially reflect stakeholders' dynamic moral judgments. Understanding such dynamics allows the planner to design effective and acceptable interventions that guide society in the appropriate direction over time. Along with the increasing recognition that moral judgments are context dependent, there has been rising interests in quantifying moral judgments. An important gap in literature is how to concretely capture moral judgments' evolution with the decision contexts. In this work, we consider a stylized setting in which a social planner or policymaker must make a sequence of decisions regarding the allocation of scarce resources in a high-stakes domain. Stakeholders' moral preferences regarding such allocation policies are influenced by various ethical/moral principles. We aim to understand these preferences to evaluate their sensitivity to the history of allocations and the expected future impacts on socially salient groups.

Drawing motivation from medical ethics, we consider several ethical principles relevant to high-stakes allocation. Each principle is reflected via the prioritization over different groups, for example, the elderly population may be prioritized to receive public health resource based on the Prioritarian principle. We propose a framework for modeling and inferring people's moral judgments and the resulting preference over allocation policies. Our framework utilizes a Markov Decision Process (MDP) model to represent sequential allocation. We suppose a stakeholder's moral judgment over alternative policies is characterized by comparing a reward each policy leads to on the MDP. The reward function captures various ethical principles depending on the underlying contexts. Based on the MDP, we design an active learning process to infer a stakeholder's moral preference by recovering the reward function. We illustrate

our model through simulation experiment and human subject study based on allocating scarce medical resources during a hypothetical viral epidemic. Our simulation experiments verify the validity of the proposed modeling and learning approach. The human subject study results reveal that participants' preferences are indeed history and impact-dependent, and there are interesting moral judgment patterns specific to medical resources.

The Health Equity Discourse of Indian Digital Health Interventions: An Application of Large Language Models

Zhixi Chen¹, Diwakar Mohan¹, Meng Zhu¹, Ligia Paina¹, Madhu Gupta², Kapil Goel², Adarsh Bansal² and Munita Jat²

1. Johns Hopkins University, 2. Postgraduate Institute of Medical Education & Research (PGIMER)

Abstract

In India, the health equity gap between slum-dwellers and marginalized groups and the rest of the population is staggering. Compared to traditional interventions, Digital Health Interventions (DHIs) provide “simple, low cost and immensely user-friendly services” and foster more inclusive dialogue and collaborative problem-solving between marginalized communities and health systems which is critical for reimagining modern urban systems that reduce health equity gaps.

This study aimed to explore the health equity discourse of Indian digital health interventions applying large language models. It used text mining to analyze news and policy content on Indian DHIs and identify the use of equity-related language in DHI discourse before and after the COVID-19 pandemic. A total of 1,496 news articles about Indian DHIs were collected through ProQuest and fed into the standard natural language cleaning pipeline for further mining. We finally identified 25 topics.

Equity-related topics were divided into two categories. The first category focuses on community-level initiatives to promote health equity for vulnerable populations through the synergy of Digital Health Interventions (DHIs) with education and economic sectors. 43 articles shed light on gender inequality and discrimination, which prevent millions of people from protecting themselves from sexually transmitted diseases and unwanted pregnancies. DHIs have been proposed as a solution to this issue. Furthermore, 44 articles discussed the initiatives taken by the Indian government to bridge the social equity divide and ensure the welfare of veterans, wherein DHIs are seen as a crucial component of these initiatives.

The second category highlights achievements and challenges in advancing health equity at the societal level, particularly in terms of awareness, accessibility, and affordability. Articles highlighted the public health awareness campaigns as one of the contributors to the Indian health market. The penetration of quality healthcare services in turn created health awareness among the community in remote areas. The development of technology in the telecom industry was also mentioned as a facilitator of the use of telemedicine which increased access to healthcare services. The COVID-19 pandemic has highlighted startling disparities in health insurance penetration among vulnerable sections of society, with affordability remaining a persistent challenge, even as accessibility is gradually improved through DHIs.

This study explores the feasibility of using Natural Language Processing to contextualize the problem prior to piloting the intervention in the target community. Findings from this analysis will be used to inform participatory research methods to improve inclusive and joint problem-solving and decrease the health equity gap in Chandigarh, India.

Regular Session 1B: Telemedicine 2 (Speed Presentations and Panel Discussion)

Developing a Framework for Remote Patient Monitoring

Jennifer Claggett¹, Amol Joshi, Eric Kirkendall, Stacie Petter and Todd Ponzio

1. N/A

Abstract

The COVID-19 pandemic placed unprecedented strains on healthcare systems in diagnosing and treating patients required to isolate in order to reduce community spread of the highly contagious novel coronavirus (Birkmeyer et al., 2020). This situation drove the U.S. federal government and private payers to revise regulations and reimbursement rules and allow greater use of new and existing technologies to care for remote patients (Mann et al., 2020). Remote patient monitoring (RPM) refers to clinicians' capabilities for assessing and adjusting their patients' plan of care by utilizing remotely-gathered data such as vital signs to proactively make medical decisions aimed at improving health outcomes and well-being (Vegesna et al., 2017). However, we still have a limited understanding of how to best develop and deploy RPM solutions to save lives, manage safety risks, and promote public health (Mueller et al., 2020).

We synthesize the literature, and suggest RPM necessitates four core elements as depicted: (1) data collection; (2) transmission and storage; (3) algorithmic analysis; and (4) presentation of remote patient data. Collection of data occurs outside the clinical setting, typically by means of wearable sensors, mobile phones, and/or portable equipment installed in a patient's home or workplace (Baig et al., 2017). We define these core elements and outline the key technologies and enablers as well as the key people and processes for each in our proposed RPM infrastructure framework. After collection, patient data is transmitted via communications networks and stored in cloud-based systems, which involves tackling problems related to network inequities (Saeed et al., 2020, p. 1) and security threats (Adashi & Thomasian, 2020, p. 1). Next, the data is screened and analyzed to find and summarize patterns and trends in individual patients and patient populations. Finally (and oftentimes ignored), the analysis must be presented to clinicians in a way that enables the remotely gathered data from their patients to be meaningfully integrated into the decision-making process for the relevant clinical care pathway (Gold et al., 2018).

By exploring the literature across these four components, we surface key research questions important to each element, which is summarized in a proposed research agenda. We believe this framework will be

useful to research studying remote patient monitoring as it provides a holistic viewpoint of the technologies involved and how those core elements may interact.

Does Telehealth Reduce Rural-Urban Care-Access Disparities? Evidence from COVID-19 Telehealth Expansion

Shujing Sun¹, Guihua Wang¹

1. University of Texas at Dallas

Abstract

Using a unique set of national healthcare claims data, we investigate the effect of telehealth expansion policy on rural-urban healthcare-access disparities. Leveraging a difference-in-differences design, we compare the total number of visits to urban and rural providers before and after telehealth expansion. We find an enlarged disparity in rural-urban healthcare access, with the rural-urban gap in total patient visits increasing by about 3.8% due to telehealth expansion. We then examine the underlying mechanisms that drive the enlarged disparity. We start by decomposing the visits into in-person and telehealth visits, finding that the urban population switches from in-person to telehealth visits and has significantly more access overall. By contrast, the rural population continues with in-person visits and has a much lower uptake of telehealth services. We then examine patient composition by rural and urban providers, and find that the enlarged disparity is due to urban patients having better access to care, especially more telehealth access. We supplement the paper with analyses from the supply and demand sides to understand the mechanisms driving the disparity gap. We also examine barriers to telehealth adoption and the effect of telehealth on care quality. Our findings are important to policymakers, healthcare providers, and researchers seeking to understand the broad implications of telehealth expansion on rural-urban disparities and further promoting and integrating remote care delivery into the healthcare system.

Tele-Follow-Up and Outpatient Care

Gu wei¹, Meng Li², and Shujing Sun³

1. University of Science & Technology Beijing, 2. University of Houston, 3. University of Texas at Dallas

Abstract

Follow-up appointment is a critical part of the healthcare process that benefits both providers and patients. However, due to various barriers, patients often have limited access to follow-up appointments. In this study, we examine the value of telemedicine in improving access to follow-up care by collaborating with a large Asian hospital that sequentially adopted the tele-follow-up service in different departments. Using a staggered difference-in-differences design, we find that the adoption of telemedicine significantly increases the follow-up volume by 54%. Moreover, telemedicine generates positive spillover effects on onsite care provision, with onsite follow-up visits increasing by 10.7% and onsite initial visits increasing by 5.7%. The mechanism tests show that the increased patient volume is jointly driven by changes in individual patients' visit frequency and the patient base. In particular, patients switch from the in-person

to the telemedicine channel for follow-up appointments, and the tele-follow-up service attracts new patients to the hospital for initial care. We also leverage patient heterogeneity to examine the differential demand elasticity and find that tele-follow-up is more valuable to patients with higher transportation costs or comorbid conditions. Finally, we show that tele-follow-up improves patient care quality, as evidenced by a significant reduction in the readmission rate, which reinforces the value of tele-follow-up applications.

Novel population health management systems: Insights into the evolution process of a telemonitoring system

Adrian Yeow¹ and **Shi-Ying Lim²**

1. Singapore University of Social Sciences, 2. National University of Singapore

Abstract

This ongoing case study documents how a large Singaporean healthcare organization launched and evolved a community telemonitoring system to serve novel uses, ranging from a kiosk-based community vital signs monitoring to mobile application-based home monitoring for different diseases. Drawing on interviews, observational, and archival data collected from the team of developers of the telemonitoring system and clinical representatives from the programs, we trace the evolution process of the system.

The first phase of this case study covers the period from May 2020 to December 2021. We note several key ‘pivoting’ events in the evolution process – the first pivot involved the conversion of the purpose-built COVID-19 telemonitoring system to kiosk-based vital signs monitoring support system. The second pivot involved converting the kiosk-based vital signs monitoring system into a mobile application vital signs monitoring and intervention system.

Our initial analysis of our case study reveals various design actions that undergird the pivot events. The first set of design actions – “cognitive design actions” – include: “reaching back in time,” “reaching in,” and “reaching out” to cognitively connect between the form and functions of digital resources. The second set of design actions – “material design actions” – include splitting, extending, and substituting components within the telemonitoring system to support the new uses. Together these design actions helped the healthcare IT team to pivot effectively and leverage available digital resources to build a highly adaptable digital healthcare platform.

This case study contributes to clinical practice and information system (IS) research. From a clinical perspective, we show how a healthcare IT team can work closely with clinical teams to shape and build digital resources for specific population health management programs. From an IS perspective, our case study provides theoretical insights into IT-based design actions and practical recommendations for developing digital healthcare innovations.

Caring for Rural Health: Socially-oriented Organisations Catalysing Healthcare Provision through Technology-enabled Social Innovations

Arti Barapatrey Jain¹, Bikramjit Ray Chaudhuri¹

1. S.P. Jain Institute of Management & Research

Abstract

Social innovation (SI), as an emerging concept, has been researched in different contexts. In healthcare, there is limited research on its relevance to strengthening healthcare service provision and improving outcomes for unserved populations. Most developing and less- developed countries have supply constraints due to inadequate health infrastructure and a shortage of human and material resources for healthcare delivery. This leads to health disparity, specifically in rural areas, due to their reliance on public healthcare systems. However, with the penetration of communication and digital technologies in these areas, technology-enabled SI initiatives hold the promise to fulfill healthcare needs/demands. Thus, this study investigates and analyses the role of technology-enabled SIs and the implied role of socially-oriented organisations driving these innovations for rural populations, in addressing institutional voids. We draw empirical evidence in the Indian context – a developing country with approximately 65% rural population and prevalent institutional voids in the public healthcare systems. Considering qualitative inquiry, we adopt the grounded theory approach of Strauss and Corbin for data analysis and the eventual theoretical framework emerging from primary data. Based on emerging themes, we propose the ‘technology-enabled social innovation ecosystem’ approach at an institutional level through collaborations, capacity building, and scalability. The proposed ecosystem can potentially strengthen the public healthcare systems to achieve universal health coverage, through various health-related benefits for rural populations – aligning with Sustainable Development Goal 3 targets. The study bridges the research gap of adopting a technology-enabled SI approach in rural healthcare while contributing to a widely researched topic of SI in the specific context of technology enablement of SI. The findings theoretically contribute to institutional theory and the ecosystem approach in strategy research by linking technology-enabled SI initiatives as a response to institutional voids in public healthcare.

When Less is More: An Analytical Analysis and Empirical Evidence of a Telemedicine Platform

Jingxuan Geng¹

1. Temple University

Abstract

We consider a telemedicine platform that allows patients to seek for multiple medical diagnoses from doctors online. We derive theoretical models and find that the pricing mechanism alone may yield a downward distortion on price to prevent the over-participation of doctors compared to a centralized benchmark. A platform can address this limitation by capping the number of diagnoses each inquiry receives (i.e., the control limit mechanism). Control limit allows the platform to charge a higher price while maintaining the appropriate number of responses from doctors. When patients are delay sensitive, interestingly, we find that a platform may benefit from a patient’s increased delay sensitivity because it

plays a similar role as the control limit mechanism by discouraging the later arrived doctors' participation. The profit improvement from imposing a control limit mechanism may decrease with the delay sensitivity. When doctors are heterogeneous in their service quality, we discover the undesirable scenario where low-quality doctors drive out all high-quality doctors. A control limit could prevent it from happening by increasing the high-quality doctors' chances of being awarded from patients. We empirically test our predictions using data from a large telemedicine platform and find consistent results.

Regular Session 2A: Equity & Fairness

When Systemic Biases Taint Algorithms: A Path to More Equitable Access in Healthcare

Ozgur Aksoy¹, Mehmet Ayyaci¹, Asunur Cezar² and Srinivasan Raghunathan¹

1. The University of Texas at Dallas, 2. Bogazici University

Abstract

Predictive algorithms can sift through vast amounts of past data and assist in critical decisions in healthcare. For example, health insurance providers use predictive algorithms to identify beneficiaries who will generate higher expenditures. The insurers can intervene by allocating preventive and therapeutic services to high-risk patients, which may result in healthcare savings. Despite its potential for efficiency gains, systemic inequities in historical data can taint algorithmic predictions. We call this phenomenon "social bias." Algorithms tainted with social bias can assign lower risk scores to patients from socially disadvantaged groups, even when these patients have similar health scores as patients from advantaged groups. Explicitly measuring this social bias via social determinants of health (SDoH) and adjusting the decision-making process can help mitigate the unintended consequences of social bias. Our research aims to formulate a bias-aware algorithmic decision-making framework that factors in the origins of social bias while also considering fairness outcomes. To address this research question, we develop a decision analytic model that uses predictive algorithms exhibiting social bias. A social planner/payer, such as the CMS, employs the predictive algorithm and aims for a more equitable allocation of resources for two groups in society: socially disadvantaged and socially advantaged. To achieve this, the social planner adjusts the algorithmic predictions using a social risk score that measures the extent to which each beneficiary is vulnerable. The adjustment can help in reducing the disparity in the allocation of resources. Yet, it may result in more or less disparity in outcomes depending on how intervention effectiveness and the uncertainty around the social risk measurement varies.

People Talking and AI Listening: How Stigmatizing Language in HER Notes Affects AI Performance and Fairness

Yizhi Liu¹, Weiguang Wang², Gordon Gao³ and Ritu Agarwal³

1. University of Maryland, 2. University of Rochester, 3. Johns Hopkins University

Abstract

Electronic medical records (EHRs) are essential for the artificial intelligence (AI) transformation in healthcare. However, clinician biases reflected in EHR notes can lead to AI models inheriting and amplifying these biases, perpetuating health disparities. This study aims to examine how racial biases in EHR notes affect AI performance and fairness. Specifically, we focus on the impact of stigmatizing language (SL) in EHR notes on mortality prediction using a Transformer-based deep learning model and explainable AI (XAI) techniques. Our findings demonstrate that SL written by clinicians impedes AI performance, particularly for African American patients, highlighting SL as a source of racial disparity in AI model development. To explore an operationally more efficient way to mitigate SL's impact, we investigate the patterns in the generation of SL through a clinicians' collaborative network using social network analysis approaches. We identify central clinicians, who have a higher centrality in the collaborative network, as having a stronger impact on AI racial disparity than peripheral clinicians. Removing SL from central clinicians is a more efficient strategy. This study provides actionable insights for responsible AI development and contributes to understanding clinician behavior and EHR note-writing in healthcare.

FAIR-DSP: Fair Dynamic Survival Prediction on Longitudinal Electronic Health Record

Xin Huang¹, Xiangyang Meng¹, Ni Zhao³, Wenbin Zhang², Jianwu Wang¹

1. University of Maryland, Baltimore County, 2. Michigan Tech University, 3. Johns Hopkins University

Abstract

Scarce medical resources and highly transmissible diseases may overwhelm healthcare infrastructure. Fair allocation based on disease progression and fair distribution among all demographic groups is demanded by society. Surprisingly, there is little work quantifying and ensuring fairness in the context of dynamic survival prediction to equally allocate medical resources. In this study, we formulate individual and group fairness metrics in the context of dynamic survival analysis with time-dependent covariates, in order to provide the necessary foundations to quantitatively analyze the fairness in dynamic survival analysis. We further develop a fairness-aware learner (Fair-DSP) that is generic and can be applied to a dynamic survival prediction model. The proposed learner specifically accounts for time-dependent covariates to ensure accurate predictions while maintaining fairness on the individual or group level. We conduct quantitative experiments and sensitivity studies on the real-world clinical PBC dataset. The results demonstrate that the proposed fairness notations and debiasing algorithm are capable of guaranteeing fairness in the presence of accurate prediction.

Regular Session 2B: Novel Techniques

Federated Causal Inference in Heterogeneous Observational Data

Ruoxuan Xiong¹, Allison Koenecke², Michael Powell³, Zhu Shen⁴, Joshua Vogelstein⁵ and Susan Athey⁶

1. Emory University, 2. Cornell University, 3. United States Military Academy, 4. Harvard University, 5. Johns Hopkins University, 6. Stanford Graduate School of Business

Abstract

We are interested in estimating the effect of a treatment applied to individuals at multiple sites, where data is stored locally for each site. Due to privacy constraints, individual-level data cannot be shared across sites. In addition, the sites may have heterogeneous populations and treatment assignment mechanisms. Motivated by these considerations, we develop federated methods to draw inference on the average treatment effects of the combined data across sites. Our methods first compute summary statistics locally using propensity scores and then aggregate these statistics across sites to obtain point and variance estimators of average treatment effects. We show that these estimators are consistent and asymptotically normal. To achieve these asymptotic properties, we find that the aggregation schemes need to account for the heterogeneity in treatment assignments and in outcomes across sites. We demonstrate the validity of our federated methods through a comparative study of two large medical claims databases.

Machine Learning-Based Precision Nudging: Evidence from a Randomized Field Experiment on a Mobile Healthcare Platform

Idris Adjerid¹, Nakyung Kyung², Sriram Somanchi³ and Yoonseock Son³

1. Virginia Tech, 2. National University of Singapore, 3. University of Notre Dame

Abstract

Despite the highly heterogeneous effects of nudges, prior studies mainly focused on assessing the average effect size of the interventions across different groups. However, not considering the heterogeneous effect of nudges can lead to adverse effects from individuals' (e.g., disengaged by interventions—snudges) and firms' (e.g., ineffective distribution of economic incentives) perspectives.

In this study, we aim to better understand the effectiveness of interventions that account for individual heterogeneities in designing the choice architecture of the intervention. Specifically, we assess the effectiveness of nudges on a mobile healthcare platform, which is an appropriate context to understand the heterogeneous effect of interventions. There is high variability in individuals' motivation (or willingness) to engage in physical activities, not only between individuals, but also within individuals across different time periods. In this notion, one of the challenges when nudging individuals to encourage their physical activities is that everyone has different daily routines, and their capacity to exercise varies on a daily basis. Hence, it is essential to customize the intervention at the individual level considering the dynamics of their exercise routines. In doing so, we employ a machine learning technique to personalize nudging to impact individuals' behaviors based on their predicted daily steps (Machine Learning-Based Precision Nudging; MLPN hereafter). Our primary interest lies in assessing the heterogeneous effect of nudging the days individuals are predicted to exercise the least/most (low/high days, hereafter) to examine whether the intervention can maintain, or enhance their exercise level. We also assess the differential effect of economic incentives to assess which days the economic incentives are most effective.

Against this backdrop, we address the following research questions: Is MLPN effective in modifying individuals' exercising behaviors? Are there heterogeneous effects of MLPN on the low and high days?

How can economic incentives be provided most effectively in complement with MLPN? Do the effects of MLPN spill over to the days that are not nudged?

In collaboration with a mobile health app provider with one of the largest user bases in Asia, we conducted a field experiment to investigate the effectiveness of MLPN. We find that MLPN is effective in increasing individuals' number of steps, compared to non-MLPN. Specifically, individuals assigned to the MLPN groups walked about 358 and 569 more steps than those assigned to the non-MLPN and control groups, respectively. Furthermore, when segmenting the MLPN groups into those nudged on the low and high days (LowDay and HighDay groups, respectively), we find that the LowDay group walked as much as the HighDay group on the nudged days. This implies that nudging on the days in which individuals are predicted to walk the least during the week can encourage them to walk as much as the days they are predicted to walk the most during the week. We also observe spillover effects of MLPN on the days that are not nudged, whereas we do not find such effects for random nudging. This implies that MLPN twice a week is sufficient to promote behavioral changes even on the days the individuals are not nudged. In regards to the differential impacts of incentives provided conditional on attaining the proposed step goal, our results show that they further motivate individuals to increase their step counts only when nudged on the low days and do not further increase the step counts when provided on the high days. However, the HighDay group is more likely to attain the step goal than the low day group when conditional incentives are offered. This implies that although offering economic incentives on the high days may seem effective in terms of goal attainment, leading to a higher redemption rate of the incentives, they do not promote additional step counts. Collectively, these results show that MLPN is effective in encouraging individuals' exercising behaviors; in particular, nudging the low days maximizes the impact of MLPN. Furthermore, our results also suggest that incentives should be offered intelligently in consideration of individuals' predicted exercise intensity. Whereas incentives can act as a motivator for individuals to exercise more when provided on the low days, individuals can exhibit a cherry-picking behavior (e.g., only exercising until the proposed goal) without being further motivated, when provided on the high days.

Telemedicine and its impact on Atrial Fibrillation Patients

Muhammad Zia Hydari¹, Rahul Telang², Eric Dueweke³ and Shalini Allam³

1. University of Pittsburgh, 2. Carnegie Mellon University, 3. UPMC Hospitals

Abstract

We study the impact of telemedicine in a large mid-Atlantic hospital system on the process and clinical outcomes of atrial fibrillation patients. Telemedicine became the primary modality for outpatient and ambulatory care at many hospital systems at the onset of the COVID-19 pandemic. As the prevalence and severity of the COVID-19 pandemic receded, hospital systems reverted to in-person encounters as the primary modality, but telemedicine encounters have stayed at levels much higher than those in the pre-COVID period. Hospital systems are thus looking for empirical evidence on the efficacy of telemedicine, especially, if there are any heterogeneities in its efficacy so the hospitals and health plans can nudge patients towards telemedicine when it is efficacious.

Atrial fibrillation (AFib) is an abnormal condition of the heart in which the muscle fibrils or fibers in the atria (upper chambers) undergo rapid and irregular contraction. This irregular contraction causes the atria to lose synchronization with the ventricles (lower chambers) and causes blood to accumulate in the atria. The blood accumulation can in turn lead to clotting and the blood clots in the left atrium may travel out of the heart toward the brain and may cause a stroke. Strokes arising from AFib have high morbidity and mortality and thus Afib patients are often treated with anticoagulant medication. An evidence-based guideline for AFib treatment is to prescribe direct-acting anticoagulants (DOAC) to patients with AFib diagnosis, which reduces the risk of blood clotting and thus reduces the risk of stroke. Currently, the compliance for DOAC prescribing is sub-optimal. Medical practitioners have surmised that telemedicine encounters may improve the medical providers' adherence to the DOAC guidelines as telemedicine can lead to more focused communication between providers and patients.

To examine this hypothesis, we obtained data on approximately 20,000 patients who were diagnosed with AFib from March 2019 to February 2023 and approximately 1,400 of whom had their first post AFib-diagnosis cardiology encounter using telemedicine. Using these data, we study the effect of telemedicine (cf. in-person encounters) on the likelihood of DOAC prescription, any anticoagulant prescription (DOAC or warfarin), stroke hospitalization, mortality, emergency department treat-and-release (EDTR), and days to first cardiology encounter from AFib diagnosis (follow-up-days). We use regression analysis that controls for patient demographic, patient's pretreatment clinical characteristics (including the CHA2DS2-VASc score widely used by clinicians) as well as provider and time fixed effects. Contrary to the hypothesis based on physician expectations, we find that telemedicine encounters resulted in a 4 percent lower likelihood of DOAC or anticoagulant prescriptions on all patients. However, the effects on stroke hospitalization, mortality, EDTR, and follow-up-days are not statistically significant. In our subgroup analysis, we find that the female patients' mortality risk rises by 3 percent for telemedicine encounters, although we do not find any statistically significant differences in the effect of telemedicine on female patients for other outcomes or black patients on all outcomes.

Regular Session 3A: Social Media

Influence of Twitter on Hydroxychloroquine Medication Prescriptions for COVID-19 Patients

Alain Pinsonneault¹, Kartik Ganju², Sai Soundarya Gorthi¹

1. McGill University, 2. University of Minnesota

Abstract

Background: In recent years, social media has been transforming the healthcare industry through its informational support. The recent COVID-19 pandemic, with social distancing measures in place, has substantially increased the dependence of users on these social media platforms to seek and share health information. However, the information on social media tends to be driven by emotional cues and can be inaccurate. Users may rely on this imprecise information to inform their subsequent behavior. To better

understand the influence of social media on users' behavior, we examine if social media discourse affected provider medication behavior during the COVID- 19 pandemic.

Data & Methods: We used data on Twitter discourse regarding Hydroxychloroquine (HCQ) for treating COVID-19 patients. We focused on this discourse as it was highly contentious, with high utilization of the drug even after the WHO recommended against using it. Through a mixed- method approach of human coding and machine learning, we estimated the stance and geolocation of users involved in the discourse. Utilizing pharmacy claims data from the Symphony dataset on the COVID-19 research database, we estimated the proportion of HCQ prescriptions prescribed to COVID-19 patients in the USA in 2020. We matched this with the geographical locations of providers to assess how the volume and stance of Twitter discourse affected the prescription behavior of Hydroxychloroquine.

Results: Our econometric analysis of the influence of volume and stance of discourse on the proportion of prescriptions of HCQ to COVID-19 patients indicates that the higher the volume of discourse supporting the utilization of the drug for treating COVID-19 patients, the higher the proportion of prescriptions. These results are robust to controlling for time-invariant location- specific metrics. Our findings suggest that social media plays a significant role in shaping healthcare providers' prescription behavior. The possible pathways through which Twitter discourse could have influenced medication prescription behavior are discussed.

Conclusion: Our study provides valuable insights into the influence of social media discourse on user behavior. We highlight the pivotal role of social media in physician care decisions and the need for critical evaluation of information shared on social media. We discuss the contribution of our study to social media in healthcare literature and the practical and policy implications of the growing influence of social media on the physician. Future research needs to explore effective strategies to mitigate the potential risks associated with social media use for health information.

The Vanishing Voice: Disappearing Local Newspapers and Public Health Outbreaks

Avinash Collis¹, Fangchen Song¹, Jason Chan¹

1. University of Minnesota

Abstract

Local newspapers disappeared in more than 1000 counties in the United States after being replaced by online websites in the competition for advertising revenue. This study explores the impact of disappearing local newspapers on the severity of public health outbreaks. Specifically, we investigate the effect of disappearing local newspapers on the spread of COVID-19 in all counties of the United States from 2020 to 2021, and the spread of food poisoning outbreaks in affected counties from 2000 to 2020. Our findings reveal that the one-unit change in the disappearing local newspapers per 100,000 people leads to an average of 20.53% increase in COVID-19 infection rates. In addition, we demonstrate the negative impact of disappearing local newspapers on food poisoning outbreaks. After addressing endogenous issues using

the instrumental variable method, we find that in areas where local newspapers have disappeared, there is a 52.48% higher rate of confirmed primary cases of food poisoning per 100 people, a 202% higher rate of confirmed total cases per 100 people, and a 48.41% higher rate of primary cases visiting healthcare providers per 100 people. These effects hold even with instrumented regression models highlighting the robustness of our findings. Overall, our study sheds light on one of the many societal challenges arising from the disappearing local newspapers.

Who is Your Doctor and Why? The Role of Information Signaling in Physician Selection

Soham Ghosh¹, Saravana Jaikumar¹ and Soumyakanti Chakraborty¹

1. Indian Institute of Management Calcutta

Abstract

Physician Review Websites (PRWs) contain information signals which may impact the choice of physician for consultation. Patient-physician interaction involves a high degree of information asymmetry due to the credence nature of healthcare services, which makes evaluation of physician quality difficult. It is noted in previous research that information from a diverse set of patients (or the “crowd”) can provide more credible information regarding physician quality compared to information from individual responses. The information signals received by prospective patients on PRWs are thus expected to be important in this context. There is a dearth of studies synthesizing signaling effects based on visual signals from images and linguistic signals from text reviews and profile information in PRWs. Our research thus seeks to study the joint effect of text and image signals on physician selection.

We extract data on 3,051 physicians from an Indian PRW, collecting information on physician ratings and physician qualifications, years of experience, number of votes, availability of video consultations, and number of clinics at which they practice. We extract the emotional content of reviews from text using a pre-trained transformer-based BERTweet model, with an accuracy of 88.92% on annotated data. We obtain six dimensions of emotion using the model and then condense them into positive or negative emotional content. We use physician profile pictures to obtain two features – i) eye contact, using a pre-trained deep convolutional neural network (CNN), which achieves a precision and recall of 0.99 and 0.87, respectively, on annotated data, and ii) smiling faces, using a pre-trained Haar cascade classifier, with an accuracy of 87.9% on annotated data. We then employ a cross-sectional regression model to assess physician selection, using the change in the number of votes received by the physician between two time periods as a lagged dependent variable. We find that while direct eye contact positively impacts physician selection, the presence of a smile has no effect. We also find that the negative emotional content of physician reviews affects physician selection negatively; however, positive emotional content does not influence selection. Using a moderation approach, we note that as the proportion of negative emotions in the reviews reduces, physicians with high ratings experience a sharper rise in selection compared to those with low ratings.

Our work contributes to the healthcare information systems literature by illustrating the effect of joint signals on physician selection from PRWs. The moderating effect of physician ratings on the emotional

content of reviews illustrates the interplay between ratings and reviews on PRWs. Our findings regarding the proportions of negative emotional content show that the role of emotions in the healthcare domain is not identical to that of e-commerce domains, one reason being the high information asymmetry on such platforms. The results can help PRWs to modify their mechanisms for ranking and/or listing physicians on the platforms by accounting for the combined effect of visual and linguistic signals. PRW managers can also reach out to physicians to explain the importance of building a patient-friendly profile page.

Regular Session 3B: Health Services

AI Adoption and Access to Healthcare Resources: An Empirical Analysis of Mental Health Therapies

Jian Ni¹, Jiang Qian², Meng Zhu¹

1. Johns Hopkins University, 2. The University of Sydney

Abstract

The distribution and allocation of healthcare resources are important indicators to address health inequalities. Telemedicine through virtual visits has become prominent since the pandemic. As an alternative to in-person office visits, virtual visits allow patients from regions with scarce health resources to access high-quality care. The convenience of teletherapy compared to office visits certainly attracts more people to choose online services. This is especially true for non-emergency care like mental health therapies, where the difference in service quality between the online and offline offerings is arguably not substantial. However, patients might face considerable uncertainty when selecting among therapists through traditional webpages that typically contain only texts and profile pictures. Many service platforms have started to offer AI tools to therapists for free. While these AI tools can help promote therapists to patients, they must be enabled by providers manually and require therapists to provide detailed background information about their services as well as themselves. On one hand, the adoption of AI could facilitate preference matching, further helping patients to choose the proper therapists. Yet, on the other hand, more efficient matching might generate increased demand, which could potentially raise the price of therapy, subsequently exacerbating inequalities in access to care. In our study, using a dataset from one of the largest mental health teletherapy platforms in the U.S., we examined over 140,000 therapists and their therapy prices. To address the endogeneity problem, we employed border strategy using the discrete nature of state AI legislation. The border strategy has been used to evaluate the impact of state-level minimum wage act on employment rate and DMA-level advertising effect (Card and Krueger 1994; Shapiro 2016, Wang et al 2022). The two identification assumptions of border strategy, that is, enough variation across states and location of border being not endogenous, are both satisfied in our dataset. We constructed the border experiment on a county-by-county basis at state-level. There are a total of 3,412 counties in the United States. Of these, 1,183 counties meet our border county condition. There are a total of 1,308 county pairs at state borders. We then examined the impact of AI adoption on therapy pricing within border counties using a difference-in-differences approach, which used the data captured

at two time periods to control for the underlying price trend over time. The results revealed that therapists' decision to adopt AI leads to a premium for the price of offline, in-person therapy. To further examine whether this effect varies across different population groups, we employed the casual forest approach to estimate heterogeneity in a random forest. We find that the lower the average county-level household income, the bigger the price increase in in-person therapy. The same pattern of results holds for other variables capturing socioeconomic status, such as education and race. Our results carry policy implications for the nuanced interplay between technology adoption and healthcare inequalities.

Do You Remember What Your Doctor Instructed? Impacts of Length, Readability and Diagnosticity of Discharge Instructions on Out of Hospital Mortality

Gaurav Jetley¹ and Shivendu Shivendu²

1. Colorado State University, 2. University of South Florida

Abstract

During care transitions, lapses in communication is common and can lead to increased likelihood of adverse events. Patient discharge is one of these care transitions where sound communication in terms of discharge instructions is important for proper management of illness post discharge and avoid errors. Written discharge instructions are directed towards the patient and contain information concerning pending clinical actions appropriate for continuity of care post-discharge. These serve as a stable reference point for patient to review later and are critical for minimizing the chances of medical errors and adverse events post discharge. Although it is clear that discharge instructions have the potential to influence the patient's health trajectory post discharge, there is a lack of understanding of the impacts that different styles of discharge instructions can have on post discharge outcomes. In this study, we investigate the impacts of different writing styles of discharge instructions and their impacts on 30-day out-of-hospital mortality. Using a large dataset of ICU stays in the US, econometric methods, host of controls and robustness checks, we make three important findings: (1) We find that increasing the length of the discharge instructions decreases the 30-day mortality rates. (2) The length and readability of discharge instructions are interlinked. Very long instructions which also are less readable (require greater number of years of education to understand and comprehend) increase the 30-day mortality rates. Very long instructions which are relatively readable lowers the 30-day mortality rates. However, less readable instructions which are shorter in nature lower the 30-day mortality rates. While the prior two outcomes are intuitive, the later indicates less readable (more complex) instructions are not always detrimental if combined with shorter instructions. (3) More diagnostic information (unambiguous medical terms) within discharge instructions is beneficial and lowers the rates of 30-day mortality. The depth and diagnosticity of discharge instructions are also interlinked. Longer instructions which are combined with greater amount of diagnostic information lowers the 30-day mortality rates. On the other hand, less diagnostic information in longer instructions increase the rates of 30-day mortality. We find our results to be very robust to a range of tests and sensitivity analysis. One interpretation from these results is that longer discharge instructions are not always beneficial for the patient's post discharge health trajectory or the physician's time. When writing these instructions, physicians can be

more succinct whilst keeping in mind the educational level of the patient as well as increasing the diagnostic information within the discharge instructions. This would not only save the physicians time in noting down the instructions but also benefit the patient's health trajectory post discharge. If the instructions inevitably need to be longer, then physicians should make them more readable as well as increase the diagnostic information within them. The results from this study have direct implications for care providers, managers in healthcare institutions and future healthcare professionals.

Do Patient Navigators Improve Health Outcomes?

Jiajia Qu¹, Raj Sharman², Ying-Chih Sun³ and Indranil Bardhan⁴

1. University of Texas Permian Basin, 2. State University of New York at Buffalo, 3. Harrisburg University of Science and Technology, 4. University of Texas at Austin

Abstract

Navigating the U.S. healthcare delivery system, such as scheduling appointments with specialists, adherence to medication regimens, and communications with healthcare providers, is often challenging. For disenfranchised communities, access to transportation for medical appointments and primary care services poses significant challenges. Medicaid and Medicare patients from poorer neighborhoods are even more at risk. Further, the number of social determinants of health (SDOH) that impact disenfranchised communities is often significant. In these communities, patient medication adherence and appointment no-show rates are problematic.

Patient navigators are professionals who help patients navigate through complex healthcare systems. Navigators inform, educate, motivate, and guide patients and their families to help them make informed decisions about their health care and provide the emotional support necessary to improve health outcomes and reduce healthcare disparities. Our research focuses on an important question, "Do patient navigators have an impact on patient health outcomes?" Specifically, we focus on three measures of health outcomes - Hb1ac, BMI, and avoidable emergency room (ER) visits. Understanding the influence of patient navigators is extremely important as it informs policy on the utility of such intervention programs.

In this study, we examine the impact of a navigator intervention on patients' health outcomes by utilizing patient-level, panel data of over 10,000 patients from a clinic situated in a poor Black and Hispanic neighborhood in Western New York affected by more than one social determinant of health. The data tracks patients from 2015 to 2021 and also includes patients from similar neighborhoods who were not part of the treatment group but were part of the same clinic. Both treatment and control groups are Medicare / Medicaid patients. We study the change in health outcomes - HbA1c, blood pressure, BMI, and ER utilization - among patients enrolled in the patient navigator program. The provider organization recommends patients with higher levels of health risk or lower socioeconomic status to be enrolled in the patient navigator program. However, patients are free to choose whether or not to participate. In our case-controlled study, 62% of patients did not enroll in the navigator program while 38% enrolled. Among enrolled patients, 23% had more than four years of experience with the navigator program.

Our analysis utilizes a matched difference-in-differences approach, controlling for patient age, gender, race, and first language, to minimize the impact of patient selection, individual patient differences, and year effects. Further, we repeated the analysis using propensity score matching (PSM) and inverse probability weighting (IPW) methods. Our analysis compared patient outcomes before and after enrollment in the navigator program with patients who never enrolled. Our results confirm the significant effectiveness of patient navigators in improving health outcomes (HbA1c: $\beta = -0.449, p < .05$; BMI: $\beta = -0.741, p < .01$; BP_Systolic: $\beta = -2.567, p < .01$; BP_Diastolic: $\beta = -1.080, p < .1$). Our study also shows that patient navigator interventions reduce ER visits ($\beta = -0.018, p < .01$), especially among high-risk patients with conditions such as coronary artery disease, congestive heart failure, peripheral vascular disease, and stroke ($\beta = -0.945, p < .01$). We also found that participation in the navigator program for at least two years resulted in significant improvements in reducing unnecessary ER visits for most patients, with stronger impact in the first year for patients with high ER visit frequency. Similarly, for patients with high levels of diabetes (HbA1c > 9), the navigator intervention was associated with an improvement in HbA1c from the first year.

Our findings demonstrate that patient navigator interventions can lead to improved patient health outcomes, especially among high-risk patients. Our findings provide valuable insights for practitioners and researchers seeking to enhance care quality among high-risk populations. Overall, our study makes a significant contribution to the literature on patient navigation and provides a foundation for further research and development of interventions aimed at improving patient outcomes. It also provides supporting evidence for the rollout of patient-focused targeted interventions to improve patient engagement, especially among disenfranchised communities, and promote greater health equity in a value-based healthcare environment.

Regular Session 4A: Social Media

Combating health misinformation: A multi-level analysis of discussion ratio, temporal, and sentiment effects of health news stories on Twitter

Bu Zhong¹, Lola Xie¹, Qian Liu²

1. The Pennsylvania State University, 2. Jinan University

Abstract

Health misinformation has become an unfortunate truism of social media platforms, where lies could spread faster than truth. Despite considerable work devoted to suppressing fake news, health misinformation persists and even increases in recent years. One promising approach to fighting bad information is studying the temporal and sentiment effects of health news stories and how they are discussed and disseminated on social media platforms like Twitter. As part of the effort of searching for innovative ways to fight health misinformation, this study analyzes a dataset of more than 1,600 objectively and independently reviewed health news stories published over a 10-year span and nearly

50,000 Twitter posts responding to them. Specifically, it examines the source credibility of health news circulated on Twitter and the temporal, sentiment features of the tweets containing or responding to the health news reports. We find that health news stories that are rated low by experts are discussed more, persist longer, and produce stronger sentiments than highly rated ones in the tweetosphere. However, the highly rated stories retained a fresh interest in the form of new tweets for a longer period. An in-depth understanding of the characteristics of health news distribution and discussion is the first step toward mitigating the surge of health misinformation. The findings provide insights into understanding the mechanism of health information dissemination on social media and the practical implications to fight and mitigate health misinformation on digital media platforms.

Gender Bias in Physician Communities of Practice: The Case of Pathologists on Twitter

Priyanga Gunarathne¹ and **Ruba Aljafari²**

1. University of Pittsburgh, 2. Virginia Tech

Abstract

Gender disparity and bias in STEM fields in general and in medical sciences continue to be a universal concern. Extant literature in offline contexts documents evidence of gender bias in clinical and academic medicine, where recognition of female physicians and the advancement in their medical careers are hampered by gender perceptions. Gender is typically viewed as a status characteristic that shapes beliefs and expectations for medical professionals, such as performance and productivity. Gender could be particularly influential when a community of practice holds distinct expectations for men and women, as they assume that men are more capable of relevant tasks/activities (e.g., diagnosis) due to their ability to access more resources, or women are more suitable for certain roles (e.g., attributions of warmth).

The emergence of social media has transformed this very idea of communities of practice, facilitating barrier-free interactions among physicians across typical organizational and geographic boundaries. Despite the persistent debate within the medical community as to whether social media is a trustworthy resource for professional development and lifelong learning, physicians are increasingly adopting social media as a platform for professional networking, community interaction, and outreach. Gender stereotypes may play an important role in the context of online physician communities of practice, where opportunities for career advancement and recognition by interacting with peers emerge. It is unclear, however, which gender expectations will prevail and influence peer interactions in physician communities of practice on social media, as such interactions involve nuances that pertain to knowledge engagement and diffusion. On the one hand, peers may interact (e.g., post more comments, likes, or shares) with male physicians more, due to assumptions about their expertise, more-agentic nature, and accessibility to resources (i.e., negative bias towards women). On the other hand, peers may interact with female physicians more, due to assumptions about their tendency to help and care about the community (i.e., positive bias towards women). Despite the lack of quantitative evidence of such phenomena in online contexts, limited effort has been devoted to understanding whether a potential gender bias exists in

online physician communities of practice and how it may differentially impact knowledge engagement and diffusion outcomes on social media.

Using a novel dataset of professional conversations among pathologists on Twitter, we investigate the impact of gender-based stereotypes on knowledge engagement and diffusion outcomes in physician communities of practice on social media. Leveraging 2,860 patient case-based tweets (i.e., case tweets) posted by pathologists on general pathology and across 24 pathology sub-specialties, we employ a holistic set of tweet- and pathologist-specific controls and a matched sample approach to find quantitative evidence of a positive bias in knowledge diffusion (in terms of the number of retweets and likes a case-tweet receives), in favor of female pathologists on Twitter. Interestingly, we do not observe such a bias in knowledge engagement outcomes, in terms of the number of replies a case-tweet receives. Our findings provide key insights for several groups of stakeholders including hospitals, physicians, healthcare providers, and platform owners. Managerial and practical implications are discussed.

Effect of Warm Glow Nudge and Privacy Control on Online Physician Reviews: Evidence from a Field Experiment

Aishwarya Deep Shukla¹, Guodong Gao², Jie Mein Goh¹, Ritu Agarwal², Tianshu Sun³

1. Simon Fraser University, 2. Carey Business School, 3. University of South Carolina Marshall School of Business

Abstract

Patients availing of healthcare services have substantial privacy concerns when writing online reviews for physicians, which may deter them from sharing information that can be welfare enhancing for others. This causes an intrinsic dilemma where patients have to trade-off between privacy and social welfare, leading to less helpful and less informative reviews. We investigate how nudging strategies and privacy controls affect reviewers' privacy decisions related to the disclosure of sensitive information and identity revelation in their reviews to better understand disclosure behavior and manage this trade-off. To address this question, we use a large-scale two-stage field experiment, together with lab experiments to establish causality and address potential confounds. Analysis of data from the field experiment for physician reviews reveals that the warm glow nudge, as compared to the altruism nudge, increases the likelihood of patients revealing sensitive medical information in their reviews by 27.0% and increases the proportion of patients who reveal their identity by 5.9%. Furthermore, positioning the privacy control up front makes patients reveal less sensitive medical information by 18.2%, but it increases the likelihood of identity revelation by 37.9%. Our results shed light on the effects of choice architecture on patient reviewers' responses to a request for public feedback and private information disclosure. The study has implications for how platforms can influence review behaviors of patients with complex motivations and privacy concerns. It contributes insights that are critical for understanding patient choice in a credence-based and privacy-sensitive market like healthcare and interventions to address the privacy-social good tradeoff.

Communication Accommodation in Online Health Communities: Implications for Patients with Opioid Use Disorder

Anton Ivanov¹, Zhasmina Tacheva², Murtaza Nasir³, Ali Dag⁴, Serhat Simsek⁵ and Michael Shaw¹

1. University of Illinois at Urbana-Champaign, 2. Syracuse University, 3. Wichita State University, 4. Creighton University, 5. Montclair University

Abstract

This study examines how linguistic behavior, or the communication style of two or more participants in online discourse, can affect health outcomes of patients with opioid use disorder (OUD) in the context of online health communities (OHC). Building on the communication accommodation theory (CAT), we theorize and empirically test how different communication accommodation strategies (i.e., approximation, interpretability, discourse management, interpersonal control, and emotional expression) influence the recovery process of users with OUD. To assess the linguistic behavior between conversation participants, we situate our study in a unique online environment on Reddit where a large group of people is dedicated to helping each other address the OUD problem. To ensure correct identification of the communication accommodation effects on health outcomes, we rely on the multi-method approach involving econometric modeling, agent-based modeling, and Amazon Mechanical Turk randomized experiment. The results of our initial empirical analysis provide suggestive evidence that of the heterogeneous effects of different strategies on patient outcomes. Based on the results, our study seeks to offer readily implementable solutions for training OUD stakeholders with respect to their communication accommodation styles for patients in online environments.

Regular Session 4B: HIT & Providers

The Effects of Physician Documentation Support

Nate Apathy¹, A Jay Holmgren¹ and Dori Cross¹

1. MedStar Health, 2. University of California - San Francisco, 3. University of Minnesota

Abstract

Documentation support approaches, including collaborative team-based documentation and the use of medical scribes, have gained popularity as ways to reduce excessive physician EHR documentation burden and curb burnout. However, most studies of these interventions take place in a single institution or practice and involve very few physicians, leaving a dearth of generalizable evidence speaking to their effects. Moreover, the few studies that have evaluated the first-order effect of documentation support on physician EHR burden have relied upon self-reported survey measures rather than now widely available EHR log data that captures objective measures of usage.

The purpose of our study was to analyze the first- and second-order effects of documentation support at a national scale. We use national physician-week level EHR use metadata that captures details of note composition – including shared authorship of notes – from Epic Systems, the largest US ambulatory EHR

vendor. Our data include all US-based ambulatory physicians using Epic between September 2020 and April 2021. We exploit variation in timing of physician adoption of collaborative documentation support, operationalized as a change from no shared note authorship to consistent shared note authorship. We measure the effects of this support on physician documentation time, EHR time in evenings and on off days, visit volume, and physician efficiency, measured as the share of visits closed within two days.

We use a staggered adoption two-way fixed effects (TWFE) difference-in-differences approach to analyze changes in our outcomes among physicians who adopted team-based documentation support (n=1,024) compared to physicians with no documentation support during the study period (n=66,972). We exclude physicians with sporadic or constant documentation support (e.g., attending physicians working consistently with residents).

Physicians that adopted documentation support realized a 9.9% decrease in weekly EHR documentation time; this held both in aggregate (-24.5 min per week; 95% CI: [-31.5,-17.5]) and per visit (-0.8 min per visit [-1.1,-0.5]). Total weekly EHR time decreased by 4.5% (-30.0 min per week [-39.5,-20.5]). Documentation support did not lead to decreases in after-hours EHR use (-3.6 min per week [-7.8,0.7]) or time spent in the EHR on off days (-0.9 min per week [-5.0,3.3]). Visit volume increased by 5.7% (2.4 visits per week [1.8,2.9]), while two-day visit closure rates did not change (-0.4 percentage points [-1.2,0.4]).

Our estimates are robust to alternate specifications using a TWFE event study approach and estimates of group-specific average treatment effects, and illustrate substantial and durable decreases in both aggregate documentation time and overall EHR time for physicians that adopt documentation support. We also find substantial increases in visit volume, which may help to explain the null effect that documentation support appears to have on after-hours and off-day EHR use. Our findings highlight an important trade-off between visit volume and after-hours EHR use, a measure correlated with physician burnout. While documentation support may facilitate increases in visit volume to “cover costs,” that increased volume may prevent reductions in salient and burdensome aspects of EHR use.

Informal Hierarchies - Ramifications For Nursing and Interprofessional Communication In German Hospitals

Kirsten Hannah Seerig¹, Maximilian Haug¹, Alexander Maier¹ and Heiko Gewald¹

1. Hochschule Neu-Ulm University of Applied Sciences

Abstract

Hospitals are strictly hierarchically organized based on power, knowledge, and prestige. Despite not being defined formally, nursing is typically subordinated to medicine. Although the effects of strict hierarchies on job satisfaction, collaboration, and error culture are partially known, research lacks an understanding of how exactly hierarchies in hospitals influence the different professions. Additionally, literature is scarce regarding potential solutions to the identified problems.

We address this research gap by examining existing and perceived hierarchies in German hospitals and attempting to derive possible practical implications. Within a qualitative research design, we conducted eleven semi-structured interviews with experts from nursing and medical professions at a German hospital concerning their perception of the hospital hierarchy.

Our findings illustrate that two forms of hierarchies, formal and informal, are prevalent in German hospitals, and these ultimately influence communication within and between the professions. The formal hierarchy is perceived as helpful, especially in emergencies demanding a clear chain of command and responsibilities. The informal hierarchy is mainly enforced by physicians having specific knowledge and implicit power over the decision-making process and consequently imposing influence on nursing. Interaction and communication between nursing and medical professions are harmed by the manifested informal hierarchy. The hospital context facilitates a harsh and sharp communication style that harms the error culture since many lower hierarchy-level employees do not dare to admit or address errors with their superiors for fear of reprisal. The higher hierarchical levels place little value on nurses' needs and sensitivities, resulting in demotivation and discontent.

As one practical implication, we call for a format where physicians and nurses can meet without the burden of hierarchy, ideally with a mediator who is not bound to either side. This way, communication issues can be resolved more easily, giving nurses a greater appreciation from which, ultimately, patients can benefit.

Documenting Complexity: A Longitudinal Textual Analysis of Electronic Medical Records

Kartikeya Bajpai¹, R. Kannan Mutharasan², David M. Liebovitz³, Steven M. LoBue²

1. University of South Florida, 2.

Abstract

In recent decades, the ubiquitous adoption of electronic medical records (EMRs) has created opportunities for improving the value, quality and accountability aspects of patient care. However, these systems are not without costs. For instance, providers increasingly find themselves spending up to an hour documenting for every hour spent on patient care. Further, overloaded providers are also liable to experience burnout, stress and lower satisfaction. While there are proposed solutions such as the redesigning of the documentation process and alternative conceptions of medical records, such ideas have tended to paint with a broad brush, treating medical records in a homogeneous manner across disciplines. We utilize a text analysis approach to explore changes in electronic medical record complexity over time. We present preliminary results from an analysis of a longitudinal dataset of cardiology patient notes (n=~8k, 12 years) and briefly outline our ongoing research.

Theory-driven Evaluation of Usefulness of Explanations in CDSS

Suparna Ghanvatkar¹ and Vaibhav Rajan¹

1. National University of Singapore

Abstract

Explainable Artificial Intelligence (XAI) techniques generate explanations for predictions from AI models. These explanations can be evaluated for (i) faithfulness to the prediction, i.e., its correctness about the reasons for prediction, and (ii) usefulness to the user. While there are metrics to evaluate faithfulness, to our knowledge, there are no metrics to evaluate the usefulness of explanations in the clinical context.

We develop a principled method to evaluate XAI explanations by drawing on theories from social science and accounting for specific requirements of the clinical context. The theories for explanation identify four key requirements from an explanation: (1) contrastive in nature, (2) select number of concepts, (3) generate causal understanding and (4) must consider the user. To account for these, we develop a novel scoring method using literature-derived biomedical knowledge graphs to estimate the association between the clinical concepts used in the explanation and prediction task. These estimates score a given explanation regardless of the technique used for explanation generation and apply to any explanation with a score for the importance of the input features to the particular prediction.

We evaluate our method on a case study of predicting the onset of sepsis in Intensive Care Units. Our analysis shows that the scores obtained using our approach corroborate with independent evidence from clinical literature. Consequently, along with the metrics for prediction performance, our metric can select between different explainable models, such as black-box models like neural networks with post-hoc XAI and interpretable models like logistic regression. As our metric is neither dependent on the data used for training nor on the algorithms used for generating the explanation, it may be extended to a diverse set of modeling strategies, making it useful for future research.

Can Information Sharing Reduce Diagnostic Bias? Evidence from a Health Information Exchange

Indranil Bardhan¹, Minghong Yuan², Wen Wen²

1. The University of Texas at Austin McCombs School of Business, 2. University of Texas

Abstract

Prior literature has documented evidence of variations in physician diagnosis, which negatively impacts some minority groups. For example, van Ryn and Burke (2000) observed a negative perception of physicians toward African Americans and those with low social-economic status. Even race-neutral physicians may be likely to stereotype racial/ethnic minority members (Burgess et al., 2004). Recent studies have shown that such implicit biases can contribute to disparities in healthcare delivery, patient perceptions of care, and health outcomes (Hall et al., 2015; McDowell et al., 2020). The literature suggests that physician biases may be influenced by information access (Ortega, 2023). In particular, when there is lack of objective diagnostic information, physicians may rely on preconceived ideas to make decisions.

The use of health information exchanges (HIE) can potentially mitigate such biases by reducing the barriers to information sharing of patient health data. Prior research has documented that HIE systems may lead to reductions in redundant testing, lower healthcare expenditures, and hospital readmissions (Lammers & McLaughlin, 2017; Vest et al., 2015; Ayabakan et al. 2017). However, little research has been conducted to explore the impact of HIE systems on the diagnostic biases associated with physician decision-making. HIEs provide physicians with access to objective health information about their patients, which may reduce the likelihood of making decisions based on preconceived notions and allow for correction of existing biased beliefs. Motivated by these observations, our research aims to answer two important questions: First, do disparities exist in physician diagnoses, and are they induced by physician biases? Second, how do biases change when physicians have more access to patient health status through access to HIE systems? By addressing these questions, our study will provide deeper insights that can inform efforts to improve healthcare delivery and promote health equity.

Building on previous research that identified systematic differences in acute myocardial infarction (heart attack) diagnosis rates between African American and other individuals in the US (Grines et al., 2021), our study focuses on physician diagnostic behavior in the emergency room (ER) setting for patients with chest pain, a common symptom of heart attack. The ER setting was chosen for the following reason. In such a setting, patients are usually sent to the nearest hospital, and therefore unlikely, to plan a visit *ex ante* to the hospital. Thus, it allows for random assignments of physicians to patients, which is crucial for model identification.

We utilize data from the Integrated Care Collaboration (ICC), a network of healthcare providers in Central Texas. The dataset covers patient visits in Texas from 2015 to 2020, where each visit record includes patient demographics, hospital and physician details, and diagnosis information. Our analysis reveals significant disparities in accurate diagnosis rates between African American patients and other patients. Specifically, African American patients who visit the ER with chest pain are more likely to be diagnosed with heart-related diseases than other patients. Furthermore, even after accounting for potential omitted variable bias in physician diagnosis based on our quasi-experimental approach, the disparities persist around 70% of physicians. In other words, a significant amount of the disparities seem to be driven by physician biases. However, we find that greater use of HIE systems can mitigate the magnitude of the biases. Specifically, we observe a reduction of 12% in the magnitude of physician diagnostic biases in hospitals that use HIEs to share patient health data.

Our research offers a new perspective on identification and measurement of implicit biases in physician decision making. We also contribute to the health IT literature by highlighting the impact of health information sharing on reducing physician diagnosis bias with respect to the treatment and diagnosis of minority patients. Our study suggests significant potentials of using HIE systems to address disparities in the treatment and care of ethnic minority groups, thereby promoting health equity.

Regular Session 5A: Machine Learning /Analytical Methods

A novel visualization framework for explaining healthcare AI decisions

Matthew Baucum¹, Meysam Rabiee² and Babak Aslani³

1. Florida State University, 2. University of Colorado Denver, 3. George Mason University

Abstract

The increased availability of high-dimensional healthcare datasets allows machine learning (ML) models to uncover complex, nonlinear patterns predicting patient outcomes. Yet rather than simply relying on models' 'black box' predictions, researchers and clinicians should be able to see and understand these ML-discovered patterns to aid in hypothesis generation. This research presents a holistic visualization framework, which combines existing and novel ML metrics and plots, that allows for efficient pattern and interaction discovery in healthcare ML models. As part of this framework, we introduce a novel set of feature importance metrics separately quantifying variables' linear, nonlinear, and interaction effects on patient outcomes, which are special cases of the commonly-used permutation feature importance metric. We also introduce an algorithm for estimating the magnitude of all pairwise interactions learned by ML models among k patient variables that runs in $O(k)$ time (rather than $O(k^2)$, as required by existing techniques), which facilitates interaction discovery even in high-dimensional datasets. We combine these novel developments with partial dependence plots (an existing ML tool for visualizing variables' marginal effects on outcomes) to introduce a set of visualizations that allow researchers and clinicians to quickly identify the most important patient variables for predicting outcomes and visualizing the functional form of these effects, even in the presence of nonlinearity and complex interactions. Finally, we validate our proposed metrics and visualizations with extensive simulation studies and with a large dataset of substance use disorder (SUD) treatment outcomes. We show that our novel feature importance and interaction metrics can quickly identify complex patterns predicting which patients are most likely to complete SUD treatment, and which patients are at most risk of treatment dropout or attrition. We conclude by discussing other use cases for medical decision making and health informatics.

Detecting signs of depression from social media content: A deep learning model enhanced by depressive emotion recognition

Fei Peng^{1,2}, Zhijun Yan²

1. Beijing Institute of Technology, Beijing, China 2. University of North Carolina at Charlotte

Abstract

Depression is a serious and popular mental illness that can significantly impact an individual's life. Early detection and treatment of signs of depression are crucial for effective management of the disease. However, traditional diagnosis of depression relies on in-person interviews and specialized questionnaires (e.g., PHQ-9), which is often ineffective due to the lack of patients with depression coming forward and proactively seeking for help from mental health professionals due to unawareness of their mental issues

and social bias, etc. On the other hand, it has been found that people with depression do use social media for sharing their experiences and thoughts as other uses, making user-generated social media posts a unique source for timely detecting depression. The emotion is one of the most popular features that have been used in the existing studies on depression detection from social media content. However, the emotions used in the current studies are also commonly expressed in posts from users without depression about their ups and downs in life. It is challenging to distinguish emotions expressed in a post as a sign of depression from a general bad mood expression of non-depressive users. In addition, these emotions usually are extracted by some existing models and lexicons from text only. But some depressive emotions that can be used to diagnose depression by doctors could not be recognized by existing methods and some features that are valuable for emotion recognition, such as POS (Part-of-speech), have not been considered by existing methods. To address these issues, we propose a novel deep learning model enhanced with emotion recognition (DLEER) to detect signs of depression from social media content. DLEER pre-trains a model based on text representation and POS representation first to identify depressive emotions defined by DSM-5 (The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition), such as depressed mood, anhedonia, feelings of worthlessness or guilt, and suicidal ideation. Then the identified depressive emotions fused with text representation are used to detect signs of depression from posts. We evaluate DLEER with a dataset collected from Sina Weibo, a social media that includes a depression community. Our proposed DLEER based on a pretrained model that integrates text representation and POS representation contributes to the recognition of depressive emotions and signs of depression. We also contribute to the deep learning community and design science literature by recognizing depressive emotions to improve the model's ability to detect signs of depression from social media content. Furthermore, our work can contribute to early diagnosis and management of depression for patients and mental health professionals.

Predicting Medical Device Recalls: An ML-Based Framework Leveraging Regulatory Submission Characteristics

Yi Zhu¹, Soumya Sen¹, Alex Everhart² and Pinar Karaca-Mandic¹

1. University of Minnesota - Twin Cities, 2. Harvard-MIT Center for Regulatory Sciences, Harvard Medical School

Abstract

We present an end-to-end machine learning (ML) framework for predicting medical device recalls, addressing a crucial need in preventing harm to patients and protecting the financial stability of manufacturers and insurers. Previous research has explored the reasons for device recalls and potential predictors, but a comprehensive framework for efficient recall prediction is lacking. We propose a novel approach inspired by the association between applicant device recalls and the characteristics of existing devices (known as predicate devices or predicates) referenced in FDA's 510(k) clearance pathway submissions.

Our framework consists of three steps. First, we use an automated text extraction algorithm to create a database of predicate devices from 35,176 510(k) applicant devices approved between 2002 and 2018.

This ensures reliable and analysis-ready predicate device data. The algorithm achieves 94% accuracy in identifying predicate devices. Second, we construct a predicate citation temporal network from the predicate device data, extracting features for recall prediction. These features include the number of predicates, age range of predicates, percentage of recalled predicates, applicant devices' manufacturing and therapeutic information, and more. Finally, we connect these computed features with device recall records and build predictive models. We validate the models on a rolling basis, predicting recalls across different time windows. Our best-performing model, applied to highly imbalanced data (2% recall observations), achieves a 76% ROC score and a 62% F1 score when considering a five-year prediction time window.

Our study offers a predictive framework with implications for the medical device and life science industries. It can improve patient safety and quality of life by enhancing device quality and enabling faster recall prediction and device replacements. Stakeholders, including manufacturers and insurers, stand to benefit from cost savings by avoiding financial losses associated with malfunctioning or recalled medical devices.

Estimating Heterogeneous Survival Treatment Effects Among Veterans with Metastatic Castrate-Resistant Prostate Cancer: An Application of Causal Inference Using Machine Learning

Deepika Gopukumar¹, Nirup Menon¹ and Martin Schoen¹

1. Saint Louis University - School of Medicine, 2. George Mason University - School of Business, 3. St. Louis Veterans Affairs Medical Center and Saint Louis University – School of Medicine

Abstract

Background: Metastatic castrate-resistant prostate cancer (mCRPC) is one of the most commonly diagnosed cancers in the United States. ARTAs such as Abiraterone and Enzalutamide are preferred over other therapies as they have tolerable safety profiles and are known to improve survival. Among ARTAs, Abiraterone and Enzalutamide have different action processes. These differences in functioning mechanisms may result in different treatment outcomes for patients. Existing studies have compared the two treatments on survival and with a few comorbid conditions such as cardiovascular and diabetes as covariates in correlational studies. However, to our knowledge, no study has used the baseline's patient characteristics and comorbidities to identify heterogeneity in treatment effects.

Objective: The main objective of this study is to identify if there is heterogeneity in treatment effects and if treating individuals with one or the other drug based on their characteristics and comorbidities would lead to longer survival time.

Methods: Our dataset consists of 5822 veterans treated for mCRPC who have been administered either Abiraterone or Enzalutamide from 2014 to 2017. The study follow-up ended in 2020. The outcome of our study is the number of follow-up days. Covariates included are age, race, prostate-specific antigen (PSA) test result, body mass index category, comorbidities such as diabetes mellitus, hypertension, kidney disease, osteoporosis, Parkinson's disease, vision, androgen deprivation therapy (ADT) ever, and

orchiectomy procedure. We use a data-driven approach to identify heterogeneous subgroups because the data is non-experimental.

Results: mCRPC patients who were administered Enzalutamide survived longer than mCRPC patients who were administered Abiraterone. The positive value of the Ranked-Weighted Average Treatment Effect (RATE) indicates that the conditional average treatment effect estimates were able to identify heterogeneity among subgroups.

Using Early Engagement Data from a Digital Health Solution to Predict Future Engagement Patterns

Junjie Luo¹, Abhimanyu Kumbara², Anand K. Iyer², Mansur E. Shomali², Guodong (Gordon) Gao¹

2. Johns Hopkins Carey Business School, Center for Digital Health and Artificial Intelligence, 2. WellDoc, Inc.

Abstract

Introduction

Optimizing glucose management for people with diabetes requires them to 1) Understand their glucose levels and 2) manage critical elements of their health, accordingly, including medication, education, diet, activity and labvital data (MEDAL). Continuous glucose monitoring (CGM) helps solve the first problem by providing critical, real time glucose data to individuals with diabetes. Welldoc's® research is focused on addressing the second area; how regulated, AI driven personalized digital health solutions can leverage the power of CGM Data to effectively help people with diabetes manage MEDAL factors and their overall health. We hypothesized that early CGM signals and engagement observations could predict both health and engagement outcomes later in the patient's journey.

Method

We reviewed real-world data from 499 patients with type 1 and type 2 diabetes, who were enrolled in the digital health coaching platform. The dataset included CGM data, along with MEDAL engagement data. The baseline observation period to train the model was defined as 30 days from when the first MEDAL and CGM reading was recorded.

The outcome (prediction) period includes days 70-90 from the baseline period. The predictor variables included CGM and MEDAL engagement features. The outcome variable was categorized as high engagement (≥ 5 MEDAL records) or low engagement (< 5 MEDAL records). The dataset was split into training set and testing set with ratio of 7:3. A Random Forest Classifier model was used to predict the engagement outcome.

Results

Among the 499 patients, 119 were classified as high engagers and 380 patients as low engagers. The Random Forest Classifier was highly accurate, with accuracy rate of 0.87, in predicting high or low MEDAL engagement in the outcome period. This model also had an AUC of 0.89.

Conclusions

Obtaining early CGM and self-management behavior data can prove to be useful in predicting both health and engagement outcomes associated with later points in the patient journey. This prediction is important in 1) building the right capabilities within

Regular Session 5B: IT Value/Outcomes

Assessing the Impact of Health Information Exchange on Hospital Data Breach Risk

Sung Choi¹, Min Chen¹ and Xuan Tan¹

1. University of Central Florida, 2. Florida International University, 3. Santa Clara University

Abstract

Objective: Widespread electronic health information exchange (HIE) across hospitals remains an important policy goal for reducing costs and improving the quality of care. Meanwhile, cybersecurity incidents are a growing threat to hospitals. The relationship between the electronic sharing of health information and cybersecurity incidents is not well understood. The objective of this study was to empirically examine the impact of hospitals' HIE engagement on their data breach risk.

Materials and Methods: A balanced panel dataset included 4,936 US community hospitals spanning the period 2010-2017, which was assembled by linking the American Hospital Association annual survey database and the Information Technology (IT) supplement and the Department of Health and Human Services reports of health data breaches. The relationship between HIE engagement and hospital data breaches was modeled using a difference-in-differences specification controlling for time-varying hospital characteristics.

Results: The percentage of hospitals electronically exchanging information has more than tripled (from 18% to 68%) from 2010 to 2017. Hospital data breaches increased concurrently, largely due to the rise in hacking and unauthorized access. HIE engagement was associated with a 0.672 percentage point increase in the probability of an IT breach three years after the engagement. Hospitals actively engaging in a health information organization and exchanging data with outside providers were associated with a higher risk of IT related breaches in the long run; however, hospitals actively engaging in HIE and exchanging data with inside providers were not associated with any significant risk of IT related breaches.

Discussion: Over time, the increasing amount and complexity of patient information being exchanged can create challenges for cybersecurity if data protection is not up to date. Additionally, data security is dependent on the weakest link of HIE, and providers with fewer resources for data governance and infrastructure are more vulnerable to data breaches.

Conclusion: Moving toward widespread health information exchange has important cybersecurity implications that can significantly impact both patients and healthcare organizations.

Value of Healthcare IT and Advances in Medical Sciences: The Role of AI Research

Ruba Aljafari¹ and **Franck Loic Soh Noume²**

1. Virginia Tech, 2. University of North Carolina

Abstract

Research on the value of healthcare IT has tackled a variety of problems that continue to burden healthcare systems, such as cost, chronic conditions, and health divides. To understand the value of healthcare IT, patients and providers are generally studied as key healthcare IT users who provide and receive data. The main assumption in this research is that healthcare IT operates in environments where medical practices are based on timely findings from medical sciences. However, healthcare systems are also expressing concerns about the lack of advancement in medical sciences, especially amid the increased complexity of health issues worldwide. Although healthcare IT offers a promising role in supporting medical sciences through accessing diverse data sources and generating useful information, less is known about the value of investments in research capabilities that leverage IT to advance medical sciences, such as analytics and artificial intelligence and, more importantly, whether and how the patient perspective matters. Continuing the journey of research on healthcare IT value, we explore approaches for identifying potential synergies between research on medical sciences that utilize analytics and artificial intelligence and healthcare IT. We identify research grants that utilize analytics and artificial intelligence and match them to U.S. hospitals as one approach to explore their synergies with healthcare IT. Our approach builds on differences between data types to unravel healthcare IT complementarities. We propose that the impact of analytics- and artificial intelligence-related research grants on hospital performance is contingent on two types of data: health records and health & fitness data. Health records are data generated during a patient visit to the hospital. Examples of health records include lab tests, visit summaries, prescriptions, referrals, etc. Health & fitness data are generated outside the hospital. Examples of health & fitness data include blood glucose, blood oxygen, blood pressure, weight, body temperature, blood pressure, heart rate, respiratory rate, etc. Devices such as smartphones and smartwatches have democratized the generation of health & fitness data. Using these two types of data, we create four configurations: (i) health records are not shared among hospitals, and health & fitness data are not available to the hospital, (ii) health records are shared among hospitals, and health & fitness data are not available to the hospital, (iii) health records are not shared among hospitals and health & fitness data are available to the hospital, and (iv) health records are shared among hospitals and health & fitness data are available to the hospital. We compare the impact of analytics- and artificial intelligence-related research grants on hospital performance across the four configurations. Using data from 2016 to 2018, our preliminary results indicate that the impact of analytics- and artificial intelligence-related research grants on hospital performance is contingent on the type of data used. Specifically, sharing health records among hospitals increases the impact of analytics- and artificial intelligence-related research grants on hospital performance while health & fitness data have the opposite effect. Our study provides important

implications on how hospitals may appropriate value from analytics- and artificial intelligence-related research.

Balancing Technology and Humanity in Palliative Care: The Limits of Digitalization

Henner Gimpel¹, Kilian Osberghaus¹, Manfred Schoch¹ and Moritz Wöhl¹

1. FIM Research Center for Information Management

Abstract

The paper explores the limits to digitalization in the healthcare sector, particularly in palliative care. While the use of digital technologies can improve accessibility, quality of care, and efficiency, there are also limits to digitalization. Identifying these limits is essential to ensure digital transformation initiatives create value and deliver benefits. We use ethnographic work and action design research to understand the current level of digitalization in a German palliative care ward and potential areas for improvement. We identify technical, economic, and social limits to digitalization, including the lack of reliable technology, the need for personal connection, and ethical concerns. Identifying and respecting these limits can prevent the implementation of digital solutions that do not create value or have adverse effects.

Health Information Technology and Patient Outcomes: Evidence from Hacker Attacks

Emilia Simeonova¹, Yaa Akosa Antwi¹

1. Carey Business School

Abstract

The adoption of health information technology (IT) in medical practice was expected to revolutionize health care delivery and improve patient health outcomes. While there is evidence that the adoption of modern health information technology increased healthcare quality for infants, reduced mortality for complex patients, improved patient safety and has a modest effect on productivity, critics argue that realized benefits are small compared to projections and investment cost. A plausible explanation for the modest estimates of the effect of health IT on hospital production and patient outcomes is the inherent difficulty of identifying a causal path from technology adoption to hospital performance. The decision to adopt technology is not random; therefore, comparing early adopters of health IT to late adopters may confound the adoption effect with other unobserved hospital characteristics that affect both the adoption decision and health outcomes.

We use instances where hospitals' IT are compromised because of a cyberattack to quantify the importance of health information technology. A distinct advantage of this methodological approach is that we can look at within-hospital within-month variation in patient outcomes and hospital operations that happens after the hospital has already adopted the technology, and thus circumvent the problem of selection into technology adoption.

We collect data on hospital cyberattacks from local and national newspapers as well as government and third party websites. We distinguish between data breaches (loss of confidential data) and ransomware (complete takeover or disabling of computer systems) attacks.

We merge information on hospital cyberattacks from 2005 to 2018 to the universe of patient emergency department (ED) and inpatient data from California. Our main outcomes of interest are related to how disruptions in hospital IT affect hospital operations (ED visits and inpatient admissions) and patient outcomes (readmission and mortality). We use differences-in-difference regression analysis to compare outcomes at hospitals that experience a data breach or ransomware attack with similar hospitals that did not experience an attack. We also explore spillover effects of attacks on nearby hospitals.

We find that data breaches have no direct operational consequence on hospitals. Ransomware attacks on the other hand, have an instant and significant effect on hospital operations. We document an immediate decline in ED visits and inpatient admissions. ED visits decline by about 10 percent in the week following a ransomware attack. The operational disruptions caused by ransomware attacks are, however, short-lived –lasting on average less than a week. We also find evidence that nearby hospitals experience an increase in visits. Finally, we estimate an increase in readmission rates with no measurable change in mortality.

Hospital IT systems play a critical role in the care delivery process. Our results suggest that disruptions to a hospital's IT system have a negative impact on operations and an adverse effect on patient outcomes. It is important for health care administrators and policy makers to increase efforts to safeguard the IT systems of health care providers.

Spillover Effects of a Trauma I Unit on Hospital Operations

Kanix Wang¹, Chase Corvin², Vanessa Buie², Priya Prakash², Daniel Adelman²

1. University of Cincinnati, 2. The University of Chicago

Abstract

Racial disparities in trauma care accessibility are a significant issue in the United States, with the African American population disproportionately affected, especially in areas like the South Side of Chicago. To tackle this issue, in May 2018, the University of Chicago Medicine established a new level-I trauma center with dedicated resources, algorithms, and efficiency measures to ensure operational and financial viability. However, whether these measures have spill-over effects on non-trauma patients and hospital operations remains unclear. This study evaluates the potential spill-over effects of the new trauma center on non-trauma patients by examining changes in efficiency measures in the operation rooms (OR) for non-traumatic surgical encounters. The analysis of unadjusted data reveals an increase in consistency across service lines and more non-trauma, non-elective patients entering the operation room within service level agreement targets. Propensity-score-matched results demonstrate that the new case classification system's flexibility led to an average reduction in time-to-OR for the most urgent cases by 99 minutes.

The dedicated resources reduced time-to-OR by an average of 144 minutes, with plastic surgery and neurosurgery showing the most significant impact due to the new system and additional resources. Overall, the updated case classification system and the dedicated general resource increase resulted in better efficiency for non-trauma, non-elective surgical patients. This study sheds light on the potential positive spill-over effects of measures that address racial disparities in improving hospital operations.

Adapting with Health Information Technologies in Joyfully Productive Ways

Roopa Raman¹

1. University of Dayton

Abstract

As healthcare providers strive for the well-being of their patients, their own well-being is often left to the wayside. Among physicians, nurses, and other allied health professionals, stress and burnout are common side-effects of providing consistently high-quality patient-care (Costa & Pinto, 2017); indeed, there is often a sense of shame and/or guilt associated with even acknowledging these side-effects (Rehder et al., 2021).

Among healthcare administrators, regulators, and policy makers, there is emerging interest in how electronic medical records (EMR) systems can improve clinician wellness (Adler-Milstein et al., 2020; Melnick et al., 2021). EMR systems are designed for efficiency and effectiveness in patient-care delivery and information management, which in turn, would ostensibly alleviate provider stress. However, the increased, more structured, and time-sensitive documentation demanded by EMR systems ends up adding to the stress of physicians as they care for their patients using EMR-enabled patient-care practices that are approved by the hospital.

Despite the interest in EMR-enabled clinician wellness, many studies on adapting to health information systems tend to focus on productivity gains as the primary outcome of successful adaptation (Bhargava & Mishra, 2014; Terry et al., 2019). In this study, we supplement existing research to investigate ways of adapting to health information systems with the dual purpose of enabling productivity and enabling joyfulness at work; we consider joyfulness as a mark of well-being. We ask the following question: How can adapting to health information systems in large healthcare organizations enable joyful productivity among healthcare professionals? Thus, the HIT adaptation process that we seek to uncover in this study would explicitly aim for increased productivity in terms of patient- and hospital-outcomes, as well as greater joy or well-being in the experience of clinicians while providing this care.

We combine two complementary theories of IT adaptation – coping theory (Beaudry & Pinsonneault, 2005) and status quo bias theory (Polites & Karahanna, 2012) – to develop the theoretical underpinning for a model that answers our research question. Coping theory presents a process model of how people adapt to a new information system or a new practice or a new intended outcome (e.g., joyful productivity) as a disruptive event. It lays out a process framework for how people appraise the triggering event, and

the kinds of cognitive and behavioral efforts they exert to adapt with or cope with the new change. Status quo bias discusses why people stick to habitual practices (e.g., solo focus on productivity without regard to provider stress/burnout) even when they are faced with change that they know will be beneficial to them.

We use the combination of coping theory and status quo bias theory to develop a process model explaining how healthcare organizations can move from the traditional approach of using EMR systems to be efficient/productive to new ways of using EMR systems to be joyfully productive.

Regular Session 6A: Deep Learning

What Symptoms and How Long? An Interpretable AI Approach for Depression Detection in Social Media

Jiaheng Xie¹, Junwei Kuang¹, Zhijun Yan²

1. University of Delaware, 2. Beijing Institute of Technology

Abstract

Depression is the most prevalent and serious mental illness, which induces grave financial and societal ramifications. Depression detection is key for early intervention to mitigate those consequences. Such a high-stake decision inherently necessitates interpretability, which most existing methods fall short of. To connect human expertise in this decision-making, safeguard trust from end users, and ensure algorithm transparency, we develop an interpretable deep learning model: Multi-Scale Temporal Prototype Network (MSTPNet). MSTPNet is built upon the emergent prototype learning methods. In line with the medical practice of depression diagnosis, MSTPNet differs from existing prototype learning models in its capability of capturing the depressive symptoms and their temporal distribution such as frequency and persistence of appearance. Extensive empirical analyses using real-world social media data show that MSTPNet outperforms state-of-the-art benchmarks in depression detection, with an F1-score of 0.851. Moreover, MSTPNet interprets its prediction by identifying what depression symptoms the user presents and how long these related symptoms last. We further conduct a user study to demonstrate its superiority over the benchmarks in interpretability. Methodologically, this study contributes to extant literature with a novel interpretable deep learning model for depression detection in social media. Our proposed method can be implemented in social media platforms to detect depression and its symptoms. Platforms can subsequently provide personalized online resources such as educational and supporting videos and articles, or sources for treatments and social support for depressed patients.

Diagnosis of Brain Tumors through MRI Images: A Novel Approach using Local Binary Patterns and Convolutional Neural Networks

Leila Amini¹, Kamil Yurtkan²

1. Florida International University, 2. Cyprus International University

Abstract

Brain tumors can affect individuals of any age, race, ethnicity, or gender. In the United States, in 2021, over 84,000 people were diagnosed with primary brain tumors. There are nearly 120 different types of primary brain and other central nervous system tumors, and almost one-third (29.7%) of these are malignant. According to official statistics in the USA, brain tumors caused 3.6 deaths per 100,000 females and 5.3 deaths per 100,000 males, making malignant brain tumors the tenth leading cause of death (National Cancer Institute, 2019). Moreover, over 28,000 people under the age of 20 are estimated to suffer from brain tumors in the United States, with primary malignant brain tumors claiming the lives of almost 18,000 people (Central Brain Tumor Registry of the United States, 2019; National Brain Tumor Society, 2019). Early detection and treatment of brain tumors can potentially save patients' lives.

In the medical field, data collection on mild and fatal diseases has been beneficial in finding new ways to diagnose and treat diseases. Medical imaging applications have significantly accelerated as digital computers became capable of analyzing digital images with human-like performance. This study aims to diagnose brain tumors using Magnetic Resonance Image (MRI) digital images. The method employed in this paper is based on the combination of Local Binary Patterns (LBP) and Convolutional Neural Networks (CNN), which is a novel approach.

To conduct a comprehensive evaluation, an appropriate database was needed that included images of both sick and healthy individuals. Most databases focus on individuals with specific illnesses such as tumors or Multiple Sclerosis (MS), and only a few websites have information on healthy individuals for comparison. An appropriate database should also include a sufficient number of test specimens and a variety of patients with tumors of different volumes. To achieve this goal, the Kaggle website was used, which involves 253 black and white MRI scans. This database contains MRI images of 98 healthy individuals and 155 patients with varying lesion volumes.

Finally, the findings of the study were compared with CNN without LBP, as well as with the CNN technique (Kunimatsu et al., 2019) and the Mask R-CNN algorithm (by Ruslan Klymentiev et al., 2019). The results illustrate that the combination of CNN and LBP techniques can achieve improved and comparable accuracies.

Did we personalize? Assessing personalization by an online reinforcement learning algorithm using resampling

Susobhan Ghosh¹, Raphael Kim¹, Prasidh Chhabria¹, Raaz Dwivedi¹, Predrag Klasnja², Peng Liao¹, Kelly Zhang¹, Susan Murphy¹

1. Harvard University, 2. University of Michigan

Abstract

There is a growing interest in using reinforcement learning (RL) to personalize sequences of treatments in digital health to support users in adopting healthier behaviors. Such sequential decision-making problems involve decisions about when to treat and how to treat based on the user's context (e.g., prior activity level, location, etc.). Online RL is a promising data-driven approach for this problem as it learns based on each user's historical responses and uses that knowledge to personalize these decisions. However, to decide whether the RL algorithm should be included in an "optimized" intervention for real-world deployment, we must assess the data evidence indicating that the RL algorithm is actually personalizing the treatments to its users. Due to the stochasticity in the RL algorithm, one may get a false impression that it is learning in certain states and using this learning to provide specific treatments. We use a working definition of personalization and introduce a resampling-based methodology for investigating whether the personalization exhibited by the RL algorithm is an artifact of the RL algorithm stochasticity. We illustrate our methodology with a case study by analyzing the data from a physical activity clinical trial called HeartSteps, which included the use of an online RL algorithm. We demonstrate how our approach enhances data-driven truth-in-advertising of algorithm personalization both across all users as well as within specific users in the study.

Detecting the Human Behaviors Associated with Depression via Sensor Signal Analysis: An Agreement-based Self-Attention Deep Learning Approach

Aijia Yuan¹, Edlin Garcia¹, Hongyi Zhu¹, Sagar Samtani²

1. Indiana University, 2. The University of Texas at San Antonio

Abstract

Mental health illnesses are a group of non-communicable diseases that have become a leading cause of disability worldwide. Depression is one of the most common and serious mental illnesses that affect millions of people each year. One of the most vulnerable populations to depression is students. Although many universities provide counseling services to monitor students' mental health status, traditional approaches usually involve interviews and self-reported assessments that require extensive human labor and financial cost. Sensors are a promising alternative to assess students' mental health conditions more objectively, holistically, and cost-effectively. In this study, we propose a novel Deep Learning (DL)-based Agreement Self-Attentive Model (ASAM) to detect depression and identify the human behaviors associated with depression. We evaluated our proposed ASAM against benchmark machine learning and DL models on the publicly accessible StudentLife dataset that contains sensor signal data for 48 students across 65 days. Overall, our proposed ASAM achieves an F1-score of 0.88 and an AUC of 0.93 and outperforms all other machine learning and deep learning-based benchmarks. Results from our experiments indicate that regular sleep patterns and properly reducing the amount of time for mobile phone usage are essential for college students to avoid depressive behaviors.

Unsupervised discovery of biomedical associations from clinical and auxiliary data

Aishwarya Jayagopal¹, Lin Jing¹, Folefac Aminkeng¹, Kee Yuan Ngiam¹, Vaibhav Rajan¹

1. National University of Singapore

Abstract

The aim of our work is to discover novel associations, in clinical data, between biomedical entities, such as patients' demographics, vitals, lab investigations, genomic profiles and diagnosed diseases. Finding associations from clinical data alone is often difficult due to two reasons: (1) high dimensionality, especially in genomic data, and (2) small sample sizes for some diseases. As a result, the signal for a potential association becomes difficult to infer. To address this problem, auxiliary data about the entities, from external databases, may be used in the knowledge discovery process.

For instance, consider two matrices from electronic medical records (EMR) of patients – one describing patients' clinical measurements and another containing the genetic alterations (mutations) found in these patients. Another matrix, from an external database, that includes known associations between mutations and diseases can be obtained and used to discover new associations between clinical measurements and diseases. Note that in this example, entities corresponding to patient matrices from EMR are directly connected while diseases, from external databases, are indirectly connected (through their association to mutations) to the patients. Data from numerous biomedical studies are increasingly available online in external databases. However, to utilize them for such challenging settings, we need methods to find associations in arbitrary collections of matrices.

To integrate such auxiliary information within our analysis, we utilize unsupervised clustering methods based on collective matrix factorization that can consider any number of matrices as inputs and cluster the row and column entities in each input matrix. These methods also output associations at the level of cluster blocks, that can be used to form cluster chains across the input collection of matrices. We use a recent deep learning based method called Deep Collective Matrix Tri Factorization (DCMTF) to identify associations between clusters of patients, diseases and genomic features by integrating data from auxiliary databases.

We used DCMTF to find associations between clinical and genomic data collected from patients diagnosed with Hodgkin's Lymphoma (a cancer originating in the immune system), and auxiliary data from clinical databases. The resulting cluster chains link patient clusters to clusters of genetic alterations, which link to disease clusters. The results of this method were compared against other clustering algorithms (that do not consider auxiliary information or do not use deep learning) in terms of (1) clustering performance and (2) clinical relevance of cluster chains formed. DCMTF was found to perform well in both aspects. DCMTF could uncover the association between a set of patients with alterations in the BRCA2 gene and hereditary breast and ovarian cancer syndrome. We verified this association using independent evidence from medical literature.

Our work illustrates how we can seamlessly integrate growing information from heterogeneous biomedical studies to inform clinical data analysis. Finding clinically relevant patient clusters and cluster-specific associations also enable targeted treatment of specific patient groups, leading to improved care. For instance, if a patient is known to be susceptible to a specific disease due to a genetic predisposition, care can be taken in advance to monitor them for relevant symptoms.

Fast Multi-Dimensional Subset Scan With Sparse Tensor Decomposition For Pattern Detection

Daniel Neill¹, Girmaw Abebe Tadesse², Skyler Speakman², Sriram Somanchi³

1. New York University, 2. IBM Inc. 3. University of Notre Dame

Abstract

We propose a novel multidimensional subset scanning (MDScan) technique to detect sub-populations with differential outcomes in the foreground data compared to background data. Specifically, we offer a model-independent algorithm based on subset scanning to computationally efficiently discover subpopulations that exhibit a concept shift. Existing techniques for discovering subpopulations suffer from low detection power due to sparsity in estimating the baselines from background data. We utilize tensor decomposition techniques to effectively estimate the baselines in sparse background data, improving our ability to detect, both in terms of detection power and accuracy, and improving our computational efficiency. Through simulations and a real-world case study using child mortality data from Ethiopia, we demonstrate the performance of the MDScan technique for discovering novel patterns.

Regular Session 6B: HIT & Patient Outcomes

Impact of telehealth on appointment adherence in ambulatory care

Masoud Kamalahmadi¹, Christos Zacharias¹ and Howard Gitlow¹

1. University of Miami

Abstract

Telehealth services grew massively popular amid the COVID-19 pandemic. Yet, their operational implications for healthcare organizations are not well-understood. We explore the impact of telehealth on two patient behaviors that result in non-adherence to designated appointment times: no-shows and unpunctuality. We analyze 280,067 in-person and video/audio telehealth appointments from a large medical system in the United States in 2020 and 2021. We find that (1) no-show probability decreases by 4.0 percentage points and late-arrival probability decreases by 10.2 percentage points over telehealth, (2) no-show probability decreases more for follow-up patients who may have a stronger preference for convenience, whereas late-arrival probability decreases more for new patients who may care more about timely access to care, (3) patient groups with the lowest odds of adherence to in-person appointments -- women, racial minorities, Medicaid patients, and younger adults -- exhibit the largest improvements in adherence over telehealth, and (4) transportation issues, competing priorities, inconvenience, and health

and safety concerns are the likely drivers of lower adherence to in-person appointments. We further assess the practical implications of our empirical findings using recent advances in stochastic mathematical models for appointment scheduling. Through computational experiments, we demonstrate that better appointment adherence leads to significant improvement in daily patient throughput and overall operational efficiency, attesting to the value of telehealth in expanding access to care and improving operations. Our study demonstrates the value of telehealth in expanding access to care, not only as an alternative option to access, but also as a latent consequence of improved adherence to appointments. Our findings imply that even in the absence of telehealth pay parity, providers should over telehealth, as the additional revenues from the higher throughput may offset the lower telehealth pay rate.

What Factors Affect the Digital Technology Adoption by the Elderly for Daily Life and Health Management - a UTAUT Analysis

Haijing Hao¹, Monica Garfield¹

1. Bentley University

Abstract

As the baby boomer population ages, the percentage of the US population that is over 55 years old continues to grow. In 2021, 30% (97.6 million) of the US population were 55 years or older, and this number will keep growing (U.S. Census Bureau, 2021). This segment of the US population has been hit the hardest by the COVID-19 pandemic, which has led to an increase in their social and physical isolation (The Center for Disease Control, 2020). One way that people have coped with such isolation is to rely on digital technologies to connect with others and seek out information and advice from reputable (and sometimes not very reputable) sources. However, little research has focused on how the elderly use digital technologies in their daily life and if they are using them to manage chronic health issues to improve their quality of life.

Previous research on technology adoption has focused on technology users who are in the workforce population (Hao et al. 2018; Rahimi et al., 2018; Salloum et al. 2019). Some studies have used TAM to study how the elderly use online patient portals (Portz et al. 2019). Only a few studies on senior's technology use and are dated (Sterns 2005; Mitzner et al. 2008).

The present study utilizes the Unified Theory of Acceptance and Use of Technology (UTAUT) framework to examine the acceptance of digital technology by seniors in the U.S. UTAUT evaluates the acceptance of technology, determined by the effects of performance expectancy, effort expectancy, social influence, and facilitating conditions. UTUAT is one of the most widely used models in information systems to examine technology adoption (Venkatesh et al. 2003).

We deployed a survey based on UTAUT to people who are 55 years or older and live in the U.S. on Amazon Mechanical Turk in Feb. 2023. After collecting the data, our preliminary data analysis shows that among

468 responses 60% are female and 88% are white, which is not a normal demographic distribution in general, but it is a consistent demographic group composition as a previous study on the Amazon Turk population (Aaron and Litman, 2020). Also, this is consistent with the observation that among retired people elder women are more socially active than elder men (Caetano et al. 2013; Steinmayr et al.2020). About 90% of the respondents said that they agree that computer or mobile apps are useful in managing their daily life, but only 47% of them said that they have used any applications for chronic health management. De facto, 95% of the senior population have at least one chronic condition and nearly 80% of them have two or more (National Council on Aging, 2021). There is a gap between the reality and the needs. One of the goals of the present study is to figure out what factors affect the senior people's technology adoption for chronic health management. We believe our study will make contributions to both information systems and aging society areas. Further UTAUT framework analysis is undergoing.

Balancing Patient Convenience and Healthcare Costs: A personalized recommendation approach using reinforcement learning

Tianjian Guo¹, Indranil Bardhan¹ and Wen Wen¹

1. The University of Texas at Austin McCombs School of Business

Abstract

Precision medicine aims to deliver personalized healthcare and provide the most effective care to patients by considering their unique genetic makeup, lifestyle, and environmental factors. While medical professionals have explored the use of dynamic treatment regimens and, more recently, reinforcement learning (RL) algorithms for personalized medicine, the extant focus has been on optimizing a single clinical objective. To the best of our knowledge, few studies have used personalized medicine approaches to balance tradeoffs among multiple objectives related to the economic value and societal impact of healthcare delivery. Often, wellness and prevention programs often recommend preventive care visits and procedures to all patients to improve long-term health and reduce associated treatment costs. However, such a strategy may not achieve the most efficient allocation of resources in a value-based healthcare environment. Thus, it is important to provide curated preventive care plans to individual patients and strike the right balance between various health, economic, and societal objectives.

In this research, we develop a novel, multi-objective RL framework to create a personalized recommendation of preventive care procedures for both diabetic and non-diabetic patients. The model seeks to minimize both patient inconvenience, as measured by the annual count of preventive care visits, and total annual cost associated with all patient treatments. Both objectives are essential for patients - improving patient convenience can ensure patient compliance and treatment efficacy, while optimizing total treatment costs can result in significant financial savings. These two objectives also have useful implications for external stakeholders - reducing a focal patient's visits can help optimize healthcare utilization, while reducing costs is important to other parties in the healthcare system such as insurance companies and employers, since patients usually only pay a portion of the total cost. Furthermore, we

seek to understand how deploying a range of preference functions with different weights assigned to these objectives can influence the types of personalized preventive care procedures recommended. We construct this problem as a Multi-Objective Markov Decision Process (MDP) and solve the MDP by utilizing an innovative multi-objective, offline reinforcement learning (RL) approach.

Specifically, we construct a novel deep Q network (DQN) that recommends the optimal set of preventive procedures. Patient characteristics, such as age and gender, along with data on patient diagnosis and treatment procedures from clinical encounters are utilized as model inputs. The DQN is optimized over the entire space of possible linear preference functions to generate multiple recommendation schemes based on both objectives. We test the effectiveness of our proposed approach and evaluate the recommendation schemes generated through off-policy evaluation (OPE) on a visit-level patient dataset from a regional health information exchange in Central Texas.

Our RL-based recommendation model suggests that changing the behavior of just 20% of our patient population can reduce the average treatment cost for all patients by 25% and the count of preventive care visits by 36% for patients under 65. Our model recommends reducing preventive care for younger and female patients while increasing preventive care treatments for older, male, and advanced diabetic patients. It's worth noting our model has a greater impact on female patients, potentially leading to a 40% reduction in costs and a 48% reduction in the count of preventive visits for female patients under the age of 65 years.

Several key insights emerge based on our RL approach. First, our results indicate the presence of trade-offs between optimization of patient convenience versus patient treatment costs. Recommendations that emphasize patient convenience are likely to recommend treatments that do not involve preventive care for patients such that they reduce the number of primary care visits. Similarly, recommendations that emphasize overall treatment costs are more likely to recommend preventive care procedures for diabetes

care. Based on real-world patient data, we observe that patients mostly optimize for their own convenience by reducing the number of preventive (primary) care visits. Furthermore, we find that elderly patients and patients with greater disease progression are more likely to behave in a suboptimal manner with respect to both objectives. Thus, our research highlights the need for programs that incentivize elderly patients and patients with advanced diabetes to change their treatment regimen to reduce total healthcare costs and number of primary care visits.

Our study contributes to the emerging research on prescriptive analytics by providing insights to manage healthcare risk through development of personalized artificial intelligence. Our findings suggest that development of personalized medicine models for long-term care management can achieve significant financial savings without impairing patient convenience. Our prescriptive approach encourages patients to change their behavior using targeted incentives that support greater utilization of preventive care, while reducing total healthcare costs by avoiding the need for costly hospital and emergency room visits.

Evaluation and Revision of a Postpartum Depression Risk Prediction Model in Preparation for Implementation

Yiye Zhang¹, Yifan Liu¹ and Rochell Joly¹

1. Cornell University

Abstract

Introduction: Postpartum depression (PPD) is estimated to affect 1 in 7 women who give birth in the US, and early detection has substantial significance for maternal and child health. This study team has developed a machine learning model to predict postpartum depression for early identification and intervention. In preparation for its electronic health record (EHR) system integration at NewYork-Presbyterian Hospital/Weill Cornell Medical Center, we evaluated the prediction model against the distributional shift, fairness, and patient safety concerns.

Methods: The intended workflow involving the model is to alert the clinicians treating pregnant patients if the risk of postpartum depression is detected. As subsequent actions, clinicians can choose to observe or identify interventions for patients, including referrals for mental health services and additional lifestyle-based support. Thus, positive prediction from the model potentially leads to higher patient resource allocation. The original model was a logistic regression model (AUC=0.96) using 32 variables from the EHR, including demographics, medical history, pregnancy complications, medication prescription, and health service utilization. Since the original model was developed using EHR data from 2015 to 2018, we used 2019 and 2020 data under the inclusion and exclusion criteria from the same study site as the evaluation dataset. The evaluation focused on the model's performance in new populations, and fairness metrics across patient subgroups, including privileged (defined as White) and non-privileged groups. Fairness metrics examined include statistical parity difference, disparate impact (DI), equal opportunity difference (EOD), average odds difference, and predictive parity difference. Based on the evaluation, we revised the model by comparing several methods to mitigate bias, including reweighting and modifying race as a sensitive feature. In addition, working with the hospital's governing body on artificial intelligence, legal/regulatory, and clinical leadership, we identified areas for prospective evaluations.

Results: Our evaluation dataset had 8007 and 9623 patients in 2019 and 2020. The original model maintained overall performance in 2019 and 2020 (AUC=0.94) and relatively equal performance across subgroups by race and insurance. However, fairness evaluation identified biases defined by DI and EOD (0.527 and -0.004). In addition, the hospital's governance body expressed concerns about race as an input variable. As a result, we took de-biasing strategies by modifying race as an input variable while 1) keeping model parameters, 2) retraining a model, and 3) reweighing the population, respectively. These methods improved DI (0.573, 0.574, 0.685) while keeping EOD (0.043, 0.034, 0.063) in the acceptable range. The de-biasing strategies did not significantly affect the predictive performance (AUC=0.92). We also identified a change in practice where postpartum depression was screened for and diagnosed with higher frequency in post-development years that affected the performance.

Conclusion: There remains to be a gap in machine learning model development and its use in clinical care, including model generalizability, fairness, and safety. A thoughtful and thorough evaluation is warranted to ensure machine learning's value in patient care. This model for postpartum depression risk prediction is currently undergoing implementation at NYPH after this retrospective evaluation.

Hybrid Artificial intelligence and Inverse Learning for Diet Recommendation

Farzin Ahmadi¹, Fardin Ganjkanloo¹ and Kimia Ghobadi¹

1. Johns Hopkins University

Abstract

Purpose: Metabolic syndrome is one of the most serious and common health conditions in the US that affects one in three adults. This syndrome includes a group of conditions that raise the risk of diabetes, cardiovascular disease, and stroke. Balanced dietary behaviors are known to play a major role in preventing it and are important from policy, public health, and individual patient points of view. Despite the well-known knowledge, the current one-size-fits-all view of dietary interventions that are recommended in public policies leads to long-term adherence challenges and does not address the individuals' dietary goals, preferences, and lifestyles. Data-driven personalized and scalable diet recommendation algorithms can address this need and complement current policy efforts that mainly focus on treatment and better access to testing and care. Recent efforts in personalizing dietary plans focus on using Artificial Intelligence (AI) on patient-reported past decisions or mechanistic and optimization models that aim to find optimal diets. While both sets of methods have achieved some success, neither has been able to find optimal personalized diets that can adapt to patient preferences. There remains a need to incorporate human-AI interaction and integrate user preferences into the algorithms to reduce algorithm aversion and improve the chances of adherence and acceptance of AI-generated recommendations by users. To this end, we propose a hybrid AI and inverse optimization method, called "Inverse Learning".

Methods: Inverse Learning is a method that integrates AI and optimization to recover underlying preferences of users based on their past decisions while also respecting knowledge-based constraints. In the diet recommendation problem, we consider the objective to be maximizing patient preference and the constraints to reflect dietitian guidelines and patient preferences. We develop AI and inverse optimization-based models that use food intake choices by patients and dietary guidelines by medical experts to arrive at optimal dietary choices tailored to patients. The integrated AI models draw patterns and insights from past food choices to dynamically adapt the shape of the optimization feasible set to patient behavior, hence, improving patient adherence to healthier diets recommended by the dietitian.

Results/Conclusion: An analysis of results based on publicly available data from NHANES and FDA nutritional value shows that the models recommend healthier alternatives that contain a higher nutritional value and lower levels of unhealthy food items compared to patients' existing behaviors. Perhaps as importantly, the models are also capable of providing the patient with a pathway of deviating

from unhealthy eating habits gradually toward healthy eating plans. A successful personalized dietary recommendation can enable integration of nutrition precision as medical treatment and reduce the public health cost of chronic diseases like diabetes. Further developments of the models can also potentially lead to decision-support tools to inform public health decision-makers of dietary requirements and habits of vulnerable populations, e.g., Medicaid patients, and reduce the cost and complexities associated with metabolic syndrome illnesses and treatments.

Regular Session 7A: AI in Action

Will Machines Take Over? Algorithms for Human-Machine Collaborative Decision Making in Healthcare

Gustavo Stolovitzky¹, Mehmet Eren Ahsen²

1. GeneDX, 2. University of Illinois

Abstract

Artificial intelligence (AI) has become an increasingly popular alternative for performing tasks that are typically performed by humans. Mammography imaging is one context in which the role of AI is growing. Some experts claim that, with recent advancements in image processing algorithms and the increasing availability of data, AI will replace radiologists. Others argue that the rise of AI will change how diagnostic tasks are allocated, eventually paving the way for human-machine collaborative decision-making. In this research, we study a healthcare system's problem of when, if, and how to use human-machine collaborative decision-making in the context of mammography imaging. To that end, we propose an optimization model for the healthcare system that minimizes costs related to mammography screening and determines whether and when a complete automation (AI alone) strategy or a delegation (collaboration between humans and machines) strategy is preferable to an expert-alone strategy. We find that the disease incidence relative to the ratio of follow-up costs against litigation costs is an important determinant of whether the delegation strategy is preferable to the automation strategy. Reductions in the cost of the algorithm could either result in delegation (sharing of work between humans and machines) or full automation depending on the algorithm's performance. In contrast, a higher litigation cost for machine decisions could alter healthcare systems' preferences away from automation to delegation or expert-only solutions. We also evaluate the use of two or more algorithms from different sources to produce an ensemble (i.e., aggregation of algorithms and their predictions) and characterize when using an ensemble is superior to using any of the algorithms alone. We use data from an AI contest to empirically validate our results and back test its performance in real life. The contest's goal was the crowdsourcing AI algorithms to meet or exceed radiologists' performance in assessing mammograms obtained from actual clinical practice. The empirical analysis suggests that using the delegation strategy is optimal, with a potential of 17.5-30.1% cost savings as compared with the expert-only strategy. Our work has significant implications beyond radiology imaging for the design of work in the AI era and the human-machine collaboration contexts.

Artificial Intelligence on Call: The Physician's Decision of Whether to Use AI in Clinical Practice

Tinglong Dai¹ and Shubhranshu Singh¹

1. Johns Hopkins University

Abstract

Physicians are increasingly able to use artificial intelligence (AI) systems to aid their medical decision-making. This paper examines a physician's decision regarding whether to use an assistive AI system when prescribing a treatment plan for a patient. Using AI helps the physician generate an informative signal that lessens clinical uncertainty. It can also change the physician's legal liability in the event of patient harm. We analyze two patient-protection schemes that determine physician liability when using AI: the prevailing patient-protection scheme uses the AI signal to enforce the current standard of care, whereas an emerging scheme proposes using the AI signal as the new standard of care. We show that in both schemes, the physician has an incentive to use AI in low-uncertainty scenarios, even if AI provides little value. Furthermore, the physician may avoid using AI in higher-uncertainty scenarios where AI could have aided in better decision-making. As AI becomes more precise, the physician may become more hesitant to use it on certain patients. A comparison of the physician's decision to use AI under the two schemes reveals that using the AI signal as the new standard of care may mitigate AI underuse (overuse) for certain patients but may boost AI underuse (overuse) for some other patients.

Love and Robots: An Empirical Investigation of Warehouse Automation and Worker Injury

Gordon Burtch¹, Brad Greenwood² and Kiron Ravindran³

1. Boston University, 2. George Mason University, 3. IE University

Abstract

Scholars have long been interested in the relationship between industrial automation and workplace safety (Sheridan et al. 1983). Recent advancements in computer vision and machine learning have led to the development of mobile, adaptive, and collaborative forms of robotics that can operate alongside human laborers. These new forms of robotics have led to renewed questions around the implications for worker safety (Badri et al. 2018). Researchers and practitioners alike have argued that these de novo robotic forms could significantly improve worker safety by obviating the need for workers to engage in repetitive or hazardous work (Brosque and Fisher 2022). However, those assertions are at odds with recent media accounts, reporting a rise in worker injury following the introduction of robotics. So, which claim is correct?

We argue that both may be true, depending on the unit of analysis. At the task level, the adoption of robotics may indeed lead to reduced injury, by removing human involvement from hazardous activities. However, when one considers broader organizational processes, there are compelling reasons to believe that robotics may have spillover consequences for workers who continue to perform non-automated

tasks; chiefly because robotics i) reduce the mix of tasks for human workers (Parker and Grote 2022) and ii) increasing the pace of work (Gutelius and Theodore 2019; Willocks 2019). We focus on Amazon's Fulfillment Center Network, considering warehouse injury reports obtained from the Occupational Health and Safety Administration (OSHA), contrasting Robotics and Legacy Fulfillment Centers. We consider injuries in general, as well as injuries by level of severity (proxied by whether the injury requires days of missed work).

Two critical findings emerge. Regarding general injury volumes, we observe no significant differences between Robotic and Legacy Fulfillment Centers. However, and strikingly, when we distinguish between severe and non-severe injury, we observe a clear substitution effect, such that robotics is associated with a significant decline in severe injury, yet a simultaneous rise in less-severe injury.

To identify better understand the underlying mechanisms, we further examine 2018 injury log data for 26 Amazon Fulfillment Centers. We evaluate the possibility that the increase in less-severe (stress-based) injuries can be attributed an increase in the pace following the implementation of robotics. We assess this via a difference-in-difference-in-differences (DDD) design (Olden and Møen 2022), wherein we contrast the rate of sprains/strains with that of other injury types, across Robotic versus Legacy Fulfillment Centers, between peak (i.e., Black Friday and Amazon Prime Day) and off-peak seasons. Consistent with expectations; we find that the volume of sprains and strains rises systematically in Robotics Fulfillment Centers during peak season.

While scholarly work in management and economics has examined robotics for decades, the focus has typically been on the antecedents of adoption (Pillai et al. 2022) or the downstream implications for labor and productivity (Bessen 2015, 2016; Dixon et al. 2021; Pasparakis et al. 2021). We thus build on past work by examining the changing nature of collaborative co-creation between humans and robots and its impact on worker safety.

Regular Session 7B: Telemedicine 1

Does Telemedicine Affect Physician Decisions? Evidence from Antibiotic Prescriptions

Shujing Sun¹, Tongil Ti¹, Guihua Wang¹

1. University of Texas at Dallas

Abstract

Telemedicine has long been of interest to the U.S. general public. Yet, despite the advent of high-speed internet and mobile device technology, telemedicine did not reach its full potential until the COVID-19 pandemic spurred its unparalleled adoption. This sudden shift in the setting of healthcare delivery raises questions regarding possible changes in clinical decision-making. Using a unique set of patient-provider encounter data from the U.S. in 2020 and 2021, we examine the effect of telemedicine on antibiotic prescription errors for urinary tract infections. We consider two types of prescription errors: prescribing when not recommended by guidelines (type I errors) and not prescribing when recommended (type II

errors). After accounting for potential endogeneity issues using provider fixed effects and an instrumental variable approach, we find a significantly lower likelihood of overall prescription errors (type I and II errors combined) with telemedicine relative to in-person encounters. We also find heterogeneous effects by a provider's patient volume and the patient-provider relationship. Further analyses show that the reduction in prescription errors is mainly attributable to type I errors, and that patient health outcomes are not compromised when care is delivered via telemedicine. Finally, we discuss managerial implications for the pharmaceutical and insurance industries, as well as policy implications for governments.

How Does the Adoption of Telehealth and Telepsychiatry Services Level in US counties Affect Community Mental Health Outcomes: Results from a Quasi-Natural Experiment

Yi Yang¹, Arun Rai¹ and Aaron Baird¹

1. George State University

Abstract

Telehealth enables health care providers to provide virtual services for patients. With accelerating scale and scope of telehealth adoption among consumers and providers, \$250 billion of US health care spending will be for care provided virtually as per a recent McKinsey Report¹.

We focus on the impact of telehealth on mental health outcomes of the community. We are motivated to do so on two fronts. First, mental health is a significant issue in health care, severely impacting the lives and livelihoods of individuals and communities (Wells et al., 2004), as well as having adverse economic impacts when not sufficiently addressed (Pearman et al., 2020; Sharac et al., 2010). There are 19.86% of adults experiencing a mental illness, equivalent to nearly 50 million Americans as per a recent survey from Mental Health America.² Vulnerable populations are the most affected, which results in significant mental health disparities (Diaz et al. 2021). Second, while we have learned a great deal about the adoption, effective use, and impacts of telehealth, there is limited understanding about the impacts of telehealth on mental health outcomes. Specially, through past work, we have learned about the determinants of telehealth adoption, discontinuation, and performance at the hospital level (Adler-Milstein et al., 2014; Baird et al., 2022); the impacts of telehealth on ameliorating patients' physical health, including heart failure (Agha, 2014), chronic obstructive pulmonary disease (Anderson et al., 2023) and pneumonia (Shigekawa et al., 2018); how telehealth adoption reduces the frequency of hospital visits and readmission risks (Bao et al. 2020); and how telehealth facilitates communication between health providers and patients (Abdolkhani et al., 2019). In general, this literature assumes telehealth services play a positive role in improving an individual's physical health conditions and enhancing the communication between health care providers and patients. However, the impact of telehealth on mental health outcomes at the community level has yet to be studied.

Motivated to understand the impact of telehealth adoptions on community level (county, in this study) mental health outcomes, the first objective of this work is to establish a causal link between the adoption of telehealth/tele psych services and mental health outcomes at the county level in the US. We construct

our dependent variable: county level mental health outcome as self-reported average number of mentally unhealthy days in the county, adopted from the County Health Ranking Data. Self-reported health status is a general and widely used measure of health-related life outcomes. What's more, reports of days of unwell mental health is a reliable estimate of recent mental health outcomes³.

We construct our dependent variable as self-reported average number of mentally unhealthy days in the county, adopted from the County Health Ranking Data. Self-reported health status is a general and widely used measure of health-related life outcomes. ⁴ We apply a quasi-natural experiment (difference-in-differences) research design to establish the causal link. We compare counties that have adopted telehealth to groups of counties with similar characteristics that did not adopt telehealth, which establishes a counterfactual. For this quasi-natural experiment, treatment counties are those that did not have hospitals-based telehealth services in 2017 but started in the subsequent years ($n = 748$). Control counties are those that did not offer telehealth in all 5 years ($n=1500$). To reduce selection bias, we calculated coarsened exact matching (CEM) to match the covariates between the treatment and control group. Our covariates consisted of county-level variables such as urban status, income, education, population, broadband access rates, etc.

Given the vast differences in socioeconomic and demographic conditions among counties, interpreting an average treatment effect of telehealth adoptions on mental health outcomes will not identify what type of counties with similar conditions would benefit the most, or least, from adopting telehealth as a solution to mental health problems. Employing the causal forest approach, a causal inference method exploits the large feature space characteristic of big data and abstract inherent heterogeneity in treatment effects, we investigate the heterogeneous treatment effects of telehealth/tele psych services adoption on community mental health levels across counties with different socioeconomic and demographic characteristics.

We constructed a panel dataset of the 2814 U.S. counties from for the 2017-2021 by integrating data from the following sources: American Hospital Association (AHA) Annual Survey and the associated IT Supplement Survey, the American Community Annual Survey, and the County Health Ranking Data.

From the difference-in-differences analysis, we find that the adoption of telehealth services by hospitals and their affiliate clinics in US counties has a positive impact, on average, on mental health of the county. From the causal forest analyses, we also find that counties with greater rural population and lower unemployment rate benefit the most from telehealth service adoption within the county. Our findings suggest that telehealth, by offering a particularly effective solution for those in rural areas, can be an effective tool to bridge the urban-rural divide in supporting mental health of communities. The findings also suggest that it is not as effective of a tool to overcome mental health problems for those who are economically disadvantaged because of unemployment, suggesting that other modes of care delivery and supportive resources need to be maintained and offered.

Does Telehealth Mitigate or Exacerbate Privacy Concerns in Mental Healthcare Delivery?

Atiye Cansu Erol¹, Lorin Hitt¹, Lynn Wu¹

1. University of Pennsylvania

Abstract

Mental health conditions are responsible for 23% of all pregnancy-related deaths in the US. Higher access to mental health care in the postpartum period is critical, since postpartum depression (PPD) is a prevalent condition in mothers in the US, with 1 in 8 mothers indicating PPD symptoms and more than half of mothers with depression not being treated. We investigate the potential of telehealth to increase access to mental health services for mothers in the postpartum period. We also analyze the potential of telehealth to overcome a specific barrier to care, privacy concerns in the community, considering the fact that privacy concerns may be especially acute for patients seeking mental health services, where even the appearance at a provider's office is sufficient to signal the presence of mental health issues (in contrast to, say, a primary care office). However, not having enough private space for telehealth at home, i.e. privacy concerns in the home, may counteract these benefits. Analyzing the effect of telehealth availability and the interplay of these two potentially countervailing privacy concerns, we find that 1) the availability of telehealth increases care utilization more in the areas where privacy concerns in the community are high (high stigma areas), 2) for mothers having high privacy concerns in the home, the benefits of telehealth availability are almost completely neutralized, and 3) the concern about privacy at home dominates telehealth's promise to protect privacy in the community for mothers experiencing both privacy forces at the same time.

Regular Session 8A: M-health

A Fully Automated and Culturally Adapted mHealth Intervention for Diabetes Self-Management Among Vietnamese Adults

Anna Nguyen¹, Zsolt Nagykalai¹, Thanh Bui¹, Sixia Chen¹, Valerie Eschiti¹, Kathleen Dwyer¹,

1. University of Oklahoma Health Sciences Center

Abstract

3 Key Takeaways of Our Research:

1. This is the first Vietnamese diabetes self-management and education support (DSMES) intervention using mHealth technology to address the barriers of language and culture differences, diabetes literacy, and transportation.
2. REDCap is a secure web application typically used for managing online surveys and databases. We used it as an mHealth platform to support online and offline data capture for research studies and operations. Our study team created the intervention in REDCap and set up the structure to support both English and Vietnamese language to be delivered simultaneously.
3. Vietnamese-speaking individuals have few opportunities to participate in randomized clinical trials about their healthcare. This study allows such an opportunity to promote self-

determination, independence, and improve self-management strategies. The stepped wedge design allows participating clinics and patients to be randomized with an opportunity to participate in the intervention period.

Background/Purpose:

Language differences, health literacy, and cultural practices may be barriers to DSMES program participation and potentially result in ineffective self-management, but these factors can be addressed with appropriate interventions. The purpose of this pilot study is to evaluate a multilevel, multicomponent, culturally tailored diabetes self-management intervention among Vietnamese Americans using mHealth technology.

Research Hypothesis and/or Research Questions and Specific Aims:

This study aims to test the acceptability, feasibility, effectiveness, and sustainability of the intervention. We hypothesized that (1) > 70% of eligible participants who are invited will participate and (2) patients who receive the intervention will have improved self-care behaviors and biological outcomes and maintain improved outcomes at each follow-up.

Methods/Methodology:

The 12-week intervention provides linguistically and culturally relevant diabetes education, health behavior change support, and emotional support through secure mHealth technology to help patients understand the disease process and gain skills to manage their condition. This multicenter, stepped wedge cluster randomized controlled trial compares the outcomes between the control and intervention period, and between baseline and follow up data. Participating clinics begin with eligible patients enrolling in the control period and receiving standard care. Each clinic is randomly assigned to an implementation date. During the intervention period, patients receive standard care plus the DSMES mHealth intervention. This study incorporates the RE-AIM framework to evaluate reach, effectiveness, adoption, implementation, and maintenance of the intervention.

Results:

We invited 10 clinics with 8 who agreed to participate (80%). Using the RE-AIM framework, we found that our tailored DSMES mHealth intervention is feasible, acceptable, and sustainable to change health behaviors and improve biological outcomes among the participants. Outcome measures include the translated and validated Summary of Diabetes Self-Care Activities (SDSCA) scale, lab results (triglyceride, LDL, HDL, HbA1c) and clinical assessments (BMI, BP, weight). We will have more data on participant self-care behaviors and biological results by May 2023. Lessons learned include the impact of culture on study design, intervention development, health outcome measures, and teaching/learning modalities.

Conclusion:

This study provides a better understanding of the implementation process, allows fine-tune of the intervention, accelerates the pace of moving evidence-based interventions to practice, and provides a replicable implementation model that can be tailored to other ethnic populations. Furthermore, this study provides guidance for future research in developing and optimizing mHealth features for intervention

effectiveness and suggests culturally-relevant approaches for successful implementation of mHealth interventions.

Covid-19 Apps - Population Health Crisis Management and Impact around the world

Deepti Singh¹ and Shivendu Shivendu²

1. California State University - Long Beach, 2. University of South Florida

Abstract

The covid-19 pandemic challenged the digital capabilities of population health management in countries around the world. Several COVID-19 mobile apps were launched as a response in a record timeframe. A systematic empirical investigation into the effect of these COVID-related apps on the pandemic itself in its early phase is much needed. This study explores what effect the population adoption and penetration of these apps had on the number of infections when the country-level and the app level factor have been controlled for. We collected data on the 100 most populated countries in the world and collected information regarding their digital and non-digital strategy to deal with the Covid-19 population health crisis during the first phase of infections (Feb 2020 – May 2020). We collected information about all the COVID-19 apps launched in different countries during this timeline from both App Store and Google Play. The number of installs, ratings, reviews and rating scores were used as indicators of adoption and penetration within the population. We collected the Covid-19 growth rate in these countries along with country-level factors like median age, GDP per capita, strictness, and swiftness of government response. Negative binomial regression and OLS (Ordinary Least Square) regression were used to analyze the data. Our analysis reveals a significant effect of app penetration in the population and sentiments associated with it. Country and app-level factors are significant contributors to the pandemic infection progression as well. Adoption and acceptance of the apps were very important during the initial phase of the pandemic. The adoption of these Covid-19 apps mediates the effect of these apps on population health crisis management.

Smartphone Use, Social Support, and Sleep Health

Shaokang Yang¹, Idris Adjerid¹ and Jiayi Liu¹

1. Virginia Tech

Abstract

Lack of sleep has been declared a public health epidemic, with more than one-third of adults in the United States being sleep-deprived. Persistent shortages of sleep can cause individuals to accumulate sleep debt over time and can deteriorate their brain and other bodily functions. Simultaneously, individuals (in particular younger demographics) are becoming increasingly addicted to digital technology and smartphones. In one study utilizing 1043 participants, thirty-nine percent of individuals between the age of 18 and 30 reported being addicted to their smartphone. According to a survey from

the United States, this number is expected to rise dramatically to 47 percent in 2022. The above phenomenon has made problematic smartphone use (PSU) a growing social problem, where PSU refers to the excessive use of smartphones accompanied by malfunctions, withdrawal difficulties, and other characteristics resembling substance addiction. The conventional wisdom, largely corroborated by a growing stream of research, suggests that these phenomena are related, with several studies suggesting that excessive smartphone use is a contributing factor in the worsening of sleep health.

However, the relationship between smartphone use and sleep health may not be as clear-cut as conventional wisdom suggests. Specifically, some scholars suggest that some uses of smartphones may result in benefits to the users' health. For example, recent work suggests that smartphone-based social interactions can improve users' mental health by preventing them from being isolated and enabling them to obtain social support from their contacts. This suggests that the conceptualization of smartphone use as "screen time" only without much thought given to the type of activities conducted on the phone is overly simplistic and a limiting factor of current work. Relatedly, recent work recommends that future studies also focus on content, context, and experience-centered research. Finally, most of the extant literature measures sleep quality using self-reported data, which have been shown to be subjective and often do not accurately reflect participants' real-time behavior patterns, such as sleep length.

In this study, we examine the impact of pre-sleep smartphone communication activities on one's sleep health, differentiated by communication type (call vs. text). We also examine heterogeneous effects across social closeness and relationship type between smartphone users and their contacts. We leverage a detailed longitudinal data set (2015-2019) on individual phone activity and sleep measured using Fitbit devices. Our results suggest heterogeneous effects of pre-sleep use of smartphones for communication activities: pre-sleep calls decrease sleep duration while text messaging increases sleep length. We also find that the effect of smartphone communication on sleep varies with the smartphone user's social closeness and relationship with the contacts. Our study highlights the nuanced effects of smartphone use on sleep health, including positive effects on sleep health.

Regular Session 8B: HIT & Industry Structure

Are FDA Enforcement Letters the Cure for Misleading Pharma Ads? An Empirical Investigation of Prescription Claims Data

Brad Greenwood¹, Jeffrey Clement², Siddharth Bhattacharya¹,

1. George Mason University, 2. University of Minnesota

Abstract

US prescription drug spending exceeded \$407 billion in 2019, outstripping all other OECD countries while simultaneously grappling with dire issues related to chronic diseases, maternal mortality, and the lowest life expectancy of any OECD nation. At the same time, the US is one of the only jurisdictions which permits

direct- to-consumer pharmaceutical advertising. This is concerning, because such marketing can negatively affect health outcomes, and increase costs, if it drives utilization that is not clinically appropriate. To police against inappropriate efforts, the Food and Drug Administration (FDA) issues both Untitled Letters and Warning Letters to firms. Yet, the efficacy of such letters in deterring marketing malfeasance is unclear. We evaluate firm reactions to such letters using a comprehensive database of US pharmaceutical claims and a single group interrupted time series analysis. Findings suggest that FDA letters slow future growth in drug utilization. However, and strikingly, they yield no drop in utilization, suggesting: i) such letters have limited deterrence effects and ii) pharmaceutical providers are able to retain rents accrued through misleading marketing efforts to consumers. We further find no evidence that markets punish firms for receiving FDA letters (i.e., there is no significant decline in stock after issuance) or impact physician payments from pharma firms (i.e., firms do not change their payments to physicians). Finally, results suggest civil legal suits filed by competing firms for misleading advertising do not prompt competitors to adopt truthful messaging, highlighting limitations of relying on private rights of legal action to protect the public. Taken in sum, these findings suggest that patients are under-protected from false marketing claims and policy makers should evaluate alternate options to deter and punish violative marketing.

Let it Ride! An Empirical Investigation of Problem Gambling and the Implications of Legalized Sports Betting on Societal Outcomes

Gorkem Turgut Ozer¹, Anand Gopal², Brad Greenwood³,

1. University of New Hampshire 2. Nanyang Technological University 3. George Mason University

Abstract

In 2018, the United States Supreme Court held that PASPA, the Professional and Amateur Sports Protection Act, unconstitutionally contravened the Tenth Amendment of the US Constitution, eliminating a near thirty- year federal ban on sports betting and paving the way for more than a dozen states to legalize such operations. The effects have been pronounced, with the market size for sports betting reaching nearly \$100b in 2022, a 72.7% increase relative to 2021. In this work, we examine the downstream implications of these changes in the form of calls to the National Problem Gambling Helpline, drug overdose deaths, and suicides.

Using a difference in differences framework to exploit the phased legalization of sports betting at the state level, results indicate that while the number of calls associated with problem gambling is uncorrelated with the legalization of physical sportsbooks, it is strongly correlated with the legalization of mobile and online sports betting. This effect is further intensified when states permit proposition bets (i.e., prop bets or bets which are about random occurrences within a game, e.g. a wager on an individual player or specific event in a game that is not directly connected to the outcome of a game) or betting parlays (a wager type that combines multiple bets for a greater payout). Results further indicate that the legalization of online sports betting is correlated with an increase in the number of drug overdoses, which remains an outcome

historically associated with problem gambling, and the number of suicides in a jurisdiction, which is also an associated social outcome.

Theoretical and practical implications are discussed within.