objectives are: i) data delivery with data governance and cohort discovery under a managed self-service model and ii) data science and analytics tool for advanced users. METHODS/STUDY POPULATION: Using existing commercial frameworks, we implemented a few pilot self-service tools. The key characteristics of the tools were i) high degrees of functionality and flexibility for data access and data governance, ii) lower cost to build and maintain, and iii) long-term organizational strategic alignment with the academic medical center. We conducted a two-phase evaluation with the pilot self-service tool: functionality-based assessment, prioritizing tools for data science users, and usability-based assessment, evaluating selected tools through customized maturity models and surveys. The evaluation study targeted a focus group study with five diverse faculties and researchers in an academic medical center seeking improved access to research resources. RESULTS/ANTICIPATED RESULTS: In evaluation phase 1, we explored seven self-service tool frameworks suitable for our research data warehouse (RDW). In phase 2, we implemented the top two tools selected from phase 1, QlikView and Palantir Foundry. Although the tool built on Palantir has higher mean and individual scores for user feedback than Qlik's, there is no statistically significant difference. Both tools had steep initial learning curve. Palantir has better feedback from qualitative responses. Our study findings highlight prioritized functionalities (efficiency, flexibility, sustainability, security, and cost reduction) for data science tool users; however features and the tool itself requires long term organizational planning and investment. DISCUSSION/SIGNIFICANCE: Academic and research medical centers strongly focus on efficient pilot data access for researchers to aid hypothesis generation. Establishing a clinical research-focused self-service data tool addresses the well-established demand for research resources and offers a model for similar organizations.

312

In vivo electrophysiology sex differences in the locus coeruleus of wild type F344 rats

Rosaria J. Rae and Lori L. McMahon
Medical University of South Carolina

OBJECTIVES/GOALS: Women are twice as susceptible to ailments like anxiety, PTSD, and Alzheimer's disease compared to men. The locus coeruleus (LC), the primary source of noradrenaline for the brain is implicated in these disorders, however physiological sex differences have never been assessed in the LC. METHODS/STUDY POPULATION: To address this gap, In vivo electrophysiology under anesthesia was used to measure single unit activity of noradrenergic LC firing patterns in 4-month-old wild-type Fischer male and female rats. Recordings measured neuronal activity under basal conditions and in response to a footshock stimulation which elicits burst firing in LC neurons. Single unit activity is sorted via automatic valley seeking scan PCA, additional manual sorting is done via line and template method. Analysis is done extracting interspike interval (ISI) and firing rate of single units, additional analysis is done to quantify properties of bursting patterns (burst duration, spikes per burst, interburst interval, etc . . .). RESULTS/ANTICIPATED RESULTS: This data shows that during LC burst firing, females have longer interspike intervals compared to males, supporting the inhibitory effect of E2 on LC firing. Additionally, females have significantly different waveform patterns than males, indicating possible differences in intrinsic properties, but further supporting sexually distinct physiology of the LC . Because female rats have been estrous cycle tracked via vaginal lavage, stratification into estrus groups and further analysis may uncover differences within females. These data suggest that estrogen acts as a potent neuromodulator of noradrenergic LC neurons, providing valuable insights into the physiology of this brain region. DISCUSSION/SIGNIFICANCE: This study is the first exploration of LC physiological sex differences. This work offers insights into a critical brain region implicated in many diseases, and may pave the way for future therapeutic approaches, particularly for women, who are at a higher risk of neurological disease developing.

313

Identification of novel plasma protein of Community Health Worker Program

Roselyne Tchoua, Kate Karam and Kelly McCabe
DePaul University

OBJECTIVES/GOALS: Thiswork is an evidential study that demonstrates the positive impactof integrating Community Health Workers (CHWs) and SocialDeterminants of Health on an important health outcome, notably in decreasing the 30-day unplanned hospital ED readmissions at Sinai Health Systems. METHODS/STUDY POPULATION: Using datafrom the Sinai Urban Health Institute (SUHI), we compare predictingthe readmissions of patients with and without data pertainingto Social Determinants of Health (SDoH). We thoroughly describe the data cleaning and data preprocessing, done in collaboration with experts in community health. We use a fundamental and ubiquitous classifier in Random Forest for its feature characterization capability in order to translate models results into insights and recommendations for the CHW program. RESULTS/ANTICIPATED RESULTS: We show that when patients are simply engaged byCHWs, regardless of the content of those conversations, we canincrease the predictive accuracy of our classifier by 5%. We usethis result to make recommendations for improving patient careand discuss limitations and future work. Importantly our workpoints directly to the human connection between patients andCHWs as an important feature in the readmission rate. DISCUSSION/SIGNIFICANCE: Our work shows that the predictive capabilities of the classifier increases with CHW logs and SDoH survey data, highlighting the benefit of collecting this information. This is the first step in early identification of such patients so that CHWs are focusing on and providing resources to patients who will most benefit from the program.

314

Large Language Model Approaches to Understand Differences Between Guidelines and Clinician Perception of Best Practices

Carrie E. Gold[1], Jorie M. Butler[1], Ithan D. Peltan[2] and Julio C. Facelli[3]
[1]Department of Biomedical Informatics, The University of Utah, Salt Lake City, UT; [2]Department of Pulmonary and Critical Care Medicine, Intermountain Medical Center, Salt Lake City, UT and [3]Department of Biomedical Informatics and Clinical and Translational Science Institute, The University of Utah, Salt Lake City, UT

OBJECTIVES/GOALS: The Clinical Implementation stage in the translational pipeline is hampered by the tension between formal evidence and clinician perceptions. For instance, when guidelines are translated into electronic clinical decision support alerts, they are often ignored. Using advances in LLMs we present a framework to quantify these discrepancies. METHODS/STUDY

https://doi.org/10.1017/cts.2024.285 Published online by Cambridge University Press

POPULATION: We hypothesize that ignoring guideline-based alerts may be driven by discordances between clinical guidelines' deterministic realities and clinician' perception of clinical reality. Until now this has been very difficult to measure using quantitative methods. We argue that advances in Large Language Models (LLM) provide an avenue for exploring this quantitatively. Here we present the method and preliminary results comparing the responses of BioBERTT from a carefully designed set of questions when the LLM is fine-tuned using either formal guidelines or transcripts of clinicians discussing guidelines and clinical care in the parallel domain. The formal "distance" between the LLM responses is evaluated using quantitative metrics like the Hamming Distance. RESULTS/ANTICIPATED RESULTS: We present a description of the architecture used to prove or disprove our hypothesis. We will present results obtained when training the architecture with data that could be used to test the limits of our hypothesis, by fine-tuning BioBERT with diverse synthetic clinical views, either in agreement or disagreement with the formal guidelines. Results comparing sepsis guideline text with transcripts of interviews with Emergency Department clinicians discussing care practices for sepsis in the ED transcripts will also be considered. Our current emphasis is on securing a wider range of transcripts of clinicians interviewed from different clinical specialties and different clinical settings. While here we focus on clinical guidelines, the framework supports any intervention in the Clinical Implementation stage. DISCUSSION/SIGNIFICANCE: Leveraging recent advances in LLMs, we develop a framework that can quantitatively measure the differences between guidelines and clinician perception of best practices. We demonstrated the functionality of this approach using synthetic data and initiated the collection of clinician transcripts to test the framework in real clinical situations.

316

Machine Learning to Predict Fluid Responsiveness in Hypotensive Children

Sarah B. Walker, Kyle S. Honegger[2], Michael S. Carroll[2], Debra E. Weese-Mayer[2] and L. Nelson Sanchez-Pinto[2]
[1]Ann & Robert H. Lurie Children's Hospital of Chicago and [2]Divisions of Critical Care Medicine and Autonomic Medicine, Department of Pediatrics, Ann & Robert H. Lurie Children's Hospital of Chicago; Stanley Manne Children's Research Institute; Northwestern University Feinberg School of Medicine Department of Pediatrics, Chicago IL

OBJECTIVES/GOALS: Fluid boluses are administered to hypotensive, critically ill children but may not reverse hypotension, leading to delay of vasoactive infusion, end-organ damage, and mortality. We hypothesize that a machine learning-based model will predict which children will have sustained response to fluid bolus. METHODS/STUDY POPULATION: We will conduct a single-center retrospective observational cohort study of hypotensive critically ill children who received intravenous isotonic fluid of at least 10 ml/kg within 72 hours of pediatric intensive care unit admission between 2013 and 2023. We will extract physiologic variables from stored bedside monitors data and clinical variables from the EHR. Fluid responsive (FR) will be defined as a MAP increase by ^{3}10%. We will construct elastic net, random forest, and a long short-term memory models to predict FR. We will compare complicated course (multiple organ dysfunction on day 7 or death by day 28) between: 1) FRs and non-FRs, 2) predicted FRs and non-FRs, 3), FRs and non-FRs stratified by race/ethnicity, and 4) FRs and non-FRs stratified by sex as a biologic variable. RESULTS/ANTICIPATED RESULTS: We anticipate approximately 800 critically ill children will receive 2,000 intravenous isotonic fluid boluses, with a 60% rate of FR. We anticipate being able to complete all three models. We hypothesize that the model with the best performance will be the long short-term memory model and the easiest to interpret will be the tree-based random forest model. We hypothesize non-FRs will have a higher complicated course than FRs and that predicted non-FRs will have a higher rate of complicated course than FRs. Based on previous adult studies, we hypothesize that there will be a higher rate of complicated course in patients of black race and/or Hispanic ethnicity when compared to non-Hispanic white patients. We also hypothesize that there will be no difference in complicated course when comparing sex as a biologic variable. DISCUSSION/SIGNIFICANCE: We have a critical need for easily-deployed, real-time prediction of fluid response to personalize and improve resuscitation for children in shock. We anticipate the clinical application of such a model will decrease time with hypotension for critically ill children, leading to decreased morbidity and mortality.

317

Clinical Informatics for Head and Cancer Patient Management*

Ricky Savjani, William Delery, Myung-Shin Sim and Robert Chin
UCLA

OBJECTIVES/GOALS: The management of head and neck cancers is complicated and allows for a variety of disparate approaches from providers. However, data from several hundred or thousands of patients is necessary to decipher the optimal decisions for populations of patients. And prospective trials would take years to accrue and often are financially not possible. METHODS/STUDY POPULATION: Instead, we collated the entire electronic medical record for all patients at our institution treated for head and neck cancers. We employed a variety of clinical informatics and natural language processing to gather text data into large data frames. We found key conclusions in the diagnostic, treatment, and surveillance of our patients. RESULTS/ANTICIPATED RESULTS: First, obtaining post-operative PET/CT changes in management in over one-third of patients, highlighting the utility of optimizing diagnostic imaging. Second, using a newer silicone-based cream instead of just a moisturizer decreased the absolute risk of grade 2+ radiation dermatitis by almost 15%. Lastly, we are deploying novel autosegmentation frameworks to better understand tissue decomposition in the head and neck to identify patients in need of further nutritional support while undergoing radiation therapy. DISCUSSION/SIGNIFICANCE: Collectively, we showcase the value and opportunity of mining oncological data for the improvement of patient care.

318

Discovering Subgroups with Supervised Machine Learning Models for Heterogeneity of Treatment Effect Analysis

Edward Xu[1], Joseph Vanghelof[2], Daniela Raicu[1], Jacob Furst[1], Raj Shah[2] and Roselyne Tchoua[1]
[1]DePaul University and [2]Rush University Medical Center

OBJECTIVES/GOALS: The goal of the study is to provide insights into the use of machine learning methods as a means to predict heterogeneity of treatment effect (HTE) in participants of randomized clinical trials. METHODS/STUDY POPULATION: Using data

https://doi.org/10.1017/cts.2024.285 Published online by Cambridge University Press