samsi
NSF·Duke·NCSU·UNC

PMED APPM Workshop
March 14-15, 2019

SPEAKER TITLES/ABSTRACTS

Jason Burke
UNC Health Care

"Evidence-based Practice vs. Practice-based Evidence: Innovations in Precision Medicine within Health Care Delivery"

Health careâ€™s ongoing evolution (e.g., accountable care, risk delegation, population health management, consumer technologies, Internet of Things) is creating new imperatives for the way both physicians and patients approach heath care delivery. Clinicians are faced with competing demands: standardize care across patients, but ensure optimal individual health outcomes. Care pathways targeting broad-based patient populations derived from clinical research (i.e., evidence-based practice) must be reconciled with real-world observations (i.e., practice-based evidence). The intersection of these two forces create innovative opportunities for transforming patient care through richer understandings of patients, clinicians, health outcomes, quality measures, processes, and costs. Capturing these opportunities requires new competencies in harnessing data sciences and analytics to make more tailored care decisions, rationalize sparse delivery resources, manage risks, and drive efficiencies that reduce the overall cost of care.

Glen Wright-Colopy
Current Health and Oxford University

Physiological variability over time contains a vast amount of clinical information. However, population-based metrics can easily lose this information in the sea of inter-patient noise. This challenges the ability (of algorithms and doctors alike) to identify an individual patientâ€™s deterioration. The loss of patient-specificity, in turn, hurts clinical performance via high false-alarm rates, missed early warning signs, and emergency readmissions.

A better understanding of patient-specific ranges and patient-specific time-series dynamics can improve early detection of deteriorating patients. However, the real-time design and inference of patient-specific models is not trivial. Nor is it trivial to communicate novel warning metrics to clinical staff.

Iâ€™ll cover how time-series techniques can supplement or completely reverse clinical intuition. More importantly, the algorithmâ€™s output is interpretable and intuitive to clinical experts.

Rachael L. DiSantostefano
Janssen R&D, LLC

"Patient Preference Studies: How they Contribute to Personalized Medicine"

Patients and patient advocacy groups have increasing input into the medical product development lifecycle including study design, endpoint selection, and acceptability of benefit-risk tradeoffs. The science of patient input has gained momentum with initiatives such as the FDAâ€™s 21st Century Cures Act, FDA and EMA pilot projects in patient preferences, and the IMI-PREFER initiative. The objective of this session is to define what is meant by a preference study and demonstrate how patient preferences inform medical product decision making and relate to personalized medicine. A case study looking at preferences in genetic testing will be reviewed to illustrate when preference studies are useful and how they are set up, conducted and analyzed.

Laura Drinkard
Lumeris

"The Opportunity of Population Health in a Fragmented Healthcare Delivery System"

Precision medicine has become an integral component of successful healthcare delivery across nearly every service line, medical intervention, and even payor contract arrangements, such as the ones found in population health. However, since population health's debut into the public marketspace, it has experienced slow growth and resistance from most payors and health systems who typically earn money from each premium dollar under a fee-for-service system. The hallmark of population health lies in alternative financial and clinical models. To make value-based arrangements profitable for both providers and their respective health systems, competitive population health services organizations (PHSO), such as Lumeris, are using a combination of aligned incentive contracting, technology, advanced analytics, and practice transformation to decrease the total cost of care, improve patient outcomes, and enhance the patient experience. As PHSOs continue to enter the market, attention has turned to platform architecture and advanced analytics. Successfully leveraging the latest technology stacks, ensuring interoperability, and integrating and transforming data into actionable insights accessible during the patient-provider clinical encounter are proven catalysts in profitable upside and downside risk arrangements. However, the legal and technical frameworks put into place that created our value-based system will fail without precision medicine. That is, without predictive analytics more accurately identifying the most appropriate intervention for a specific disease or acute event, at the right time, we will continue to see rising healthcare costs, poor health outcomes, and unsatisfied patients.

Jessilyn Dunn
Duke University

"Precision Cardiometabolic Care through Multi-scale Biomedical Data Integration"

Recent technological advancements make it possible to closely and continuously monitor individuals on multiple scales in real time while also incorporating genetic, environmental, and lifestyle information. We are collecting and using this multi-scale biomedical data to gain a more precise understanding of health and disease at molecular and physiological levels and developing actionable, predictive health models for improving cardiometabolic outcomes. We are simultaneously developing tools for the digital health community, including the Digital Biomarker Discovery Pipeline (DBDP), to facilitate the use of mobile device data in healthcare.

Madalina Fiterau
University of Massachusetts, Amherst

"Hybrid Machine Learning Methods for the Interpretation and Integration of Heterogeneous Multimodal Data"

The prevalence of smartphones and wearable devices and the widespread use of electronic health records have led to a surge in multimodal health data that is noisy, non-uniform, and collected at an unprecedented scale. This talk focuses on machine learning techniques that learn expressive representations of multimodal, heterogeneous data for biomedical predictive models designed to interact with domain experts. In the first part of the talk, the focus is on techniques for partitioning data and leveraging low-dimensional structure to enable visualization and annotation by humans. The latter part addresses the construction of hybrid models that combine deep learning with random forests, and the fusing of structured information into temporal representation learning. This array of methods obviates the need for feature engineering while improving on the state of the art for diverse biomedical applications. Use cases include the classification of alerts in a vital sign monitoring system, the prediction of surgical outcomes in children with cerebral palsy, and forecasting the progression of osteoarthritis from subjects' physical activity. Finally, I will present the use of weak supervision for the classification of rare aortic valve malformations from unlabeled cardiac MRI sequences.

Jeff Fuller
UNC Health Care

"Care Variation Analytics at UNC Health Care System"

High Value Care is the goal. This is achieved when we have balance between the ideal outcomes that matter most to the patients and the costs required to achieve those outcomes. Variability within the value equation is intrinsic to healthcare. Variability among clinicians is always present and is often an acceptable way to achieve high value care. However, with ever-present variation in clinical practice, there are infinite opportunities to reduce waste. Advanced analytics techniques were developed at the UNC Health Care System to engage clinicians in meaningful value improvement; targeting capabilities to establish homogeneous cohorts where opportunities to minimize unwarranted variation would be targeted to empower improvement efforts.

Sophie Guo
Deep 6 Al

"Leveraging Data Science in Clinical Trial Matching and Recruitment"

Clinical trials are an integral component of drug development and is a gatekeeper for cutting-edge treatments to make it to the market. The status quo of trial recruitment is a very manual, time-consuming process, resulting in over 70% of clinical trials suffering delays. Beyond the financial impact on pharmaceutical companies from increased trial budget and reduced patent pools, clinical trial delays result in delayed treatments making it to the general population. Deep 6 AI uses cutting-edge data and engineering techniques to find more, better-matching patients for clinical trials in minutes, not months. In this talk, I will discuss the challenges of mining insights from unstructured medical data and how we address the challenges and build a machine learning system that is scalable with growing numbers of patient records.

Telba Irony
Food and Drug Administration

"The Value of Bayesian Approaches in the Regulatory Setting: Lessons from the Past and Perspectives for the Future"

The use of Bayesian approaches for the regulation of medical products has been advocated for the past 20 years, but its rather limited implementation in the US could be expanded. The emerging importance of benefit-risk determinations to inform the approval of medical products presents an exceptional opportunity to advance the use of Bayesian methods because they are uniquely suitable for decision making. While the main challenge of benefit-risk determinations for medical product approval is to combine information with values, the essence of the Bayesian approach is to collect data and update information to merge it with stakeholder values for rational decision making.

In this talk we will discuss lessons learned from the use of Bayesian approaches for the regulation of medical devices, explaining what has worked and what has not worked, aiming to advance the use of Bayesian methods for the regulation of medical products in general.
We will argue that the use of Bayesian approaches in the regulation of medical products will have its highest value when combined with Decision Analysis where patient preferences and the context in which the medical product is evaluated are considered and integrated into the process.

Alistair Johnson
MIT

"Open Data to Accelerate Medical Research"

Fields such as computer vision and natural language processing have seen unprecedented advances in a number of practical applications. Key to these advances, in addition to new techniques, was the existence of large databases to develop models. In medicine, such databases are few and far between. In this talk I will argue that openly available data is a necessary first step for the development of robust and clinically accepted tools. I will discuss the impact of our publicly available database, MIMIC-III, and show recent results demonstrating the need for more initiatives of the sort.

Dukka B. KC
North Carolina A&T University

"Protein Sulfenylation Site Prediction using Support Vector Machine (SVM)"

Cellular homeostasis is maintained through changes in proteome and metabolome. Post-translational modification have important roles as switches of cellular homeostasis and are critical for pathogenesis. Hence, various types of PTMs have been used as clinical markers of diseases. Protein *S*-sulfenylation is a type of post-translational modification (PTM) that involves the covalent binding of a hydroxyl group to the thiol of a cysteine amino acid. Recent evidence has shown the importance of *S*-sulfenylation in various biological processes, including transcriptional regulation, apoptosis, cytokine signaling, metabolism, inflammation and disease processes. Moreover, S-sulfenylations affect regulators of phosphorylation, acetylation and ubiquitylation, which suggests regulatory crosstalk between redox control and signaling pathways.

Determining the specific sites of *S*-sulfenylation is fundamental to understanding the structures and functions of *S*-sulfenylated proteins. However, the current lack of reliable tools often limits researchers to use expensive and time-consuming laboratory techniques for the identification of *S*-sulfenylation sites. In this talk, I will discuss SVM-SulfoSite, a novel sulfenylation prediction tool that uses support vector machines (SVM) to identify key determinants of sulfenylation among five feature classes: binary code, physiochemical properties, k-space amino acid pairs, amino acid composition and high-quality physiochemical indices. Using 10-fold cross-validation, SVM-SulfoSite achieved 95% sensitivity and 83% specificity, with an overall accuracy of 89% and Matthew's correlation coefficient (MCC) of 0.79. Our method represents a robust and complementary technique for advanced exploration of protein S-sulfenylation. Our method should find application in understanding the mechanisms of Sulfenylation as well as in drug discovery.

Warren Kibbe
Duke University

Data, Analytics and Precision Oncology

Changes in technology, computing, and modeling are radically transforming our understanding of biology, cancer and our ability to measure, predict and intervene as individuals go through their life trajectories. Specific examples in applying machine learning techniques to healthcare and opportunities for wearables and sensors in understanding patient trajectories as well as population health will be discussed.

Heather Kopetskie **Derek Lawrence**
Rho, Inc. Rho, Inc.

"Combating Data Source Diversity and Complexity with Expanded Skills and Teamwork"

The number of sources of data in a clinical trial is increasing and the data coming from those sources are more complex than ever. What does that mean for the future of data handling and interpretation? It is a broad problem that touches more areas than just traditional clinical data management. This presentation will address how skill sets across multiple clinical research disciplines including bioinformatics, statistics, project management and clinical operations will need to broaden, and discuss why greater teamwork and cross-functional collaboration are needed.

Rebecca Krouse
Rho, Inc.

"Modernizing Clinical Research with Interactive Open Source Tools"

Clinical trial research is highly regulated and notoriously slow moving. As a result, modern data analysis tools and techniques that have become commonplace in other environments have struggled to achieve broad or consistent adoption. Rho's Data Visualization team strives to modernize key clinical trials processes by building interactive web-based tools using open source technology. These tools allow clinical trial researchers to gain insight on data related to research subject safety (vital signs measures, lab results, adverse event tracking) as well as other operational metrics. The power of interactivity is the ability to concurrently monitor protocol-wide trends in subject wellbeing as well as the status of individual subjects in the study. Furthermore, these tools are free, reproducible, customizable, and easy to use.

Ilya Lipkovich
Eli Lilly and Company

"Overview of Methods for Subgroup Identification in Clinical Trials"

This talk will provide a review of a broad class of statistical methods dealing with exploratory subgroup analysis in clinical trials as one of the key components of personalized medicine. This includes methods that can be applied in both early and late-phase clinical trials, as well as to observational studies (for some of the methods). A broad taxonomy of existing approaches to subgroup/biomarker identification will be given illustrating the key elements of principled data-driven subgroup evaluation using data from a case study.

Carol Mansfield
RTI Health Solutions

"Patient Preference Studies: Applications of Patient Preference Information to Personalized Medicine"

Patient preference information can be useful in a variety of settings related to personalized medicine. This talk will cover examples of how the analysis of patient preference data can be used to understand differences across patients. Examples will be provided that present how the analysis of patient preference data can be used to assess subgroups and individual-level preferences and stratify patients by risk tolerance. A second example will focus on a project that used patient preference data to estimate individual-level p-values that reflect a patient's risk tolerance. Much of the discussion around personalized medicine focuses on developing treatments that target the unique characteristics of each patient, and patient's preferences will play an important role in the acceptance and success of personalized treatments.

Christopher McCann
Current Health

“Personalized Patient Vital-Sign Monitoring in Home, Hospital, and Clinical Trial Settings”

The continuous monitoring of patientsâ€™ vital-signs is an ideal setting to showcase the success of personalized medicine. A patientâ€™s vital signs (such as heart rate, respiratory rate, oxygen saturation) provides invaluable context in acute care settings as well as real-world homecare.

It is widely accepted that current clinical practice (population-based approach to monitoring vital-signs) is hindered by high inter-patient variability. However, a personalized approach to data collection is even more fundamental, and requires analysis in its own right.

This presentation will showcase patient-specific approaches to monitoring (i) data quality and (ii) a patientâ€™s adherence to protocol. From this we trust our patient-specific insights for clinical interventions.

Jeffery L Painter, JD
Jivecast

“Crowd Sourcing for Patient and Physician Medical Insights”

Crowd sourcing is a novel method for reaching both patients and physicians to gain real time feedback and insights into a variety of medically relevant questions.

In this talk, I will highlight the pros and cons identified through implementing several projects where crowd sourcing was used. Examples include (1) targeting difficult to find patient populations (e.g. pregnant women with lupus), (2) understanding treatment options and patient opinions regarding UTI and gonorrhea, (3) understanding physician hurdles to enrolling patients into a pregnancy registry, and more.

Paolo Piraino
Bayer AG

“Subgrouping of Heart Failure Patients using Functional Parameters Collected in Real Life: a medical device perspective”

A variational autoencoder was designed for non-linear dimensionality reduction of medical monitor data collected in real life from 500 subjects affected by heart failure, either with preserved (HFpEF) or reduced (HFrEF) ejection fraction. In multiple visits, subjects were continuously monitored for one week for activity intensity, posture, hearth and respiration rate, and fluid status, while their quality of life was assessed according to standard validated questionnaires at each visit. The lower dimensional representation of the monitor data learned by the autoencoder was used in cluster analysis. For each group, average daily profiles of the functional parameters were computed and interpreted together with clinical profiles and the observed changes in quality of life across visits. The latent information extracted from the monitor data, its clinical relevance and potential application in clinical trials are discussed.

Russell Reeve
IQVIA

“Precision Dosing Based on Population PK Models to Minimize Blood Draws”
Precision dosing algorithms aim to dose a patient with a dose that achieve a target for each individual patient. As such, this dosing improves the efficiency of study medication and possibly reduces the adverse events associated with a therapeutic agent. Many drugs with narrow therapeutic indices require a prÃ©cising dosing approach; but other agents can also benefit from this approach. Furthermore, clinical trials built using precision dosing, for instance randomized concentration-controlled trials, are more efficient than even adaptive designs using conventional non-precision doses. However, finding the right dose for each patient can be challenging without the appropriate pharmacokinetic models. We discuss the background for precision dosing, and describe an approach we have taken that uses population pharmacokinetic models, and software to develop precision dosing applications.

Xiling Shen
Duke University

“Precision Cancer Therapies Targeting Metastasis”

Cancer metastasis accounts for the majority of cancer-related deaths and remains a clinical challenge. Metastatic cancer cells may be influenced by the microenvironment of the organs they colonize, providing new therapeutic strategies that have not been explored. Based on meta-analysis of extensive clinical datasets and in vivo metastasis models, we show that colorectal, breast, and pancreatic cancer cells undergo metabolic reprogramming after they metastasize. In particular, metastatic cells in the liver up-regulate the fructose metabolism enzymes to fuel tumor cell proliferation. Targeting the fructose metabolism enzymes or reducing dietary fructose significantly reduces liver metastatic growth. Collaboration with Pfizer enables us to test the first potent small molecule inhibitor against the fructose enzymes in our patient-derived preclinical liver metastasis models. Furthermore, working with the Duke Woo Center for Big Data and Precision Medicine and with industrial partners, we are launching new, first-of-its-kind organoid-based precision clinical trials for metastatic colon cancer aided by a data driven approach.

Ben Villard
Nagoya University

Colorectal cancer is one of the leading cause of cancer related deaths among industrialized countries with increasing prevalence. Due to the increase in colorectal cancer screening programs, there has recently been a decrease in polyp differentiation into malignant cancer cases. However, if not treated appropriately, metastases can form, complicating the chances of survival. The Japanese Medical Research and Development ministry has been working in conjunction with companies to developing a computer aided diagnosis tool to help clinicians in the diagnosis and prediction of colorectal cancer.

As part of a clinical study, clinicopathologic and demographic characteristics of 984 colorectal cancer patients were obtained and statistical analysis was performed to select the most discriminant features. An ANN was then used to predict the likelihood of metastasis occurrence and classify whether the patient was at risk. An analysis of the neural network architecture was performed as well as the optimal use of various data analytic methods such as one hot encoding and standardization types. Due to the unbalanced nature of the pathology, the training and testing set were randomly generated to reflect the appropriate unbalance. In order to evaluate our models, the original dataset was randomly partitioned into 100 balanced training/testing datasets. Each network was trained on each training set and validated on the relevant testing set. This allowed us to assess the accuracy of our models' designs.
Our final model is able to predict metastases occurence with a mean balanced accuracy, sensitivity and specificity of 77.1%, 82.5% and 71.7%, respectively.

Herbert Weisberg
Causalytics LLC

“Eliminating the Irrelevant: The HARVEST Algorithm”

Feature selection with high-dimensional data and a small proportion of relevant features poses a severe challenge to standard statistical methods. We have developed a new approach (HARVEST) that is straightforward to apply, albeit somewhat computer-intensive. This algorithm can be used to pre-screen a large number of features to exclude those that are truly irrelevant. The basic idea is to evaluate each feature in the context of many random subsets of other features. HARVEST is predicated on the plausible assumption that an irrelevant feature can add no real predictive value, regardless of which other features are included in the subset. Conversely, a relevant feature will have predictive value in at least some of the random subsets. Empirical analyses and simulations produced so far indicate that the HARVEST algorithm is highly effective for predictive analytics.

Co-authors: Victor Pontes, Mathis Thoma

Richard Zink
TARGET PharmaSolutions Inc

"From Real World Data to Real World Evidence: A Case Study of Direct-Acting Antivirals for the Treatment of Hepatitis C Infection"

The 21st Century Cures Act was signed into law in December of 2016. One of the provisions of the act charged the U.S. Food and Drug Administration (FDA) with developing a framework for how real world data (RWD) may be used to support the approval of new indications for currently marketed products or to satisfy post-approval requirements. Compared to traditional forms of evidence (i.e. clinical trials), the medical product industry is still learning how and when RWD may be utilized to answer questions of clinical and regulatory significance. In this presentation, we describe a case study for direct-acting antivirals used to treat hepatitis C infection that illustrates how RWD may be used in lieu of, or in conjunction with, data obtained from clinical trials to provide crucial insight into patient management and care. We provide specific examples of questions that have been, or are in the process of being, answered for the FDA and the European Medicines Agency. We discuss the advantages and disadvantages of RWD compared to other sources, and describe some novel methodologies for identifying patients with enhanced response.