Johns Hopkins inHealth

Precision Medicine Symposium 2023
Addressing the Challenges of Precision Medicine Data Management
Vision
• Understand the factors that underlie the pathobiology of COVID-19, the progression to severe illness or death, and the effectiveness of therapeutic interventions in order to provide personalized care to patients infected with SARS-CoV-2.

Mission
• Improve the care of patients infected with SARS-CoV-2 by learning about COVID-19 pathobiology, likelihood of disease progression, and impact of specific therapeutic interventions. We aim to provide this information to clinicians, patients and family members at the point of care.

Research Aims
• Furthering precision medicine by reducing the costs and barriers associated with managing precision medicine data.
Addressing the Challenges of Precision Medicine Data Management

Focus
- EMR data require considerable processing to produce analysis-ready data.
- Sharing of data management code is complicated by lack of standardization across data projections.
- Development of a flexible, readable, and well-documented workflow suitable for sharing is necessary.

Method and Analytics
- Workflow code was developed in an inHealth Github repository, designed for the Crunchr containers with R versions 4.0 and 4.2.
- Credentials were managed using the Linux Secret Service and `keyring` package in R.
- Varying table and column names were addressed using an abstraction layer: names were retrieved from project-specific configuration files managed by the `config` package, and inserted into `dplyr` queries using ‘non-standard evaluation’ in R.
- Metadata tables were used to avoid ‘hard-coding’ measure names and values of interest. Range checks (high/low) were implemented using metadata tables.
- Comorbidities were extracted using the `comorbidity` package in R. Height, weight, and BMI values were cleaned using robust Z-scores to identify potential data entry errors in values or units.
- Labs, vitals and meds, were extracted from flowsheets and CCDA provisioned tables. Regular expressions were used to standardize unit labels.
- SQL Server command line tools were used to bulk copy processed data to the database.

Results and Highlights
Data are extracted from CCDA curated tables and flowsheets. Queried data are cleaned, and then saved to the scratch projection. Cleaned data are further processed for Q4H/Q6H/Q24H summaries from admission and initiation of mechanical ventilation: min, median, mean, max, count, first, and last. Scores (WHO COVID-19 severity, SOFA, etc.) are calculated from summaries. Reports visualize quality and completeness.

Configuration & Workflow:
1. Update and install Linux/R packages
2. Secure credentials in keyring
3. Create project-specific configuration
4. Run .Rmd reports and check results

```r
# Example code
library(tidyverse)
library(config)
```

Tables 1-2: Examples of using metadata tables to select, group, and filter raw values. Data processing can be modified just using the tables without editing the underlying code.

<table>
<thead>
<tr>
<th>meas_id</th>
<th>variable</th>
<th>low</th>
<th>high</th>
</tr>
</thead>
<tbody>
<tr>
<td>301250</td>
<td>map_arterial</td>
<td>0</td>
<td>300</td>
</tr>
<tr>
<td>301360</td>
<td>map_cuff</td>
<td>0</td>
<td>300</td>
</tr>
<tr>
<td>30460914</td>
<td>map_cuff</td>
<td>0</td>
<td>300</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>o2_device</th>
<th>category</th>
<th>rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasal cannula</td>
<td>nasal cannula</td>
<td>2</td>
</tr>
<tr>
<td>RAM nasal cannula</td>
<td>nasal cannula</td>
<td>3</td>
</tr>
<tr>
<td>Venturi Mask</td>
<td>mask</td>
<td>3</td>
</tr>
<tr>
<td>Face Mask</td>
<td>mask</td>
<td>3</td>
</tr>
</tbody>
</table>

Conclusion
- Data management remains a significant challenge, even using a common data model.
- R Studio naturally integrates with GitHub, allowing for version control and easy access to its software development tools.
- R Markdown combines R, SQL, and Python in a single notebook or report, combining their strengths and facilitating collaboration.
- ‘Non-standard evaluation’ in R allows for flexibility in code evaluation that is not possible in SQL and Python.
- Metadata can serve as documentation of code and control how code behaves.
- Bulk copy significantly reduces the time necessary to write data to a database.
- Moving towards a shared workflow for common data elements can improve data quality and reduce data management costs.

Next Steps
- The existing repository is based on the CCDA and Epic data model.
- Further work is necessary to query the primary data from OMOP format data projections.
Improving Accuracy and Efficiency of Medication Data Classification in Electronic Health Record Data

Jack Bitzel, Heather Thiessen-Philbrook, Steve Menez, Chirag Parikh, Prakash Nadkarni

Kidney Precision Medicine Center of Excellence
Vision

• Efficiently and effectively identify prescription medication records from electronic medical record data

Mission

• Develop a standardized and convenient process to programmatically extract medication records for use in both research and clinical care

Research Aims

• Create a linkage to an existing medical language system to facilitate the use of pre-existing and standardized relationship hierarchies
• Design a user-friendly interface for the extraction of medications and complex medication classes
• Compare brute-force approaches to medication classification to a programmatic approach

Interested in Collaboration?

Contact: Jack Bitzel
jbitzel1@jh.edu

Kidney Precision Medicine Center of Excellence
Improving Accuracy and Efficiency of Medication Data Classification in Electronic Health Record Data

Jack Bitzel, Heather Thiessen-Philbrook, Steve Menez, Chirag Parikh, Prakash Nadkarni

**Background**

The accurate categorization of medications is a crucial requirement when utilizing electronic health record (EHR) data in research and clinical settings.

Johns Hopkins’ EHR stores medication data at a detailed level, complicating the identification and extraction of medication data from a high classification level.

**Specific Aims**

- Design a classification process that will provide Principal Investigators with accurate medication information from any level of classification
- Minimize manual involvement in the classification of medication records

**Methods**

- Establish a link between JHU medication records and unified medical language system (UMLS) concepts
- Query UMLS hierarchical relationships to identify concepts of interest
- Obtain JHU medication data

**Results**

To test our process, we focused on a single, large class of medications: non-steroidal anti-inflammatory drugs (NSAIDs).

**NSAID Linkage example of JHU Medical records and UMLS concepts**

<table>
<thead>
<tr>
<th>medication_name</th>
<th>simple_generic</th>
<th>generic_name</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXCEDRIN EXTRA STRENGTH ORAL</td>
<td>aspirin/as acetaminophen/caffeine</td>
<td>aspirin as acetaminophen caffeine or oral</td>
</tr>
<tr>
<td>FELDNE ORAL</td>
<td>piroxicam</td>
<td>piroxicam oral</td>
</tr>
<tr>
<td>PLOCAPINE 2 % EPINEPHRINE EYE DROPS</td>
<td>plicarcaine HCl/epinephrine</td>
<td>plicarcaine 2 % epinephrine eye drops</td>
</tr>
</tbody>
</table>

**UMLS relationship hierarchy given NSAIDs as initial concept**

The UMLS hierarchy approach extracted 800 additional medication records that the string search approach had missed, such as bromfenac, valdecoxib, and nepafenac.

**Conclusions and Next Steps**

Our hierarchical approach provides a reliable, exhaustive method for classifying medication data stored in EHR. Future research includes creating a process suitable for integration into a production environment.
A Comparison of Frequency Oddball and Visual Stroop Event-Related Potentials in Treatment Resistant and Ultra-Treatment Resistant Schizophrenia
Vision:
• Improve the outcomes for schizoaffective patients by establishing disease subtypes based on treatment response.

Definition:
• Schizoaffective disorder is broadly defined to include schizophrenia with or without prominent mood symptoms.

Research Aims:
1. Subtype the clinical phenotype of schizoaffective disorder based on treatment response.
2. Subtype schizoaffective disorder based on EEG and neuroimaging signature.
3. Use biological samples to establish the neurobiological distinction among schizoaffective disorder subtypes.
4. Establish precision therapeutics for schizoaffective disorder.
5. Use data analytics of the Johns Hopkins EMR to establish a learning health system for care of schizoaffective disorder patients.

Interested in Collaboration?

Contact PI Dr. Russell Margolis

Phone: (410) 227-3660

Email: rmargoli@jhmi.edu
Schizoaffective disorder, defined as schizophrenia with or without prominent mood symptoms, is a heterogeneous disease that affects ~1% of the population. 30% of patients do not respond to standard antipsychotic medicines (treatment-resistant schizophrenia, TRS), and ~30% of TRS do not respond to the atypical agent clozapine (treatment-resistant schizoaffective, UTRS). We hypothesize that neurobiological differences underlie these subtypes of schizoaffective disorder. We are exploring these differences with multiple tools, including EEG to measure differential electrophysiological responses to auditory and visual stimuli. We selected event-related potential (ERP) amplitude changes to assess preattentive and attentional control processes.

Methods and Analytics

Patients were recruited from Johns Hopkins Bayview Community Psychiatry Program. All 18 schizoaffective disorder patients were on clozapine, after failing two or more typical antipsychotic medications. Treatment response to clozapine was defined based on PANSS total score (TRS ≤ 58, N=11, Age = 34.6 ± 14.7, F/M = 1/10; UTRS ≤ 58, N=7, Age = 31 ± 6.5, F/M = 3/4). All patients had normal hearing. Patients with visual impairment were glasses throughout EEG experimentation. EEG data in both schizoaffective disorder groups were recorded using Brain Vision Recorder software. 32 electrodes were mounted in a fitted cap with a standard 10-20 layout. Electrodes on the cap were initially referenced to the electrode at the vertex (Vz). Average electrode impedances was <20 kΩ. The EEG data were sampled at 1000 Hz. In EEG preprocessing steps, electrodes were averaged central-posterior windows to assess differences in amplitude. Multiple regression analysis was performed on clinical (PANSS) and electrophysiological measurements. This study was approved by the Johns Hopkins IRB.

This study was supported by NIMH T32MH015339, the Abramson Fund, and Johns Hopkins iHealth.

Results and Highlights

Figure 1: Visual Stroop Experiment (P200)

- 1st row, TRS in Blue, UTRS in Green, 2nd row sig. P-values.
- Topographic map of TRS (1st row) vs. UTRS (2nd row), P-value is in the 3rd row.
- Diminished P200 (120-170ms) amplitude in both congruent and incongruent trials in UTRS compared to TRS patients (F = 7.60, p = 0.018, r² = .32) at left frontal electrodes.

Figure 2: Visual Stroop Experiment (P300)

- Diminished P300 (310-430ms) amplitude in both congruent and incongruent trials in UTRS compared to TRS patients (F = 11.77, p = 0.004, r² = .44) at central-parietal electrodes.
- Attention-congruency more impaired in UTRS.

Figure 3: Passive Frequency Oddball Experiment (P200)

- Diminished P200 (130-160ms) amplitude during deviant oddball trials in UTRS compared to TRS patients (F = 5.93, p = .027, r² = .27) at central-parietal electrodes.
- Also, a trend towards significance (F = 4.05, p = .051, r² = .20) at left frontal electrodes.

Figure 4: Negative correlation between total PANSS and averaged amplitude for visual Stroop at P200 (frontal electrodes) and P300 (parietal electrodes) time-windows

Figure 5: Negative correlation between total PANSS and averaged central-parietal amplitude for deviant frequency oddball trials at P200 time-window

Conclusion and Next Steps

Taken together, these preliminary ERP findings suggest that deficits in preattentive sensory processing and attentional control in both visual and auditory modality may distinguish UTRS from TRS. These results support expansion of this study, including adding patients who respond to standard antipsychotics and healthy controls. Longitudinal studies before and after starting clozapine will be necessary to distinguish whether the ERP differences observed here reflect severity of symptoms or unique neurobiological differences between TRS and UTRS subtypes of schizoaffective disorder.
Hindbrain organoids from patient-derived iPSCs to evaluate pharmacological treatment for neuropsychiatric symptoms in Alzheimer’s disease
The Richman Family Precision Medicine Center of Excellence in Alzheimer’s Disease

- **Vision**
  Provide early detection, prognosis, and/or predicting response to available medications; develop new blood tests to better monitor treatment response; and, using blood-derived stem cells, test whether patients will be helped by emerging medication treatments.

- **Mission**
  Translate research into clinical care by defining patients who respond optimally to currently available and emerging therapies. Identify new therapeutic targets based on refined and granular understanding of disease mechanisms.

- **Research Aims**
  Generate patient-derived iPSCs and differentiate them into hindbrain organoids to study drug responses targeting neuropsychiatric symptoms in Alzheimer’s Disease.

Interested in Collaboration?

Contact us at:

- **Phone:** (410)-258-0926
- **Email:** vmachai1@jhmi.edu
Results and Highlights

Focus
Alzheimer's disease (AD) is a neurodegenerative disorder and the most common cause of dementia. Almost all AD patients eventually suffer from neuropsychiatric symptoms (NPS; e.g., agitation, depression) whose emergence correlates with dysfunctional serotonergic systems. Our aim is to generate hindbrain organoids containing serotonergic neurons using induced Pluripotent Stem Cells (iPSCs) from healthy volunteers or AD patients with and without NPS. The organoids can be used to study AD, NPS, and to evaluate individual differences in disease progression, NPS development, and/or pharmacological treatment response.

Method & Analytics
- iPSC lines were generated from peripheral blood mononocytes (PBMCs) of 3 healthy individuals and from 3 AD patients, two of which were diagnosed with NPS.
- The 6 iPSC lines were differentiated into hindbrain organoids. The presence of serotonergic neurons was confirmed by quantitative RT-PCR, flow cytometry, detection of released serotonin in the extracellular environment, and confocal fluorescence microscopy.
- Endothelin receptor (selective serotonin receptor inhibitor) was used on the organoids to evaluate treatment response by ELISA.

Conclusion and Next Steps
We successfully generated hindbrain organoids from human iPSCs. Organoids from different people respond differently to the application of escitalopram in vitro, possibly revealing distinct subgroups of AD patients. We propose that this 3D platform might be effectively used for drug screening purposes to predict patients with NPS most likely to respond to treatment with escitalopram.

Figure 5: Serotonin release from 5-HT-organoids upon 1 h treatment with different concentrations of escitalopram oxalate (mean ± SD, n=3). Response to the drug was dose-dependent when present. Furthermore, our 3D in vitro platform indicated that human iPSC-derived hindbrain organoids from different people respond differently to the same dose of drug. For individuals healthy_1 and AD_3 there was no observable change in extracellular 5-HT with treatment, whereas for the others there was a 2 to 3-fold increase in 5-HT. While 1h of treatment was sufficient to elicit such a response, the higher 100 µM dosage was necessary.

Table: 5-HT levels [ng/mL] / 5-HT levels at 0 µM treatment [ng/mL]

<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>5-HT levels at 0 µM treatment</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 µM</td>
<td>5-HT levels</td>
<td></td>
</tr>
<tr>
<td>10 µM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100 µM</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4: Characterization of 5-HT-organoids.
All 6 iPSC lines were successfully differentiated into organoids containing serotonergic neurons as evidenced by the detection of TUJ1 and, more specifically, TPH2 by flow cytometry (a) (mean ± SD, n=3). The secretion of 5-HT from two representative samples of the organoids was measurable above the lower limit of detection (LOD=0.49 ng/mL) for over a month after the initial 42 d required for differentiation (b) (mean ± SD, n=1). Detailed analysis of different markers by qRT-PCR for all 6 cell lines further validated the differentiation (c) (mean ± SD, n=3).

Figure 3: Immunocytochemistry for the 5-HT organoids. Confocal images of the organoids show distinct regions of differentiated cells (scale bar: 100 µm).

Figure 2: Reprogramming of PBMCs into iPSCs. Fluorescence microscopy imaging upon staining of iPSCs with pluripotency markers Oct4, Nanog and TRA-1-60 confirms the successful reprogramming (a). The enrichment of TRA-1-60 positive cells was additionally validated by flow cytometry (b).

Figure 1: Generation of 5-HT-organoids from iPSCs. Schematic overview of the differentiation protocol (a) and representative brightfield (BF) images of cells and the organoids at different stages (adherent iPSCs, organoids at days 1, 21 and 42 (b).
Multiple Sclerosis PMCOE: “Filling Research Gaps with Routine Clinical Imaging”
Multiple Sclerosis PMCOE

Goals:
The Johns Hopkins MS PMCOE was launched in April 2017 with two primary goals:
1. Identify clinical, imaging, and blood biomarkers of long-term disability risk
2. Translate evidence to clinical trials aiming to identify new therapeutic strategies to prevent disability and promote repair in people with MS

Research Aims:
- Technology-enabled tracking of neurologic functional performance and systematic clinical data capture at every clinic visit via an internally-developed Smartform
- Baseline and annual non-invasive imaging of optic nerve damage using optical coherence tomography (OCT)
- Collection of blood at biannual clinic visits for research to identify biomarkers of prognosis and treatment response
- Standardization of annual surveillance brain magnetic resonance imaging (MRI) across (and beyond) the Johns Hopkins Health System
- Home-based collection of data regarding environmental and lifestyle exposures that may be relevant to the prognosis of MS.

COE Achievements:
- The COE includes 13 MS neurologists as well as experts in MS epidemiology, neuroimaging, neuropsychology, and neurorehabilitation
- Over 2,000 people with MS have participated in the PMCOE.
- The Johns Hopkins MS Smartform is being widely adopted across other prominent institutions throughout the US.
- Smartform is enabling efficient collation of information relevant to MS state and is being leveraged to graphically project the individual disease course for a given patient.
- Smartform data, with other Epic-derived data, is being used to more specifically determine if a given person will likely benefit from brain MRI scan at a given timepoint (early estimates suggesting ~60% of patients are correctly classified as not requiring a new scan).

Interested in Collaboration?
emowry1@jhmi.edu
Focus

- Imaging is a core component of the MS PMCOE
- Analysis of routine clinical images is hampered by logistical, technical, and regulatory challenges
- PMAP allows the MS PMCOE to collect, curate, and analyze images stored in clinical systems
- Using routine clinical imaging extends imaging follow-up for research studies
- Harmonization methods allow for more images to be considered for research analyses

Method and Analytics

- The MS PMCOE MR imaging projection includes >5,700 patients and >15,600 MR imaging sessions that are ready for volumetric analysis
- Analysis-ready scans include:
  - Multiple contrasts
  - >=1 3D-acquired image
- Analysis tasks include whole-brain segmentation, lesion segmentation, and cortical reconstruction

Results and Highlights

- Rapid increase in scan availability in recent years is shown in Figure 1
  - Nearly 2,000 analysis-ready scans acquired in both 2021 and 2022 (2023 projected to be similar)
  - Increase is due to improving technology and standardizing of high-quality scans for people with MS
  - The potential for longitudinal analysis is shown in Figure 2
  - Many patients have little or no follow-up imaging
  - There are 2,543 patients with some follow-up, 1,100 with ≥5 years, 300 with ≥10 years

- Analysis-ready scans can be used to enrich research cohorts (with IRB approval) to lengthen follow-up or enroll more subjects (can’t or wouldn’t commit to a research-only MRI)
- Multiple studies have extended (or are looking to) their cohorts (or cohort follow-up) using MS PMCOE data
  - Two examples (HEAL-MS and MS-Genetics) are shown in Figure 3

- Analysis-ready scans can be used to enrich research cohorts

Conclusion

- The MS PMCOE has a defined imaging cohort for investigation of neuroimaging outcomes in people with MS
- The cohort continues to acquire 1000s of analysis-ready scans per year with history of nearly 20 years
- Most patients in the cohort have follow-up imaging needed for longitudinal analysis
- Research studies that enroll participants from the MS Center at JHMI can leverage routine clinical images to enrich their cohorts

Next Steps

- Ongoing analysis of all analysis-ready images from the MS PMCOE
- Generating quantitative results for ongoing and future studies
- Investigating spinal cord imaging in MS PMCOE
  - ~1.5x increase in scans
- Improving harmonization techniques to include even more routine clinical images (e.g., low-resolution 2D images)
  - ~5x increase in scans

Acknowledgements: We would like to thank the PMCOE providers, coordinators, staff, and collaborators who made this work possible.
Expansion of Neighborhood-level Social Determinants of Health Available in PMAP
Vision
• To allow for the consideration of social determinants of health (SDoH) as a core component of precision medicine

Mission
• To expand the availability of social determinants of health (SDoH) by building and maintaining a dataset available to CoEs through PMAP

Research Aims
• Phase 0 – Consolidate and Enhance Current Census Data in PMAP
• Phase 1 – Develop plan to get the right data moving from broad phenotypes (area level exposure)
• Phase 2 – Move to individual level
• Phase 3 - Capture all SDoH elements that seem beyond reach
Expansion of Neighborhood-level Social Determinants of Health Available in PMAP

Team: Kathryn C. Fitzgerald, Aparna Natarajan, Sharon Penttinen, and Theodore I. Iwashyna

Focus

- Social determinants of health (SDOH) are upstream social and economic conditions, related to where people are born, grow, live, work, and age that impact individual and group differences in health status (Figure 1).
- SDOH incorporate relevant factors that are measured at both an individual and neighborhood-level.
- The Precision Medicine Analytics Platform (PMAP) currently includes information for each patient related to several neighborhood-level indicators of SDOH that are derived the census block geographical unit.
- The currently available resource is limited to a relatively limited set of seven indicators and is largely restricted to patients residing in census blocks groups located in Maryland.

Method & Analytics

- Collated information from publicly available sources for various SDOH domains
  - Census block group
  - Census tract
- Sources of information included in the updated table
  - American Community Survey for 2011-2015 and 2016-2020
  - Area Deprivation Index (ADI) for 2011-2015 and 2016-2020
  - US Census for 2010 and 2020
  - USDA indicators of food accessibility
  - Index of relative rurality (IRR)
  - Persistent poverty
  - Historical redlining
  - Rural-Urban Commuting Area (RUCA) codes
  - Green space (Washington, DC and Baltimore)
  - Superfund sites indicating hazardous waste targeted for clean-up by EPA
- Expanded a code library to allow investigators to derive census block groups from geocodes available in Epic
  - Allow derivation of census block groups beyond Maryland

Results and Highlights

- Developed a PMAP accessible and linkable table (and data dictionary) in which >500 neighborhood-level SDOH indicators are available for the entire US.
  - Most indicators available at census block group (Figure 2) or census tract
  - Table and relevant measures available using 2010 and 2020 census geographies
- This resource extends the scope of neighborhood-level measures to additional census measures as well as other neighborhood characteristics including food accessibility, rurality, historical redlining, green space, toxic waste site location, among others.
  - Some of the selected indicators are summarized in Figure 3.
- Interested investigators can link to this table through two methods: (1) leveraging patient geocodes; and (2) using the existing block group identifiers included as a part of the existing census table.
- We also provide a shared code repository that will allow investigators to perform the geocode to census-block group linkage, thereby extending the availability of neighborhood level indicators of SDOH beyond Maryland (Figure 4).

Conclusion

- We provide an expanded neighborhood-level table of indicators of SDOH available to all PMAP users for use in ongoing or future research studies.
- This resource will allow for more detailed consideration of SDOH for studies across PMAP projects.
- The table will be available as a part of the SQL server where PMAP is hosted (i.e., similar to PMAP dictionaries).

Next Steps

- We are working to add further measures including indicators of crime, additional environmental hazards among others.
- Any questions about the table or provided code, please email Kate (fitzgerald@jhmi.edu)
- Office hours with introduction to the table will be held in October

Figure 1. A sample of SDOH targeted for inclusion in the updated SDoH table.

Figure 2. Geographic levels included for the updated table. The red boxes denote the selected geographies.

Figure 3. Selected indicators included in the updated table.

Figure 4. Schematic of provide code as one of the ways investigators can link to the new table

References:

- Social determinants of health (SDOH) are upstream social and economic conditions, related to where people are born, grow, live, work, and age that impact individual and group differences in health status (Figure 1).
- SDOH incorporate relevant factors that are measured at both an individual and neighborhood-level.
- The Precision Medicine Analytics Platform (PMAP) currently includes information for each patient related to several neighborhood-level indicators of SDOH that are derived the census block geographical unit.
- The currently available resource is limited to a relatively limited set of seven indicators and is largely restricted to patients residing in census blocks groups located in Maryland.
Pediatric Genetic Syndromes with Aortopathy
Pediatric Genetic Syndromes with Aortopathy

• **Vision**
The creation of a comprehensive and dynamic database that integrates granular information regarding diagnosis, disease gene, underlying mutation, modifying genetic or environmental variation, age, gender, personal medical history, results of imaging, family history, physical manifestations, and response to prior interventions to recognize patterns within patient subgroups with strong predictive value.

• **Mission**
Our main goal is the creation of a polyfactorial risk score that takes into account clinically accessible factors basing the decision to proceed with surgery on objective criteria, avoiding unnecessary surgery in stable patients and hastening surgery in patients for whom delay would possibly result in aortic dissection with its attendant morbidity and mortality.

• **Research Aims**
1) Polyfactorial risk score: As described above, visualized as a graph showing risk of aneurysm/dissection/surgery over the coming months.

2) Automated calculation of arterial tortuosity. Development of a machine learning algorithm that identifies the vertebral arteries and calculates the tortuosity index.

3) Evaluate patterns in strain measurements in the aorta, left atrium, and left ventricle in patients with vascular EDS potentially providing a novel indicator of the risk of aortic complications in these patients.

4) Evaluate if corneal topography can distinguish Marfan Syndrome from other connective tissue disorders and be used as a diagnostic criteria.
Focus
Our center focuses on connective tissue disorders that are characterized by aortic aneurysm and risk of vascular rupture (ex: Marfan Syndrome, Loewy-Dietz syndrome, vascular Ehlers-Danlos). We know that the overall phenotype in inherited presentations of aortic aneurysm is wide, both within and between specific diagnoses. As it stands, we currently use limited information about the underlying diagnosis to make critical decisions about their prognostic counseling, the frequency and extent of imaging, the screening of family members, the use of medications, and which additional features of the phenotype to screen for and treat.

In addition to vascular disease, some patients with Loewy-Dietz Syndrome are prone to osteopenia and osteoporosis with fractures after minimal to no trauma. Identification of those patients who require additional screening and treatment may prevent significant morbidity in those susceptible to fracture while avoiding unnecessary screening in others.

Method & Analytics
Patients underwent clinical evaluation:
- History
- Physical exam
- DXA scan
- Blood work

Whole body DXA score
- Patients <18yo
  - Adjusted for height and age.
- Used adjusted whole body DXA score (whole body minus head)
- Patients ≥18 yo,
  - Unadjusted whole body DXA score

Skeletal features used in the predictive models were assessed by clinicians with LDS expertise.

Results and Highlights
Who has low DXA scores? (Figures A and B)
- 9.1% of adults and 50% of children had Whole body DXA Z-scores < -2
- Type 5 LDS appeared to have higher Z-scores but limited sample size
- 60% of patients experienced ≥1 lifetime fracture. Median number of lifetime fractures was 1 (range: 0-6)
- Fractures were most common in children with types 1 and 2 LDS but were more evenly distributed across LDS types in adults

What clinical features are associated with low DXA scores? (Figures C and D)
- Skeletal abnormalities in LDS include scoliosis, pes planus, arachnodactyly, spondyloptosis, and camptodactyly
- Patients were scored for the number of skeletal features they manifested
- Whole body DXA Z-score was inversely correlated with the total number of skeletal features in adults and children
- Number of fractures per year of life correlated with skeletal count in children

Can we develop an algorithm to help clinicians identify those at risk? (Figure E)
- Only 2 clinical and laboratory features were significant in predicting whole body Z-score: skeletal features count and binary age (adult vs. child)

Conclusion
- 33% of adults and 59% of children with LDS exhibited poor bone health (whole body DXA Z-score ≤ -2)
- No significant differences in DXA scores across the different types of LDS
- Children with LDS types 1 and 2 had significantly more fractures/year than those with types 3-5 (p<0.02)
- 61% of LDS patients experienced ≥ 1 fracture in their lifetime
- Children exhibited worse bone health compared to adults with LDS. We strongly suspect this represents a referral and/or survival bias and not that DXA scores improve with age.
- The number of skeletal features a patient exhibits strongly predicts their likelihood of having a low DXA score and fractures.
- A decision tree using the number of skeletal features and binary age (adult vs. child) predicts those at risk for skeletal morbidity

Next Steps
- We recommend that bone density assessment be performed as part of routine care in all patients with LDS.
- Patients should be counseled on:
  - Dietary and lifestyle interventions to optimize bone health
  - Maintaining adequate calcium intake
  - Monitor vitamin D levels and supplement if low
  - Promoting safe weight-bearing physical activity
  - Avoiding additional risks such as smoking and alcohol use.

We have now started treating patients with bisphosphonates and are working to identify those that respond to this treatment.
Hyperacute Prediction Of Targeted Temperature Management Outcome After Cardiac Arrest

Jocelyn Hsu$^{1,2}$, Han B. Kim$^{1,2,3}$, Kirby Gong$^{1,2,3}$, Tej Azad$^{4}$, Robert D. Stevens$^{1,3}$

1 Laboratory of Computational Intensive Care Medicine
2 Department of Biomedical Engineering, Whiting School of Engineering
3 Department of Anesthesiology and Critical Care Medicine
4 Department of Neurosurgery
Johns Hopkins University School of Medicine
Motivation: Targeted temperature management (TTM) has been widely regarded as a promising therapy in the resuscitation of cardiac arrest patients primarily because it is believed to improve neurological function. However, the benefits of TTM have not been consistently observed across trials. Moreover, TTM is resource intensive and can lead to serious adverse effects.

Mission: Our goal is to develop a machine learning model capable of identifying patients most likely to benefit from TTM would have a unique data signature. Such a model could enable greater precision in the selection of potential TTM candidates, establishing a baseline for personalized cardiac arrest treatment, and enable cohort enrichment for future clinical trials.

Research Aims:
1. Predict the discharge outcome of postcardiac arrest patients receiving TTM by training a ML model with physiological data and TTM treatment data.
2. Identify specific features that are associated with favorable outcome following TTM.
Focus
Targeted temperature management (TTM) has been widely regarded as a promising therapy in the resuscitation of cardiac arrest patients primarily because it is believed to improve neurological function [1, 2, 3, 4]. However, the benefits of TTM have not been consistently observed across trials [5, 6]. Moreover, TTM is resource intensive and can lead to serious adverse effects [7].

The aims of this study were twofold. First, to predict the discharge outcome of postcardiac arrest patients receiving TTM by training a ML model with electronic data and TTM treatment data. Second, to evaluate the potential of using TTM treatment data to identify specific features that are associated with favorable outcome following TTM. We hypothesized that patients most likely to benefit from TTM would have a unique baseline physiology. Such a model could enable greater precision in the selection of potential TTM candidates, establishing a baseline for personalized cardiac arrest treatment, and enable cohort enrichment for future clinical trials.

Method and Analytics
Data were extracted from the multicenter Philips eICU Clinical Research Database [8]. Patients were included if they were adults admitted to ICU after cardiac arrest. Since TTM is not consistently recorded in eICU, we analyzed temperature time series recorded in the first 24 hours after ICU admission. We designated as TTM any instance of a patient whose body temperature decreased to <36°C after admission and remained below that threshold for >12 h in the first 24 h after admission. All patients with potentially erroneous recordings, as signified by inflexible body temperatures less than 20°C included in the database, were excluded. A total of 444 patients meeting these criteria were identified. Among the latter, neurological outcome data was available for 310 patients. The two main variables of interest were discharge and neurological function at discharge. Neurological function was defined using the motor subscore of the Glasgow Coma Scale (mGCS). A mGCS of 6 signifies a patient who is awake and able to follow verbal commands, while an mGCS < 6 likely occurring. Patient data included demographics, medical history, whether TTM was initiated or not, as well as other features that are associated with favorable outcome following TTM. We hypothesized that patients most likely to benefit from TTM would have a unique baseline physiology. Such a model could enable greater precision in the selection of potential TTM candidates, establishing a baseline for personalized cardiac arrest treatment, and enable cohort enrichment for future clinical trials.

Neurological Outcome Prediction
For neurological function prediction, GLM performed slightly better than RF, with an AUROC of 0.75 ± 0.03, accuracy 0.69 ± 0.08, specificity 0.64 ± 0.10, precision 0.57 ± 0.07, and F1 0.61 ± 0.04. Similar to mortality predictions, top predictive features were identified via forward selection with 0.5 threshold, and models improved significantly, with GLM achieving an AUROC of 0.85 ± 0.03, accuracy 0.77 ± 0.03, sensitivity 0.71 ± 0.10, specificity 0.80 ± 0.12, precision 0.73 ± 0.08, and F1 0.71 ± 0.04 across 10 cross-validations (Fig. 2a).

We also used SHAP to rank features for predicting neurological outcome. We observed some similar top predictive features such as last pH measured, which also demonstrates positive correlation with mGCS. Other features include minimum monocyte count, which also highly contributes towards mGCS of 6 predictions, and last FiO2 levels and first baseline (lab test, both of which contribute towards patient outcomes with mGCS < 6 (Fig. 2b).
Clinical Implementation of Actionable Pharmacogenomics Testing at Johns Hopkins
Pharmacogenomics

• **Vision**
  Clinical care and research studies will be informed by high-quality pharmacogenomics testing alongside practical and timely decision-support tools

• **Mission**
  Provide targeted genotyping of metabolic genes for clinical care and research investigations

• **Research Aims**
  Elucidate the impact of clinical pharmacogenomics testing on patient care, adverse drug events, and therapeutic efficacy; Enable increased accessibility of pharmacogenomics testing and education for all patients and providers; Support investigations of new and emerging gene-drug pairs in precision medicine research.

Interested in Collaboration?

Contact us at:

Phone:
(443) 287-2018
(443) 287-7516

Email:
cknezev1@jhu.edu
mmarzin1@jhu.edu
isnyder6@jhu.edu
Clinical Implementation of Actionable Pharmacogenomics Testing at Johns Hopkins

**Isabel Snyder, Jonathan Merran, Mark Marzinke, Claire Knezevic**

**Background**
Drug response is a key factor in the success of treatment regimens; applying genomic data to inform prescribing decisions is a proactive form of precision medicine. Proactive pharmacogenomic (PGx) testing identifies patient-specific drug-gene interactions and provides evidence-based guidance for appropriate drug selection and dosing before initial drug administration. PGx-guided prescribing mediates the variability of drug response and reduces the probability of adverse drug-related events. Previously, Johns Hopkins patients could only have PGx testing performed by third-party laboratories, and results were available only as pdf uploads into the patient chart.

**Method and Analytics**
The pharmacogenetic panel is simple for patients and providers: it requires one blood draw, and the data is available to view within 3-7 days. Genomic DNA is extracted from whole blood, combined with designated PCR primers and two targeted fluorescent probes, each probe corresponding to either the reference or variant allele. The TaqMan® Genotyper software determines the assay call is determined as either homozygous reference, heterozygous, or homozygous variant. From this data, metabolic phenotypes can be assigned: poor, intermediate, normal, rapid and ultrarapid.

**Results and Highlights**
The JH Actionable Pharmacogenomics Panel was validated for clinical testing. The PGx panel consists of 39 variant-specific assays covering 11 different genes with over 70 actionable drug pairings based on Clinical Pharmacogenetic Implementation Consortium guidelines. With seamless integration into EPIC, genotypes and metabolic phenotypes are reported directly within Epic Labs tab. Genotypes are also presented as Genomic Indicators, an Epic functionality that highlights clinically actionable variants and their associated medications. These Genomic Indicators enable clinical decision support tools to be deployed during prescribing activities to guide selection of drug dosage and type.

**Conclusion**
Pharmacogenomics supports the goals of precision medicine by giving the correct drug at the optimal dose from the start of treatment, improving patient care and reducing costs. A recent study with over 5,000 participants observed a $7,000 reduction in direct medical charges per person, a 7% reduction in emergency department visits, and a 15% reduction in inpatient days after PGx panel testing and pharmacist medication review [1]. An multicenter European study also demonstrated an almost 30% drop in adverse drug reactions after the use of a 12-gene PGx panel [2]. With in-house PGx testing at Johns Hopkins, results will follow the patient through their lifetime to provide stable, consistent care, with clinical decision support tools that can be continually updated to reflect the most current medical literature and treatment guidelines. Clinical implementation of the Johns Hopkins PGx Panel is a key component of our work to incorporate metabolic profiles into patient care to avoid adverse pharmacological events and improve treatment efficacy.

**References**
1. Jarvis et al. 2022, 10.3390/jpm12030421
2. Savin et al. 2023, 10.1016/S0140-6736(22)01841-4

**PGx Panel Genes and Affected Medications**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Affected Medications (CPIC Level A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCG2</td>
<td>rosuvastatin</td>
</tr>
<tr>
<td>CYP286</td>
<td>dalfenin, setraline, navarapine</td>
</tr>
<tr>
<td>CYP2C19</td>
<td>clopidogrel, voriconazole, several antidepressants, proton pump inhibitors</td>
</tr>
<tr>
<td>CYP2C9</td>
<td>warfarin, phenylid, fluvacitin, several NSAIDs</td>
</tr>
<tr>
<td>CYP2D6</td>
<td>atorvastine, codeine, tramadol, several antidepressants, ondansetron, tamofoxin</td>
</tr>
<tr>
<td>CYP3A4</td>
<td>tacrolimus</td>
</tr>
<tr>
<td>DPD</td>
<td>florouracil, capecitabine</td>
</tr>
<tr>
<td>NUDT15</td>
<td>azathioprine, mercaptopurine, thioguanine</td>
</tr>
<tr>
<td>SLCO1B1</td>
<td>simvastatin, atorvastatin, rosuvastatin, lovastatin, fluvacitin, piltavastatin, pravastatin</td>
</tr>
<tr>
<td>TPMT</td>
<td>azathioprine, mercaptopurine, thioguanine</td>
</tr>
<tr>
<td>VKORC1</td>
<td>warfarin</td>
</tr>
</tbody>
</table>

**Figures**
Figures created with Biorender.com and taken from TaqMan Genotyper Software (ThermoFisher Scientific).

**Methods**
- PCR primers and two targeted fluorescent probes
- TaqMan® Genotyper software
- Genomic Indicators in Epic Labs tab
- Clinical Pharmacogenetic Implementation Consortium guidelines
-无缝集成到EPIC
Characterizing Individuals with Diabetes in Johns Hopkins Medicine from 2016 to Present: Demonstration of Use of OMOP and ATLAS with the APC-COE Registry
Primary Care PMCOE

• **Vision**
  Healthy people and communities enabled by the delivery of the best interventions for health promotion and disease prevention, diagnosis and treatment for each patient in primary care.

• **Mission**
  To accelerate discovery and translation of research into primary care practice. We will achieve this with precision data and analytics that enables a learning health system, working in partnership with patients, families, clinicians, researchers and communities.

• **Research Aims**
  To learn about the profile and care processes of patients with diabetes using data in the OMOP Common Data Model and the Atlas tool.

Interested in Collaboration?
Contact us at:
Phone: (410) 955-9866
Email: jsegal@jhmi.edu
Primary Care PMCOE: Characterizing Individuals with Diabetes in Johns Hopkins Medicine from 2016 to Present: Demonstration of Use of OMOP and ATLAS with the APC-COE Registry

Shanshan Lin, Jodi B. Segal

Background: The extract, transform, and load (ETL) process restructured the Electronic Health Records (EHR) data to the OMOP Common Data Model (CDM), which lays the groundwork for prospective large-scale network analysis.

Objectives:
1. To learn about the profile and care processes of our patients with diabetes in primary care across Johns Hopkins Medicine (JHM);
2. To demonstrate the use of data in the OMOP CDM with the Atlas tool.

Data source: The Adult Primary Care Center of Excellence (APC-COE) projection includes all EHR data from patients who saw a primary care clinician in any of our 42 clinics since 2016.

Study design: Retrospective cohort analysis. We established three cohorts:
1. Incident diabetes (Figure 1)
2. Diabetes (without diabetes-free period compared to cohort for incident diabetes)
3. All patients (patients with observation start date after 2016-01-01).

Diabetes definition: Diagnosis code: ≥ 1 inpatient visit diagnosis or ≥ 2 outpatient visit diagnosis. Diagnosis codes include 250.x, 357.2, 366.41, 362.01–362.07; Prescriptions: ≥ 1 prescription of glucose-lowering drug; Measurements: having either HbA1c at ≥ 6.5%, fasting plasma glucose at ≥ 126 mg/dL, or random plasma glucose at ≥ 200 mg/dL.

Covariates: Characteristics were generated with the built-in features in Atlas, adapted according to our research interests.

Results and Highlights

Product 1. A standardized process of constructing the concept sets for health condition, medication, and measurement of interest

A. Health condition

B. Medication

C. Measurement

Note: light blue: standard concepts; dark blue: classification concepts; mapping: the process of translating source code to standard code.

Table 1. Baseline characteristics of individuals from the three cohorts

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Incident diabetes</th>
<th>Diabetes</th>
<th>All patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &gt; 64</td>
<td>45%</td>
<td>52%</td>
<td>47%</td>
</tr>
<tr>
<td>Gender Female</td>
<td>55%</td>
<td>50%</td>
<td>53%</td>
</tr>
<tr>
<td>Race White</td>
<td>70%</td>
<td>72%</td>
<td>71%</td>
</tr>
<tr>
<td>Education</td>
<td>15%</td>
<td>12%</td>
<td>13%</td>
</tr>
</tbody>
</table>

Figure 1. A graphical depiction of the study design for patients observed to have an incident diabetes occurrence

Figure 2. A sunburst plot showing the trajectory of the prescription history of glucose-lowering medication for patients observed as having incident diabetes

Figure 3. The counts and percentages of coded comorbidities 6 months prior and after the index date for patients with diabetes (yellow: diabetic health conditions)

Figure 4. The counts and percentages of patients prescribed 11 classes of glucose-lowering medications after their index dates

Conclusion
1. The diabetes prevalence in our health system, at 16.9%, exceeds the national estimate of 11.3% [2].
2. The cohorts of individuals with diabetes have the expected clinical characteristics.
3. The use of data in OMOP CDM with the Atlas tool minimizes researchers’ effort in data cleaning. Nevertheless, additional support, such as SQL programming and R packages in HADES library are necessary to facilitate research with more sophisticated study designs.

Limitations
1. Data availability: There is no detailed information available about medication use, such as day supply and quantity, but it can be imported.
2. Interpretation: a. Sunburst plot (Figure 2): it describes the prescription pattern of physicians rather than the medication receipt by patients. b. Bar chart (Figure 3): The default feature output in the characterization analysis displays the coded comorbidities. To more meaningfully describe comorbid conditions, we need to customize a feature by aggregating the SNOMED codes.
3. Validation: There is a validated algorithm for creating a cohort of individuals with diabetes, but this is not often the case.

Next Steps
1. Address remaining questions with SQL programming
2. Prepare for research involving incidence rate analysis, causal estimation, or prediction;
3. Prepare for a large-scale network analyses.

The Telemedicine Challenge
Primary Care PMCOE

- **Vision**
  Healthy people and communities enabled by the delivery of the best interventions for health promotion and disease prevention, diagnosis and treatment for each patient in primary care.

- **Mission**
  To accelerate discovery and translation of research into primary care practice. We will achieve this with precision data and analytics that enables a learning health system, working in partnership with patients, families, clinicians, researchers and communities.

- **Research Aims**
  To inform the Office of Telemedicine at Johns Hopkins Health System about conditions for optimal use of telemedicine in primary care.

Interested in Collaboration?

Contact us at:

Phone: (410) 955-9866

Email: jsegal@jhmi.edu
# Adult Primary Care Center of Excellence

## The Telemedicine Challenge

### Results

#### Table 1. Patient Characteristics by Physicians’ Virtual Usage Category 2020-2021*  

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Quartile 1*</th>
<th>Quartile 2</th>
<th>Quartile 3</th>
<th>Quartile 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n=171 physicians)</td>
<td>(n=171 physicians)</td>
<td>(n=171 physicians)</td>
<td>(n=169 physicians)</td>
<td></td>
</tr>
<tr>
<td>Unique patterns (N)</td>
<td>8638</td>
<td>56,327</td>
<td>76,420</td>
<td>64,315</td>
</tr>
<tr>
<td>Total visits (median [IQR])</td>
<td>521 [246, 254]</td>
<td>4681 [3746, 5486]</td>
<td>4949 [3865, 5275]</td>
<td>4175 [3080, 1751]</td>
</tr>
<tr>
<td>Age (mean years, s.d.)</td>
<td>88.7 [8.7]</td>
<td>52.4 [17.5]</td>
<td>51.3 [17.3]</td>
<td>50.2 [17.2]</td>
</tr>
<tr>
<td>Female gender, N (%)</td>
<td>4879 [66]</td>
<td>31313 [66]</td>
<td>41084 [66]</td>
<td>40116 [66]</td>
</tr>
<tr>
<td>Black non-Hispanic, N (%)</td>
<td>2784 [66]</td>
<td>15489 [48]</td>
<td>18097 [77]</td>
<td>18015 [78]</td>
</tr>
<tr>
<td>English as primary language, N (%)</td>
<td>6125 [66]</td>
<td>51605 [60]</td>
<td>69170 [68]</td>
<td>63414 [68]</td>
</tr>
<tr>
<td>Maryland resident, N (%)</td>
<td>8174 [66]</td>
<td>62208 [66]</td>
<td>66462 [66]</td>
<td>65037 [66]</td>
</tr>
<tr>
<td>MS/DC/VA/DD/PA resident, N (%)</td>
<td>8459 [66]</td>
<td>54662 [66]</td>
<td>66977 [68]</td>
<td>63375 [68]</td>
</tr>
</tbody>
</table>

* category definition: lower virtual (lower a virtual visits in 2020-22); high virtual (virtual percentage is above median percentage 2020-22).

### Methods and Analytics

- **Design**: Mixed-design etiologic study using data from patients receiving primary care from January 2020 through December 2021.
- **Participants**: Eligible patients had at least one primary care contact. Eligible physicians had 10 or more patient contacts.
- **Exposures**: Quartile of virtual visits per physician per month calculated as the percentage of total visits conducted by phone or video (Q1 is lowest).
- **Main Measures**: Six metrics used by institution for value-based reimbursement, measured monthly at physician-level.
- **Models**: Generalized linear mixed models, using Poisson or negative binomial distributions; random intercepts for physicians nested in clinics and with a first-order autoregressive structure on the residuals by month; controlled for physician-level patient covariates by month.

### Limitations
- This was an ecologic analysis, although we also had detailed individual-level data that was used for adjustment.
- Analyses answer the question of the contextual effect of heavy virtual care delivery on the whole population in primary care.
- Confounding and effect modification by physician is possible.
- There are many other outcomes of importance for primary care delivery including patient satisfaction measures.

### Conclusions
- Variation in use of virtual care across clinics but modest relative to variation by month.
- Some quality metrics were modestly impacted by high virtual primary care usage; the absolute differences in rates were small.
- Provides reassurance to physicians and their health systems that telemedicine use may not adversely impact quality metrics.
CICL – Classification of Intra-Cranial Pressure Waveforms
to identify clamped external ventricular drain signals
by applying Machine Learning using
the PMCOE-NCC Data Repository
Neurosciences Critical Care PMCOE

• **Vision:**
Leverage high-resolution multi-dimensional data, mechanistic hypotheses, and advanced modeling (including artificial intelligence) to effectively and efficiently enhance the care and outcome of critically ill neurological patients

• **Mission**
To establish and scale a data-driven research program in neurocritical with an emphasis on three types of modeling:
  • Predictive modeling (Prognostic enrichment)
  • Treatment response modeling (Predictive enrichment)
  • Health economics modeling (efficiency, cost, generalizability)

• To create an institutional resource for research in Precision Therapeutics
• To secure funding via federal agencies (NIH, NSF) and industry partnerships

• **Research Aims**
  • **Aim 1.** To develop and validate a highly accurate risk prediction score for post-operative neurosurgical complications.
  • **Aim 2.** To develop and validate a highly accurate risk prediction index of neurological (and physiologic) deterioration for patients admitted to the NCCU (Aim 2A). To predict safe discharge of patients from the NCCU (Aim 2B).
  • **Aim 3.** To determine the efficacy and cost-effectiveness of the NCCU triage score and neurologic deterioration index by testing them in clinical trials evaluating feasibility safety, clinical outcome, and health-economic endpoints.

Interested in Collaboration?

Contact us at:

**Phone:**
410-955-7481

**Email:**
jsuarez5@jh.edu
Background
Management of elevated intracranial pressure (ICP) is important in preventing secondary brain injury. In patients with acute brain injuries admitted to the intensive care unit, External Ventricular Drains (EVDs) are the current gold standard in invasive ICP monitoring, and generate a continuous ICP waveform, from which a mean ICP value is generated. A major unmet need is to create better predictive models of impending elevations in ICP. However, the use of the ICP waveform to develop predictive models is limited by the fact that EVDs produce accurate waveforms only when they are clamped (i.e. drainage is turned off). Typically, this is intermitently done for variable intervals every hour. In large retrospective datasets, information about whether the EVD is clamped or draining or if there is noise, is not captured or labelled. This project showcases an innovative means of labelling a large dataset of physiological waveforms and demonstrates how it is used to train a machine learning classification algorithm to segment and classify ICP segments from an EVD as clamped or not.

Methods
Initial Patient Selection
5 patients with acute brain injuries that underwent ICP monitoring with an EVD, were identified through a search of the NCC Data Repository (PMCOE-NCC), and the entire duration of their ICP waveforms was extracted, totaling 1488 hours (approx. 62 days) of high frequency ICP waveform data, sampled at 120Hz.

Generation of a Labelled and Validated Dataset
The labelling process first involved segmenting the ICP waveform time series from each patient into epochs using the PELT change-point detection algorithm. PELT identifies and segments the time-series where the statistical properties of the time-series change. The data was divided to create a training dataset, and a validation dataset. Next, multiple statistical features were extracted for each segment, and K-Means unsupervised clustering was used to generate individual clusters of segments, each with similar statistical properties. Next, using a uniform methodology, segments from each cluster were extensively sampled and labelled by neuro-intensivist. This process involved visualization of the segment along multiple different time axes, allowing for visualization of the segment and its neighboring segments for minutes and hours. These labelled and validated epochs were then used to train supervised machine learning algorithms. The best algorithm based on accuracy and efficiency metrics was selected.

Results
After 10-fold cross validation, and testing various machine learning Classification models, the confusion matrix, accuracy statistics and Receiver Operating Curve (ROC) of our best performing model are presented in the figure below.

Confusion matrix

<table>
<thead>
<tr>
<th></th>
<th>Clamped ICP</th>
<th>Unclamped</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clamped ICP</td>
<td>35839</td>
<td>519</td>
</tr>
<tr>
<td>Unclamped</td>
<td>930</td>
<td>244843</td>
</tr>
</tbody>
</table>

Area under ROC curve = 99.10% ± 0.001
Recall = 96.57% ± 0.002
Positive predictive value = 97.47% ± 0.003
F1 score = 98.02% ± 0.002

(10 fold cross validation)

Summary and Next Steps
CICL is a machine learning-based classification algorithm that rapidly and accurately identifies clamped ICP waveform segments.

This algorithm paves the way for rapid high throughput processing of ICP waveform data which will allow for their use in prediction model development.

Next steps include validating patient generalizability and studying the application of this algorithm prospectively in real time.
PMCOE-NCC - Automated Identification of Patients Who Developed Critical Care Needs After Elective Neurosurgical Intracranial Tumor Resection
Vision

• Leverage high-resolution multi-dimensional data, mechanistic hypotheses, and advanced modeling (including artificial intelligence) to effectively and efficiently enhance the care and outcome of critically ill neurological patients.

Mission

• To establish and scale a data-driven research program in neurocritical with an emphasis on three types of modeling:
  • Predictive modeling (Prognostic enrichment)
  • Treatment response modeling (Predictive enrichment)
  • Health economics modeling (efficiency, cost, generalizability)
• To create an institutional resource for research in Precision Therapeutics
• To secure funding via federal agencies (NIH, NSF) and industry partnerships

Research Aims

• Aim 1. To develop and validate a highly accurate risk prediction score for post-operative neurosurgical complications.
• Aim 2. To develop and validate a highly accurate risk prediction index of neurological (and physiologic) deterioration for patients admitted to the NCCU (Aim 2A). To predict safe discharge of patients from the NCCU (Aim 2B).
• Aim 3. To determine the efficacy and cost-effectiveness of the NCCU triage score and neurologic deterioration index by testing them in clinical trials evaluating feasibility safety, clinical outcome, and health-economic endpoints.
PMCOE-NCC - Automated Identification of Patients who developed critical care needs after Elective Neurosurgical Intracranial Tumor Resection

Rohan Mathur1,2, Vishank Shah1,2,3,4, Tamas Budavari1, Niteesh R. Potu1,2, Peter H. Dziedzic1,2, Lin Cheng1,2,3,4, Eusebia Calvillo1,2,4, Bill King, William S. Anderson*, Robert D. Stevens1,2,3,4, Jose I. Suarez1,2,3,4
1Department of Anesthesiology & Critical Care Medicine, JHSOM, 2Department of Neurosciences Critical Care, JHSOM, 3Department of Neurology, JHSOM, 4Department of Neurosurgery, JHSOM,
*Department of Anesthesiology & Critical Care Medicine, JHSOM, Dept. Of Applied Mathematics and Statistics, Whiting School of Engineering, JHU

BACKGROUND

Patients diagnosed with intracranial tumors, after consultation with a neurosurgeon, present to the hospital at a scheduled date for surgical resection of their intracranial tumors. Post-operatively, they are routinely admitted to the Neurosciences Critical Care Unit for monitoring.

The Problem

- Only some of these patients develop clinical issues that require care that can only be provided in the NCCU. Most receive care that may be able to be delivered in less resource-intensive environments within the hospital.
- Scarce Resource such as intensive care provision and monitoring can be better allocated by identifying those patients that will truly develop ICU needs.

THE QUESTION

In a retrospective analysis of the PMCOE-NCC data repository, how do you identify those patients undergoing elective neurosurgical resection of their brain tumors that received care that can only be delivered in the ICU environment?

METHODOLOGY

Data Source

The Precision Medicine Center of Excellence in Neurocritical Care Data Repository contains clinical EMR data, physiological waveform data and imaging data from all patients that have been admitted to the Neurosciences Critical Care Units at Johns Hopkins. Through querying and screening these patient records, research questions of interest can be explored.

Inclusion Criteria for patient selection

- Adult Patient (18 years or older)
- Diagnosis of a brain tumor based on ICD-10 code screening
- Elective admission to the hospital for resection of the brain tumor with Neurosurgery. Excluded patients who went to the OR after being first admitted to the hospital in an acute symptomatic presentation
- Directly admitted to the NCCU after their surgery
- Information was stored in the PMCOE Data Repository

Screening Criteria for ICU Needs

Consensus driven criteria were developed by a team of four neuro-intensivists. These include:
- Administration of specific medications that can only be administered in an ICU
- ICU Specific procedures including intubation; high-flow nasal canula; non-invasive positive pressure ventilation; central venous catheter placement; ventiliculostomy management
- ICU specific diagnoses
- Restricted the time frame for development of critical care needs to 72 hours post-op

RESULTS

The PMCOE-NCC Data Repository was screened including patients from July 1, 2016, to August 31, 2023. 2582 eligible patients were identified. Among these, the algorithm identified 437 patients that demonstrated needs that could only be addressed in the Neurocritical Care Setting. Of these, 341 patients were administered critical care medications, 106 had critical care orders, and 136 had critical diagnoses codes entered.

DISCUSSION

- The Precision Medicine Center of Excellence in Neurocritical Care (PMCOE-NCC) aims to effectively identify patients who would most likely benefit in the resource-intensive neurocritical care environment from those who can fare equally well in a less resource-intensive environment such as intermediate care units or the floor.
- Identifying patients that developed critical care needs in the first 72 hours after elective surgical resection of their brain tumors is the first step in developing a model that will ultimately use pre-operative and intra-operative data to triage patients to the appropriate level of post-operative care. Such a model could potentially improve resource allocation and improve patient care.
Myositis PMCOE
Vision

• To use our multidisciplinary approach, trajectory analysis, and novel subgroup identification to tailor the monitoring and treatment of the disease to the individual.

Mission

• To leverage the longitudinal nature of our clinical cohort, coupled with prospectively collected biospecimens, to better classify unique phenotypes of IIM.

Research Aims

1) Identifying distinct patient trajectories related to muscle, lung, joint, and skin involvement.

2) Determining clinical and biological predictors of response to different treatment regimens.

3) Tailoring cancer screening recommendations to distinct subgroups of myositis patients.

Interested in Collaboration?

Contact us at:

Email: cmecoli1@jhu.edu
Background of Myositis

Idiopathic Inflammatory Myopathies (IIM, commonly referred to as simply ‘myositis’) comprise a group of rare chronic autoimmune diseases affecting multiple organ systems that can lead to substantial morbidity and mortality. While many of the underlying mechanisms of IIM remain unknown, the disease expression can impact the muscles, lungs, joints, skin, and heart. Current therapies for IIM are generally nonspecific, are not targeted to individual disease pathologies, and often are prescribed on a ‘trial and error’ basis.

Launch of Myositis PMCOE

The Johns Hopkins Myositis Precision Medicine Center of Excellence was launched in 2018 with the vision of increasing the efficiency of healthcare delivery to patients with IIM and to use our multidisciplinary approach, trajectory analysis, and novel subgroup identification to tailor the monitoring and treatment of the disease to the individual. Our mission is to leverage the longitudinal nature of our clinical cohort, coupled with prospectively collected biospecimens, to better classify unique phenotypes of IIM. To date we have automated electronic consenting, biospecimen management including DNA, RNA, sera, and PBMCs, and developed myositis-specific SmartForms within EPIC.

Ongoing Projects and Highlights

(1) We are leveraging the common data model ‘Observational Medical Outcomes Partnership’ (OMOP) to perform network studies within the international myositis and Observational Health Sciences and Informatics (OHDSI) communities [Figure 1].
   • Within the International Myositis Assessment and Clinical Studies Group (IMACS), we are developing both minimum and optimum datasets in the structure of OMOP. This multi-site international network will facilitate clinical trial recruitment, enable comparative effectiveness studies, and provide adequate statistical power for genetic association studies
   • Over 30 sites in >10 countries around the world are engaged in this initiative

(2) Our myositis longitudinal cohort within PMAP, paired with biospecimens, has enabled commercial partnerships in 2022-2023.
   • Identifying novel drug targets in dermatomyositis, necrotizing myopathy, and inclusion body myositis
   • Developing and validating commercial biomarkers (myositis-specific autoantibody panels) to facilitate rapid diagnosis and management decisions in both the inpatient and outpatient settings

(3) Tailoring cancer screening recommendations to distinct subgroups of myositis patients.
   • We have validated several novel biomarkers to improve risk-stratification for contemporaneous cancer, a complication occurring in up to 10% of IIM patients. We are now in the process of building a cancer prediction tool to inform clinical decision-making in collaboration with Stanford University and the University of Pittsburgh. This tool will be embedded within EPIC to facilitate a future multi-site clinical trial assessing the tool’s value in a real-world setting

(4) Reducing FTEs required to effectively run the Myositis FTE.
   • Through use of SmartForms and PMAP, we have saved the equivalent of 3 research coordinator FTEs annually
   • Study feasibility and clinical trial eligibility have gone from a process taking weeks to queries that take minutes
Bleeding complications after native kidney biopsy associated with aspirin use

Steven Menez, David Hu, Heather Thiessen Philbrook, Jack Bitzel, Yumeng Wen, Mohamed G Atta, Derek M Fine, J Manuel Monroy-Trujillo, John Sperati, Samir C Gautam, Celia P. Corona-Villalobos, Chirag R Parikh
Vision
• Expanding native kidney biopsy procedures while maintaining safety

Mission
• Examine potential risk factors for bleeding complications and improve clinical decision making

Research Aims
• Examine the association of aspirin use prior to the time of kidney biopsy and
  • Drop in hemoglobin level within 72 hours of kidney biopsy
  • Need for packed red blood cell transfusion within 7 days following kidney biopsy

Interested in Collaboration?
Contact us at: Steven Menez
smenez1@jh.edu

Kidney PMCOE
Bleeding complications after native kidney biopsy associated with aspirin use

Focus

Significant bleeding complications after native kidney biopsies include the following:

• Drop in hemoglobin (Hgb)
• Need for blood (pRBC) transfusion
• Persistent bleeding requiring embolization

Anti-platelet agents such as aspirin are often discontinued for 7 days or more prior to kidney biopsy, with the goal to reduce the risk of bleeding complications. However, such a wait in biopsy timing leads to a delay in kidney disease diagnosis and may have a negative impact on time-sensitive therapeutic decision-making.

Method and Analytics

Results and Highlights

The mean Hgb drop post-biopsy was 0.6 g/dL. A total of 183 (26%) patients had a Hgb drop over 1 g/dL, while 43 (6.0%) patients experienced a Hgb drop over 2 g/dL.

<table>
<thead>
<tr>
<th>Time of Last Aspirin Use</th>
<th>Within 3 days of biopsy (n=86)</th>
<th>3-6 days before biopsy (n=112)</th>
<th>6-365 days before biopsy (n=112)</th>
<th>No history within 1 year (n=418)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>64 (16)</td>
<td>62 (15)</td>
<td>60 (14)</td>
<td>52 (17)</td>
</tr>
<tr>
<td>Female</td>
<td>40 (47%)</td>
<td>37 (39%)</td>
<td>50 (45%)</td>
<td>217 (52%)</td>
</tr>
<tr>
<td>BMI</td>
<td>28.9 (7.4)</td>
<td>29.6 (7.9)</td>
<td>29.8 (8.6)</td>
<td>29.2 (8.5)</td>
</tr>
<tr>
<td>Hemoglobin (g/dL)</td>
<td>9.51 (1.89)</td>
<td>9.51 (1.96)</td>
<td>8.85 (1.65)</td>
<td>9.39 (1.96)</td>
</tr>
<tr>
<td>Platelet count (10^3/L)</td>
<td>240 (81.3)</td>
<td>252 (123)</td>
<td>249 (125)</td>
<td>250 (123)</td>
</tr>
</tbody>
</table>

Compared to patients with no prior aspirin use within a year, those who stopped aspirin within 3 days of kidney biopsy did not have a significantly higher risk of either Hgb drop ≥ 1 g/dL or ≥ 2 g/dL.

Conclusion

• Discontinuation of aspirin within 3 days of biopsy was not associated with increased odds of significant bleeding complications.
• These risks could be discussed when consenting patients where urgent biopsy diagnosis is necessary for timely therapeutic decision-making.

Next Steps

We plan to replicate this analysis in a prospective observational cohort of patients undergoing kidneys biopsies at Johns Hopkins (Novel Approaches in the Investigation of Kidney Disease (NAIKiD) Study) to examine the association between aspirin usage and adjudicated post-biopsy outcomes that require detailed chart abstraction:

• Blood transfusion directly related to biopsy
• Hematoma formation
• Jet formation
• AV fistula formation
Kidney Precision Medicine
Center of Excellence

Director: Chirag Parikh, MD, PhD
Vision

- Transform the care of patients with kidney disease, from prevention to diagnosis to treatment, strive to provide patient-centered and best-practice care, and provide diverse opportunities for patients to participate in cutting-edge research to potentially alter the landscape of kidney disease.

Mission

- Change how we care for patients at risk of and suffering from kidney disease and modernize kidney research.

Research Aims

- Measure clinical outcomes in patients with both acute kidney injury and progressive chronic kidney disease, applying risk assessments in real-time.
- Improve clinical phenotyping of patients with acute kidney injury.
- Identify barriers to optimal clinical care and target areas for quality improvement initiatives.
- Identify key patient subgroups to enrich enrollment in prospective observational studies and clinical trials.

Interested in Collaboration?

Contact us at:

Kidney_PMCOE@live.johnshopkins.edu
Focus
The primary goal of the Kidney PMCOE is to transform the care of patients with kidney disease.

Current Initiatives include
• Curation of high-quality renal focused data
• Integration of evidence-based medicine into clinical care
• Facilitating of clinical research
• Supporting quality improvement initiatives in clinical care

Curation of high-quality data
KPMCOE strives to curate high-quality data for use in research and clinical care. We have ~100 standardized definitions of labs, diagnoses, and procedures.

We recently developed an approach for mapping detailed EPIC medication records to a hierarchical medication classification framework, allowing for improved identification of prescribed medications. One application of this approach is to identify prescriptions by medication class to display in a patient dashboard (Patient Insight).

Supporting Evidenced-based Medicine in Clinical Care
In 2022, the kidney failure risk equation (KFRE) was added to the EPIC storyboard utilizing the curated data from KPMCOE. In the first half of 2023, roughly 25% of nephrology outpatient progress notes included the KFRE score.

Facilitating Clinical Research
In a prospective cohort of outpatients undergoing clinically indicated kidney biopsies, patients with obesity did not have a larger drop in hematocrit compared to those without obesity (Long Q, Menez S et al. Kidney360 2023; 4(1):98-101).

Quality Improvement in Clinical Care
KPMCOE has several quality improvement initiatives to improve efficiency of patient care. A recent initiative created a dashboard with details on kidney biopsies.

Next Steps
• Continue support of ongoing research projects and newly created disease-specific sub-studies (Acute Kidney Injury, Kidney Transplantation, Kidney Biopsy and Chronic Kidney Disease)
• Support and support clinical trials and new research initiatives
  • O’Brien Kidney Center Grant on Health Equities in Kidney Disease
  • QI metric development
• Implement patient dashboards for Kidney transplant patients

As a pilot test, one nephrology provider developed and successfully disseminated an electronic patient-reported outcome measure in nephrology clinic to assess patient health-related quality of life. (Patel D, Thavarajah S, Bitzel J et al. CJASN 2023)
Pulmonary Embolism (PE)
PMCOE
Vision

• Our vision is to improve outcomes for patients at risk of PE by preventing the life-threatening consequences that occur in more severe forms of this condition due to delayed detection.

Mission

• We aim to develop a computational model for prediction of PE in critically ill patients that leverages clinical data reflecting their physiological complexity.

Research Aims

1) PE Prediction: As described above, we aim to develop a computational model for prediction of PE in critically ill patients.

2) Feature discovery through the modeling process, identifying physiological markers of PE not currently known.

Interested in Collaboration?

Contact us at:

Phone: (410) 955-2611

Email: rstevens@jhmi.edu kgong1@jhu.edu
Focus

• Pulmonary embolism (PE) is a frequent and life-threatening complication in hospitalized patients whereby a thrombus occludes blood flow in the pulmonary artery or its branches.

• Current Pulmonary Embolism (PE) scoring systems (e.g., Modified Wells Scoring System, Revised Geneva Scoring System, and the Pulmonary Embolism Rule Out Criteria), have limited predictive accuracy, in part because they may not consider the physiological complexity of acutely ill patients.

• There is an unmet need for more accurate methods that can forecast the likelihood for developing PE.

Methods and Analytics

• Employed a large, multicenter database (eICU) containing >200,000 ICU admissions from 208 hospitals across the US

• Gathered and analyzed 2,799 unique ICU stays where a diagnosis of PE is recorded

• Due to a lack of a reliable determination of when the PE truly took place, we stratified our time-dependent data into three difference observation windows (12hrs, 24hrs, and 48hrs) for our PE and non-PE patients.

• Aggregated clinical data contains laboratory, demographic, vital sign, comorbidities, medications, and more.

• Trained multiple different machine learning (ML) models: decision tree, random forest (RF), gradient boosting (XGBoost, CatBoost, and GBoost), generalized linear models (GLM), support vector machine (SVM), and artificial neural network (ANN)

• Compared models to current PE risk scoring models (Wells and Geneva risk scoring models). Risk scoring models are modified to adapt to the limitations of the dataset.

• Figure 3. demonstrates inclusion & exclusion criteria

Results and Highlights

• These models stood out for their superior performance: XGBoost, CatBoost, ANN, and Logistic Regression.

• AUROCs for logistic regression and the risk scoring models are in Figure 1.

• Figure 2. presents the feature importance plot of the best performing model (by AUROC) across all time periods and models (12hr Logistic Regression).

• Figure 1:

![AUROC Scores by Model and Time Period](image)

- **Model**: Logistic Regression, Geneva Scoring System

<table>
<thead>
<tr>
<th>Time (hours)</th>
<th>AUROC Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>0.9</td>
</tr>
<tr>
<td>24</td>
<td>0.8</td>
</tr>
<tr>
<td>48</td>
<td>0.7</td>
</tr>
</tbody>
</table>

- **Feature Importance Plot**

  - Quinolones Taken Within 12 hours
  - Fluid and Electrolyte Disorders
  - Maximum O2 Saturation Value
  - Standard Deviation of Sodium Values Within 12 hours
  - Presence of Coagulopathy
  - Diuretics Taken Within 12 hours
  - Crystalloid Fluids Taken Within 12 hours
  - History of Congestive Heart Failure
  - Standard Deviation of Glasgow Coma Scores Recorded Within 12 hours
  - Admitted to ICU from Acute Care/Floor
  - Admitted to ICU from Step Down Unit (SDU)
  - Admitted to ICU from SDU
  - Mean Respiratory Rate
  - First Recorded Calcium Value in 12 hours

Conclusion

• The following results demonstrate the potential of these models to predict the onset of PE in critically ill ICU patients

• The model further allows us to understand additional variables of interest that are currently ignored by PE risk scoring models.

Next Steps

• Following external validation, this study aims to validate the model prospectively to determine its performance in a real-world setting.
Development and Assessment of Novel Continuous Glucose Monitoring (CGM) Metrics Using Time Series Analysis
Methods: Calculating the Glucose Color Index (GCI)

We developed our novel metric, the Glucose Color Index (GCI), using time-series analysis—see methods.

Within-person consistency assessed through test-retest $r^2$.

HYPNOS: N=141 adults with type 2 diabetes; no insulin.

Measured every 15 min for wear period of 14 days.

Decomposition of signal variance into components at the fundamental frequencies—defined by sampling schema to be $ω_i = (2iπ) / n$ for total samples $n$ and integers $i$.

Regression interpretation of $p_{ω_i}$ for time series $x_t$:

$$x_t = \sum_{i=1}^{n} \beta_i \cos(ω_i t) + \alpha_i \sin(ω_i t)$$

where $p_{ω_i} = \alpha_i^2 + \beta_i^2$.

Development and Reliability

We developed our novel metric, the Glucose Color Index (GCI), using time-series analysis—see methods.

Within-person consistency assessed through test-retest $r^2$.

HYPNOS: N=141 adults with type 2 diabetes; no insulin.

2 CGM wear periods, 3 months apart.

Median age of 61 yrs.

Clinical Utility

As associations with each comorbidity derived as odds ratio (OR) per standard deviation from logistic regression model.

ARIC: N=301 older adults with diabetes.

1 CGM wear period and comorbidity indicators.

Median age of 82 yrs.

Results

Table 1: Comparative Assessment of Reliability

<table>
<thead>
<tr>
<th>Metric</th>
<th>$r^2$</th>
<th>$r^2$ Adjusted for Other Existing Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Glucose</td>
<td>0.73</td>
<td>0.24</td>
</tr>
<tr>
<td>Time in Range</td>
<td>0.73</td>
<td>0.23</td>
</tr>
<tr>
<td>Coefficient of Variation</td>
<td>0.73</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Test-Retest Correlations of Novel Metric

<table>
<thead>
<tr>
<th>Metric</th>
<th>$r^2$</th>
<th>$r^2$ Adjusted for Existing Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCI</td>
<td>0.75</td>
<td>0.55</td>
</tr>
</tbody>
</table>

Adjustment accomplished through calculation of linear model residuals.

Figure 1: Assessment of Clinical Utility

The GCI is a reproducible new CGM metric based upon the time series periodogram which is, above and beyond existing CGM measures, consistent within persons and strongly associated with major comorbidities.

Conclusion

Development and Assessment of Novel Continuous Glucose Monitoring (CGM) Metrics using Time Series Analysis

Joseph Sartini¹, Mike Fang², Mary R. Rooney³, R. Nisha Aurora⁴, Naneth Punjabi⁴, Elizabeth Selvin⁴, Joe Conchel⁵, Scott Zeger²

¹Johns Hopkins Bloomberg School of Public Health, ²Rutgers School of Medicine, ³University of Miami, ⁴1. Johns Hopkins Bloomberg School of Public Health, 5. Virginia Commonwealth University

1. Collect and pre-process raw CGM data.
2. Periodogram calculated for each CGM wear period.
3. Fit piece-wise linear model over 3 frequency bands (long L, intermediate I, and short periodicity S), extract slopes S and midpoints MP to approximate log-periodogram.
4. 1-number summary Glucose Color Index (GCI) derived using Canonical Correlation Analysis (CCA)

• Assumes the 6 shape parameter values are i.i.d.

$$\bar{\omega} = \arg\max_{\omega} \text{Cov}(w^T X_t, w^T X_{t-1})$$

$$= \arg\max_{w \in \mathbb{R}^k} \sqrt{w^T \text{Var}(X_t) w} \sqrt{w^T \text{Var}(X_{t-1}) w}$$

5. Verify the reliability and clinical utility of the GCI.
Pharmacogenetics eConsult for JHHS Clinicians
Vision
• To use germline pharmacogenetic information in routine clinical practice to improve medication outcomes.

Mission
• Build evidence for pharmacogenetic associations through discovery research, demonstrate clinical utility in pragmatic studies, and develop infrastructure to seamlessly integrate pharmacogenetic data into routine clinical workflows.

Research Aims
• Identify ideal candidates for pharmacogenetic testing
• Demonstrate improved clinical outcomes through pharmacogenetics-guided care

Interested in Collaboration?
Contact us at:
Phone: (248)396-9113
Email: Stevenson@jhmi.edu
Pharmacogenetics eConsult for JHHS Clinicians

Background
Pharmacogenetics refers to the influence on an individual's genetics on medication efficacy or adverse drug reactions. Pharmacogenetics can be useful in current clinical practice. Here, we are specifically referring to germline pharmacogenetics rather than somatic/tumor genetics. For example, germline variation in genes encoding drug metabolizing enzymes are known to affect dose requirement or likelihood of efficacy or adverse drug reactions for dozens of medications.

However, many healthcare providers have received inadequate training in the field. eConsults are a mechanism for provider-to-provider consultation with an expert in a specialized field.

Implementation
The Epic electronic health record includes functionality to deliver asynchronous eConsults. In the first quarter of 2023, the Office of Telemedicine and Division of Clinical Pharmacology collaborated to launch a Pharmacogenetics eConsult service. The service provides a method for timely, but asynchronous clinical expertise when considering testing or after results are available. The eConsult is available at all mid-Atlantic JHHS facilities regardless of whether the testing has been performed within the health system or with a third party.

Results and Highlights
As intended, the Pharmacogenetics eConsult has been used by clinicians to discuss patients in which testing is being considered as well as patients with existing pharmacogenetic results. Turnaround time is <2 days. There is currently no charge to patients for the Pharmacogenetics eConsult.

Conclusion
The Pharmacogenetics eConsult is a convenient way for JHHS providers to seek pharmacogenetics expertise in an efficient and timely manner within the electronic health record.

Next Steps
We are currently considering automated ways to incorporate the option for a pharmacogenetics eConsult when new orders are placed for pharmacogenetic testing, such as the Hopkins-based Actionable Pharmacogenomics Panel.

Table 1. Examples of pharmacogenetics eConsults

<table>
<thead>
<tr>
<th>Timing</th>
<th>Provider</th>
<th>Context/question</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>Primary care physician</td>
<td>36 y/o F with anxiety and ADHD has multiple treatment failures due to efficacy and ADRs</td>
<td>Explained testing options, costs, and which medications it could shed light on</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test</td>
<td>Primary care physician (geriatrics)</td>
<td>67 y/o F with hypertension, Ehlers Danlos, and potential dysautonomia. Wondering about MTHFR gene results from 23andMe data</td>
<td>Clinical use of MTHFR genotype is not recommended. Third-party interpretation of 23andMe data is not appropriate for clinical use.</td>
</tr>
<tr>
<td></td>
<td>Ambulatory care pharmacist (internal med)</td>
<td>37 y/o M with hx of Stevens-Johnson syndrome from albuterol</td>
<td>PGx is most useful for explaining dose-dependent toxicity. PGx can predict SJS from certain medications, but not albuterol.</td>
</tr>
</tbody>
</table>
Streamlined Extracellular Vesicle Loading: Harnessing Protein-Specific Subpopulations Through Direct Electroporation
In this study, we present a novel method for encapsulating protein-specific subpopulations of nanoscale extracellular vesicles with miRNA directly through electroporation, eliminating the need for laborious follow-up purification and sorting procedures.

Our study introduces a promising avenue for advancing the production of genetic cargo-loaded extracellular vesicles, paving the way for more effective and targeted gene therapies for a wide range of diseases.
Streamlined Extracellular Vesicle Loading: Harnessing Protein-Specific Subpopulations Through Direct Electroporation

Corinna Torabi and Soojung Claire Hur
Department of Mechanical Engineering | Johns Hopkins University | Baltimore, MD

**EV-Mediated Delivery of Gene Therapy Agents**

- Extracellular Vesicles (EVs) Isolation from Cells
- mRNA Loading via Electroporation
- Loaded EV Purification

**Accurate Detection of Electroporation-Loaded miRNA**

- polyA cRNA reduces non-specific RNA binding
- Freezing destroys unloaded residual miRNA
- Purified EVs are loaded with miRNA

**EV Characterization**

- Exosomal Tetraspanins
- Endogenous Gene

**EV Integrity**

- CD81* EV-microbeads have enriched levels of CD81*.
- Consistent expression of EV endogenous control after freezing for all conditions.

**Conclusions**

We have designed a novel protein-specific EV electroporation process that provides significant advantages over existing methods:

- Simplified workflow
- Improved purity
- Accurate detection of EV miRNA contents
- Eliminate the need for post-electroporation EV selection
- Opportunity to integrate with microfluidic technology for automated microscale electroporation and multiplexing of protein targets

**Acknowledgements**

Thomas Piontek, Ph.D. | Johns Hopkins University, Institute for Nanoelectronics Technology
Michael Paulaitis, Ph.D. | Johns Hopkins University School of Medicine, Center for Nanomedicine at the Wm. W.singular Institute
David Lytton, Ph.D. | W. J. Carolina Medical College, Department of Pediatrics
Department of Cell and Developmental Biology

(CD81 expression monitored using FlowMethod's CellTiter-Glo®)
Consumer Health Information Technology to Engage and Support ADRD Care Partners
Vision

• Persons with cognitive impairment and other disabilities may benefit from managing their health virtually, but they and their care partners have been largely excluded from research on consumer health information technology (CHIT) interventions.

Mission

• To increase evidence regarding the role and use of CHIT and of novel, scalable, high-impact CHIT-supported interventions to improve Alzheimer's Diseases and Related Dementias (ADRD) assessment, care, and management.

Research Aims

• Aim 1: To produce new knowledge of individual, contextual, and organizational factors that affect patient portal use and evidence regarding the effects of access and use of such technologies on ADRD care quality and outcomes.

• Aim 2: To understand how persons with ADRD (mild through severe), caregivers, clinicians, and other relevant stakeholders (e.g., case managers, direct care workers, residential care staff) perceive and use consumer health information technology and identify novel technology-supported interventions that hold promise to improve ADRD care and management across the spectrum of residential and care delivery settings.
Focus

• Persons with cognitive impairment and other disabilities may benefit from managing their health virtually but have been largely excluded from research on patient portal interventions.

Project Goal: Increase evidence regarding the role and use of consumer health information technology (CHIT) and CHIT-supported interventions to improve assessment, care, and management of Alzheimer’s Disease and Related Dementias (ADRD).

Method and Analytics

Data Source

Sample Characteristics
Established JHHS patients 65 yrs and older (n = 49,382)
- Mean age: 76.6 years old
- 57.3% female
- 67.2% White, 23.4% Black, 3.3% Asian or Pacific Islander, and 6.1% Another Race
- 6.4% of patients had a dementia diagnosis

Methods
*Portal activity metric: ratio of number of portal sessions to number of clinical encounters
Dementia diagnosis: based on ICD-10 codes and patient’s problem list

Results and Highlights

71.5% of patients were registered for the portal
Portal registration did not differ by dementia diagnosis (with vs. without) for patients.

Conclusion

Our work highlights the need to:
✓ Better identify, engage, and support care partners of persons with ADRD via the patient portal.
✓ Understand the potential of embedded interventions delivered via the portal.

Next Steps

Ongoing analyses using focus on:
• Portal use during home health episodes
• Portal use and the Medicare annual wellness visit
• Portal-based deprescribing intervention
• Accuracy of natural language processing to identify authors of portal messages
• Portal use to better support care partners

For patients with dementia, care partners authored most messages, most commonly while logged in with patient identity credentials.

This work has been published in:
Neonatal Precision Medicine Center of Excellence
Neonatal PMCOE

Vision
• To identify, quantify, and define early predictors of clinical outcomes and lower barriers to research and personalized medicine for all neonates

Mission
• Develop the infrastructure and expertise that will allow for personalized evaluation, counseling, management, and prevention of neonates at risk for adverse outcomes

Research Aims
• Can we build a model to identify premature infants most at risk for high morbidity and mortality using their response to different treatments and available data?
• Can this model be developed into an objective EHR tool that could facilitate dynamic individualized counseling in the Neonatal Intensive Care Unit?

Interested in Collaboration?
Contact us at:
Phone: (410) 614-3829
Email: kaziz5@jh.edu
NICU PMCOE: Analyzing Premature Infants' Response to Surfactant Therapy

Flora Yazigi MD, Sai Manasa Kalyanam MEng, Kartikeya Makker MD, Jen Park MS, Hrishikesh Kambli MS, Khyzer B Aziz MD

Background
- Advances in neonatal care are enabling higher survival rates of extremely premature (<29 weeks gestation) and extremely low birth weight infants (<1000g).
- A vast majority of these infants suffer from neonatal respiratory distress syndrome (NRDS), commonly due to surfactant deficiency, which is crucial for preventing lung alveoli collapse.
- Surfactant administration and respiratory support, both invasive and non-invasive, are standard interventions for NRDS.
- The response to surfactant treatment, especially in terms of oxygen and respiratory support needs, is inconsistent among neonates.
- The neonatal sequential organ failure assessment (nSOFA) score measures patient severity of illness, including respiratory support needs & degree of organ dysfunction.

Objectives
- To assess in detail the surfactant response by tracking changes in organ dysfunction from birth to either death or discharge, factoring in gestational age and time, particularly in extremely preterm infants.
- Hypothesis: Distinct modifiable and non-modifiable maternal, fetal, and neonatal factors will differentiate surfactant responders from non-responders.

Method and Analytics
- A retrospective cohort study at a single center assessed inborn, extremely preterm infants with very low birth weight from July 2016 to June 2023 using the NICU PMAP Registry (~35,000 records). Hourly neonatal sequential organ failure assessment scores were tracked until discharge or death. We developed replicable phenotype definitions using SQL and analyzed the data with Python.

Results and Highlights

Figure 1. Distribution of Patients Based on Number of Surfactant Doses. Histogram of number of patients and doses of surfactant administered. Total of 332 patients.

Figure 2. Longitudinal nSOFA trajectory by doses of surfactant administered. Mean nSOFA scores with 95% CIs for the first 72hrs of life for all patients. Higher nSOFA trend observed in infants given ≥2 surfactant doses (green band) compared to one (blue) and none (red). Peak scores around 1-5hrs, reflecting typical post-birth surfactant administration; secondary peak at ~12hrs of life coinciding with usual second dose administration. Distinct differentiation among the three groups at these intervals.

Figure 3. Example patient level nSOFA profiles for non-survivors relative to matched epoch. Mean nSOFA scores with 95% CIs for the first 72hrs of life for all patients. Dashed line represents non-survivor. Complex, shared decisions often dictate care for extremely premature and extremely low birth weight infants. Situations where no surfactant is given due to parental preference, equipment size constraints, and/or major anomalies (A). Cases where surfactant administration results in complications such as pneumothorax or pulmonary hemorrhage, preventing repeat dosing (B). Some patients face severe pulmonary hypoplasia that is not responsive to surfactant (C).

Conclusion and Next Steps
- Using birth weight and gestational age, we discerned nSOFA profiles based on surfactant doses (none, one, or ≥2).
- Next, we’ll examine the NICU PMAP registry features to test our hypothesis and identify maternal, fetal, and neonatal factors influencing key outcomes.
- Our main analysis will target survivors treated with surfactant who develop bronchopulmonary dysplasia (BPD), aiming to create a BPD prediction model and personalizing existing care paradigms (therapies) for improved outcomes.
Validation of Calculated Sequential Organ Failure Scores in a COVID-19 Registry
The Johns Hopkins Precision Medicine Center of Excellence for COVID-19

Vision
• Understand the factors that underlie the pathobiology of COVID-19, the progression to severe illness or death, and the effectiveness of therapeutic interventions in order to provide personalized care to patients infected with SARS-CoV-2.

Mission
• Improve the care of patients infected with SARS-CoV-2 by learning about COVID-19 pathobiology, likelihood of disease progression and impact of specific therapeutic interventions. We aim to provide this information to clinicians, patients and family members at the point of care.

Research Aims
• To characterize the longitudinal trajectories of patients infected with SARS-CoV-2 to inform research questions and clinical care.
• To develop prediction models of COVID-19 severe disease and death that can be used at the point of care to inform treatment decisions, resource allocation and discussions with patients and families.
• To examine the impact of different treatment interventions on COVID-19 outcomes.
• To explore racial and social inequities in COVID-19 diagnosis and management.

Interested in Collaboration?
Contact us at:
Email: bgariba1@jhmi.edu
The Sequential Organ Failure Assessment (SOFA) score is a key component of the JH ASR.

It is unclear if SOFA scores derived from electronic health record (EHR) data are comparable to manually abstracted scores.

### Results and Highlights

- Abstractors were able to calculate a SOFA score for 191 patients.
- Completeness of EHR SOFA subscores ranged from 14% (95% CI: 10-20%) for respiratory to 63% (95% CI: 56-69%) for coagulation; completeness in EHR total SOFA score was 88% (95% CI: 83-92%).
- Completeness of CROWN SOFA subscores ranged from 91% (95% CI: 86-94%) for cardiac to 99% (95% CI: 96-100%) for neurologic; completeness in CROWN total SOFA score was 87% (95% CI: 81-91%).
- EHR and manually abstracted SOFA scores differed by 2 or more points in 76% of cases (95% CI: 69-82%).
- CROWN and manually abstracted SOFA scores differed by 2 or more points in 27% of cases (95% CI: 20-34%).
- EHR SOFA scores were on average 2.1 points lower than manually abstracted scores (95% CI: 1.8, 2.5).
- CROWN SOFA scores were on average 0.3 points higher than manually abstracted scores (95% CI: 0.1, 0.6).
- The ICC between abstractors and moderators was 0.82 (95% CI: 0.69, 0.92).

### Method and Analytics

- The JH-CROWN registry includes all patients admitted to the Johns Hopkins Health System with a laboratory-confirmed diagnosis of COVID-19.
- 250 patients who underwent mechanical ventilation (MV) were randomly selected from CROWN.
- SOFA scores were manually calculated by trained abstractors in the 24 hours preceding MV, with 10% of cases reviewed by two expert moderators.
- Intraclass correlation (ICC) between abstractors and moderators was assessed using a linear mixed effects model with random effects for case and reviewer.
- EHR SOFA scores, SOFA scores calculated from data elements in JH-CROWN (CROWN SOFA) and manually abstracted scores were compared.

### Table 1: Characteristics of patients with a SOFA score able to be abstracted.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Sample: N=191</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Years: Median [IQR]</td>
<td>68 [57, 76]</td>
</tr>
<tr>
<td>Sex</td>
<td>Female</td>
<td>127 (66.3%)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>64 (33.5%)</td>
</tr>
<tr>
<td>Race</td>
<td>White</td>
<td>69 (36.1%)</td>
</tr>
<tr>
<td></td>
<td>Black</td>
<td>72 (37.7%)</td>
</tr>
<tr>
<td></td>
<td>Asian</td>
<td>9 (4.7%)</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>38 (19.9%)</td>
</tr>
<tr>
<td></td>
<td>NA. Not obtained</td>
<td>3 (1.6%)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Not Hispanic/Latino</td>
<td>158 (82.7%)</td>
</tr>
<tr>
<td></td>
<td>Hispanic/Latino</td>
<td>32 (16.8%)</td>
</tr>
<tr>
<td></td>
<td>NA. Not obtained</td>
<td>1 (0.5%)</td>
</tr>
<tr>
<td>Admission to MV</td>
<td>Days: Median [IQR]</td>
<td>2.3 [0.6, 4.2]</td>
</tr>
<tr>
<td>SOFA: Abstracted</td>
<td>Median [IQR]</td>
<td>5 [4, 7]</td>
</tr>
<tr>
<td>SOFA: Epic</td>
<td>Median [IQR]</td>
<td>3 [1, 5]</td>
</tr>
<tr>
<td>Missing</td>
<td>N (%)</td>
<td>23 (12.0%)</td>
</tr>
<tr>
<td>SOFA: JH-CROWN</td>
<td>Median [IQR]</td>
<td>5 [4, 7]</td>
</tr>
<tr>
<td>Missing</td>
<td>N (%)</td>
<td>25 (13.1%)</td>
</tr>
</tbody>
</table>

- **Validation of Calculated Sequential Organ Failure Scores in a COVID-19 Registry**
- **The Johns Hopkins Precision Medicine Center of Excellence for COVID-19**
- **Team:** Elizabeth K. Zink, PhD, RN, Joshua F. Betz, MS, Natalie Wang, Brian T. Garibaldi, MEHP, FACP, FRCP (E)

### Method and Analytics

- The JH-CROWN registry includes all patients admitted to the Johns Hopkins Health System with a laboratory-confirmed diagnosis of COVID-19.
- 250 patients who underwent mechanical ventilation (MV) were randomly selected from CROWN.
- SOFA scores were manually calculated by trained abstractors in the 24 hours preceding MV, with 10% of cases reviewed by two expert moderators.
- Intraclass correlation (ICC) between abstractors and moderators was assessed using a linear mixed effects model with random effects for case and reviewer.
- EHR SOFA scores, SOFA scores calculated from data elements in JH-CROWN (CROWN SOFA) and manually abstracted scores were compared.

### Table 1: Characteristics of patients with a SOFA score able to be abstracted.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Sample: N=191</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Years: Median [IQR]</td>
<td>68 [57, 76]</td>
</tr>
<tr>
<td>Sex</td>
<td>Female</td>
<td>127 (66.3%)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>64 (33.5%)</td>
</tr>
<tr>
<td>Race</td>
<td>White</td>
<td>69 (36.1%)</td>
</tr>
<tr>
<td></td>
<td>Black</td>
<td>72 (37.7%)</td>
</tr>
<tr>
<td></td>
<td>Asian</td>
<td>9 (4.7%)</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>38 (19.9%)</td>
</tr>
<tr>
<td></td>
<td>NA. Not obtained</td>
<td>3 (1.6%)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Not Hispanic/Latino</td>
<td>158 (82.7%)</td>
</tr>
<tr>
<td></td>
<td>Hispanic/Latino</td>
<td>32 (16.8%)</td>
</tr>
<tr>
<td></td>
<td>NA. Not obtained</td>
<td>1 (0.5%)</td>
</tr>
<tr>
<td>Admission to MV</td>
<td>Days: Median [IQR]</td>
<td>2.3 [0.6, 4.2]</td>
</tr>
<tr>
<td>SOFA: Abstracted</td>
<td>Median [IQR]</td>
<td>5 [4, 7]</td>
</tr>
<tr>
<td>SOFA: Epic</td>
<td>Median [IQR]</td>
<td>3 [1, 5]</td>
</tr>
<tr>
<td>Missing</td>
<td>N (%)</td>
<td>23 (12.0%)</td>
</tr>
<tr>
<td>SOFA: JH-CROWN</td>
<td>Median [IQR]</td>
<td>5 [4, 7]</td>
</tr>
<tr>
<td>Missing</td>
<td>N (%)</td>
<td>25 (13.1%)</td>
</tr>
</tbody>
</table>

### Results and Highlights

- Abstractors were able to calculate a SOFA score for 191 patients.
- Completeness of EHR SOFA subscores ranged from 14% (95% CI: 10-20%) for respiratory to 63% (95% CI: 56-69%) for coagulation; completeness in EHR total SOFA score was 88% (95% CI: 83-92%).
- Completeness of CROWN SOFA subscores ranged from 91% (95% CI: 86-94%) for cardiac to 99% (95% CI: 96-100%) for neurologic; completeness in CROWN total SOFA score was 87% (95% CI: 81-91%).
- EHR and manually abstracted SOFA scores differed by 2 or more points in 76% of cases (95% CI: 69-82%).
- CROWN and manually abstracted SOFA scores differed by 2 or more points in 27% of cases (95% CI: 20-34%).
- EHR SOFA scores were on average 2.1 points lower than manually abstracted scores (95% CI: 1.8, 2.5).
- CROWN SOFA scores were on average 0.3 points higher than manually abstracted scores (95% CI: 0.1, 0.6).
- The ICC between abstractors and moderators was 0.82 (95% CI: 0.69, 0.92).

### Conclusion

- Automated EHR tools can process large amounts of clinical data to inform processes such as resource allocation.
- The importance of clinical and contextual verification cannot be overstated.
- CROWN SOFA scores (as compared to EHR SOFA scores) demonstrated superior agreement with manually abstracted SOFA scores.
- EHR SOFA domain scores had a high degree of missigness, resulting in a score of 0, causing underestimation of illness severity, particularly in respiratory domain.
- SOFA Scores calculated by EPIC should be modified to avoid underestimating severity in research and practice.
- This analysis will assist in the development of better methods to predict mortality, allocate scarce resources, and address public concerns regarding equity during crises.

### Next Steps

- Refinement of specifications for the cardiac and neurologic domains of the CROWN SOFA score will be made to improve the agreement between CROWN and abstracted scores.
- Recommendations to improve the coding specifications in the EHR for SOFA will be made.
Management of Nonalcoholic Fatty Liver Disease in a Primary Care Setting
Management of Nonalcoholic Fatty Liver Disease in a Primary Care Setting

**Vision**

- Prove there is a need for further education on the outpatient management of NAFLD among primary care physicians.

**Mission**

- Improve management of fatty liver disease within the primary care setting

**Research Aims**

- To better understand how PCPs are diagnosing and managing NAFLD in their practices.

Interested in Collaboration?

Contact us at:

Phone: (856)-577-9013

Email: jgips1@jhmi.edu
Management of Nonalcoholic Fatty Liver Disease in a Primary Care Setting

Julia R. Gips, MD; Lisa Yanek, MPH; Jiajun Wu, MS; Tinsay A. Woreta, MD; James P. Hamilton, MD; Jeannine M. Clark, MD, MPH

Divisions of General Internal Medicine & Gastroenterology and Hepatology, Department of Medicine
Johns Hopkins University, Baltimore, Maryland, USA

BACKGROUND

- The prevalence of nonalcoholic fatty liver disease (NAFLD) has grown exponentially in recent years.
- NAFLD is quickly becoming one of the leading causes of cirrhosis in the United States.
- Primary care physicians (PCPs) are often the first to encounter patients with NAFLD, as it is most often diagnosed incidentally.
- Several recent guidelines have been published detailing optimal NAFLD management.
- How PCPs are diagnosing and managing NAFLD in practice is unknown.

HYPOTHESIS

Few patients with NAFLD in the primary care setting are receiving evidence-based evaluation and follow-up designed to prevent the progression of NAFLD to NASH or cirrhosis and its complications.

METHODS

Dataset:
- 370,000 patients in the Adult Primary Care Center of Excellence (APCCOE) patient registry with a primary care appointment between July 1, 2016 – September 17, 2023

Study Population:
- 10,334 Adults who had at least one primary encounter during this time where NAFLD or NASH was billed.

Extraction:
- Completed by the core for clinical research data acquisition within Microsoft SQL Server Management Studio and RStudio
- NAFLD and NASH population extracted using ICD-10 codes K75.1 or K76.0
- Those with advanced fibrosis or cirrhosis were identified as those who also had ICD-10 codes K74.60 or K74.02

Descriptive data analyses:
- Frequencies and percentages for categorical variables
- Means and standard deviations or medians and interquartile ranges for continuous variables.

RESULTS

![Image of a table summarizing patient demographics and comorbidities at the time of NAFLD diagnosis.]

- **Diagnosis**
  - NAFLD
  - NASH
  - Mean Weight: 210.0 ± 8.6 lbs

- **Mean Age**: 52.8 years

- **BMI Category**
  - Underweight
  - Normal Weight
  - Overweight
  - Obesity Class I
  - Obesity Class II
  - Obesity Class III
  - Mean BMI: 33.2 ± 3.9 kg/m²

**Comorbidity Present in Patients**

- Hypertension: 67.4%
- Hyperlipidemia: 49.9%
- Diabetes Mellitus: 36.9%
- Coronary Artery Disease: 15.5%

**Treatment Received**

- Prescriptions
  - ACE-I
  - OATP
  - Vitamin E
  - Pioglitazone
  - Thiazolidinediones
  - Referred to a Nephrologist
  - Referred to a Nutritionist

**Imaging Studies**

- Liver ultrasound
- Liver biopsy
- CT or MRI of the liver

**Mean Laboratory Values at First NAFLD Visit**

- Alanine Aminotransferase (ALT): 34.7 U/L
- Aspartate Aminotransferase (AST): 44.1 U/L
- Total Bilirubin: 1.1 mg/dL
- Albumin (G3): 3.6 g/dL
- Uric acid: 5.8 mg/dL

**Conclusions**

- In general, PCPs are not screening, stratifying, referring, or managing patients with NAFLD according to guidelines.
- While the Fib-4 score has been shown to be an excellent predictor of fibrosis, 15% of patients did not have AST or ALT values, and 55% did not have a platelet count.
- Less than 1% of all patients and 15% of those with advanced fibrosis/cirrhosis were referred to hepatology.
- Follow-up imaging, or biopsies was done in < 4% of high-risk patients, despite recommendations for repeat fibrosis screening every 1–3 years.

**Limitations**

- NAFLD management has been evolving with recommendations published at various points throughout the study period.
- We were not able to access provider notes for further details on counseling or patient engagement with care.
- Data was only obtained from one healthcare system, limiting our knowledge of outside referrals or results.

**Implications**

There is a strong need to enhance education around NAFLD and its management particularly within primary care and preventative settings.
Standardizing Diabetes Medication Data in the Adult Primary Care PMCOE
Vision

• Healthy people and communities enabled by the delivery of the best interventions for health promotion and disease prevention, diagnosis and treatment for each patient in primary care.

Mission

• To accelerate discovery and translation of research into primary care practice. We will achieve this with precision data and analytics that enables a learning health system, working in partnership with patients, families, clinicians, researchers and communities.

Research Aims

Our research aims to improve patient outcomes, by:

• Providing access to curated data to support research on primary care delivery
• Conducting research on common conditions in primary care, including hypertension, diabetes, and heart failure
• Understanding the impact of community and social determinants of health through predictive models and spatial analysis
• Identifying barriers to optimal care

Interested in Collaboration?

Contact us at:

Phone: (949) 981-8710

Email: mchiu17@jhu.edu
Focus

- Medication data must be standardized before they can be used in studies
- RxNorm codes provide a method of standardization
- The APCCOE medications failed to match RxNorm codes for analysis
- We used active ingredients to identify the missing RxNorm codes

Method and Analytics

- Data sources were the Epic Clarity (n=173,352) and Adult Primary Care Center of Excellence medications from First Data Bank therapeutic class "ANTHYPERGLYCEMICS" that failed to match to RxNorm codes (n=2,449)
- The branded drug data set was generated using RxNav
- We used two algorithms to extract active ingredients and drug classes in both data sets (Figure 1)
- Entries in both data sets were then matched to RxNorm codes using active ingredients and exploratory analyses were performed

Results and Highlights

- Active ingredients matched more unique entries to RxNorm codes than medication ID in both the Epic Clarity and APCCOE data sets (Figure 2)
- Almost all identified medications were monotherapies (Figure 3)
- Four medications constituted about 70% of all prescriptions in the APCCOE data for both monotherapy and combination therapies (Figure 4)
- These exploratory analyses validated the drug information extracted using the algorithms

Conclusion

- APCCOE prescription data were matched to RxNorm codes using active ingredients to provide completeness for data analysis
- Health system medication data may require novel methods to identify and classify them using standard terminology
- The methods and results require validation
- Standardized data can now be included in epidemiologic studies, which allows the APCCOE to gain insights into prescription practices

Next Steps

- Additional analyses will be performed to assess prescription of GLP-1 receptor agonists and SGLT2 inhibitors in diabetes patients
- Standardized data can be combined with data from other health systems to gain population-level insights on prescription practices
- Prescription practices can be carefully analyzed to ensure that patients are receiving the best of available treatments

Figure 1. The workflow to filter, extract, and match the data. The top left quadrant describes the process to select non-insulin diabetes from the raw Epic Clarity data set. The bottom left quadrant shows the steps using the algorithms to extract active ingredients and drug classes from each entry. The right half illustrates the procedure to match RxNorm codes using active ingredients.

Figure 2. The proportion of unique entries that successfully matched to an RxNorm code in each data set.

Figure 3. All combinations contain two active ingredients.

Figure 4. The four most common therapies are depicted in the bar graphs for each type of therapy and the rest are categorized as “Other.”
Prescribing of evidence-based medications for the prevention of adverse cardiovascular events and progression of chronic kidney disease among patients with diabetes
Vision
• Healthy people and communities enabled by the delivery of the best interventions for health promotion and disease prevention, diagnosis and treatment for each patient in primary care.

Mission
• To accelerate discovery and translation of research into primary care practice. We will achieve this with precision data and analytics that enables a learning health system, working in partnership with patients, families, clinicians, researchers and communities.

Research Aims
Our research aims to improve patient outcomes, by:
• Providing access to curated data to support research on primary care delivery
• Conducting research on common conditions in primary care, including hypertension, diabetes, and heart failure
• Understanding the impact of community and social determinants of health through predictive models and spatial analysis
• Identifying barriers to optimal care
Focus

- Glucagon-like peptide-1 receptor agonists (GLP-1 RA) and sodium-glucose transport protein-2 inhibitors (SGLT2i) reduce adverse cardiovascular events and progression of kidney disease among high-risk patients with diabetes.
- In prior studies, few eligible patients received a GLP-1 RA or SGLT2i, with racial disparities in their use.
- The proportion of eligible patients in primary care at Johns Hopkins with diabetes who have been prescribed a GLP-1 RA and SGLT2i is unknown.

SGLT2i and GLP-1 RA indication by condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>SGLT2i</th>
<th>GLP-1 RA</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASCVD or CV high risk*</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Chronic kidney disease (CKD)*</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Heart failure (HF)</td>
<td>✔️</td>
<td>✔️</td>
</tr>
</tbody>
</table>

Results and Highlights

We identified 26,478 patients with diabetes, of whom 7,837 (29.6%) had ASCVD, 6,506 (24.6%) had CKD, and 3,837 (14.5%) had HF. (Table 1)

Among patients with diabetes:

- 30% of patients with CKD or ASCVD had an active medication order for an SGLT2i or GLP-1 RA (Figure 1).
- 20.1% of patients with HF had an active medication order for an SGLT2i (Figure 2).
- 17% of patients with CKD or ASCVD had an active medication order for an SGLT2i (Figure 2) or GLP-1 RA (Figure 3).
- An additional 15–20% of patients had at least one prior medication order for an SGLT2i or GLP-1 RA.

Method & Analytics

- Identified patients in the Adult Primary Care COE with:
  - Type 2 diabetes, determined by at least one ICD-10 diagnosis code (E11 – E14).
  - At least one visit to a Johns Hopkins primary care location in the prior 12 months.
- Further identified patients with additional conditions by ICD-10 code:
  - ASCVD (I00 – I30).
  - HF (I50).
  - CKD (N18).
- Defined patients as having a medication order for a GLP-1 RA or SGLT2i if they had at least one medication order identified:
  - Active medication order: Order is less than 12 months old and has ended or been discontinued.
  - Prior medication order: Order is identified but not active.

Table 1: Demographics of patients with diabetes in primary care in the past 12 months

<table>
<thead>
<tr>
<th>Demographic characteristic</th>
<th>n (%) - except as noted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean +/- SD</td>
<td>64.1 +/- 13.9</td>
</tr>
<tr>
<td>Female</td>
<td>14,104 (53.3%)</td>
</tr>
<tr>
<td>White</td>
<td>12,845 (48.5%)</td>
</tr>
<tr>
<td>Black</td>
<td>9,759 (36.9%)</td>
</tr>
<tr>
<td>Asian</td>
<td>1,198 (4.5%)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>977 (3.7%)</td>
</tr>
</tbody>
</table>

Conclusions

- Fewer than 1/3 of patients with diabetes and CKD or ASCVD had an active medication order for an SGLT2i or GLP-1 RA.
- Only 1/5 of patients with diabetes and HF had an active medication order for an SGLT2i.
- No more than 1/2 of patients with diabetes had an indication for evidence of any medication order (active or prior) for an SGLT2i or GLP-1 RA.
- Increasing the use of these evidence-based medications would reduce the risk of adverse cardiovascular events and progression of kidney disease in patients with diabetes.

Limitations

- We analyzed all SGLT2i and GLP-1 RA medications, not just those with proven benefits for each condition.
- We relied on ICD-10 codes, which may not accurately identify the study population.
- We did not include glomerular filtration rate or proteinuria in the definition of CKD, which are criteria for the use of SGLT2i.

Next Steps

Additional study is necessary to:

- Refine the definitions of the target populations (e.g., CKD with estimated glomerular filtration rate <60 or proteinuria).
- Examine differential prescribing by race, ethnicity, and language preference.
- Understand barriers to prescribing to guide interventions to improve the use of SGLT2i and GLP1 RA among high-risk populations with diabetes.
Identifying and classifying medications for hypertension
Vision

• Healthy people and communities enabled by the delivery of the best interventions for health promotion and disease prevention, diagnosis and treatment for each patient in primary care.

Mission

• To accelerate discovery and translation of research into primary care practice. We will achieve this with precision data and analytics that enables a learning health system, working in partnership with patients, families, clinicians, researchers and communities.

Research Aims

Our research aims to improve patient outcomes, by:

• Providing access to curated data to support research on primary care delivery
• Conducting research on common conditions in primary care, including hypertension, diabetes, and heart failure
• Understanding the impact of community and social determinants of health through predictive models and spatial analysis
• Identifying barriers to optimal care

Interested in Collaboration?

Contact us at: Samantha Pitts

Phone: (415)608-9750

Email: spitts4@jh.edu
Ching-Huan Wang, John Scott, Lisa Yanek, Samantha Pitts

Adult Primary Care COE

Focus
- Extracting relevant medications from clinical datasets is challenging due to the number of medication names and formulations for each ingredient.
- ATC® and EPC® are standard classification systems which are mapped to RxNorm codes for medication ingredient and formulation identification.
- Epic uses a proprietary system (FDI®) to classify medications and not all medications can be mapped to RxNorm codes.
- We sought to develop a standards-based approach to identify and extract anti-hypertensive medications from Epic.

Method & Analytics

Phase 1: Identification and classification of medications in the Clarity medication table (Fig. 1).

Active ingredient extraction:
- Extracted active ingredients from medications’ generic name (variable, GENERIC_NAME).
- Next, extracted ingredients from simple generic code (variable, SIMPLE GENERIC C) if generic name was missing or invalid.
- Pharmacologic class assignment:
- Assigned pharmacologic classes to extracted ingredients using the PHARMACEUTICAL_SUBCLASS of medications with only one ingredient.

Phase 2: Identification of angiotensin converting enzyme inhibitors (ACEI) via NLM vs. Epic pathways:
- Queried the NLM RxNav API® using ATC and EPC classes to identify all unique ACEI ingredients and their corresponding RxNorm codes.
- Identified medications in the Clarity medication table by ACEI ingredient RxNorm codes and by a string search of GENERIC_NAME and NAME for ingredients – "NM pathway".
- Queried the COE for medications identified as ACEI by the NLM pathway, and compared this to using FDO pharmaceutical subclasses – "Epic pathway".

Results and Highlights

Phase 1:
- Using generic name, we extracted at least one ingredient from 2,314 (91.9%) unique medications, with failures due to missing or invalid generic names (Fig. 2, Table 1).
- Adding a search of the simple generic code decreased failures in ingredient extraction by 77.5%, to 46 (1.8%) unique medications.
- 2,398 (96.7%) of unique medications matched to at least one pharmacologic class.

Phase 2:
- The NLM pathway for ACEI identified 570 unique medications in the Clarity medication table, including all 371 identified by the Epic pathway (Fig. 3).
- This resulted in a small increase in medications orders identified in the COE dataset – 27 unique medications and 818 medication orders.

Conclusions

The NLM pathway for ACEI:
- Identified all ACEI found by the Epic pathway.
- Used standard classification systems and nomenclature for identification of medications.
- Identified ingredients rather than medications.
- As a result, required SME review of a small set of ingredients – rather than a longer list of medications – with a small additional review of medications.
- These characteristics lend it to automation with reduced need for curatin by an SME.
- Using generic name and simple-generic codes to identify medication ingredients is feasible.
- Some gaps remain due to missing or invalid information.

Next Steps:
- Compare performance of NLM and Epic pathways with additional medication classes.
- Abbreviations and acknowledgments:
  We would like to thank Brant Chee for sharing his work using ATC, EPC and RxNorm for medication data and Michael Chiu for his parallel work on diabetes medications.

1ATC: Anatomical Therapeutic Chemical classification system
2EPC: Established Pharmacologic Class
3FDI: First Data Bank
4NLM: National Library of Medicine
RxNav: https://rhlnc.nlm.nih.gov/RxNav/
API: application programming interface
Psychosis PMCOE
The Johns Hopkins Schizophrenia Center
Psychosis PMCOE

Vision
• Our PMCOE wishes to establish a new, more effective method of precision patient care for patients with psychosis in a wide range of medical conditions.

Mission
• We aim to develop a new precision care for patients with psychosis. A novel biomarker and medicine will be applied to patients with psychosis in a wide range of medical conditions, resulting in a mechanism-driven targeted approach with less side effects than traditional antipsychotics.

Research Aims
• To dissect the pathophysiology of psychosis with accompanying cognitive deficits, particularly focusing on lysosomal deficits and augmented cellular autofluorescence (AF).
• To apply this biological knowledge to a wide range of medical conditions that accompany psychosis beyond the boundary of psychiatry and other medical specialties. We hypothesize that the lysosomal deficits underlie a subset of patients who show psychotic manifestations in a cross-disease manner.
• To apply a novel high-throughput method of detecting cellular AF for biospecimens from a wide range of medical conditions that accompany psychosis, hopefully that this method will be a tool in conducting clinical trials with an ongoing developing compound that targets the lysosomal deficits.

Interested in Collaboration?
Contact Dr. Akira Sawa, Director of the Johns Hopkins Schizophrenia Center, at:

Phone: (443)846-1409
Email: asawa1@jhu.edu
Psychosis PMCOE

Focus
1) To dissect the pathophysiology of psychosis with accompanying cognitive deficits, particularly focusing on lysosomal deficits and augmented cellular autofluorescence (AF).
2) To apply this biological knowledge to a wide range of medical conditions that accompany psychosis beyond the boundary of psychiatry and other medical specialties. We hypothesize that the lysosomal deficits underlie a subset of patients who show psychotic manifestations in a cross-disease manner.
3) To apply a novel high-throughput method of detecting cellular AF for biospecimens from a wide range of medical conditions that accompany psychosis, hopefully that this method will be a tool in conducting clinical trials with an ongoing developing compound that targets the lysosomal deficits.

Background
Our group originally discovered and studied lysosomal deficits, featured with augmented cellular autofluorescence (AF) in neuronal cells from patients with psychotic disorders. The level of AF is also associated with specific domains of cognitive deficits associated with psychosis (NIH RO1 grant: MH107730). The cellular deficits resemble the cellular pathology reported in lysosomal storage disorders, which are accompanied by cognitive and mental symptoms. In collaboration with a pharmaceutical industry, we have conducted this phenotype-based compound screening and the final candidate compounds are under toxicology and safety testing prior to clinical trials. We are also optimizing a high-throughput method of detecting cellular AF from blood samples.

Study design
Psychosis PMCOE

Data collection
General landscape
Patient subgroup for precision care

Data collection
- Epic Medical Record
- Inclusion criteria
- Diagnosis
- Medication
- PMAP registry

General landscape
- The use of antipsychotics in different types of diseases in the entire hospital system
- Reported adverse events of antipsychotics and their manifestation in different types of diseases

Patient subgroup for precision care
- Antipsychotic use
- Diagnosis
- Adverse events
- Dimensional/Stratified
- Blood data (AF)
- Machine learning

Federal support / industrial collaboration

Next step
1) We will investigate a hospital-wide landscape of psychosis, paying an attention to antipsychotic use and its adverse events.
2) In accordance with the development and refinement of a novel high-throughput method of detecting cellular AF and a novel compound that targets the lysosomal deficits with psychosis, we will define a subgroup of patients who show psychotic manifestations in a cross-disease manner.

Conclusion
1) By defining lysosome deficits-mediated psychotic manifestation, we will investigate a novel landscape of psychosis in a cross-disease manner beyond the boundary of psychiatry and other medical specialties.
2) Based on the precision medicine-guided approach, we will stratify patients with psychosis, mainly using a novel high-throughput method of detecting cellular AF.
3) We hope that the efforts of this PMCOE project, in combination with other projects under federal support/industrial collaboration, will lead to a new precision care for patients with psychosis.

Members
- Investigators: Nicola Cascella, Jennifer Coughlin, Koko Ishizuka, Akira Sawa (PI), Kun Yang (data science leader)
- Consultants and advisors: William Eaton, Donald German, Laurent Younes
- Staff members: Meevi Schumacher (lead of this poster)
- Research manager: Yukiko Lema
Predicting Imminent Conversion to Neovascular Age-related Macular Degeneration Using Optical Coherence Tomography Imaging and Deep Learning
Vision
• Harness the power of artificial intelligence, multimodal ophthalmic imaging and big data to provide previously-unavailable stratification, prognostication and treatment recommendations for patients with ophthalmic diseases.

Mission
• Usher in an era of personalized ophthalmic care and transform the field of ophthalmology using artificial intelligence.

Research Aims
• Provide fine-grained predictions for conversion to wet age-related macular degeneration (AMD) in patients with dry AMD
• Identify the optimal intravitreal anti-VEGF medication for patients with retinal vascular diseases at treatment initiation

Interested in Collaboration?
Email: tliu25@jhmi.edu
Predicting Imminent Conversion to Neovascular Age-related Macular Degeneration Using OCT Volumes and Deep Learning

T. Y. Alvin Liu, M.D. 1,2; Yuxuan Liu, M.S. 1; Neslihan Dilruba Koseoglu, M.D. 1; Craig K. Jones, Ph.D. 1,2,3

1. Wilmer Eye Institute, National Institutes of Health, Bethesda, MD, USA
2. Johns Hopkins University School of Medicine, Baltimore, MD, USA
3. Johns Hopkins Precision Medicine Analytics Platform, Baltimore, MD, USA

Introduction

• Age-related macular degeneration (AMD) is the most common cause of irreversible central vision loss in patients over age 50.1
• Most AMD patients lose vision due to neovascular AMD (NVAMD).
• Average risk for progression over 5 years can be estimated using AREDs criteria.

The goal of the current study is to analyze 3D optical coherence tomography (OCT) volumes using deep learning (DL) and to create a model that will predict imminent conversion (within 6 months) from non-neovascular to NVAMD. We chose a 6-month time frame, as it is a more actionable time frame as compared to 5 years.

Objective

• The goal of the current study is to analyze 3D optical coherence tomography (OCT) volumes using deep learning (DL) and to create a model that will predict imminent conversion (within 6 months) from non-neovascular to NVAMD. We chose a 6-month time frame, as it is a more actionable time frame as compared to 5 years.

Materials and Methods

• 33,189 3D OCT scans from 2,084 patients with AMD were split into a training (70%) and validation (20%) and hold-out test set (10%). Data partition was performed at the patient level.
• The conversion date was defined as the first anti-VEGF injection date based on previous work.2
• For training, all scans were center cropped into 496 x 512 x 24 (width x height x # line scans per volume). Then, the signal intensity was scaled with zero mean and unit standard deviation.
• Data augmentation with transformations and random flipping along vertical axis were applied with a probability of 50%.
• A 3D ResNet-50 model was used as backbone, along with a Stochastic Gradient Descent optimizer (learning rate = 1e-4) and a learning rate scheduler.
• Best model parameters were selected based on best balanced accuracy during validation.
• All processing done in Azure Machine Learning Studio within the Precision Medicine Analytics Platform.

Future work will include incorporating tabular clinical data into model training and testing.

Results

The hold-out test set contained 2051 OCT volumes from 136 patients.

On the hold-out test set, the model achieved a 68.8% balanced accuracy and 0.75 Area Under Curve (AUC). When set to a “sensitive” mode with a sensitivity of 80%, the model achieved a specificity of 45%. When set to a “specific” mode with a specificity of 80%, the model achieved a sensitivity of 57%.

Table 1. Test set confusion matrix: “yes” is yes conversion within 6 months and “no” is no conversion within 6 months; threshold = 0.5.

Table 2. Confusion matrix for subgroups within the test set: first converted eye in patients with NVAMD OU (left), second converted eye in patients with NVAMD OU (middle), first converted eye in patients with unilateral NVAMD (right). Threshold = 0.5.

Table 3. Model subgroup performance with threshold set at 80% sensitivity and 80% specificity.

Table 3. Model subgroup performance with threshold set at 80% sensitivity and 80% specificity.

Fig 2. Sample correct predictions by our model in the hold-out test set. Ground truth: no conversion within 6 months (top); yes conversion within 6 months (bottom)---note the presence of a double layer sign.

Conclusion and Future Work

• Our model showed reasonably robust performance in predicting imminent conversion to NVAMD using only a single OCT volume as input.
• In contrast to previous work2 that only included patients with known NVAMD in one eye, our model was trained and tested with OCT images from both patients with known NVAMD in one eye and patients without previous history of NVAMD.
• Future work will include incorporating tabular clinical data into model training and testing.

References

Scleroderma Precision Medicine Center of Excellence
Scleroderma Precision Medicine Center of Excellence

Vision
- Scleroderma is a heterogeneous disease with significant variability in a patient’s disease expression and clinical trajectory. Our precision medicine program in scleroderma will lead to improved understanding of a patient’s disease course, provide opportunities for early detection of complications and development of more targeted treatment approaches.

Mission
- The mission of The Johns Hopkins Scleroderma Center is to provide outstanding, holistic clinical care while seeking to better understand the causes and best treatment approaches for scleroderma. The Scleroderma Precision Medicine Center of Excellence is harnessing the revolution occurring in big data and computational science to neutralize disease complexity and variability.

Research Aims
- Identify patients at high risk of progressive disease across a broad spectrum of complications: interstitial lung disease, pulmonary hypertension, cardiomyopathy, among others.
- Detect emerging complications at an earlier stage of disease using novel biomarker, ambulatory device and quantitative imaging strategies.
- Identify patients who are most likely to respond to different treatment strategies.
- Prospectively define whether immunologically distinct subgroups predict clinical outcomes and treatment responsiveness.
- Construct an individual level predictive model of a patient’s likely trajectory and outcome.

Interested in Collaboration?

Contact us at:

Phone: (410)550-7715
Email:
Ami Shah: Ami.Shah@jh.edu
Laura Hummers: lhummers@jhmi.edu
Ji Soo Kim: jkim478@jhu.edu
Probabilistic clustering of longitudinal trajectories of scleroderma patients

Ji Soo Kim*, Yizhen Xu*, Rachel Wallwork, Laura Hummers, Ami Shah†, Scott Zeger†
Joint first authors*, joint senior authors†

Focus
• Autoantibodies are used clinically as tools to identify scleroderma patients with distinct risks of disease progression and complications
• For example, anti-topoisomerase 1 (anti-topo) positive patients have higher risk of interstitial lung disease, yet we lack tools to identify patients who are likely to progress from those likely to remain stable over time (Figure 1)
• Can scleroderma patients be further separated into subgroups based upon their longitudinal biomarkers?

Method and Analytics
• We used the population of anti-topo positive patients that, a priori, we would expect to be more homogeneous
• We use a multivariate linear mixed model so that the subgroups can be identified by the joint trajectories of the two measures, FVC and DLCO, while accounting for the correlation between them
• The model also accounts for baseline characteristics including age of onset, sex, race, and cutaneous subtype

Results and Highlights

- 289 anti-topo positive patients
  - “stable” group (151 patients): both biomarkers changed little from the date of disease onset (Figure 1 left panel)
  - “progressor” group (138 patients): a steep decline in both measures (Figure 1 right panel)

- The two-subgroup model was statistically significantly better than the others – the model with one group and models with more than two groups

- 289 anti-topo positive patients
  - “stable” group (151 patients): both biomarkers changed little from the date of disease onset (Figure 1 left panel)
  - “progressor” group (138 patients): a steep decline in both measures (Figure 1 right panel)

Figure 1. Estimated average trajectories of FVC and DLCO in standardized scale for each identified cluster among anti-topo positive patients. The values are quantile normalized so that observations within each measure follow the standard normal distribution. Cluster 1 (left) consists of patients with stable trajectories for both measures, and cluster 2 (right) with faster progressing trajectories.

- Our algorithm produces an updated individualized probability of belonging to the stable group at any given time (Figure 2)
- At baseline when no FVC and DLCO values are observed, the probability is calculated based on only clinical and demographic characteristics
- More precise estimations are obtained as more longitudinal observations become available

Conclusion
• Using our model, we can further risk-stratify homogeneous subgroup of patients sharing same baseline characteristics

Next Steps
• By identifying patients where there is tremendous uncertainty at baseline, we can now study biomarkers that may enhance risk prediction beyond known parameters
• Investigate variables highly associated with fast progressors among anti-topo positive patients with high uncertainty at baseline that can possibly lead to new biomarker discovery
Distinct scleroderma autoantibody profiles stratify patients for cancer risk at scleroderma onset and during disease course.


Focus

- Are scleroderma autoantibodies, alone or in combination, useful tools for cancer risk stratification?
- An array of scleroderma immune responses examined to study association with risk of cancer

Method and Analytics

- Scleroderma cancer and controls from Johns Hopkins and University of Pittsburgh Scleroderma Centers were studied
- Cancer cases identified using data on PMAP
- Sera assayed by Lineblot and ELISA for autoantibodies against centromere, topoisomerase I, RNA polymerase III (POLR3), PM/ScI, Th/To, NOR90, U3RNP (Fibrillarin), Ku, Ro52, U1RNP, and RNPC3 (U11/U12 RNF)

Logistic regression models constructed to examine whether distinct autoantibodies associated with overall cancer at any time and cancer-associated scleroderma
- ORs estimated from logistic regression models including each autoantibody as a predictor, adjusting for age at scleroderma onset, sex, race, cutaneous subtype, and history of smoking
- The effects of having >1 autoantibody on cancer further examined using random forest analysis

Results and Highlights

Study Population

Distinct immune responses associate with increased or decreased risk of cancer

<table>
<thead>
<tr>
<th>Autoantibody</th>
<th>Adjusted OR (95% CI)</th>
<th>Cancer-associated scleroderma</th>
<th>Overall cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>POLR3</td>
<td>2.28 (1.33-3.91)**</td>
<td>1.47 (1.03-2.11)**</td>
<td></td>
</tr>
<tr>
<td>Centromere</td>
<td>0.39 (0.20-0.74)**</td>
<td>0.69 (0.51-0.93)**</td>
<td></td>
</tr>
<tr>
<td>Monospecific Ro52</td>
<td>2.58 (1.05-6.30)**</td>
<td>2.19 (1.29-3.74)****</td>
<td></td>
</tr>
<tr>
<td>U1RNP</td>
<td>0.32 (0.11-0.93)*</td>
<td>0.63 (0.43-0.93)**</td>
<td></td>
</tr>
<tr>
<td>Th/To</td>
<td>0.41 (0.18-0.97)**</td>
<td>0.72 (0.48-1.08)</td>
<td></td>
</tr>
</tbody>
</table>

- Anti-POLR3 and monospecific anti-Ro52 associated with an increased overall cancer risk, whereas anti-centromere and anti-U1RNP associated with lower risk

Distinct combinations of 5 scleroderma immune responses may stratify patients for cancer risk

- Anti-POLR3 or anti-Ro52 alone associated with a higher predicted probability for cancer
- Patients with anti-Ro52 who also have anti-Th/To have a moderate cancer risk, as do patients who are negative for anti-POLR3, anti-centromere and anti-Ro52
- Anti-centromere or the combination of anti-Ro52 and anti-U1RNP have the lowest risk

Figure 1. Summary tree of the random forest model showing top five autoantibodies for estimating cancer risk in our population. Positivity or negativity of an autoantibody dictates a decision rule at each tree node, classifying patients into two subgroups. The number and percentage of cohort patients meeting criteria and their predicted cancer risk are shown at each terminal node.

Conclusion

5 distinct scleroderma immune responses, alone or in combination, may be useful tools to stratify scleroderma patients’ cancer risk

Next Steps

Further study examining cancer risk in autoantibody subgroups relative to the general population is warranted.
Cystic Fibrosis
Precision Medicine Center of Excellence
Vision

• To improve the outcomes and quality of life for all individuals with Cystic Fibrosis (CF) by identifying subgroups with similar disease complications and trajectories to inform individualized diagnostics and treatments.

Mission

• To fully harness all available clinical, genetic, social, and patient-derived information to deliver safe, beneficial, and cost-effective treatments to all individuals with CF. To facilitate collaboration with the CF Foundation and with pharmaceutical companies in early phase development and clinical trials of novel precision treatments for CF.

Research Aims

• Aim 1: To establish human nasal epithelial (HNE) cell culture from people with CF (pwCF) and CF carriers for in vitro assessment of cellular response to treatment with CFTR modulator therapies.

• Aim 2: To return insights from in vitro HNE testing to clinical practice to inform treatment decisions for pwCF.

• Aim 3: To establish a biorepository for primary HNE cells to allow for in vitro study of various molecular strategies for restoration of CFTR function.
The CF PMCOE studies a cohort of ~1000 individuals of all ages with CF and with CFTR-related disorders, utilizing a variety of different data streams including clinical (EPIC), genomic, environmental, and laboratory data. This project focuses on use of primary human nasal epithelial cell cultures (HNE) to serve as a reliable in vitro system for evaluation of response of CFTR variants that are not yet approved for CFTR-targeted therapies. Functional analysis of these cultured cells reflects the contribution of both alleles to CFTR function and considers the patient’s entire genomic context.

**Method and Analytics**

To date, 80 HNE samples have been collected from the CF individuals, as well as their parents and healthy controls (IRB# 00116966).

**Results and Highlights**

- **Q359R Sibling 1**: Symptoms completely resolved. Sweat chloride decreased to <10mmol/L from 55 mmol/L after initiation of therapy. BMI increased to normal range.
- **Q359R Sibling 2**: Sweat chloride decreased to 14mmol/L from 60mmol/L after initiation of therapy. Spirometry improved to an FVC of 101%, FEV1 101%, and FEV1/FVC 100%.
- **G480S/3120+1G>T**: Significant clinical improvement observed. Sweat chloride concentration dropped to 46mmol/L from 119mmol/L. Initial PFTs were FVC 47%, FEV1 29%, and FEV1/FVC 83%; which increased to FVC 70%, FEV1 50%, and FEV1/FVC 72%.

- We emphasize the importance of using HNE cells to determine which modulators, either individually or in combination, result in maximum functional restoration. E.g. Ivacaftor alone to Q359R, Trikafta to G480S.
- R334W variant was previously believed to be unresponsive, but our studies demonstrate a clinically significant increase in function to 14% with Trikafta treatment.
- Through RNA analysis, we can assess the impact of variants on splicing, such as c.1679G>C (R560T) and c.2490+1083C>T, and predict the response to modulators.

**Conclusion**

- We emphasize the importance of using HNE cells to determine which modulators, either individually or in combination, result in maximum functional restoration. E.g. Ivacaftor alone to Q359R, Trikafta to G480S.
- R334W variant was previously believed to be unresponsive, but our studies demonstrate a clinically significant increase in function to 14% with Trikafta treatment.
- Through RNA analysis, we can assess the impact of variants on splicing, such as c.1679G>C (R560T) and c.2490+1083C>T, and predict the response to modulators.

**Next Steps**

To establish a primary cell repository that will help:
- Investigate molecular genetics of CF-causing variants.
- Assess CFTR function recovery through modulator treatment.
- Evaluate antisense oligos ASOs for splicing restoration.
- Identify a suitable method for delivering gene editing materials.