Machine Learning in Medicine

INDIANA CTSI Clinical and Translational Sciences Institute

PURDUE UNIVERSITY Engineering Medicine

INDIANA UNIVERSITY SCHOOL OF MEDICINE

POSTER BOOKLET

May 9th, 2025

Indianapolis, IN

TABLE OF CONTENTS

Machine Learning in Medicine

Poster Number 1

Indiana CTSI: Catalyzing Research

Sharon Moe

Medicine/Nephrology and CTSI, CTSI

Information poster

Indiana CTSI: Catalyzing Research

S. Moe, MD | S. Wiehe, MD, MPH | G. Claxton MPH, RD | D. Lynch, BFA

The Indiana CTSI is a collaborative effort among Indiana's three major research-intensive academic institutions—Indiana University (IU), Purdue University (PU) & the University of Notre Dame (UND)—along with the Regenstrief Institute & numerous non-academic partners. These partners include community organizations, healthcare providers, governmental bodies & industry leaders. Together we catalyze research.

1 Connect & Initiate

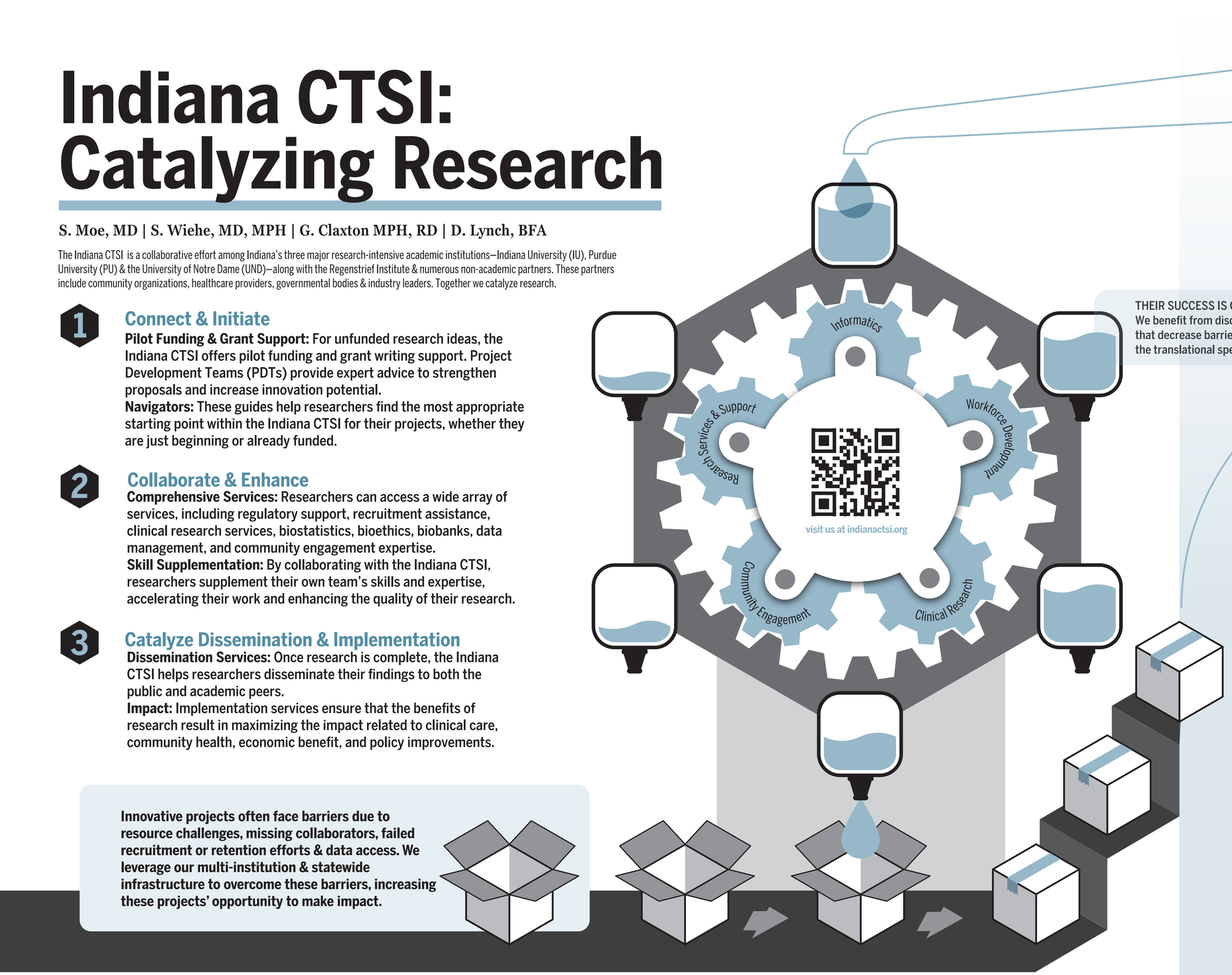
Pilot Funding & Grant Support: For unfunded research ideas, the Indiana CTSI offers pilot funding and grant writing support. Project Development Teams (PDTs) provide expert advice to strengthen proposals and increase innovation potential.
Navigators: These guides help researchers find the most appropriate starting point within the Indiana CTSI for their projects, whether they are just beginning or already funded.

2 Collaborate & Enhance

Comprehensive Services: Researchers can access a wide array of services, including regulatory support, recruitment assistance, clinical research services, biostatistics, bioethics, biobanks, data management, and community engagement expertise.
Skill Supplementation: By collaborating with the Indiana CTSI, researchers supplement their own team's skills and expertise, accelerating their work and enhancing the quality of their research.

3 Catalyze Dissemination & Implementation

Dissemination Services: Once research is complete, the Indiana CTSI helps researchers disseminate their findings to both the public and academic peers.
Impact: Implementation services ensure that the benefits of research result in maximizing the impact related to clinical care, community health, economic benefit, and policy improvements.



Innovative projects often face barriers due to resource challenges, missing collaborators, failed recruitment or retention efforts & data access. We leverage our multi-institution & statewide infrastructure to overcome these barriers, increasing these projects' opportunity to make impact.

Mechanisms to Catalyze Translational Research

Clinical Research	Streamlining the clinical research process, giving researchers capacity to conduct high-quality clinical trials.
Informatics	Supporting efficient management analysis & utilization of data in clinical & translational science.
Workforce Development	Cultivating skilled & knowledgeable leaders to advance translational research.
Research Services & Support	Providing access to specialized services that increase researchers' capacity to do high-quality studies.
Community Engagement	Fostering community & university partnerships in all phases of research.

Resources to Catalyze Translational Research

- Data Access
- Regulatory Support
- Recruitment
- Dissemination
- Training & Mentorship
- Facilities
- Grantsmanship
- Advice & Funding
- Statewide Reach
- Partnership Development

INDIANA CTSI
Clinical and Translational Sciences Institute

THEIR SUCCESS IS OUR SUCCESS
We benefit from discoveries bubbling up that decrease barriers for projects across the translational spectrum.

Notable Successes

CONNECT

Enhancing Pedestrian Safety: The Indiana CTSI collaborates with the Indiana Department of Health to fund community-engaged research projects. These awards, capped at $25,000, catalyze significant change. Health by Design, IU School of Medicine, and Riley Hospital for Children, focused on reducing traffic-related injuries and fatalities. By expanding data collection and analysis on pedestrian-involved crashes, this initiative guided the City of Indianapolis to improve road and crosswalk engineering. **The city secured over $19.9 million in federal funding for these improvements and enacted a no-turn-on-red law, enhancing pedestrian safety.**

COLLABORATE

Advancing Kidney Transplantation: IU School of Medicine physician scientist received advice, financial support, and a critical connection to a biomedical engineer from Indiana CTSI Project Development teams. **He developed a device to increase transplantation rates and reduce unused kidneys, founded AKICept, which is commercializing the first intravascular treatment procedure for acute kidney injury.**

CATALYZE

Economic Impact of Untreated Mental Illness: The Indiana Family and Social Services Administration (FSSA) requested, from the Indiana CTSI, an assessment of the economic impact of untreated mental illness in Indiana. This work showed an estimated cost of over $4.2 billion annually. **These results supported the Indiana Behavioral Health Commission's recommendations for policy and programs proposed in the 2023 Senate Enrolled Act 1, "Behavioral Health Matters" a comprehensive legislation enhancing access to mental health care and protections across the state.**

This project was funded with support from the Indiana Clinical and Translational Sciences Institute which is funded in part by Award Number UM1TR004402 from the National Institutes of Health, National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Poster Number 2

Emergency Room Overcrowding and Timely Sepsis Care

Kay Alderfer

Nursing, Purdue University

Title of presentation: Emergency Room Overcrowding and Sepsis Care

Background: Sepsis is a deadly disease with a global mortality rate just under 20% (Rudd et al., 2020). Timely administration of antibiotics has lasting impacts for septic patients (Peltan, Brown, et al., 2019). For every hour of delayed antibiotics mortality can increase from 4-9% (Kumar et al., 2006)(Seymour et al., 2017)(Liu et al., 2017). Overcrowded Emergency Departments can precipitate deadly delays for a septic patient for whom time is of the essence (Niederman et al., 2021).

The National Emergency Department Overcrowding Scale (NEDOCS) measures the number of ED beds, hospital beds, patients in the ED, patients on the ventilator, admissions in the ED, admission times, and waiting room wait times (Weiss et al., 2004). The aim of this study was to see if there is a correlation between NEDOCS scores and the time from door to antibiotics for patients diagnosed with sepsis.

Methods: This study is a retrospective cohort analysis that was conducted on data from three community hospital emergency rooms in Indiana. Right skewed administration times were analyzed using a generalized linear model with a gamma distribution within XLSTAT. Covariates included site, shift, order priority, order set usage, arrival route, acuity, trauma cases, ICU care, sex, census, and door to provider times. Also considered were season, weekend shifts, race, ethnicity, language, hospice care, age, provider role, and discharge disposition.

Results: For every increase in the NEDOCS score, the average wait time for antibiotics increased by 4.2 seconds (p=0.018). In a facility with a severe state of overcrowding (NEDOCS=180) the wait for antibiotics increases by 12.6 minutes. The primary way that overcrowding impacts antibiotic timing is via changes in the time from door to provider. The biggest delays were seen amongst sepsis patients with low acuity levels assigned at triage (219 minutes, p=0.014). Differences in location, order set usage, arrival method, weekend, gender, season, and ICU admission also had statistically significant impacts.

Conclusions: Overcrowding can slow antibiotic administration by a marginal amount. The primary way that overcrowding impacts antibiotic timing is via changes in the time it takes to be seen by a provider. Societal factors that affect antibiotic timing include sex, season, and weekends. Community factors include hospital location, order set usage, and arrival route. Physiological factors such as acuity and ICU needs also impacted the timeliness of antibiotics. Further research is needed to design and implement interventions.

Emergency Room Overcrowding and Timely Sepsis Care

K. Alderfer, MSN, RN, CNL, CPPS
Clinical Nurse Quality Coordinator – IU Health
PhD Student – Purdue University School of Nursing
kalderfer@iuhealth.org
kalderfe@purdue.edu

Introduction

Sepsis is a deadly disease with a global mortality rate just under 20%.[1] Timely administration of antibiotics has long lasting impacts for septic patients.[2] For every hour of delayed antibiotics mortality can increase from 4-9%.[3,4,5] Overcrowded Emergency Departments can precipitate deadly delays for a septic patient for whom time is of the essence.[6] Often symptoms are vague and there is no single immediate test for sepsis, further delaying intervention.[7]

Background

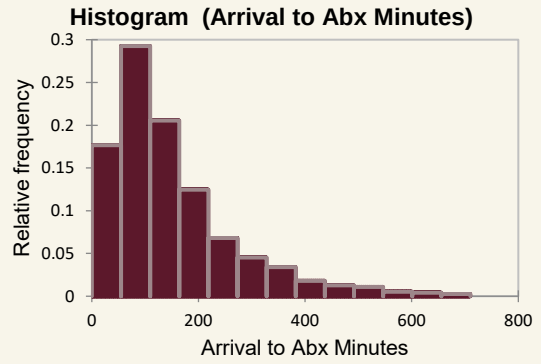
The National Emergency Department Overcrowding Scale (NEDOCS) is a calculated way to measure Emergency Department Over Crowding. Developed by the University of New Mexico, this measure considers the number of ED beds, hospital beds, patients in the ED, patients on the ventilator, admissions in the ED, admission times, and waiting room wait times.[8] This metric is commonly used to quickly convey how overcrowded an Emergency Room is for setting safety decisions. Resource allocation can be adjusted based on the level of point in time overcrowding.[9,10]

Purpose

The aim of this study is to see if there is a correlation between NEDOCS scores and the time from door to antibiotics for patients diagnosed with sepsis. This study will consider confounding variables through the lens of the Society to Cells Framework.[11]

Methods

This study is a retrospective cohort analysis that was conducted on data from three community hospital emergency rooms in Indiana. Data was merged to analyze the correlation between NEDOCS scores and time from arrival to antibiotic administration. Right skewed administration times were analyzed using a generalized linear model with a gamma distribution within XLSTAT software. Covariates included site, shift, order priority, order set usage, arrival route, acuity, trauma cases, ICU care, sex, census, and door to provider times. Also considered were season, weekend shifts, race, ethnicity, language, hospice care, age, provider role, and discharge disposition.



Conclusions

Overcrowding can slow antibiotic administration by a marginal amount. The primary way that overcrowding impacts antibiotic timing is via changes in the time it takes to be seen by a provider, Societal factors that affect antibiotic timing include sex, season, and weekends. Community factors include hospital location, order set usage, and arrival route. Physiological factors such as acuity and ICU needs also impacted the timeliness of antibiotics.

Data/Results

Variables	Minutes	P Value
NEDOCS	+0.07	0.018
Society		
Male	-9	< 0.0001
Quarter-2	-9	0.003
Weekend	-10	< 0.0001
Community		
Order Set Use	-46	< 0.0001
Order Priority STAT	-18	< 0.0001
EMS Arrival	-10	< 0.0001
Order Priority-Routine	+17	< 0.0001
Trauma Alert	+13	0.030
Facility B	+37	< 0.0001
Facility C	+37	< 0.0001
Physiology		
ICU Level of Care	-9	0.001
Acuity Level 3	+58	< 0.0001
Acuity Level 4	+70	< 0.0001
Acuity Level 5	+219	0.014

Implications

Based on these results, further research is needed to design and implement interventions aimed at all society to cells framework levels. Below are an outline of further studies to be done.

- Evaluate order set usage and its role in antibiotic delivery
- Studies on the modeling of sepsis improvement programs and other differentiating factors between the sites should be conducted.
- Further analyze the role of gender within sepsis care and treatment in the Emergency Room setting.

References

Poster Number 3

AI for Surgical Scene Understanding: Current Progress and Next Steps

Matthias Carstens

Weldon School of Biomedical Engineering, Purdue University

Artificial Intelligence (AI)-based Surgical Scene Understanding (SSU) is a rapidly evolving field that seeks to enhance intraoperative decision-making and ultimately improve patient outcomes. However, despite notable technological progress, the clinical integration of neural network models into the operating room remains minimal, and supportive applications are still under development.

To explore the reasons behind this limited translation into practice, a systematic review was conducted. Studies were included if they analyzed intraoperative data from minimally invasive abdominal surgeries in humans, developed computational models for SSU, and reported formally validated, trainable models with performance metrics.

Out of 3,011 screened publications from six literature databases, 188 studies met the inclusion criteria. Key findings highlight several challenges impeding clinical adoption. First, most studies relied on small, single-center datasets, often involving fewer than 100 patients and lacked essential patient metadata such as demographics or disease characteristics. Furthermore, research heavily depended on a few public datasets, particularly laparoscopic cholecystectomy videos like Cholec80, which severely limited the topical diversity of the models. Technically, the majority of models were based on convolutional neural networks and employed fully supervised training approaches, frequently neglecting real-time processing requirements or temporal modeling techniques such as optical flow. The review also revealed poor reporting practices, especially concerning the intended clinical goals, rationale for data selection, model uncertainty and error estimation, and the availability of code and trained models. Critically, only a small proportion of studies assessed their models in real-life surgical scenarios, and most research efforts remained focused on descriptive tasks like object detection or anatomical localization without offering deeper clinical insights or assistance.

After over a decade of neural network development, the field must move beyond algorithmic innovation toward clinically useful applications. Yet, real-world SSU translation remains rare. Research has largely followed data availability rather than clinical needs, with very limited procedural coverage. Future progress requires proactively collecting diverse, high- quality data and metadata, improving validation standards, bridging the gap between technical performance and clinical utility, and establishing clear ethical, regulatory, and clinical frameworks.

AI for Surgical Scene Understanding: Current Progress and Next Steps

Matthias Carstens, Shubha Vasisht, Zheyuan Zhang, Iulia Barbur, Annika Reinke, Lena Maier-Hein, Daniel A. Hashimoto, Fiona R. Kolbinger

PURDUE UNIVERSITY
Weldon School of Biomedical Engineering

BACKGROUND

AI-based Surgical Scene Understanding (SSU) is an active research field aiming to support intraoperative decision-making and improve patient outcomes. Despite progress, clinical use of neural networks in the OR remains rare, and assistive features are still under development.

Why? Are technical barriers still too high or are other factors delaying clinical translation?

METHODS

A **systematic review** was conducted. Studies met the following criteria: (1) analysis of intraoperative data from minimally invasive abdominal surgeries in humans, (2) development of computational models for SSU, and (3) reporting of trainable models with formal validation procedures and performance metrics. Relevant data were extracted and analyzed (see Fig. 1).

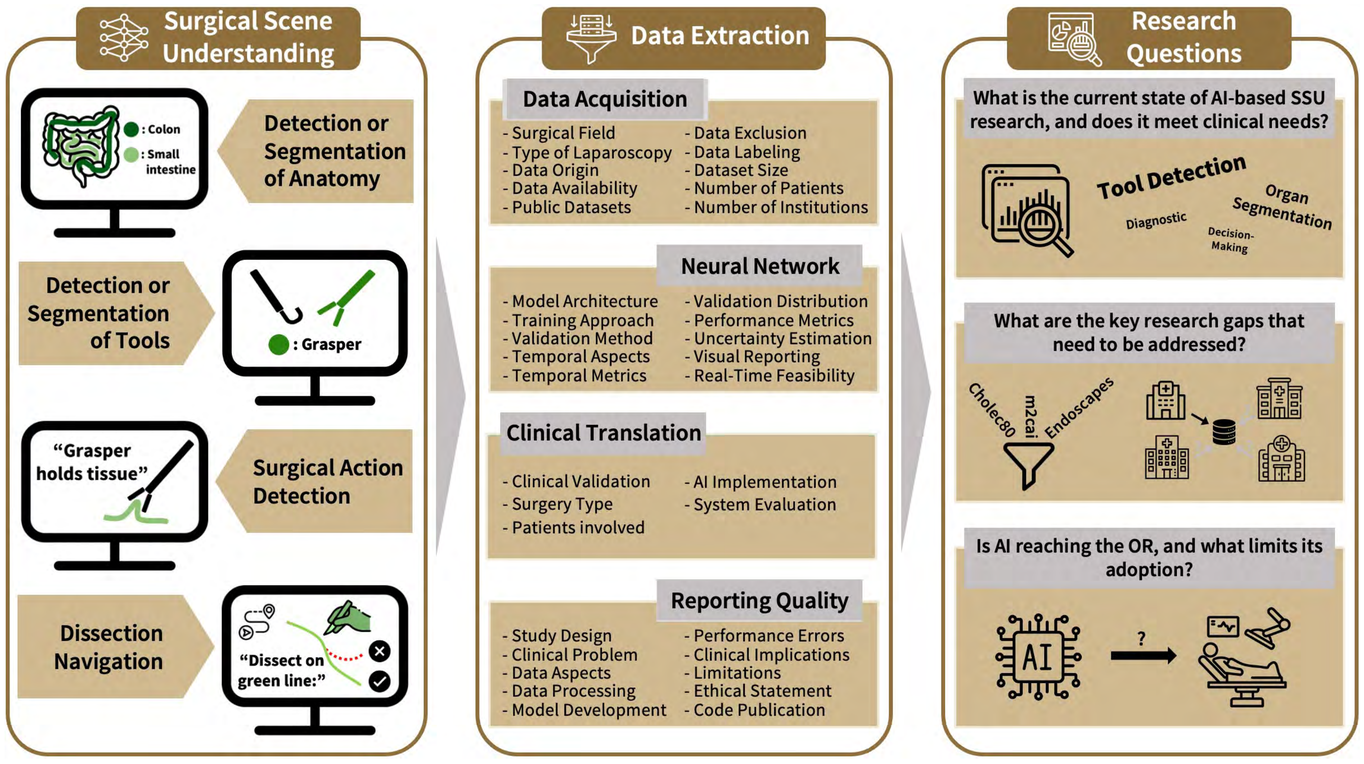


Figure 1: Main components of this systematic review.

FINDINGS

188 papers of 3,011 in 6 literature databases met the inclusion criteria.

Major findings were (see Fig. 2):

- Primarily the use of **small, single-center datasets** (often <100 patients), without patient metadata (e.g., demographics, disease parameters)
- **Heavy reliance on few public datasets** (mainly laparoscopic cholecystectomy videos, such as Cholec80 (56.2%)) leading to limited topical diversity
- Predominantly CNN-based (88.3%) **fully supervised** (82.4%) model training, often **without addressing real-time processing** needs (48.9%) or temporal aspects such as optical flow (27.1%)
- **Poor reporting quality** regarding: Clinical goal and potential for clinical assistance, data selection, assessment of model uncertainty and error estimation, code and model availability
- **Minimal focus on clinical translation** with only few studies (5.9%) tested it in real-life scenarios.
- Most works remain **limited to descriptive tasks** such as object detection or anatomical localization without further clinical interpretation (see Fig. 3).

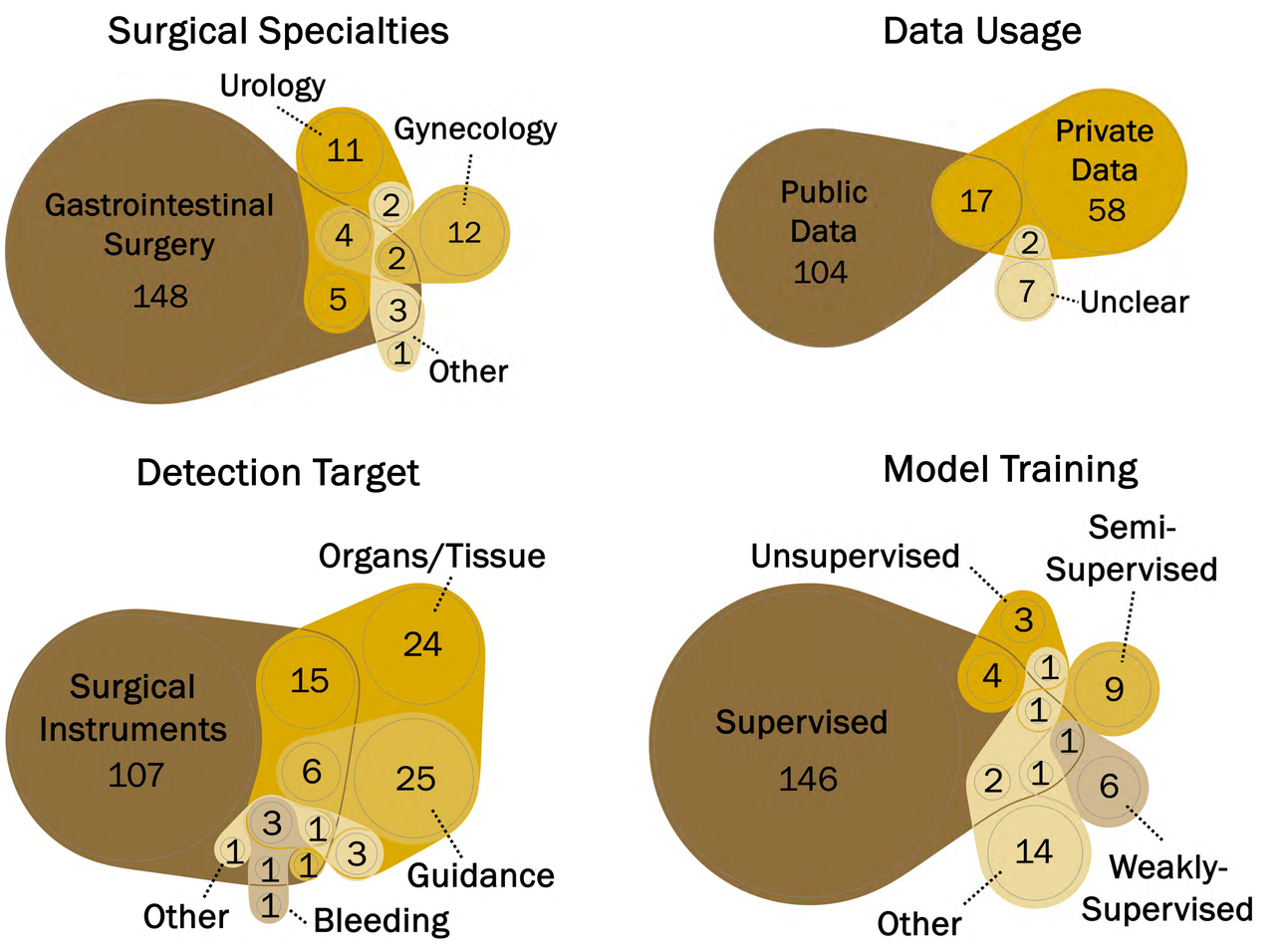


Figure 2: Examples of our results. The area size corresponds to the absolute frequency of reporting, with overlapping regions indicating studies that addressed multiple aspects simultaneously (Numbers within each section represent the exact count of papers).

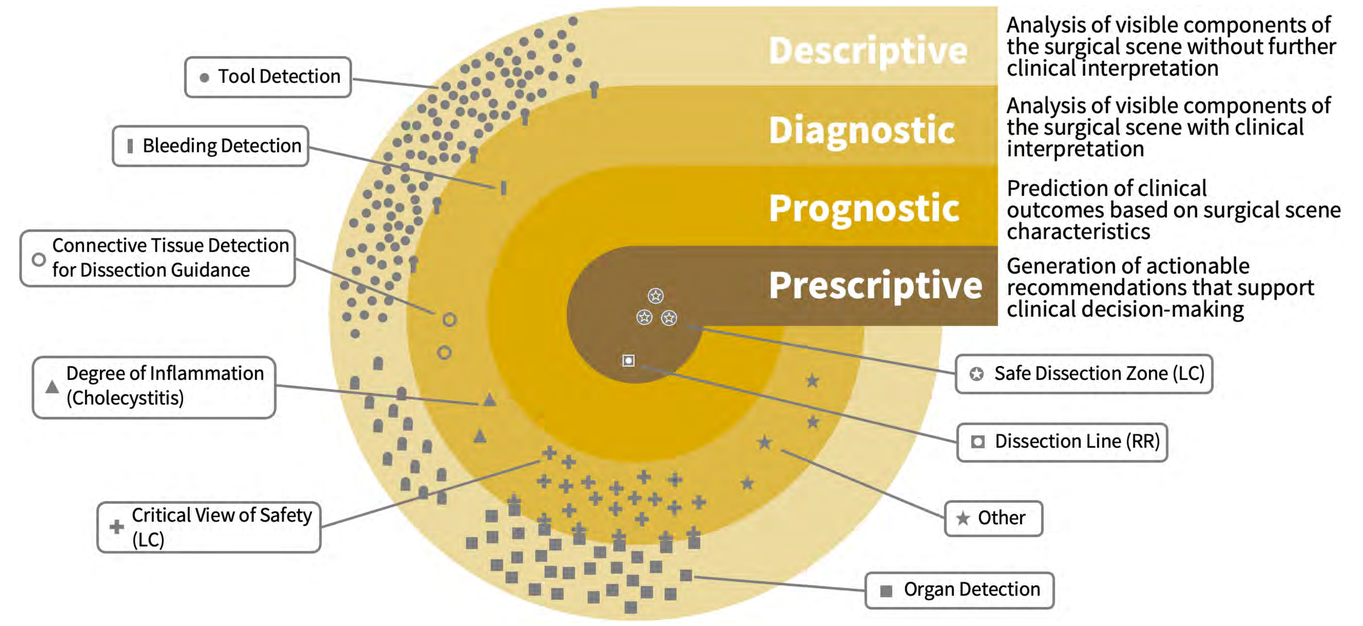


Figure 3: Distribution of Surgical Scene Understanding publications regarding their potential depth and significance for clinical translation of the evaluated AI models. (Abbr.: LC: Laparoscopic Cholecystectomy, RR: Rectal Resection)

CONCLUSION

After over a decade of neural network development, the field must move beyond algorithmic innovation toward clinically useful applications. Yet, real-world SSU translation remains rare.

Research has largely followed data availability rather than clinical needs, with very limited procedural coverage.

Future progress requires **proactively collecting diverse, high-quality data and metadata**, **improving validation** standards, bridging the gap between technical performance and clinical utility, and establishing clear ethical, regulatory, and clinical frameworks.

Matthias Carstens (MD), carstens2@purdue.edu

May 9, 2025

INDIANA CTSI Clinical and Translational Sciences Institute
PURDUE UNIVERSITY Engineering Medicine
INDIANA UNIVERSITY SCHOOL OF MEDICINE

Poster Number 4

Domain Adaptive Deep Learning for Motion Correction in High Resolution Peripheral Quantitative Computed Tomography

Farhan Sadik

Biomedical Engineering, Purdue University

High-resolution peripheral quantitative computed tomography (HR-pQCT) facilitates the detailed microstructural analysis of bone at peripheral sites, such as the distal tibia and radius, achieving a resolution of 60 Î¼m. However, this high-resolution imaging is prone to microstructural distortions caused by patient motion, commonly referred to as motion artifacts. The conventional method for mitigating these artifacts involves rescanning the patient, which is both resource-intensive and costly. Furthermore, rescanning often fails to eliminate motion artifacts in patients with tremors or involuntary movements, resulting in persistently degraded image quality. Recent advancements have explored the use of deep learning to address rigid motion artifacts, employing models trained on simulated motion-corrupted data. However, these models exhibit limited generalizability to real-world motion-corrupted data due to the inherent distribution shift between simulated and real-world datasets. In this study, we propose a domain adaptation framework to bridge the distribution gap between real-world motion-affected scans and simulated data. A shifted windows transformer is subsequently trained in an adversarial manner on the adapted images and their corresponding ground truth to effectively correct motion artifacts. Qualitative evaluation using visually graded scores (VGS) demonstrates that the proposed method achieves a minimum improvement of 2 Â± 0.5 points in artifact correction. Additionally, bone geometry measurements confirm the methodâ€™s efficacy in enabling accurate quantitative bone assessments in motion-affected scans.

Domain Adaptive Deep Learning for Motion Correction in High-Resolution Peripheral Quantitative Computed Tomography

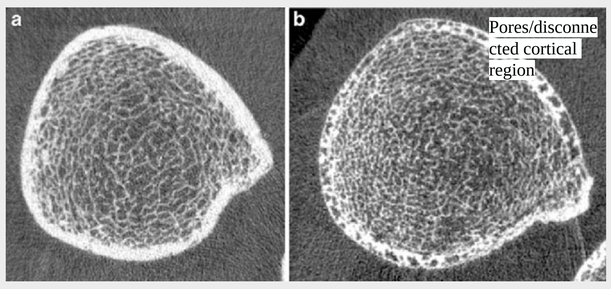
[1]Farhan Sadik, [2]Anika Mathur, [3]Michael Wan, [4]Christopher L. Newman, [5]Stuart J. Warden, [1]Rachel K. Surowiec

[1]Weldon School of Biomedical Engineering, Purdue U.; [2]College of Science, Purdue U.; [3]Institute for Experiential AI, Northeastern U.; [4]Dept. of Radiology and Imaging Sci., IUSM; [5] Dept. of Physical Therapy, School of Health and Human Sci., IU.

PURDUE UNIVERSITY
Weldon School of Biomedical Engineering

1. INTRODUCTION

- High-resolution peripheral quantitative computed tomography (HR-pQCT) is an advanced imaging modality specifically designed to visualize the **microarchitectural features** of bone. It enables the early detection of skeletal abnormalities associated with conditions such as osteoporosis and the skeletal manifestations of chronic kidney disease.
- Compared to conventional stationary CT, HR-pQCT offers significantly higher spatial resolution (~**60.7 µm**), enabling detailed visualization of bone microstructural changes, particularly within cortical and trabecular compartments. In cases of **osteoporosis**, features such as cortical porosity progression and trabecular disconnection can be detected, which are only discernible through high-resolution modalities like HR-pQCT.

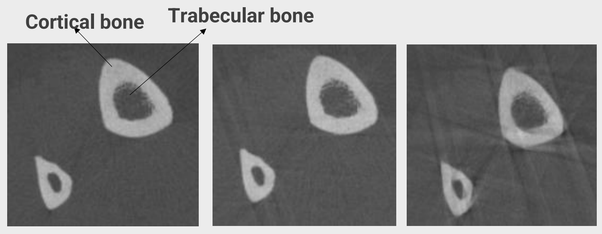


High-resolution peripheral QCT (HR-pQCT) images of the distal tibia in (a) without fractures, and (b) with osteoporotic fragility fractures.

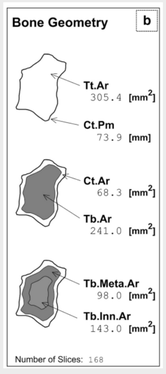
However, high-resolution comes with a cost - ***motion artifacts.***

2. Consequence of motion artifacts

a.



b.



[1]GAZZOTTI et al.

Direct Consequences

- Cortical surface area is disturbed leading to incorrect calculation of
  - Total area **Tr.Ar**
  - Cortical perimeter **Ct.Pm**
  - Cortical Area **Ct. Ar**
  - Trabecular Area **Tb.Ar**
  - Trabecular Metaphyseal Area **Tb.Meta Ar**
  - Trabecular Innermost Area **Tb.Inn.Ar**

About **23%** of the first-generation scans are motion corrupted ([2]Pialat et. al.)

The motion correction pipeline can be a valuable tool in research facilities, serving as a standalone software to **reduce the need for time-consuming rescans**

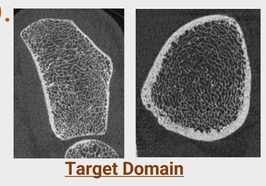
3. Current solution

a.

Physics-based motion simulation → Deep Learning-based correction → Quantitative geometry analysis / Re-scan

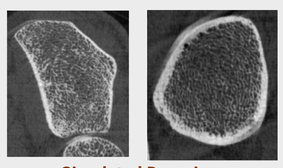
Real-world scans (Target Domain) | Simulated corrupted scans (Source Domain) | Predicted scan (Source domain)

b.



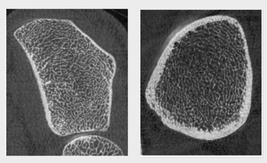
Target Domain

- Motion-corruption localized to specific sections in the scan
- Some sections are motion-free



Simulated Domain

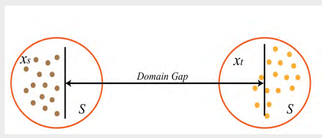
- Consistently blurry throughout the image
- Not exact to the real world



Motion Compensation in the simulated domain

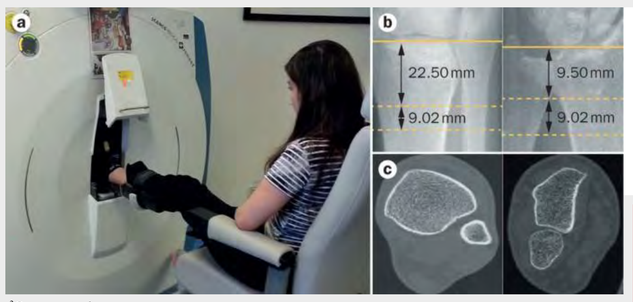
- Resolves cortical breaks.
- Effective Deblurring performance.
- Poor performance on the target domain data.

d.

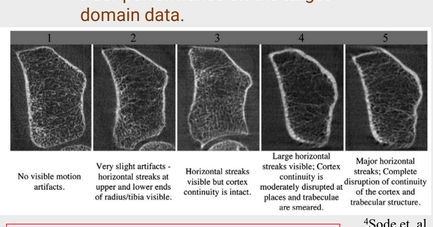


As the model is trained on the **source** domain, it does not perform well on the **target** domain.

c.



[3]Geusens et. al



[4]Sode et. al

The primary objective is to achieve motion correction such that the Visual Grading Score (VGS) improves from 5 to 1 or 2, thereby rendering the images suitable for quantitative geometric analysis.

Data Collected from FIT Core, ICMH

INDIANA UNIVERSITY SCHOOL OF MEDICINE
Indiana Center for Musculoskeletal Health

4. The key component – Self-Attention

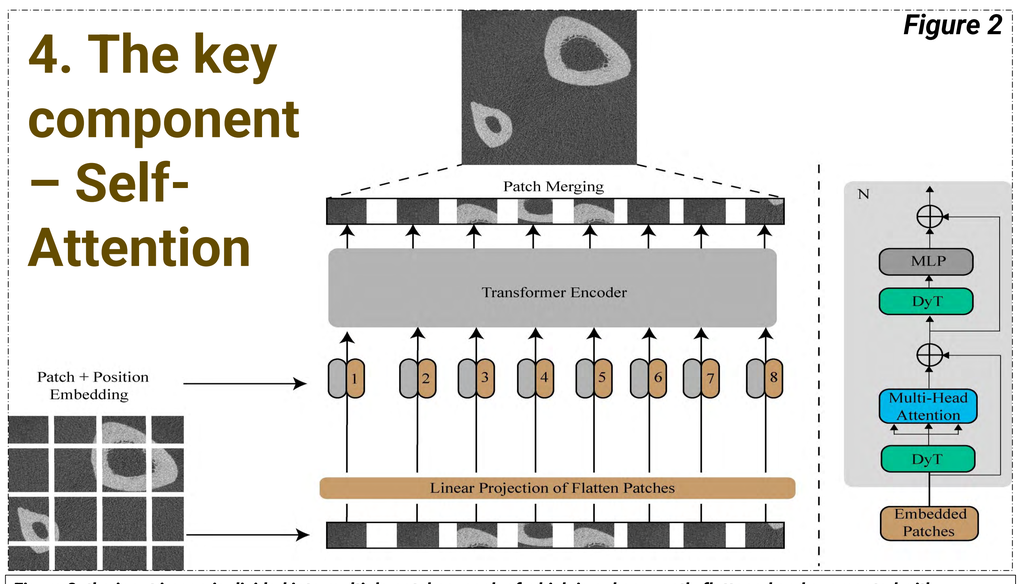


Figure 2: the input image is divided into multiple patches, each of which is subsequently flattened and augmented with corresponding positional encodings prior to being fed into the Transformer Encoder. The Transformer Encoder leverages a multi-head self-attention mechanism to model the inter-patch dependencies. In the final stage, a patch merging operation is employed to reconstruct the complete image. This architecture enables the utilization of global contextual information, in contrast to convolutional neural networks (CNNs) that rely primarily on local features, thereby facilitating more effective motion correction.

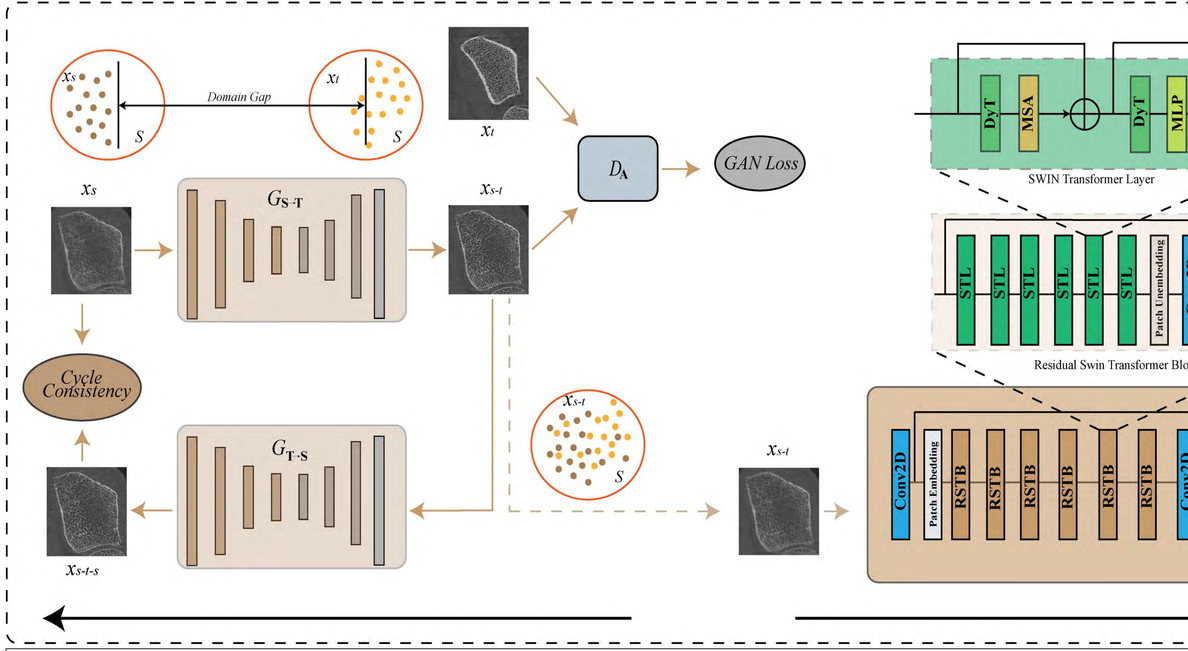


Figure 1

Figure 1: we employ two generator networks, $G_{S\to T}$ and $G_{T\to S}$, which map the source image to the target domain and the target image to the source domain, respectively. Specifically, the generator $G_{S\to T}$ projects the representation of x_s into the target feature space and reconstructs it as x_{s-t}. This reconstructed image is then compared to an unpaired target domain image, x_t, by the discriminator DA. The discriminator's objective is to distinguish between the source and target domains, and its success is measured by its inability to do so. The objective function can be formulated as follows

$$\mathcal{L}_{\text{GAN}}(G_{S\to T}, D_A, X_T, X_S) = \mathbb{E}_{x_t\sim X_T}[\log D_A(x_t)] + \mathbb{E}_{x_s\sim X_S}[\log(1 - D_A(G_{S\to T}(x_s)))]$$

However, directly adhering to the first cycle (from $G_{S\to T}$ to DA) may cause the network to generate images that are structurally similar to the target domain (unpaired mapping), which contradicts the objective of domain adaptation. To preserve the structural and functional consistency, we introduce a reverse cycle, where the generator $G_{T\to S}$ maps the generated image x_{s-t} back to the source domain, producing x_{s-t-s}. A cycle consistency loss is then computed between the original source image, x_s, and the reconstructed image. A shifted windows vision transformer is utilized to correct motion in the domain-adapted image, x_{s-t}.

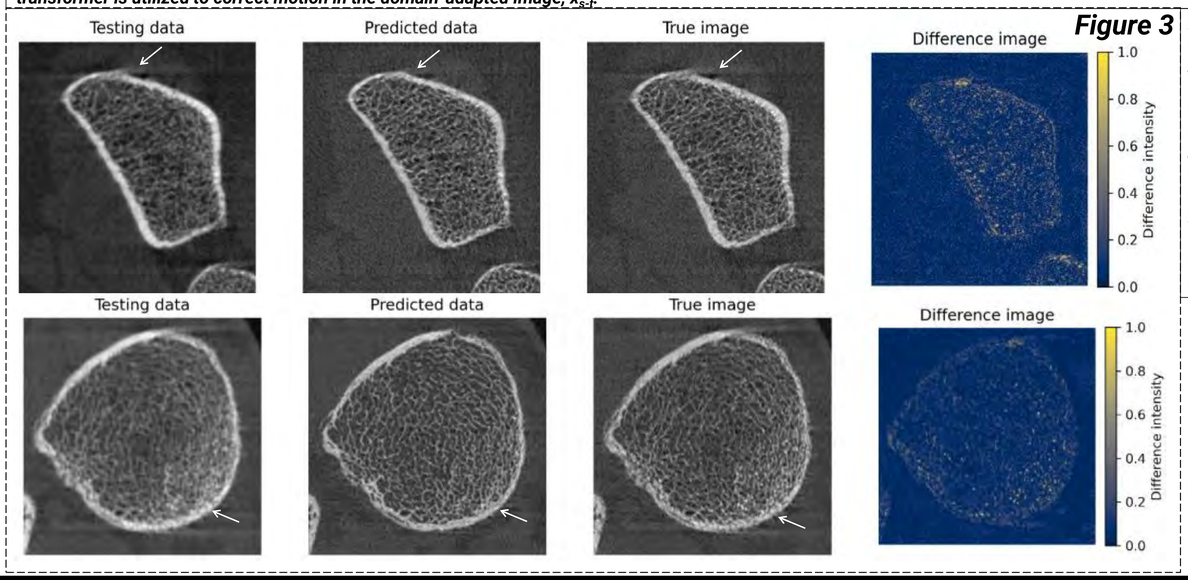


Figure 3

Figure 3: The top row displays images of the distal radius, while the bottom row depicts images of the distal tibia. The testing data were processed through the domain-adaptive network, followed by the transformer-based motion correction module (middle column). As the data were acquired from a real-world scanner, ground truth images are unavailable; however, the corrected outputs demonstrate qualitative improvements over the original real-world images.

Quantitative Biomedical Imaging and Spectroscopy Lab

5. Results and conclusion

- Figure 2 presents a qualitative analysis demonstrating the correction of cortical breaks in the distal radius and distal tibia using data acquired from a real-world scanner.
- Future work will focus on how the motion correction technique **improves segmentation performance**, as well as bone geometry measurement, such as **cortical thickness**, **trabecular number**, and derived parameters such as **Bone Mineral Density (BMD)**.
- This motion correction technique can be integrated with the scanner as an external software package, reviving about ~ 60% of motion-affected scans.

6. REFERENCES

1. Gazzotti et. al. Br J Radiol. 2023
2. J. Pialat et. al. Bone, Jan.2012
3. Geusens et. al.,. Nat Rev Rheumatol. 2014
4. Sode et al., Bone. 2011

Contact:
Farhan Sadik, fsadik@purdue.edu
Rachel K. Surowiec, rsurowie@purdue.edu

Indiana University School of Medicine
PURDUE UNIVERSITY Engineering Medicine
INDIANA CTSI Clinical and Translational Sciences Institute

Poster Number 5

Quantification of Extracellular Volume Fraction in Cardiac MRI without Blood Sampling Using Multi-Stage Training Deep Learning

Zhuoan Li

Biomedical Engineering, Purdue University

Motivation:
Blood sampling for hematocrit (HCT) measurement limits clinical use of extracellular volume fraction (ECV) in diagnosing myocardial diseases.
Goals:
Develop and evaluate a deep learning (DL) model to predict HCT from cardiovascular magnetic resonance (CMR) data, exploring if additional features improve predictability across multiple centers.
Approach:
Trained a multi-stage DL model using multi-center CMR T1 values and clinical features to predict HCT without blood sampling.
Results:
The DL model identified native blood-pool T1 and gender as optimal features, achieving higher correlation with true HCT (R=0.65) than linear regression (R=0.59) and strong agreement between synthetic and true ECV (R=0.95).
Impact:
This study shows that incorporating additional features in a DL model enhances HCT prediction from CMR data, eliminating the need for blood sampling. This advancement could streamline ECV measurement, making it more accessible for diagnosing myocardial diseases in clinical settings.

Quantification of Extracellular Volume Fraction in Cardiac MRI without Blood Sampling Using Multi-Stage Training Deep Learning

Zhuoan Li, Khalid Youssef, Mehdi Amian, Dilek M. Yalcinkaya, Venkateshwar Polsani, Michael Elliott, Rohan Dharmakumar, Robert Judd, Dipan J. Shah, Orlando P. Simonetti, Matthew S. Tong, Behzad Sharif

PURDUE UNIVERSITY
Weldon School of Biomedical Engineering

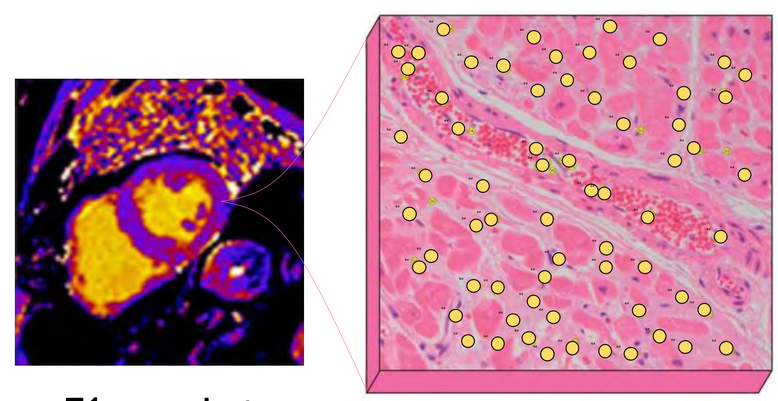
OVERVIEW

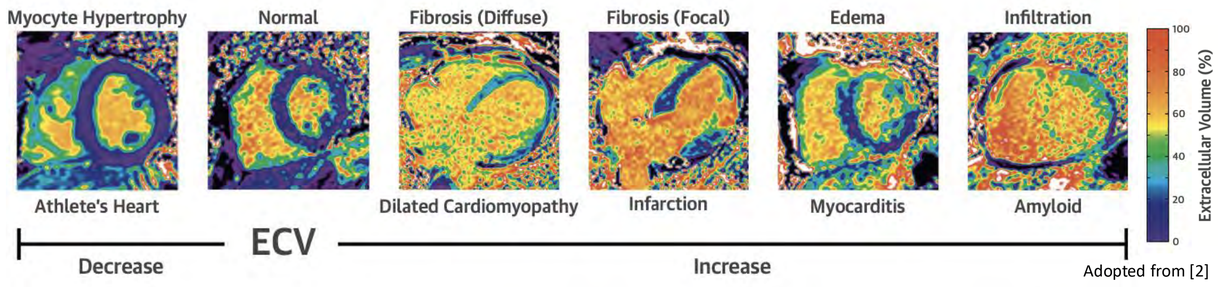
Motivation: Blood sampling for hematocrit (HCT) measurement limits clinical use of extracellular volume fraction (ECV) in diagnosing myocardial diseases.

Goals: Develop and evaluate a deep learning (DL) model to predict HCT from cardiovascular magnetic resonance (CMR) data, exploring if additional features improve predictability across multiple centers.

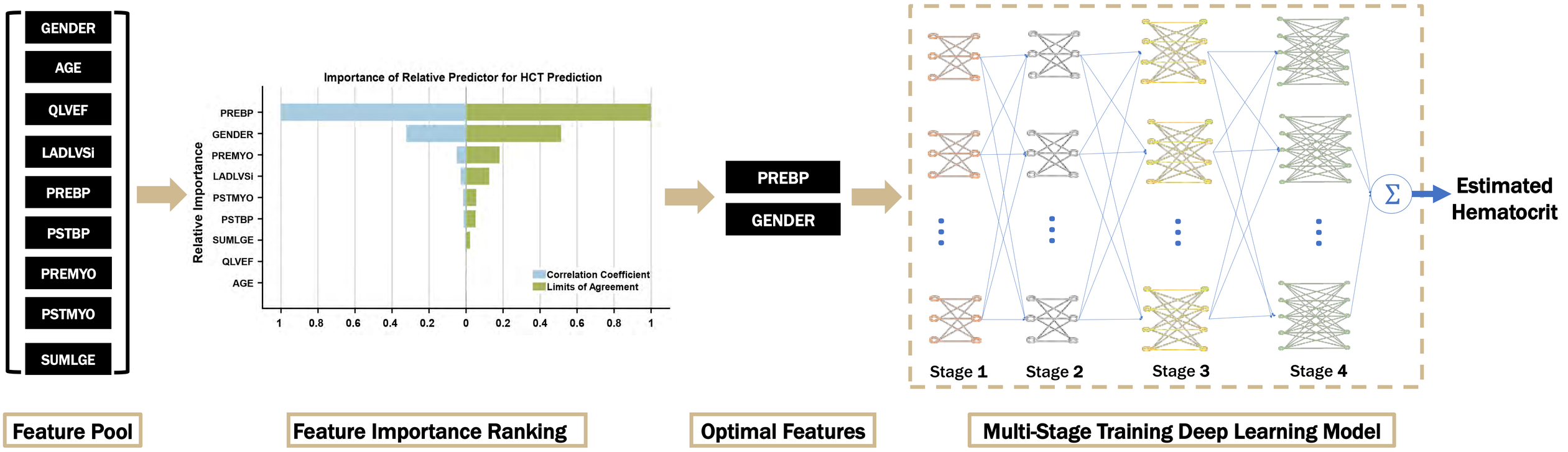
Approach: Trained a multi-stage DL model using multi-center CMR T1 values and clinical features to predict HCT without blood sampling.

Results: The DL model identified native blood-pool T1 and gender as optimal features, achieving higher correlation with true HCT (R=0.65) than linear regression (R=0.59) and strong agreement between synthetic and true ECV (R=0.95).

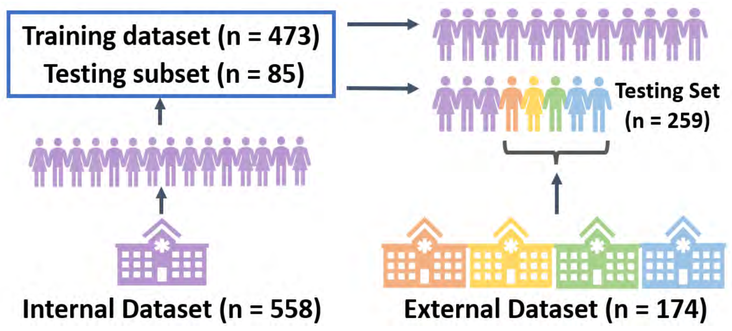




Methods: Proposed Pipeline



Multi-Center Dataset



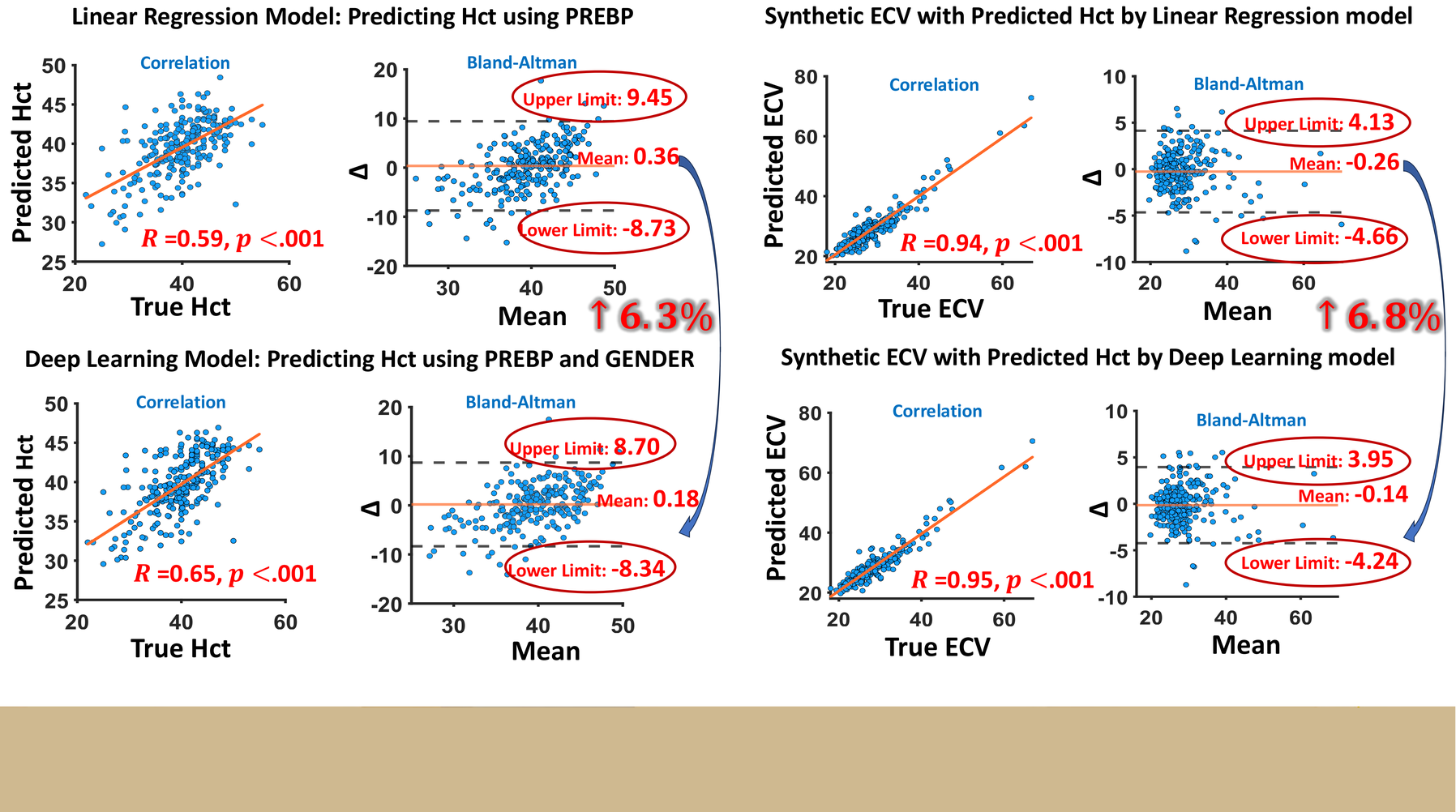
Validated across 5 centers:
- Feasible to estimate Hct from CMR without blood sampling.

First DL model:
- Applied deep learning for synthetic Hct.

Improved ECV precision:
- Tighter LOA than linear regression.

Preliminary Insights:
- Multi-feature models enhanced synthetic Hct.

Results



1. Treibel TA, et al. JACC Cardiovasc Imaging 2016
2. Robinson AA, et al. JACC Cardiovasc Imaging 2019
3. Youssef K et al. *Commun Earth Environ* 2023.
4. Tong MS et al. *J Cardiovasc Magn Reson* 2024

Poster Number 6

All of Us Data & Machine Learning

Netsanet Gebregziabher

Indiana CTSI, IU - IUB - IUI

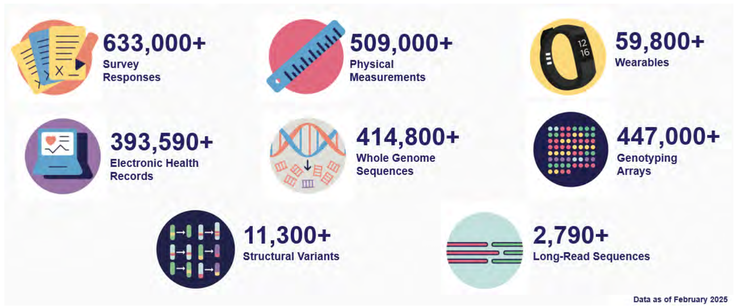
The All of Us Research Program is an initiative by the National Institute Health with the aim of building one of the largest biomedical data resources by enrolling one million participants. By gathering information on a large cohort of participants, a wide range of conditions and inclusion
of diverse participants, the program aims to advance precision medicine. The database currently has 413,457+ participants.
â€The All of Us dataset can help researchers develop complex models, powered by data from a diverse pool of participants.â€
The All of Us dataset empowers AIâ€driven health research through:
Multimodal data: EHRs, genomics, surveys, and wearables
Diverse cohort: Facilitates inclusive, biasâ€aware models
Scalable platform: Cloudâ€native tools for largeâ€scale AI training
The data management and analysis team at Indiana University supports researchers at institutions
in partnership with Indiana CTSI, including IU, Purdue, and Notre Dame.

All of Us Data & Machine Learning

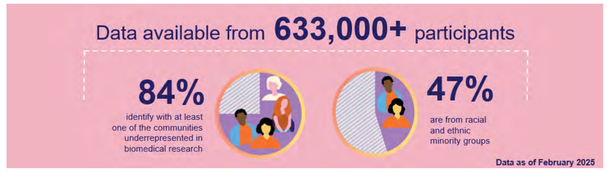
What Is It?

The All of Us Research Program is an initiative by the National Institute Health with the aim of building one of the largest biomedical data resources by enrolling one million participants. By gathering information on a large cohort of participants, a wide range of conditions and inclusion of diverse participants, the program aims to advance precision medicine. The database currently has 413,457+ participants.

- A large cohort of participants.
- A wide range of conditions.
- Inclusion of diverse participants.
- Aims to advance precision medicine.



Diversity



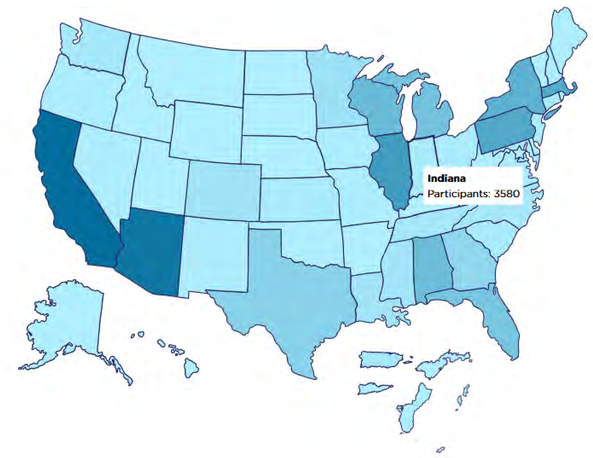
All of Us Research Hub Data Access Types

Public Tier: Aggregate data with identifiers removed. Available to anyone.

Registered Tier: individual-level data, available only to approved researchers on the **Researcher Workbench**. EHR, wearables, and surveys, as well as physical measurements.

Controlled Tier: registered tier + genomics data. Available only to approved researchers.

Distribution of All of Us Participants: National Level and Indiana



All of Us + AI: Unlocking New Possibilities

"The All of Us dataset can help researchers develop complex models, powered by data from a diverse pool of more than 413,000 participants."

The All of Us dataset empowers AI-driven health research through:

- **Multimodal data**: EHRs, genomics, surveys, and wearables
- **Diverse cohort**: Facilitates inclusive, bias-aware models
- **Scalable platform**: Cloud-native tools for large-scale AI training

The data management and analysis team at Indiana University supports researchers at institutions in partnership with Indiana CTSI, including IU, Purdue, and Notre Dame.

Supporting AI Research with All of Us

- **Infrastructure setup**: cloud compute environments and workbench onboarding
- **Data preparation**: cohort building, variable selection, data preprocessing
- **Modeling workflows**: classical ML, NLP, deep learning
- **Evaluation guidance**: bias mitigation, subgroup performance, fairness auditing

AI Research Leveraging All of Us

RCATE Framework for Treatment Effects: Researchers developed a unified, robust framework for estimating heterogeneous treatment effects using EHR data. By converting the problem to a supervised learning task and applying models like gradient boosting, random forests, and neural networks, the approach improved accuracy under nonlinearity and data irregularities. *Li R, Wang H, Zhao Y, Su J, Tu W. (2021). Communications in Statistics—Simulation and Computation.* [DOI: 10.1080/03610918.2021.1974883]

iPsRS Model for Social Risk in Diabetes: A machine learning pipeline was developed to identify T2D patients at high risk of hospitalization based on individualized polysocial risk scores (iPsRS). The model incorporates explainable AI and fairness optimization across race and ethnicity. *Huang Y, Guo J, Donahoo W, et al. (2024). Nature Communications, 15, 8653.* [DOI: 10.1038/s41467-024-52960-9]

Social Determinants and Acidosis Risk: Using a matched case-control study of 13,310 participants, this JAMIA study examined how social determinants—such as income, housing instability, and insurance type—impact the risk of severe acidosis. Results showed that these SDoHs contribute significantly, even after adjusting for demographics and clinical factors. *Gatz AE, Xiong C, Nguyen CM, et al. (2024). JAMIA, 31(12):2932–2939.* [DOI: 10.1093/jamia/ocae256]

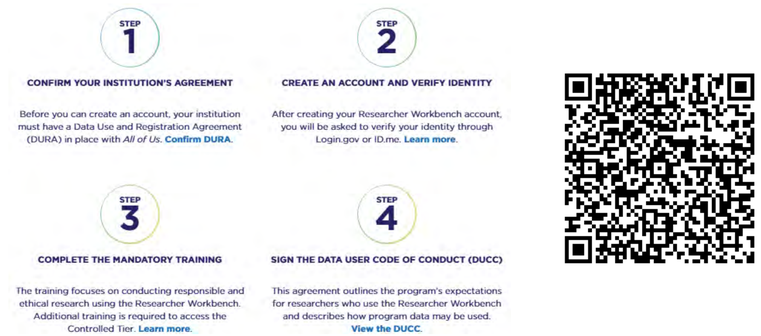
Cross-Dataset Comparison for Policy Research: Researchers compared All of Us with the Medical Expenditure Panel Survey (MEPS) to assess data completeness for health care access, cost, and demographic indicators. The study informs downstream applications of AoU data in health economics and policy analysis. *Shone H, Simon K, Ziedan E. (In progress).*

Conclusion

All of Us is a cornerstone for ethical, inclusive, and impactful AI research in healthcare.

- Enables deep investigation of multimodal health data
- Supports scalable and reproducible machine learning research
- Offers a robust foundation for fairness, generalizability, and real-world deployment

How to Register



Disclaimer: This poster was made with support from the Indiana Clinical and Translational Sciences Institute, funded in part by award UL1TR002529 from the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIH.

BIOSTATISTICS AND HEALTH DATA SCIENCE
SCHOOL OF MEDICINE
RICHARD M. FAIRBANKS SCHOOL OF PUBLIC HEALTH

Irsay Institute
INDIANA UNIVERSITY

INDIANA CTSI Clinical and Translational Sciences Institute
PURDUE UNIVERSITY Engineering Medicine
INDIANA UNIVERSITY SCHOOL OF MEDICINE

Poster Number 7

Deep Learning-enabled Approach to Eliminate Motion-induced Dark-rim Artifacts in Stress First-pass Perfusion Cardiac MRI

Hazar Benan Unal

Weldon School of Biomedical Engineering, Purdue University

Introduction: The subendocardial dark-rim artifact (DRA) remains an important challenge in the routine clinical use of stress first-pass perfusion (FPP) cardiac MRI since it reduces the diagnostic accuracy by mimicking the appearance of subtle perfusion defects. DRA is primarily induced by Gibbs ringing and cardiac motion. A method to optimize the image reconstruction (â€œreconâ€□) pipeline to automatically eliminate or minimize motion-induced DRAs for routine Cartesian-sampled studies is lacking. To this end, we propose a deep-learning-enabled method with validation against invasive testing and inline implementation.
Methods: In Step 1, conventionally acquired accelerated stress FPP data is first reconstructed into a fully encoded k-space. Step 2 performs two different recons by first applying a one-sided temporal footprint (TF) reduction (Fig. 1) to modify the effective TF of the acquired data followed by partial Fourier recon. In Step 3, a spatiotemporal deep neural network is employed to segment the free-breathing time frames and to extract the segment-wise signal intensity (S.I.) for both the left- and right-sided recons. Next, the slope of the line connecting the left- vs right-sided S.I. for each segment is computed to generate a bullâ€™s eye map (endo-epi sectors) of dark-rim TF variability. Finally, the TF variability is processed to automatically decide the optimal TF truncation for recon of DRA-suppressed images series for each slice. The method was validated on a retrospective cohort of patients (n=10) with normal invasive testing (no stenosis >20% diameter, coronary flow reserve > 2.5, negative LGE). The performance was evaluated based on a visual 0-to-4 DRA severity score (0: no artifact). Further, the feasibility of inline implementation (Siemens FIRE framework) was prospectively tested (n=8).
Results: Fig. 2A shows representative results where the scanner default recon has noticeable subendocardial DRA that can affect the assessment given the thin transmural extent; however, the automatically generated optimal recon (green boxes) effectively eliminates the motion-induced DRA by selecting the optimal TF truncation approach. Fig. 2B shows inline implementation example demonstrating feasibility in clinical routine. Fig. 3 summarizes overall performance for default vs proposed approach demonstrating significant improvement in DRA severity (2.2Â±0.9 vs 1.5Â±0.7, $p < 0.005$) and DRA prevalence (47% vs 7%, $p < 0.01$).
Discussion and Conclusions: The proposed deep learning-enabled recon method for stress FPP studies results in significant reduction in DRA severity based on a rigorously selected evaluation. Coupled with inline implementation, this technique may provide an easily deployable DRA-reduction technique in routine clinical studies.

MACHINE LEARNING IN MEDICINE 2025

Deep Learning-enabled Approach to Eliminate Motion-induced Dark-rim Artifacts in Stress First-pass Perfusion Cardiac MRI

Hazar Benan Unal, PhD [1,2], Khalid Youssef, PhD [2], Abdul Ahmed, PhD [3], Kelvin Chow, PhD [3], Xiaoming Bi, PhD [3], Luis F. Zamudio, MSc [2], Dilek Mirgun Yalcinkaya, PhD [1,2], Ronald Mastouri, MD [2], Janet Wei, MD [4], C. Noel Bairey Merz, MD [4], Rohan Dharmakumar, PhD [2], Behzad Sharif, PhD [1,2]

[1]Laboratory for Translational Imaging of Microcirculation, Purdue University, Indianapolis, IN. [2]Indiana University School of Medicine, Indianapolis, IN. [3]Siemens Medical Solutions, USA. [4]Cedars-Sinai Medical Center, Los Angeles, CA.

PURDUE UNIVERSITY

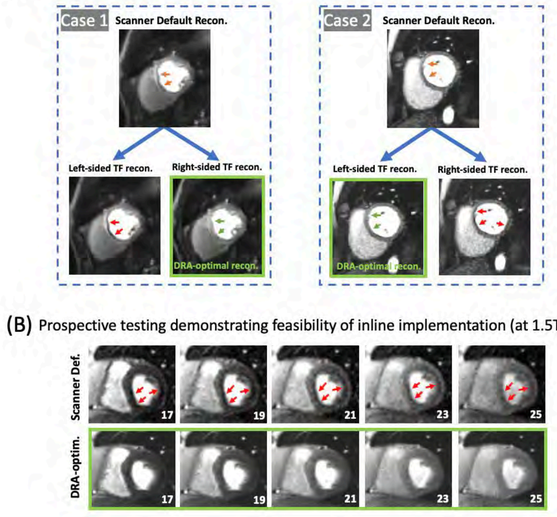
Weldon School of Biomedical Engineering

TIM Lab

OVERVIEW

Dark-rim artifact (DRA) mimics perfusion defects in myocardial first-pass perfusion (FPP) cardiac magnetic resonance imaging (MRI), affecting the diagnostic accuracy of detecting ischemia. We propose a new technique to automatically eliminate DRA that is enabled by myocardial segmentation with 2D+time U-net based deep neural network. Our results showed significant reduction in severity and presence of DRA in rigorously selected healthy dataset.

Fig 2: Representative cases



INTRODUCTION

- ❑ Dark-rim artifact (DRA) remains an important challenge in stress first-pass perfusion CMRI.
- ❑ It mimics the appearance of perfusion defects and reduces diagnostic accuracy.
- ❑ DRA is primarily induced by Gibbs ringing and cardiac motion.[1,2]
- ❑ Several methods were proposed to suppress Gibbs-ringing-induced DRA. [4-6]
- ❑ However, a method to automatically eliminate motion-induced DRA for routine clinical studies is lacking.

- ❖ We propose a deep learning-enabled method with validation against invasive testing and show the feasibility of inline implementation.

METHODS

Step 1: Conventionally acquired accelerated stress FPP data is first reconstructed into a fully encoded k-space (e.g., using the default scanner recon routine).

Step 2: Perform two different recons by first applying a one-sided (left or right) temporal footprint (TF) reduction (Fig. 1) to modify the effective TF of the acquired data followed by partial Fourier recon.

Step 3: Spatiotemporal deep neural network is employed to segment the free-breathing time frames and to extract the segment-wise signal intensity (S.I.) for both the left- and right-sided recons. Next, the slope of the line connecting the left- vs right-sided S.I. for each segment is computed to generate a bull's eye map (endo-epi sectors) of dark-rim TF variability.[7]

Finally, the TF variability is thresholded and combined with a "majority voting" approach to automatically decide the optimal TF truncation (i.e., left vs right) for recon of DRA-suppressed images series for each slice position.

- ❖ The method was validated on a retrospective cohort of patients with normal coronary angiograms (no stenosis >20% diameter) and normal invasive coronary function testing (normal microvascular reactivity including endothelial function and coronary flow reserve > 2.5) and negative LGE (n=10 stress studies).
- ❖ In this "truly normal" cohort, the performance was evaluated based on a visual 0-to-4 DRA severity score (0: no artifact). Further, the feasibility of inline implementation (Siemens FIRE framework)[8] was prospectively tested (n=8 stress studies).

Fig 1: Proposed methodology

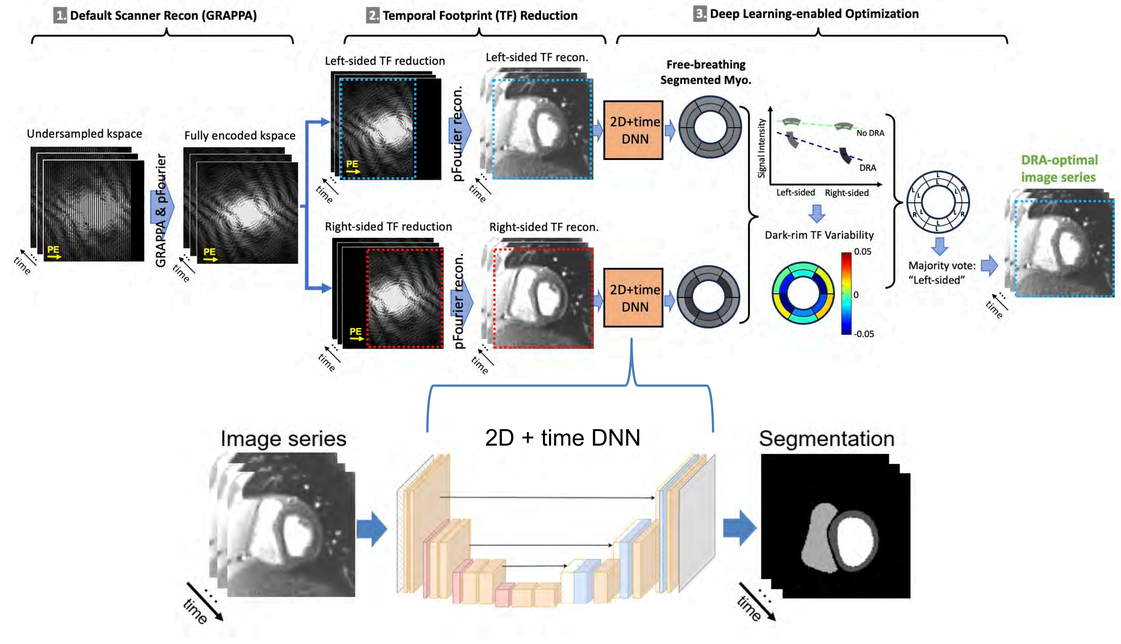
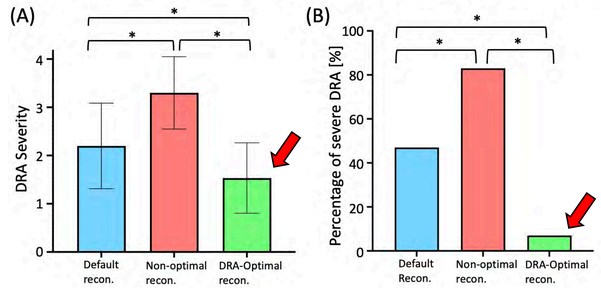


Fig 3: Overall performance



RESULTS

- ➢ **Fig. 2A**: Truly normal patient studies. In both cases, the scanner default recon has noticeable subendocardial DRA that can affect the assessment given the thin transmural extent (diastolic frame); however, the automatically generated optimal recon (green boxes) effectively eliminates the motion-induced DRA by selecting the optimal TF truncation approach.
- ➢ **Fig. 2B**: Inline implementation examples. Prospective testing at 1.5T scanner demonstrates the feasibility of proposed method in routine clinical studies.
- ➢ **Fig. 3**: Performance of default vs proposed approach. Proposed method shows significant improvement in DRA severity (2.2±0.9 vs 1.5±0.7, $p < 0.005$) and DRA prevalence (47% vs 7%, $p < 0.01$).

CONCLUSIONS

The proposed deep learning-enabled recon method for stress FPP studies results in significant reduction in DRA severity based on a rigorously selected evaluation. Coupled with inline implementation, this technique may provide an easily deployable DRA-reduction technique in routine clinical studies, both with standard and high-resolution perfusion CMR protocols.[9,10]

REFERENCES

1. Di Bella EV et al. Magn Reson Med. 2005;54:1295-1299. doi: 10.1002/mrm.20666
2. Storey P et al. Magn Reson Med. 2002;48:1028-1036. doi: 10.1002/mrm.10314
3. Kellner E et al. Magn Reson Med. 2016;76:1574-1581. doi: 10.1002/mrm.26054
4. Zhou Z et al. J Magn Reson Imaging. 2017:542-555. doi: 10.1002/jmri.25400
5. Salerno M et al. Magn Reson Med. 2013;70:1369-1379. doi: 10.1002/mrm.24569
6. Sharif B et al. Magn Reson Med. 2014;72:124-136. doi: 10.1002/mrm.24913
7. Unal HB et al. Proc IEEE EMBC 2021;2021:4079-4085. doi: 10.1109/EMBC46164.2021.9630270
8. Chow K et al. SCMR. Virtual Scientific Sessions.; 2021:838972.
9. Kellman P et al. J Cardiovasc Magn Reson 2017;19:43. doi: 10.1186/s12968-017-0355-5
10. Scannell CM et al. Eur Heart J Imaging Methods Pract 2024;2(1):qyae001. doi: 10.1093/ehjimp/qyae001

Poster Number 8

A zebrafish functional screen identifies hits from FDA-approved drugs for treating retinitis pigmentosa

Beichen Wang

Biological Sciences, Purdue University

Retinitis Pigmentosa (RP) is a genetically inherited retinal degeneration that leads to the death of rod photoreceptors and loss of vision. RP can be caused by mutations in phototransduction genes, including RHODOPSIN (RHO). One RHO mutation, Q344X, can cause a severe form of RP but has no treatment. Hence, we aimed to discover new drugs for Q344X RP by repurposing FDA-approved drugs, which is faster and cheaper than de novo drug discovery. Thus, we screened an FDA-approved drug (FDA) library with 1430 drugs using a transgenic Q344X RP zebrafish model. The screening utilized the visual-motor response (VMR), a startle response to drastic light changes in zebrafish. We previously detected a reduced VMR in Q344X larvae during light offset (light-off), providing a functional readout for drug effect. We aimed to identify drugs that induced Q344X to display a light-off VMR similar to that of the WT controls, theorizing that these drugs may improve the vision of the Q334X mutant. To this end, we first exposed Q344X larvae to 10 ÂµM drugs from 5 to 7 days post-fertilization (dpf). We began the treatment at 5 dpf because Q344X rods begin to degenerate substantially at this stage. We ran the VMR at 7 dpf because untreated Q344X larvae displayed a significantly reduced light-off VMR compared to WT. Using this treatment period would maximize our chance to identify drugs for Q344X RP. We eliminated 191 toxic drugs (13.4%) and then screened 1239 (86.6%) non-toxic drugs by VMR. The positive and negative controls were drug-carrier-treated WT and Q344X, respectively. When analyzing the VMR data, we used an unbiased approach to detect similarities in VMR profiles by unsupervised machine learning. We combined the results of three clustering algorithms: hierarchical, K-means, and Gaussian Mixed Model with Expectation Maximization (EM-GMM). Since each algorithm has unique strengths in detecting similarities, combining their results will maximize our chance to discover drugs that induced a WT-like behavior in Q344X larvae. Out of the 1239 non-toxic drugs, 26 hits were identified by at least one algorithm (positive rate = 2.1%). One hit was identified by all three algorithms. The Q344X retina treated by this hit also contained more rods than the Q344X controls in the histology. In summary, we identified functional hits for Q344X RP from the FDA library. These hits can be further developed and repurposed for human Q344X RP treatment.

A zebrafish functional screen identifies hits from FDA-approved drugs for treating retinitis pigmentosa

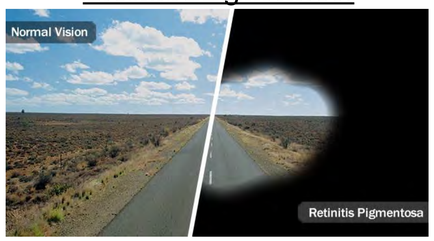
Beichen Wang[1], Logan Ganzen[1,2*], Motokazu Tsujikawa[3], Yuk Fai Leung[1,4,5]

[1]Department of Biological Sciences, Purdue University, West Lafayette, IN
[2]Purdue Interdisciplinary Life Sciences Program, Purdue University, West Lafayette, IN
[3]Department of Ophthalmology, Graduate School of Medicine, Osaka University, Osaka, Japan
[4]Purdue Institute for Drug Discovery, Purdue University, West Lafayette, IN
[5]Purdue Institute of Integrative Neuroscience, Purdue University, West Lafayette, IN
*Current address: Department of Molecular & Cell Biology, University of California, Berkeley, Berkeley, CA

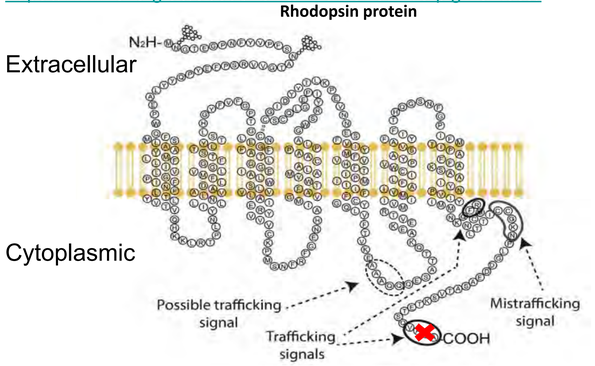
Contact Email: wang4537@purdue.edu
LinkedIn: www.linkedin.com/in/bwang258

Background

Retinitis Pigmentosa



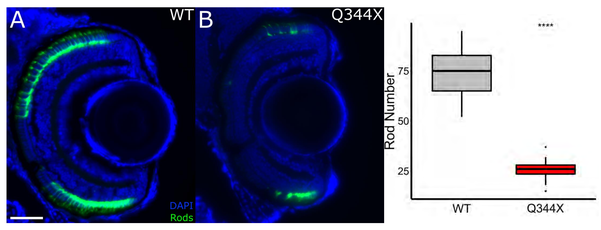
Retinitis pigmentosa (RP) is a form of retinal degeneration. RP patients suffer from tunnel vision and severe vision impairment. RP can be caused by various gene mutations, including those on the *RHO* gene. However, there are no treatments for RP caused by *RHO* mutations. Figure adopted from the Lions Eye Institute: https://www.lei.org.au/about/news/what-is-retinitis-pigmentosa/



To discover treatments for RP caused by *RHO* mutations, we focused on one mutation, Q344X. In the Q344X mutation, the codon encoding glutamine (Q) at the 344th amino acid location mutates to the premature stop codon (X). This mutation disrupts the trafficking signal at the C-terminal of the rhodopsin protein. Consequently, rhodopsin is mislocalized, leading to rod degeneration. Figure adapted from Nemet et al (2015).

Methods

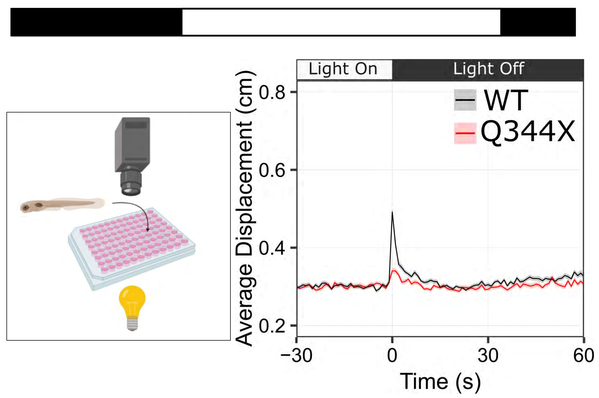
The Q344X RP Zebrafish Model



To identify potential drugs for treating Q344X RP, we used a transgenic zebrafish model expressing the human *RHODOPSIN* gene with Q344X mutation: *Tg(rho:Hsa.RH1_Q344X)* (Nakao et al., 2013). We conducted the number of rods in the Q344X and WT on their retinal cryosection. The rods were labeled by a rod GFP reporter line, *Tg(-3.7rho:EGFP)*. The Q344X showed a significantly reduced number of rods in the retina at 7 days post fertilization (dpf) compared to WT. (Welch two-sample test; t = 14.945; df = 16.989; p-value = 3.315×10^{-11}. Sample size: WT, n = 16; Q344X, n = 25). Scale bar = 50 μm. Slice thickness = 10 μm.

Methods (cont.)

Visual-Motor Response (VMR) Assay

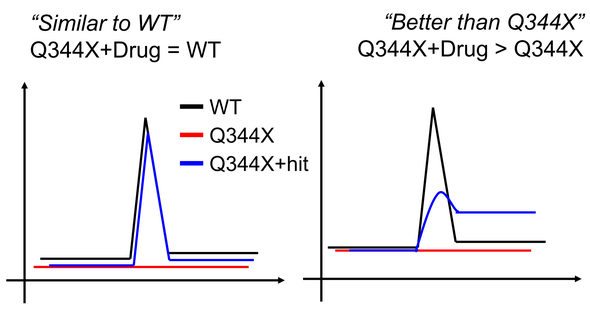


To screen drugs in Q344X larvae, we used a VMR assay in which we placed larvae in 96-well plates and recorded their behaviors toward light-off stimulation. To elicit rod-specific responses, the light intensity was tuned for dim light stimulation(Venkatraman et al., 2020). Light intensity = 0.00389 μW/cm². The 7-dpf Q344X larvae showed a reduced average displacement upon light offset compared to the WT, particularly in the first second. The ribbon represents the standard error of the mean (SEM). The sample size was 48 larvae per genotype per biological replicate. Eighteen biological replicates were collected.

Screen a SelleckChem FDA-approved Compound Library

Using the VMR assay, we screened a SelleckChem FDA-approved compound library. We exposed each drug at 10 μM to 24 larvae from 5 dpf to 7 dpf and eliminated 191 toxic compounds (13.4%). We screened the remaining 1239 non-toxic compounds with the same treatment scheme and measured their VMR at 7 dpf. Positive and negative controls were WT and Q344X larvae treated with the same amount of drug carrier (water or DMSO), respectively. Each drug was tested once.

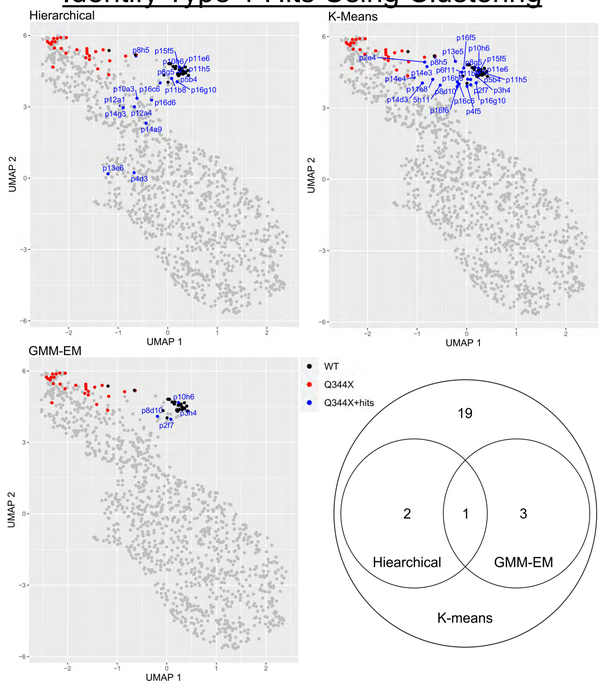
Type 1Hits and Type 2 Hits



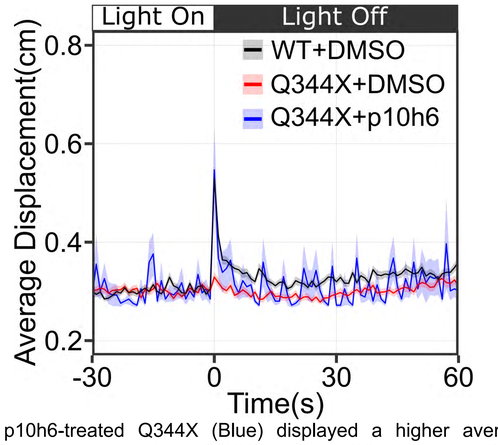
In the screening, we aimed to identify two types of hits: 1) Type 1 hits—drugs that increased the Q344X VMR to the WT level after the treatment; 2) Type 2 hits—drugs that increased the Q344X VMR to some extent after the treatment, particularly in the first second after light offset. Based on these two goals, we applied two distinct analysis strategies, clustering and hypothesis testing, respectively.

Results

Identify Type 1 Hits Using Clustering



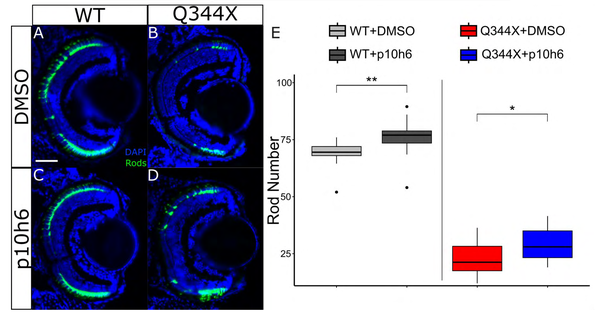
To identify drugs eliciting WT-like VMR in the Q344X mutants, we used three clustering algorithms: Hierarchical, K-Means, and Gaussian Mixture Model with Expectation Maximization (GMM-EM) clustering to analyze the data. The input data used the VMR profiles for 30 seconds after light offset. A drug was defined as a hit when the VMR of the Q344X treated by this drug was grouped into the WT cluster. The results were visualized by Uniform Manifold Approximation and Projection (UMAP). Q344X VMR (Red) was distinct from the WT VMR (Black). Twenty-five hits were identified to induce "WT-like" VMR in Q344X (Blue) by at least one clustering algorithm. One hit, p10h6, was identified by all three algorithms.



The p10h6-treated Q344X (Blue) displayed a higher average displacement than the DMSO-treated Q344X did (Red), particularly in the first second after the light offset. The average displacement trace of p10h6-treated Q344X (Blue) was similar to that of DMSO-treated WT (black). The ribbon represents the standard error of mean (SEM). Sample size: WT+DMSO, n = 432; Q344X+DMSO, n = 432; Q344X+p10h6, n = 24.

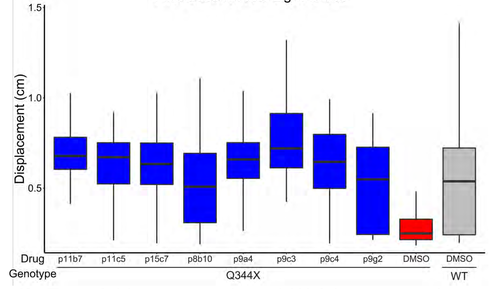
Results (cont.)

Effect of p10h6 on Rod Number



The effect of p10h6 on rod number was analyzed by histology. We collected 10-μm cryosections of retinas in different treatment groups and counted the number of rods labeled by the rod GFP reporter. The number of rods in the p10h6-treated Q344X (Blue) retina was significantly higher than that in the DMSO-treated Q344X (Red). (Welch two-sample test. Sample size: WT+DMSO, n = 11; Q344X+DMSO, n = 14; Q344X+p10h6, n = 18; WT+p10h6, n = 15). Significant level: *, p-value < 0.05; **, p-value < 0.01. Scale bar = 50 μm.

Identify Type 2 Hits Using Hypothesis Testing



To identify Type 2 hits, we conducted Welch two-sample tests to compare the mean displacements in the first second between the drug-treated and drug-carrier-treated Q344X. Eight drug-treated Q344X mutants (Blue) had significantly higher mean displacements than the DMSO-treated Q344X group (Red) in the first second after light offset (Bonferroni adjusted p-value < 0.05).

Conclusions

- Our screen identified 34 positive hits for RP. The positive rate was 2.01%.
- We characterized one Type 1 hit, p10h6, by histology. The WT and Q344X larvae had more rods after exposure to the p10h6.

Acknowledgments

Beichen Wang, Logan Ganzen, and Yuk Fai Leung were partially supported by research grants from Purdue Institute for Drug Discovery. Beichen Wang and Yuk Fai Leung were partially supported by the Clinical and Translational Science Pilot Grant Program from Indiana Clinical and Translational Sciences Institute. Logan Ganzen was supported by Grant Numbers TL1 TR001107 and UL1 TR001108 (A. Shekhar, PI) from the National Institutes of Health, National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award. Yuk Fai Leung was partially supported by grants from the Purdue Research Foundation and the International Retinal Research Foundation.

References

1. Nemet et al. *Prog. Mol. Biol. Transl. Sci.* **132**, 39–71 (2015).
2. Nakao et al. *PLoS One* **7**, e32472 (2012).
3. Venkatraman et al. *Invest. Ophthalmol. Vis. Sci.* **61**, 11 (2020).

Poster Number 9

Test-Time Model Selection for Robust Deep Learning-Based Analysis of Multi-Center Myocardial Perfusion MRI

Arian Mollajafari Sohi

Biomedical Engineering, Purdue University

Introduction. Quantitative analysis of first-pass perfusion (FPP) cardiac magnetic resonance (CMR) imaging is essential for evaluating myocardial ischemia. Deep neural networks (DNNs) have enabled automatic segmentation of myocardial tissue, a prerequisite for perfusion quantification. However, deploying these models across multi-center datasets remains challenging due to domain shifts caused by differences in scanner vendors, magnetic field strengths, and acquisition protocols. Recent approaches have leveraged uncertainty quantification to guide test-time model selection from an ensemble of DNNs. Such methods can yield subtle yet clinically significant segmentation errors that compromise perfusion estimates. This work introduces a two-stage model selection framework that complements uncertainty ranking with a novel kinetics-based selection step.

Methods. We trained an ensemble of 50 DNN models using an internal motion-corrected dataset comprising 80 stress/rest FPP cardiac MR studies. For external validation, we used data from 120 patients drawn from four institutions within the SCMR registry. All models had the same spatiotemporal patch-wise U-Net architecture but were initialized with different random seeds. For every test case, each DNN generated a candidate segmentation. Two state-of-the-art (SOTA) uncertainty quantification methods were independently applied to prune the top-10 most certain outputs. Signal intensity curves were extracted from the myocardial regions defined by each candidate mask. The kinetics-based selection stage evaluate these curves, identifies outlier pixels with implausible behavior, and computes a composite error score for each mask. The final mask was chosen as the one with the lowest score. We compared the performance of the two-stage selection method against standalone uncertainty-based selection.

Results. The two-stage selection approaches outperformed uncertainty-based methods across multiple criteria. Using semi-quantitative perfusion analysis, we assessed the myocardial perfusion reserve index (MPRI) in 12 myocardial subsegments. Our method achieved a significantly higher correlation with manual MPRIs ($p < 0.05$) compared to uncertainty-based selection in regions prone to error. Furthermore, the proposed approach reduced segmentation failuresâ€”defined as blood pool inclusion, non-contiguous masks, or the inclusion of non-cardiac pixels by over 75%. Specifically, failure rates decreased from 6.0% and 7.0% (for the two uncertainty methods) to 1.3% and 1.6%, respectively ($p < 0.0005$).

Conclusion. We present a two-stage model selection framework that combines uncertainty-based pruning with kinetics-based selection to improve segmentation robustness in multi-center perfusion MRI. We incorporated myocardial signal behavior into the selection process to enhance reliability compared to uncertainty-only methods. The framework is broadly applicable to other AI tasks that require signal-based validation for quality control.

Test-Time Model Selection for Robust Deep Learning-Based Analysis of Multi-Center Myocardial Perfusion MRI

Arian M. Sohi[1], Dilek M. Yalcinkaya[1], Khalid Youssef[2], Luis Zamudio[2], Michael D. Elliott[3], Venkateshwar Polsani[4], Rohan Dharmakumar[2], Robert M. Judd[5], Matthew S. Tong[6], Dipan J. Shah[7], Orlando P. Simonetti[6], Behzad Sharif[1,2]

[1]Purdue University · [2]Indiana University · [3]Atrium Health · [4]Piedmont Heart · [5]Duke University · [6]Ohio State University · [7]Houston Methodist

PURDUE UNIVERSITY
Weldon School of Biomedical Engineering

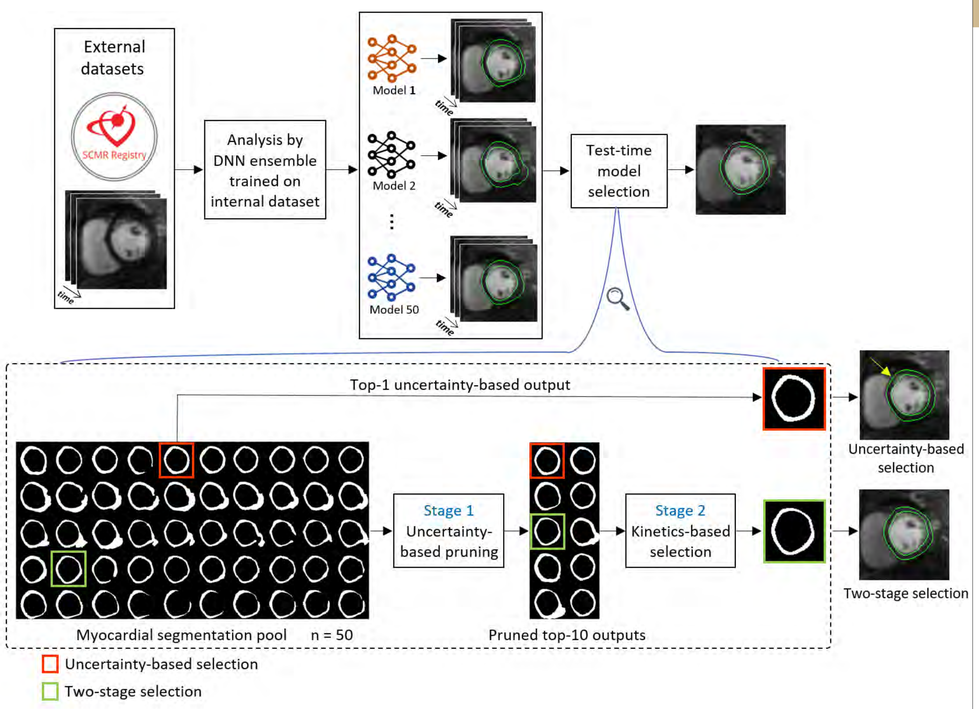


Fig 1. Two-stage DNN model selection: uncertainty pruning + signal-based selection

Introduction

- Perfusion cardiac MRI measures blood flow in the heart muscle after contrast injection
- Quantitative analysis requires accurate segmentation of the myocardium [1]
- Deep neural networks (DNNs) automate segmentation but struggle with multi-center data [2]
- Domain shifts from different scanners, field strength, and protocols reduce model reliability

Dataset		No. of subjects	Scanner Platform	No. of females	Age (years)	BMI (kg/m²)	Healthy Controls (%)
Internal Dataset	Train	n = 60	Siemens Verio (3T)	54 (90%)	56.8 ± 11.8	27.0 ± 5.4	12%
	Validation	n = 10	Siemens Verio (3T)	9 (90%)	55.8 ± 10.7	25.7 ± 4.7	10%
	Test	n = 10	Siemens Vida (3T)	9 (90%)	62.7 ± 7.1	32.0 ± 9.8	0%
External Datasets	Site #1	n = 50	Siemens Avanto (1.5T)	24 (44%)	51.6 ± 16.2	29.7 ± 6.0	0%
	Site #2	n = 30	Siemens Verio (3T)	12 (41%)	49.7 ± 14.4	30.7 ± 8.2	0%
SCMR Registry	Site #3	n = 20	Siemens Sola (1.5T)	5 (23%)	69.2 ± 9.0	30.8 ± 7.0	0%
	Site #4	n = 20	Siemens Avanto (1.5T)	8 (40%)	61.7 ± 9.9	30.5 ± 6.7	0%

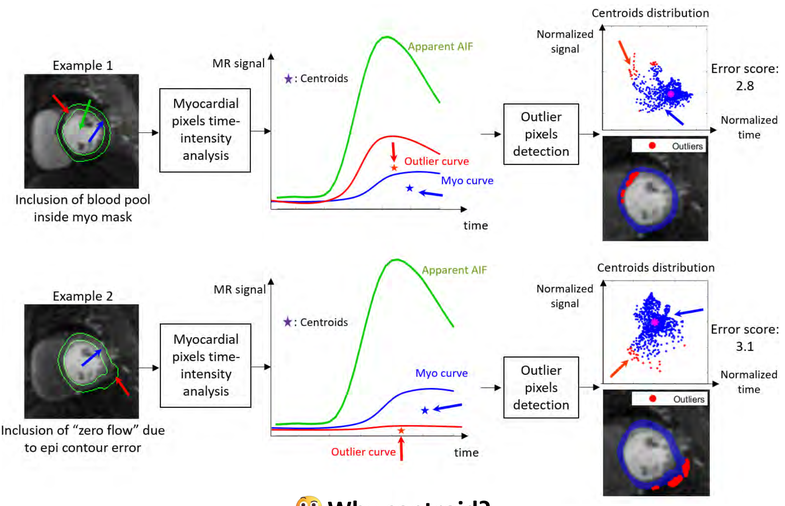
Table 1. summary of 200 multi-center stress/rest perfusion MRI studies

Methods

- Train a pool of 50 spatiotemporal U-Net models on internal, motion-corrected stress/rest perfusion cardiac MRI data
- At test time, 50 candidate outputs are generated for the selection pipeline (fig 1)
- Stage 1: Uncertainty-based pruning selects the top-10 outputs using two methods:
  - **DAUGS**: mean per-pixel uncertainty from patch-based predictions [2]
  - **QCD**: predicted segmentation quality based on ensemble agreement [3]
- Stage 2: Kinetics-based selection using myocardial time-intensity curve evaluation
- Assign an error score to each candidate and select the one with the lowest score
- Compare two-stage selection with standalone uncertainty-based methods

❓ What happens at Stage 2

It detects segmentation errors using myocardial signal curves.



🤔 Why centroid?

It captures the timing and strength of contrast uptake by myocardium — outliers deviate from normal myocardial behavior.

Results

- Performance evaluated using Myocardial Perfusion Reserve Index (MPRI), a semi-quantitative metric calculated as the stress-to-rest ratio of signal enhancement across 12 myocardial subsegments.
- Validated on 120 multi-center stress/rest perfusion MRI studies (SCMR Registry)
- Two-stage methods show higher correlation (R) with manual MPRI with significant gains ($p<0.05$) in subsegments prone to error (fig 3):
  Two-stage DAUGS vs. DAUGS: 1, 4, 8, 9, 12 | Two-stage QCD vs. QCD: 1, 4, 5, 9, 10

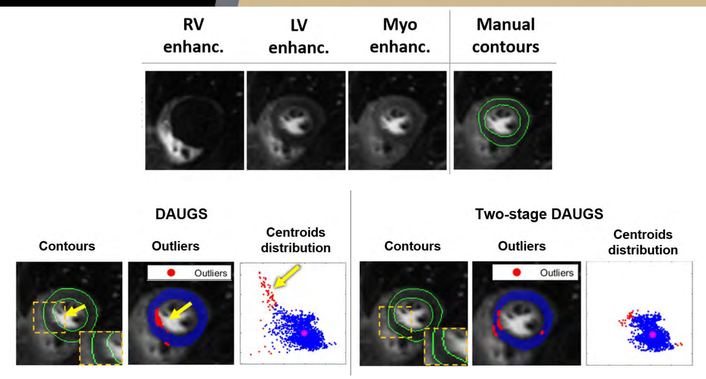


Fig 2. Example case: Comparing two methods on the same case

Conclusion

- Proposed a two-stage model selection framework combining uncertainty-based pruning with kinetics-based selection.
- Improves robustness of DNN-based segmentation across multi center perfusion MRI datasets.
- Framework generalizable to other AI applications requiring signal based validation for quality control.

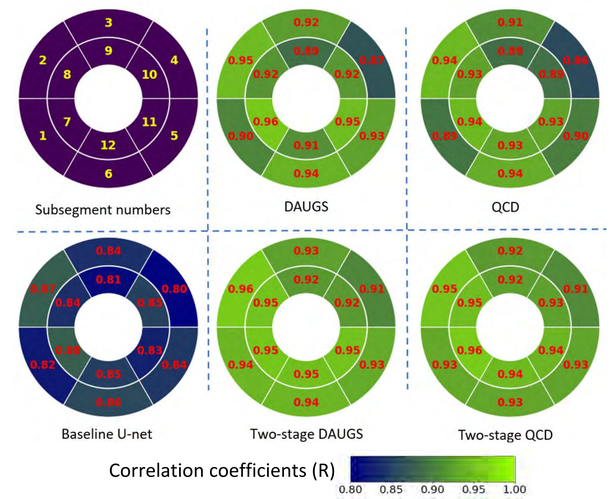


Fig 3. Correlation with expert-derived MPRI

References

1. Scannell CM et al. JMRI 2020, doi: 10.1002/jmri.26983
2. Yalcinkaya et al. JCMR 2024, doi: 10.1016/j.jocmr.2024.101082
3. Hann et al. MEDIA 2020, doi: 10.1016/j.media.2021.102029

Poster Number 10

The Data Mine: Experiential Industry Practicums in Data Science

Maggie Betz

The Data Mine, Purdue University

The Data Mine, currently in its seventh year, enables more than 2000 interdisciplinary graduate and undergraduate students with hands-on experience in data science. Based on the principles of learning by doing, teamwork, and real-world data science, this model is used by learners at more than 60 colleges annually. Additionally, The Data Mine enables approximately 100 projects with Corporate Partners across many types of domains, including aerospace, agriculture, manufacturing, pharmaceutical science, etc. This model has proven to be a very effective method for colleges and companies to quickly and easily build relationships that create genuine value for partners and students alike. The newest Data Mine location in Indianapolis is a successful example of this model to rapidly scale and return a strong institutional investment. This poster will briefly explain why The Data Mine has become pervasive as a model for data science research across institutions of varying profiles.

The Data Mine

Experiential Industry Practicums in Data Science

ABOUT THE DATA MINE

- Residential, experiential learning community
- **Data Science for All**
- By the numbers:
  - **133 majors** represented
  - **2,100 students** in AY 2024-25
    - ~150 students at MSIs
    - ~100 graduate students
  - 130+ teaching assistants
  - Launched in Indianapolis fall 2024
  - **50+ unique partners in industry**

SEMINAR (1 CREDIT)

- 1 credit hour per semester
- 8 levels (4 years)
- All content in The Examples Book
  - Publicly available at the-examples-book.com
- R, Python, SQL, bash, HPC, TB of data

CORPORATE PARTNERS

- **Undergraduate and graduate** students from **all majors and computing backgrounds** work in teams directly with partners in industry
- Teams of 5-15 students collaborate on data intensive projects via experiential learning
  - *"We're designing something our Corporate Partner will use! And we get to name it! It's a unique opportunity and I am grateful for the experience of working on something [company] will likely use often." ~student*
- Students are **mentored by industry professionals** during weekly team meetings
- Teams are led by a paid team leader (typically an undergraduate student)
- Students earn **3 credit hours** per semester and spend 8-10 hours per week working on the projects
- Projects last the **entire academic year (August – May)**
- More than 80 projects in AY 2024-25 covering many domains:
  - Digital agriculture
  - Aerospace
  - Manufacturing
  - Life sciences
  - Technology
  - Finance
- All projects conclude with a poster presentation
- ***"It's literally like working with colleagues when working with you [students] – your ideas are so great, your thought processes are so wise." ~Corporate Partner Mentor***

Students learn about Agile project management via Legos during the first lab with a visit from President Mung Chiang (Fall 2024)

Students have lunch with faculty and industry mentors (Fall 2024)

STUDENT QUOTES

*"The research project with The Data Mine **has no doubt been the highlight of my undergraduate career** here at Purdue." ~senior*

*"The opportunity to work on **real-world** projects has not only enhanced my practical **knowledge** but has also provided a sense of **purpose** in apply data analytics to solve complex **problems**." ~student*

*"I had the chance to interact with multiple teammates, iterate over instances ..., and develop testing mechanisms. ...**this project was a holistic learning experience for me."** ~student*

Students during group interviews for TA positions (Spring 2022)

Allison Transmission teams visit headquarters in Indianapolis (Spring 2025)

Students visit Corteva labs (Spring 2025)

GRANTS

- $10 million Lilly Endowment grant for expansion within Indiana
- $1.5 million NSF grant for MSI students nationwide
- Additional Grants
  - National Science Foundation (NSF) grants 0939370, 1246818, 2005632, 2118329, 2123321)
  - Foundation for Food and Agriculture Research (FFAR) grant 534662
  - National Institute of Food and Agriculture (NIFA) grants 2019-67032-29077 and 2020-70003-32299
  - Sandia National Laboratories; Society Of Actuaries grant 19111857
  - Cummins Inc. grants for the Indiana Digital Crossroads

Students visit Cummins headquarters in Columbus, IN (Spring 2025)

RESOURCES & MEDIA

Homepage: datamine.purdue.edu

Email: datamine@purdue.edu

The Examples Book: the-examples-book.com

Introducing The Data Mine video: https://youtu.be/R_kqpIMyhR4

Posters available via the Corporate Partners Symposium: datamine.purdue.edu/symposium

Download our recent paper: *The Data Mine model for accessible partnerships in data science* https://wires.onlinelibrary.wiley.com/doi/10.1002/wics.1642

Presenter: Margaret Betz, *Managing Director of The Data Mine in Indianapolis*
Full staff roster: *https://datamine.purdue.edu/about/about-welcome/*

Poster Number 11

Tangent space functional reconfigurations in individuals at risk for alcohol use disorder

Mahdi Moghaddam

Edwardson School of Industrial Engineering, Purdue University

Human brain function dynamically adjusts to ever-changing stimuli from the external environment. Studies characterizing brain functional reconfiguration are, nevertheless, scarce. Here, we present a principled mathematical framework to quantify brain functional reconfiguration when engaging and disengaging from a stop signal task (SST). We apply tangent space projection (a Riemannian geometry mapping technique) to transform the functional connectomes (FCs) of 54 participants and quantify functional reconfiguration using the correlation distance of the resulting tangent-FCs. Our goal was to compare functional reconfigurations in individuals at risk for alcohol use disorder (AUD). We hypothesized that functional reconfigurations when transitioning to/from a task would be influenced by family history of AUD (FHA) and other AUD risk factors. Multilinear regression models showed that engaging and disengaging functional reconfiguration were associated with FHA and recent drinking. When engaging in the SST after a rest condition, functional reconfiguration was negatively associated with recent drinking, while functional reconfiguration when disengaging from the SST was negatively associated with FHA. In both models, several other factors contributed to the functional reconfiguration. This study demonstrates that tangent-FCs can characterize task-induced functional reconfiguration and that it is related to AUD risk.

Tangent space functional reconfigurations in individuals at risk for alcohol use disorder

Mahdi Moghaddam[1,2], Mario Dzemidzic[3,4], Daniel Guerrero[1,2], Mintao Liu[1,2], Jonathan Alessi[3,4], Martin H. Plawecki[4,5], Jaroslaw Harezlak[4,6], David A. Kareken[3,4,5], Joaquín Goñi[1,2,4,7]

[1]Edwardson School of Industrial Engineering, Purdue University
[2]Purdue Institute for Integrative Neuroscience, Purdue University
[3]Department of Neurology, Indiana University School of Medicine
[4]Indiana Alcohol Research Center, Indiana University School of Medicine
[5]Department of Psychiatry, Indiana University School of Medicine
[6]Department of Epidemiology and Biostatistics, Indiana University Bloomington
[7]Weldon School of Biomedical Engineering, Purdue University

Published at Network Neuroscience 2025, Vol 9

Scan to read the article

PURDUE UNIVERSITY

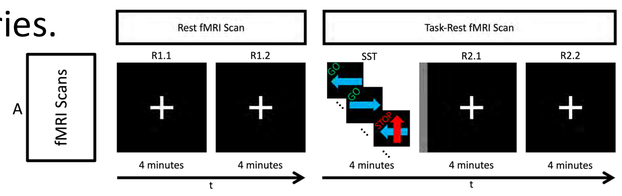
CONN*plexity* Lab

Summary and Hypothesis

Functional connectivity has been associated with behavioral attributes with the goal of understanding the relationship between individual differences and brain function. It is often estimated at a pair-wise level using Pearson's correlation coefficient of two brain regions' blood oxygenation level dependent (BOLD) time series, resulting in a symmetric correlation matrix: the **functional connectome (FC)**. Recent studies have proposed the use of tangent space projections of FCs (**tangent-FCs**); i.e., an application of Riemannian geometry. Tangent-FCs have been shown to carry a more precise "fingerprint", yielding more accurate predictors of individual or demographic traits and conditions compared to regular FCs. We present a novel framework to quantify **Functional Reconfiguration (FR)** of participants as behavioral demands changed (from rest to task and from task to rest) as the correlation distance between their corresponding tangent-FCs. Our goal was to explore the relationship between FR and **family history of alcohol use disorder (FHA)**. We hypothesized that FR would be influenced by FHA and other alcohol use disorder (AUD) characteristics.

Functional MRI Data

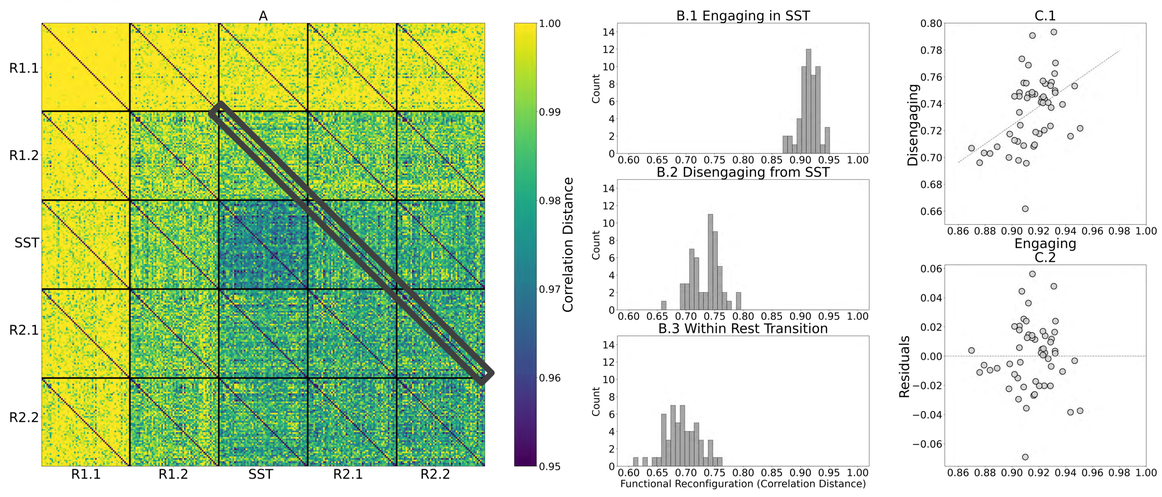
Functional MRI data of 54 participants were acquired with a BOLD contrast sensitive sequence. During the first BOLD fMRI scan (8:00 min; 400 volumes) participants were at rest and instructed to fixate their gaze on a central white crosshair. The next fMRI scan (12:12 min, 610 volumes) included 4 min of the **stop-signal task (SST)** performance followed by 8 min of rest. Both scans were divided into 4-minute segments (the duration of SST) and labeled as R1.1, and R1.2 for scan 1, and , SST, R2.1, and R2.2 for scan 2. FCs of each segment were estimated from their corresponding preprocessed BOLD time-series.



Functional Reconfiguration: engaging and disengaging

We measure **Functional Reconfiguration (FR)** of each participant from rest to task (engaging in task; R1.2 to SST) and task to rest (disengaging from task; SST to R2.1) by calculating the correlation distance of their corresponding tangent-FCs. This is the equivalent of the diagonal entries in the meta-identifiability matrix as highlighted in the figure below in Panel A. The histograms of FR vectors in Panels B.1 to B.3 show that individuals functionally reconfigure more when engaging in the SST as compared to disengaging from it. Furthermore, an individual's FR when engaging in and disengaging from SST with respect to other individuals can differ (Panel C).

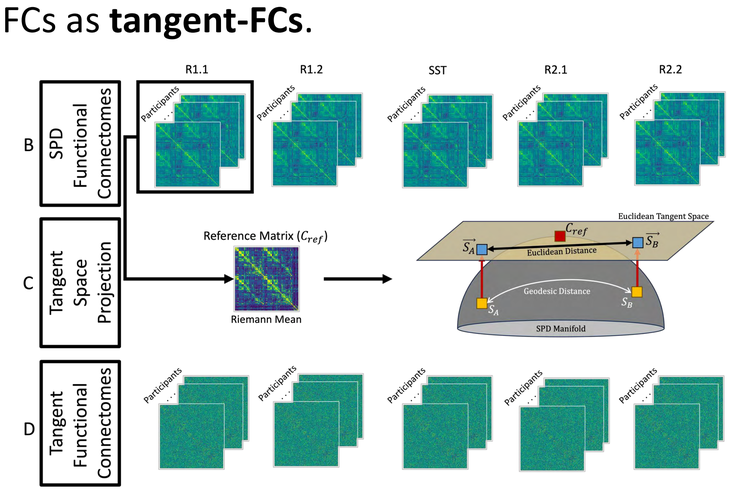
(A) Meta-identifiability matrix with functional reconfiguration vectors of R1.2 to SST (engaging), SST to R2.1 (disengaging), and R2.1 to R2.2 (within-rest) highlighted in gray. (B.1 to B.3) Histograms of functional reconfiguration values of engaging in the SST, disengaging from the SST, and within-rest post-SST, respectively. We tested functional reconfiguration using paired T-tests among Engaging, Disengaging and Within-Rest model and observed significant differences ($p < 0.001$). (C.1) Scatter plot of engaging versus disengaging functional reconfiguration vectors, showing that the two are unrelated (r=0.44). (C.2) scatter plot of residuals when predicting disengaging from engaging.



Tangent Space Projection

FCs are correlation matrices, and thus symmetric and positive definite. Therefore, they belong to the symmetric positive definite (SPD) manifold.

Tangent space projection is a mapping technique in which the FCs are projected into a Euclidean space that is tangent to a reference point, C_{ref}, on the manifold. For any reference point belonging to the SPD manifold, its tangent space is a collection of vectors that are the derivatives of the curves crossing that matrix on the manifold. We refer to the tangent space projections of SPD FCs as **tangent-FCs**.



Predicting Functional Reconfiguration

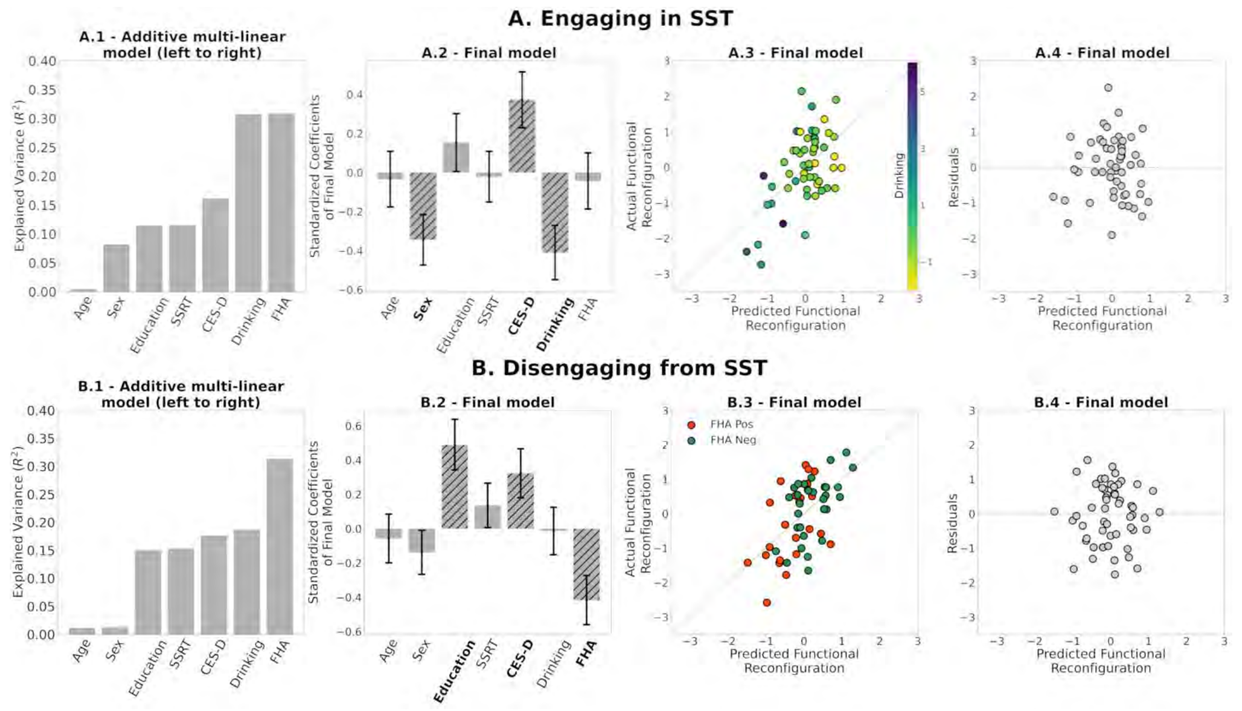
Engaging and disengaging functional reconfiguration served as the response variable in multi-linear regression models. The predictors were sex (1 male, 0 female), age, education, SSRT*, drinking behavior**, CES-D***, and FHA (1 positive, 0 negative). The extent of reconfiguration when engaging in the SST, was significantly explained by sex, recent drinking, and depressive symptoms. On the other hand, variables significantly contributing to functional reconfiguration when disengaging from the SST were education, depressive symptoms, and FHA.

Functional reconfiguration was negatively associated with recent drinking when engaging and with FHA when disengaging from SST.

* SSRT: Stop Signal Reaction Time (in milliseconds)

** Drinking behavior variable is the first principal component of AUDIT, drinking days, drinks per week, and drinks per drinking day, which explains 65% of variance.

*** CES-D: Center for Epidemiologic Studies Depression scale. Scores of 16 or greater indicate a risk for clinical depression



References and funding resources:

Moghaddam, M., Dzemidzic, M., Guerrero, D., Liu, M., Alessi, J., Plawecki, M. H., ... & Goñi, J. (2025). Tangent space functional reconfigurations in individuals at risk for alcohol use disorder. Network Neuroscience, 1-23.

NIH CTSI CTR EPAR2169, NIH R21 AA029614, NIH R01 AA029607, Indiana Alcohol Research Center P60AA07611, 2023 Bottorff Fellowship

PURDUE INSTITUTE FOR INTEGRATIVE NEUROSCIENCE
Where science meets engineering

Edwardson School of Industrial Engineering

rethink

ALCOHOL RESEARCH CENTER

Poster Number 12

Functional Connectome Fingerprinting through Tucker Tensor Decomposition

Vitor Farias Costa de Carvalho

Industrial Engineering, Purdue University

The human brain functional connectome (FC) represents the functional couplings between brain regions derived from blood oxygen level-dependent (BOLD) signals. In this study, we aim to maximize FC fingerprinting via a tensor decomposition technique. For a cohort of participants, their FC matrices can be arranged into a 3D array or tensor. Tensor decomposition techniques can then be used to produce a lower dimensional estimation of the participantsâ€™ FCs, which enables preserving the high-dimensional structure of the data while separating the data into three modalities. The first two modalities capture brain region related patterns whereas the third modality captures participant-specific information. With the information from the third modality, we compute the matching rate (a fingerprinting metric) for 426 unrelated participants from the Human Connectome Project (HCP) Young Adult dataset under eight conditions (emotion, language, gambling, motor, relational, social, resting-state, and working memory). We analyze fingerprinting in within and between-condition settings.

Functional Connectome Fingerprinting through Tucker Tensor Decomposition

PURDUE UNIVERSITY
CONN*plexity* Lab

Carvalho, V.[1,2], Liu, M.[1,2], Harezlak, J.[3,4], Estrada Gómez, A. M.[1], Goñi, J.[1,2,5]

[1]Edwardson School of Industrial Engineering, Purdue University
[2]Purdue Institute for Integrative Neuroscience, Purdue University
[3]Institute of Mathematics, University of Wroclaw
[4]Department of Epidemiology and Biostatistics, Indiana University Blooomington
[5]Weldon School of Biomedical Engineering, Purdue University

Summary

The human brain functional connectome (FC) represents the functional couplings between brain regions derived from blood oxygen level-dependent (BOLD) signals. In this study, we aim to maximize FC fingerprinting via a tensor decomposition technique. For a cohort of participants, their FC matrices can be arranged into a 3D array or tensor. Tensor decomposition techniques can then be used to produce a lower dimensional estimation of the participants' FCs, which enables preserving the high-dimensional structure of the data while separating the data into three modalities. The first two modalities capture brain region related patterns whereas the third modality captures participant-specific information. With the information from the third modality, we compute the matching rate (a fingerprinting metric) for 426 unrelated participants from the Human Connectome Project (HCP) Young Adult dataset under eight conditions (emotion, language, gambling, motor, relational, social, resting-state, and working memory). We analyze fingerprinting in within and between-condition settings.

Goals & Challenges

Goal:

- Study the effectiveness of data-driven decomposition methods in FCs fingerprinting

Challenges:

- Data is inherently high dimensional and complex
- Volatility of functional connectivity causes significant intra-participant variability across different scanning sessions, which makes it hard to extract reliable features that enable uniquely identifying participants
- Lack of a golden standard to analyze functional connectivity data

Data & Preprocessing

Data:

- Human Connectome Project (HCP) Young Adult Dataset
- Functional magnetic resonance imaging (fMRI) data
- 426 unrelated healthy participants
- Two data acquisition sessions (test and retest)
- Eight conditions: emotion processing, gambling, language, motor, relational processing, resting state, social cognition, and working memory

Preprocessing:

- In addition to the preprocessing from the HCP, we:
  1. Regressed out gray global matter signal from voxel time courses
  2. Applied a first-order Butterworth bandpass filter in the forward and reverse directions
  3. Z-scored and average, per brain regions, the voxel time courses, excluding any outlier time points falling outside three standard deviation from the mean

Parcellation:

- Schaefer parcellation functional brain atlas of the cortex
- For completeness, 14 subcortical brain regions were added to each parcellation, as provided by the HCP release
- Here, we use parcellation granularities ranging from 114-914 in steps of 100

Matching Rate

- Matching rate is the metric in the range [0,1] used to quantify fingerprinting
- Correct matches are obtained when a participant's test data is most highly correlated with their own retest data (and vice-versa) for a given identifiability matrix

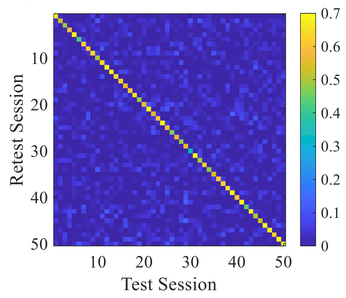


Figure 1. Sample Identifiability Matrix

Methodology

- To analyze the high dimensional data, we:
  1. Structure it into a tensor (or multidimensional array) for each of the fMRI conditions
  2. Decompose the tensor via Tucker Decomposition (ST-HOSVD)
  3. Estimate the retest session matrix based on the decomposition of the test matrix
  4. Obtain an identifiability matrix by computing pairwise Pearson's correlation between the columns of the test and retest session factor matrices
  5. Compute the matching rate

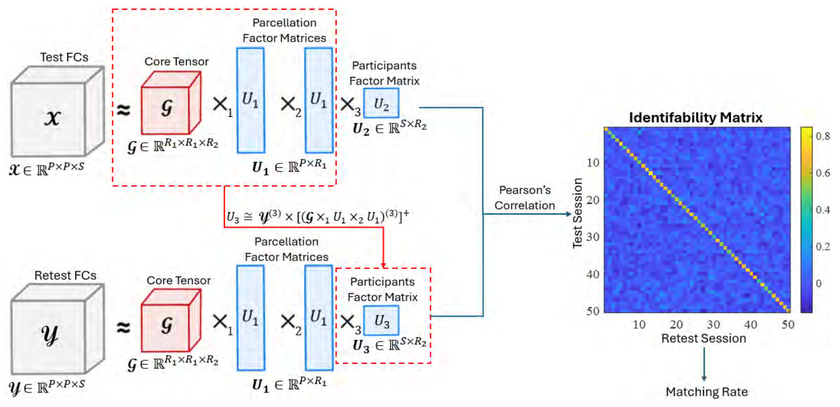


Figure 2. Fingerprinting Framework via Tucker Decomposition

Within-Condition Fingerprinting

- In A, matching rates obtained via ST-HOSVD is compared to those obtained with FCs directly
- Given the mismatch of duration between resting-state and tasks, in panel B a comparison between resting-state and tasks is done when resting-state time series are shortened to match the length of tasks

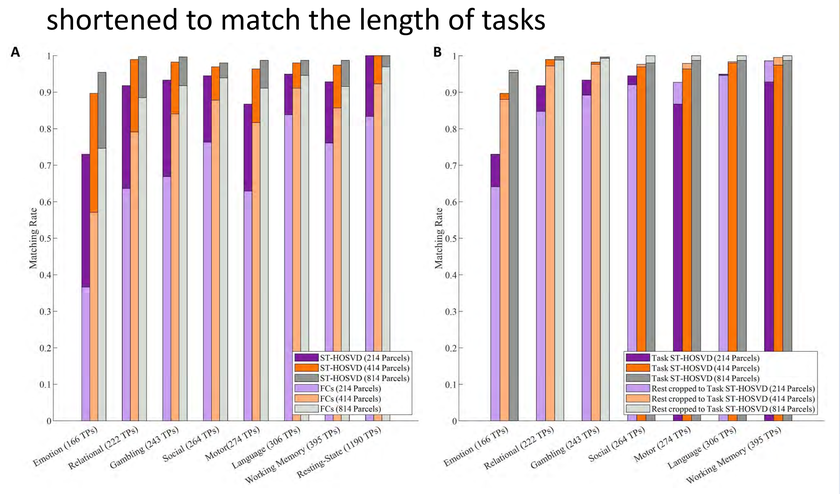


Figure 3. Within-Condition Fingerprinting Results

Rank Effects in Within-Condition Fingerprinting

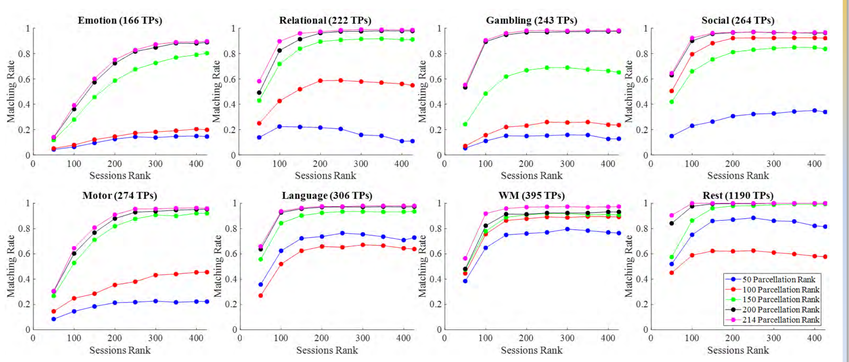


Figure 4. Within-Condition Fingerprinting Results Across Parcellation and Participant Ranks

Between-Condition Fingerprinting

- Between-condition fingerprinting expresses our ability to match subjects' resting state test sessions to their task fMRI retest sessions

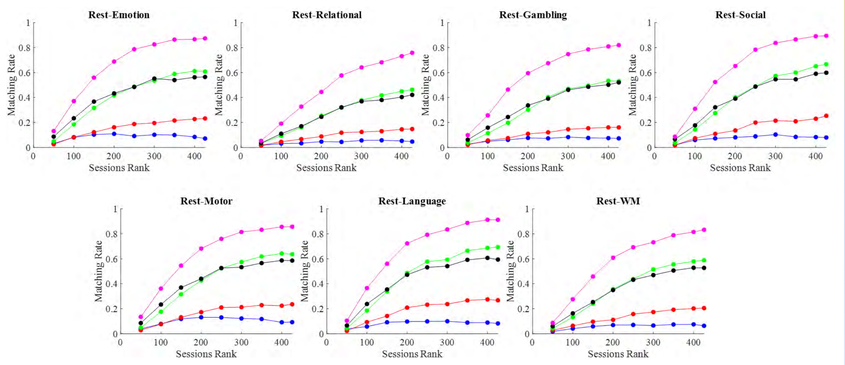


Figure 5. Between-Condition Fingerprinting Results

Time-Series Sampling Effects in Between-Condition Fingerprinting

- Given the duration mismatch between resting-state and tasks, resting-state FCs were computed with the same amount of time points as tasks. Different sampling strategies were used to select the time points used for computing resting-state FCs
- In A, resting-state FCs were obtained by randomly choosing an initial time point, and selecting the consecutive time-points
- In B, resting-state time points were randomly sampled

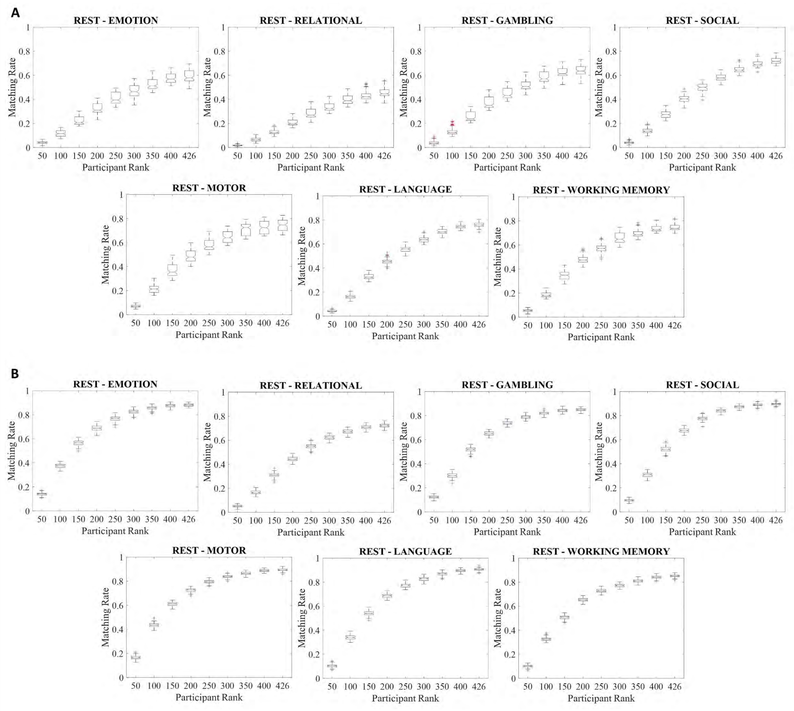


Figure 6. Matching Rate Results for Different Time Point Sampling Strategies

Assessment of Fingerprinting

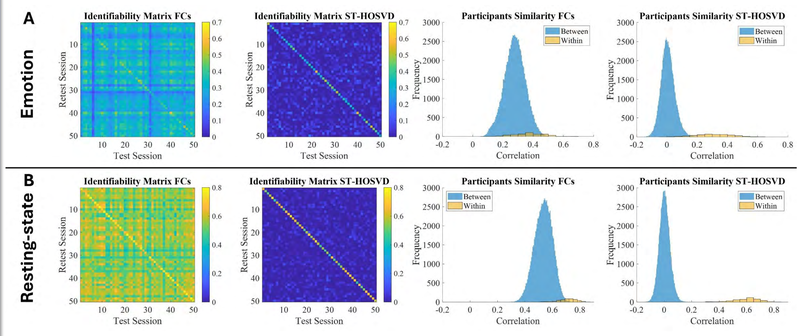


Figure 7. A and B display for the Emotion and resting-state conditions and parcellation granularity of 414, the identifiability matrices obtained via ST-HOSVD and via FCs directly along with their respective within- and between-participant distributions

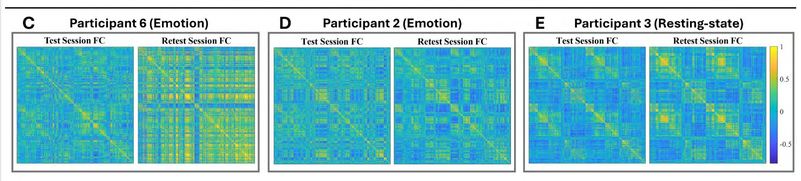


Figure 8. C, D, and E, display respectively: test and retest FCs of participants that were incorrectly matched using FCs directly and ST-HOSVD (C), FCs of a participant that was incorrectly matched via FCs, but correctly matched via ST-HOSVD (D and E)

Limitations

- Interpreting the factor matrices under a neuroscientific perspective is non-trivial due to the existence of the core tensor, which captures interactions among the various components
- The proposed fingerprinting framework does not allow for incremental updates to the dataset. Therefore, if a new participant were introduced, the entire framework would need to be recomputed
- For extremely large datasets, the computation of Tucker decomposition could become computationally prohibitive

Carvalho, V.; Liu, M.; Harezlak, J.; Estrada Gómez, A.M.; Goñi, J. Functional Connectome Fingerprinting Through Tucker Tensor Decomposition. *Appl. Sci.* **2025**, *15*, 4821. https://doi.org/10.3390/app15094821
Access the article by scanning the following QR code:

Poster Number 13

Decoding Cognitive States of Alcohol Seeking in Rats through Dynamic Functional Connectivity

Yujing Zhang

Industrial Engineering, Purdue University

Alcohol use disorder (AUD) is characterized by impaired cognitive control and dysregulated decision-making, yet the neural dynamics underlying these deficits remain poorly understood. Dynamic functional connectivity (dFC) provides a framework for capturing time-varying brain interactions that may reflect latent cognitive processes. In this study, we aimed to decode cognitive states associated with alcohol-seeking behavior in rats by interpreting the similarity of dFC patterns as evolving cognitive representations. We constructed affinity matrices to quantify the similarity between dFC patterns across time and identify discrete modules interpreted as cognitive states. To link these states to behavior, we used principal component analysis (PCA) to extract trial-level affinity curves, revealing a strong correspondence between cognitive dynamics and seeking intensity. This relationship enables a cognitive-level interpretation of behavior and highlights altered state organization in AUD-prone (P) rats compared to controls. Based on these findings, we concluded that P rats exhibit impaired cognitive control, reflected in a diminished ability to sustain goal-directed cognitive states, potentially underlying maladaptive behavioral patterns.

Decoding Cognitive States of Alcohol Seeking in Rats through Dynamic Functional Connectivity

CONN*plexity* Lab

Yujing Zhang[123], Nicholas M. Timme[4], Mintao Liu[12], Christopher C. Lapish[35], Joaquín Goñi[1236]

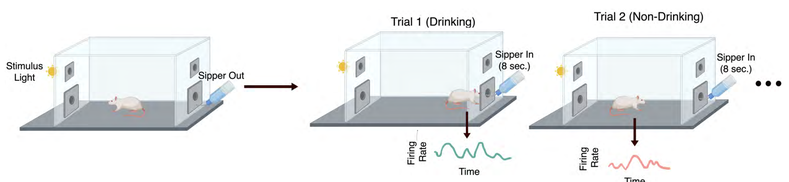
[1]Edwardson School of Industrial Engineering, Purdue University, West Lafayette, IN 47907, USA
[2]Purdue Institute for Integrative Neuroscience, Purdue University, West Lafayette, IN 47907, USA
[3]Indiana Alcohol Research Center, Indiana University School of Medicine, Indianapolis, IN, USA
[4]College of Medicine, Department of Pharmacology, Physiology, and Neurobiology, University of Cincinnati
[5]Stark Neuroscience Institute, Indiana University School of Medicine, Indianapolis, IN 46237, USA
[6]Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN 47907, USA

Summary and Hypothesis

Alcohol use disorder (AUD) is characterized by impaired cognitive control and dysregulated decision-making, yet the neural dynamics underlying these deficits remain poorly understood. **Dynamic functional connectivity (dFC)** provides a framework for capturing time-varying brain interactions that may reflect **latent cognitive processes**. In this study, we aimed to decode cognitive states associated with alcohol-seeking behavior in rats by interpreting the similarity of dFC patterns as evolving cognitive representations. We constructed **affinity matrices** to quantify the similarity between dFC patterns across time and identify discrete modules interpreted as cognitive states. To link these states to behavior, we used **principal component analysis (PCA)** to extract trial-level affinity curves, revealing a strong correspondence between cognitive dynamics and seeking intensity. This relationship enables a cognitive-level interpretation of behavior and highlights altered state organization in AUD-prone (P) rats compared to controls. Based on these findings, we concluded that P rats exhibit impaired cognitive control, reflected in a diminished ability to sustain goal-directed cognitive states, potentially underlying maladaptive behavioral patterns.

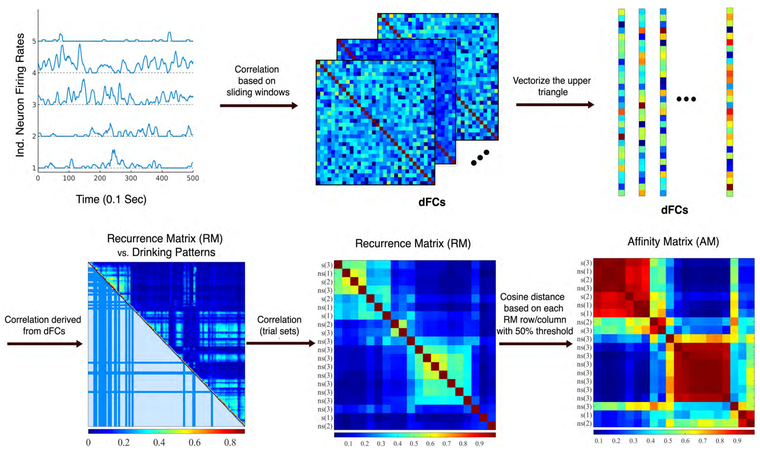
Neural Recording Data

Electrophysiological recordings were obtained from P and Wistar during a cue-triggered alcohol self-administration task **(2-CAP)** with 96 trials per session. Trials included CS+ (alcohol available) and CS- (no access), indicated by distinct visual cues. In CS+ trials, rats either drank or refrained, depending on their response to sippers containing **alcohol** or **alcohol mixed with quinine**. Local field potentials (LFPs) were recorded from dorsal medial prefrontal cortex. In total, 74 sessions were analyzed to compare neural dynamics across rat types and conditions.

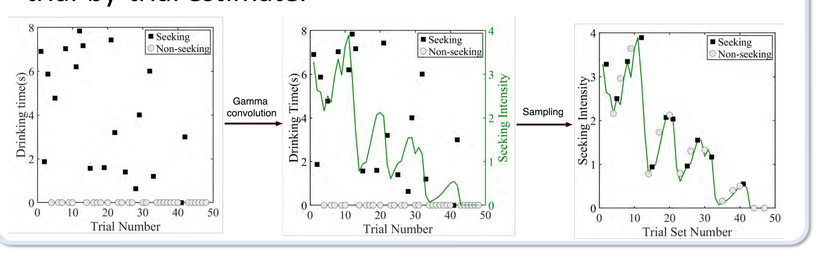


Affinity Matrix and Seeking Intensity Model Construction

Neural firing rates were binned and smoothed to generate time-series data. A sliding window approach was used to compute dynamic connectivity (dFC) matrices, capturing neuron-to-neuron correlations over time. These matrices were vectorized to build a **recurrence matrix**, measuring similarity between time windows. To better reveal cognitive state structure, we computed an **affinity matrix** by applying cosine similarity to rows of the recurrence matrix. This second-order similarity makes **cognitive states** more distinct.



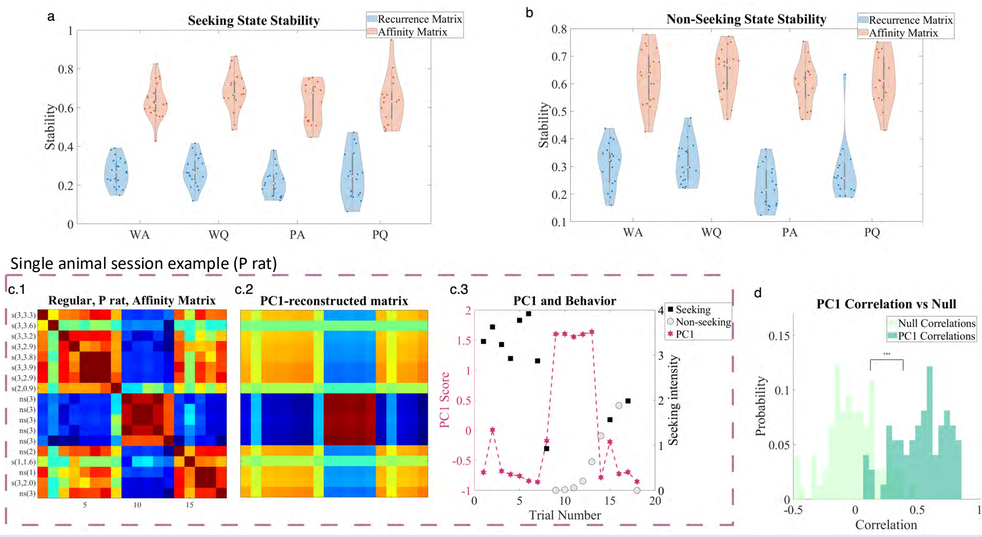
Drinking time was convolved with a **gamma function** to generate a continuous **seeking intensity model**, reflecting both current behavior and its influence on future actions. To reduce variability, the curve was resampled over 50-second trial sets for a more stable trial-by-trial estimate.



PCA-Based Decoding and Validation of Cognitive States

Violin plots (top) show that affinity matrices yield more distinct and stable cognitive states than recurrence-based representations across groups. PCA was applied to the affinity matrix to extract dominant trial-level cognitive patterns. The first principal component (PC1) was used to reconstruct a low-rank approximation (bottom middle), preserving key structure and highlighting a dominant non-seeking state. The PC1 score showed a strong negative correlation with seeking behavior (bottom right), reflecting its non-seeking **cognitive interpretation**. This correlation significantly exceeded the null distribution, supporting the behavioral relevance of the extracted cognitive dynamics.

(a) Violin plots showing seeking-state stability across groups; affinity matrices yield more stable states than recurrence matrices. (b) Violin plots for non-seeking-state stability, with similar trends favoring affinity-based representations. (c.1) Affinity matrix from a representative P rat session showing trial-by-trial cognitive similarity. (c.2) Low-rank matrix reconstructed from the first principal component (PC1), highlighting dominant non-seeking structure. (c.3) PC1 scores plotted alongside seeking behavior; the negative correlation reflects PC1's encoding of non-seeking dynamics. (d) Histogram of absolute PC1–behavior correlations compared to null correlations were used as PC1 may reflect either seeking or non-seeking dynamics. Real correlations significantly exceeded null ($p < 0.001$).



Disrupted Cognition–Behavior Coupling in rat model of AUD

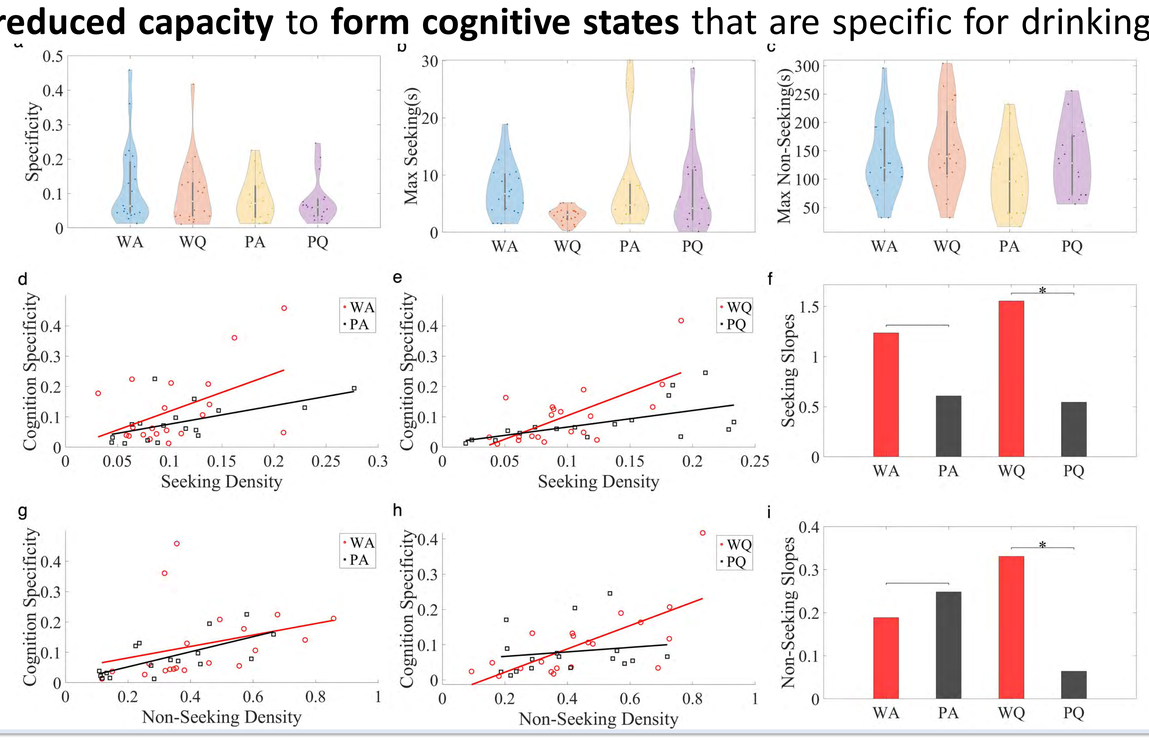
Cognition specificity was modeled as a function of behavioral stability, defined as the ratio of the longest seeking duration to total session time. In both seeking and non-seeking conditions (d–i), Wistar (WQ) showed steeper slopes than P rats (PQ) under quinine, with similar trends in alcohol sessions. P rats, particularly under aversive conditions, exhibited significantly reduced slopes (panels f, h).These patterns suggest a breakdown in the relationship between behavior and cognitive specificity in rats with genetic risk for excessive drinking (P rats), reflecting impaired cognitive adaptation and **reduced capacity** to **form cognitive states** that are specific for drinking behaviors.

Functional Cognition specificity was positively associated with behavioral stability in Wistar rats, whereas this association was markedly weaker in P rats under aversive conditions, suggesting impaired ability to accumulate goal-directed cognitive states

* Violin plots illustrate that P rats exhibit both reduced cognition specificity and lower behavioral stability compared to Wistar rats.

** Bar plots on the right show that Wistar rats exhibit significantly steeper cognition–behavior slopes, indicating stronger alignment between cognitive states and behavioral persistence.

*** The slope differences between Wistar and P rats are statistically significant ($p < 0.05$), supporting the conclusion that P rats have impaired cognition accumulation, especially under conflict conditions.



References and funding resources:

NIH CTSI CTR EPAR2169, Indiana Alcohol Research Center P60AA07611, 2024 Bottorff Fellowship.

PURDUE INSTITUTE FOR INTEGRATIVE NEUROSCIENCE
Where science meets engineering

School of Industrial Engineering

rethink

ALCOHOL RESEARCH CENTER

Poster Number 14

Data-Driven Causal Model Discovery and Personalized Prediction in Alzheimer's Disease

Guang Lin

School of Mechanical Engineering, Purdue University

With the explosive growth of biomarker data in Alzheimerâ€™s disease (AD) clinical trials, numerous mathematical models have been developed to characterize disease-relevant biomarker trajectories over time. While some of these models are purely empiric, others are causal, built upon various hypotheses of AD pathophysiology, a complex and incompletely understood area of research. One of the most challenging problems in computational causal modeling is using a purely data-driven approach to derive the modelâ€™s parameters and the mathematical model itself, without any prior hypothesis bias. In this paper, we develop an innovative data-driven modeling approach to build and parameterize a causal model to characterize the trajectories of AD biomarkers. This approach integrates causal model learning, population parameterization, parameter sensitivity analysis, and personalized prediction. By applying this integrated approach to a large multicenter database of AD biomarkers, the Alzheimerâ€™s Disease Neuroimaging Initiative, several causal models for different AD stages are revealed. In addition, personalized models for each subject are calibrated and provide accurate predictions of future cognitive status.

Data-Driven Causal Model Discovery and Personalized Prediction in Alzheimer's Disease

Haoyang Zheng[1], Jeffrey Petrella[2], Murali Doraiswamy[2], Guang Lin[1], Wenrui Hao[3]

[1] Purdue University [2] Duke University [3] Penn State University

Problem Formulation

Objective

- Develop a data-driven, causal modeling framework for Alzheimer's disease (AD) that captures biomarker progression dynamics and supports personalized prediction.

Key Biomarkers:

A_β (amyloid-beta)
τ (tau protein)
N (neurodegeneration)
C (cognitive score)

Motivations

- AD is clinically and biologically heterogeneous, with no effective cure or broadly applicable therapy.
- Existing models are often hypothesis-driven, limiting their flexibility and personalization.
- AD biomarker trajectories vary significantly across individuals.
- Need for robust, interpretable models that can personalize predictions and uncover disease mechanisms directly from data.

Contributions

Population and Personalized Models:

- A novel framework to learn the causal structure of biomarker evolution using sparse learning and ODE systems.

Proposed Models:

- Population-level model captures shared disease dynamics.
- Personalized models adapt key sensitive parameters for individual forecasting.

Sensitivity Analysis:

- Quantifies contributions of parameters to cognitive decline;
- helps select parameters for personalization.

Superior Predictive Accuracy:

- Outperforms sigmoid fitting in both population-level smoothness and subject-specific precision.

Methodology

Model Structure:

- Causal relationships among biomarkers are modeled via a system of ODEs:

$$\frac{dA_\beta}{dt} = \sum_{\ell=0}^{m} w_{1,\ell}\phi_\ell(A_\beta); \quad \frac{d\tau}{dt} = \sum_{|\boldsymbol{\ell}|\le m} w_{2,\boldsymbol{\ell}}\psi_{\boldsymbol{\ell}}(A_\beta,\tau);$$

$$\frac{dN}{dt} = \sum_{|\boldsymbol{\ell}|\le m} w_{3,\boldsymbol{\ell}}\psi_{\boldsymbol{\ell}}(\tau,N); \quad \frac{dC}{dt} = \sum_{|\boldsymbol{\ell}|\le m} w_{4,\boldsymbol{\ell}}\psi_{\boldsymbol{\ell}}(N,C),$$

Population model:

- Population parameters w, DPS parameters (α,β) fitted via Algorithm 1.

Algorithm 1 Population model calibration algorithm to compute the population parameters $\boldsymbol{w}^{(1)}$ and DPS parameters ($\boldsymbol{\alpha}$, $\boldsymbol{\beta}$). See details in Methods section.

Input $\boldsymbol{y} = \{y_{ijk}\}_{ijk}$, $\boldsymbol{t} = \{t_{ij}\}_{ij}$.
Initialize $\boldsymbol{\alpha}^0$, $\boldsymbol{\beta}^0$.

for l=1 to L **do**
  for $k \in \{A,\ T,\ N,\ C\}$ **do** ▷ Population parameter calibration
    $\boldsymbol{w}_k^l = \text{argmin}_{\boldsymbol{w}_k} \sum_{(i,j)\in\mathcal{I}_k} \left(y_{ijk} - f_k\left(\alpha_i^{l-1}t_{ij} + \beta_i^{l-1};\ \boldsymbol{w}_k\right)\right)^2$.
    $\sigma_k^l = \frac{1}{|\mathcal{I}_k - 2I - 4|} \sum_{(i,j)\in\mathcal{I}_k} \left(y_{ijk} - f_k\left(\alpha_i^l t_{ij} + \beta_i^l;\ \boldsymbol{w}_k^l\right)\right)^2$.
  end for

  for i=1 to I **do** ▷ Update DPS parameters
    $(\alpha_i^l,\ \beta_i^l) = \text{argmin}_{\alpha_i,\beta_i} \sum_{(j,k)\in\mathcal{I}_i} \frac{1}{\sigma_k^l} \left(y_{ijk} - f_k\left(\alpha_i t_{ij} + \beta_i;\ \boldsymbol{w}_k^l\right)\right)^2$.
  end for
end for

Output $\boldsymbol{w}^L$ as the population parameter $\boldsymbol{w}^{(1)}$, $\boldsymbol{\alpha}^L$, $\boldsymbol{\beta}^L$.

- DPS aligns longitudinal biomarker progression for individuals: $s_i(t) = \alpha_i \cdot t + \beta_i$

Personalized model:

- For subjects with ≥4 biomarker time points and monotonic progression.
- Algorithm 2 optimizes sensitive parameters individually for each patient.

Algorithm 2 Personalized model calibration algorithm. The personalized parameters are initialized by the population model. The personalized models are applied for subjects who meet the requirement denoted as $i \in \Omega$.

Input longitudinal biomarker data $\{y_{ijk}\}$ at $\{t_{ij}\}$ with $i \in \Omega$;
Input the DPS parameter values (α_i, β_i) for each subject $i \in \Omega$;
Input the population parameter values $\boldsymbol{w}^{(1)}$ ($\boldsymbol{w}$ for simplicity);
Input sensitivity threshold, TOL.

for m=1 to 21 **do** ▷ First order sensitivity.
  $S_m(z) = \frac{\text{Var}_{w_m}\left[\text{E}_{\text{w}_{\sim\text{m}}}(z \mid w_m)\right]}{\text{Var}(z)}$.
  if $S_m(z) \ge$ TOL **then**
    set w_m as a personalized parameter and denote as $w_m^{(2)}$ else
    keep w_m as a population parameter.
  end if
end for

for i=1 to $|\ \Omega\ |$ **do** ▷ Personalized model calibration.
  for $k \in \{A,\ T,\ N,\ C\}$ **do**
    Denote the personalized parameters in k-th equation as $\boldsymbol{w}_k{}^{(2)}$.
    ▷ Select parameters to calibrate.
    $\boldsymbol{w}_k{}^{(2)} = \arg\min_{\boldsymbol{w}_k{}^{(2)}} \sum_{j=1}^{M-1} \left(\hat{y}_{ijk} - f_k\left(\alpha_i t_{ij} + \beta_i;\ \boldsymbol{w}_k{}^{(2)}\right)\right)^2$.
    $PA_{ik} = \frac{\hat{y}_{iMk} - f_k\left(\alpha_i t_{(iM)} + \beta_i;\ \boldsymbol{w}_k{}^{(2)}\right)}{\hat{y}_{iMk}} \times 100\%$.
    ▷ Compute prediction accuracy.
  end for
end for

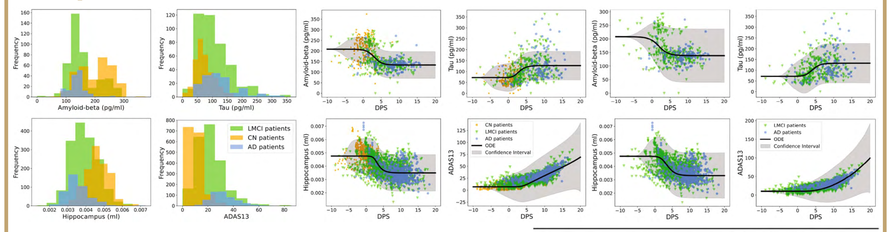
Output PA_{ik} for $i \in \Omega$ and $k \in \{A,\ T,\ N,\ C\}$.

Experiments

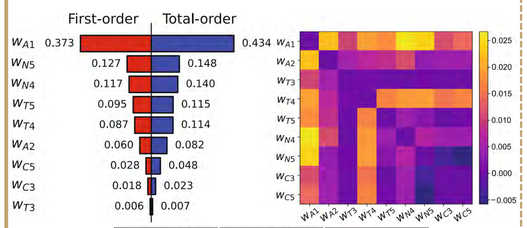
Dataset:

	CN	LMCI	AD
N	229	398	192
Age	75.72 ± 4.86	74.52 ± 7.22	75.29 ± 7.41
M / F	119 / 110	257 / 141	101 / 91
MMSE	28.82 ± 1.78	25.54 ± 4.16	21.52 ± 4.59
ADAS-13	10.29 ± 6.44	21.71 ± 10.99	32.49 ± 10.42
CSF Ab42 (pg/ml)	201.74 ± 55.15	159.37 ± 51.53	139.79 ± 35.87
CSF Total Tau (pg/ml)	72.69 ± 31.69	104.65 ± 58.28	122.01 ± 58.30
CSF Ptau (pg/ml)	29.57 ± 16.10	38.96 ± 21.09	43.91 ± 20.97
Hipp Volume (ml)	7045.38 ± 971.27	6163.12 ± 1179.50	5488.95 ± 1132.57

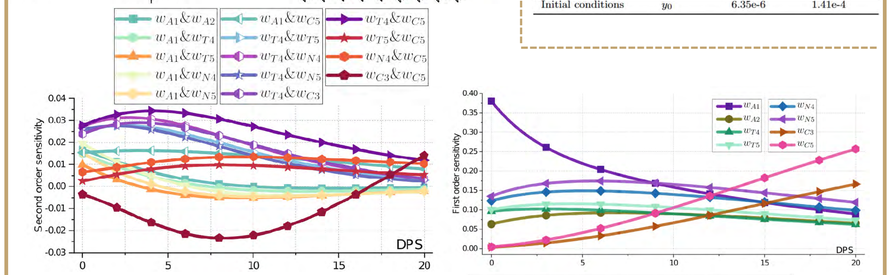
Population model:



Personalized model:



Biomarkers	Parameters	Included Subjects CN, LMCI, AD	Included Subjects LMCI, AD
A_β	w_{A0}	0	0
	w_{A1}	0.917	0.745
	w_{A2}	-0.873	-0.749
τ	w_{T0}	0	0
	w_{T1}	0.788	0.689
	w_{T2}	-0.246	-0.679
	w_{T3}	0.002	0.000
	w_{T4}	3.066	0.185
	w_{T5}	-3.650	-0.101
N	w_{N0}	0	0
	w_{N1}	1.627	0.899
	w_{N2}	-1.253	-0.927
	w_{N3}	0.018	0.554
	w_{N4}	2.342	1.792
	w_{N5}	-4.015	-2.127
C	w_{C0}	0	0
	w_{C1}	0.159	0.134
	w_{C2}	0.202	-0.067
	w_{C3}	0.010	0.004
	w_{C4}	0.019	0.007
	w_{C5}	-0.176	-0.008
Initial conditions	y_0	6.35e-6	1.41e-4



PseudoIDs (n)	DPS Diff	Model	Accuracy CSF Abeta42	CSF tTau	HIPPv	ADAS13
1 (4)	0.13	ODE	98.3%	93.6%	99.4%	92.6%
		Sigmoid	74.0%	79.8%	70.5%	84.8%
2 (4)	3.00	ODE	99.8%	93.2%	98.7%	93.0%
		Sigmoid	93.9%	61.5%	90.7%	80.4%
3 (5)	0.52	ODE	86.6%	98.8%	95.9%	85.3%
		Sigmoid	90.3%	82.6%	71.1%	56.9%
4 (5)	0.59	ODE	98.8%	96.1%	88.3%	96.6%
		Sigmoid	76.8%	76.9%	86.1%	66.7%
5 (5)	0.39	ODE	97.8%	90.0%	99.7%	94.8%
		Sigmoid	84.3%	79.5%	79.9%	81.9%
6 (4)	0.46	ODE	96.3%	93.6%	90.9%	92.7%
		Sigmoid	75.4%	91.1%	91.2%	84.0%
7 (4)	0.55	ODE	99.8%	88.2%	98.7%	90.3%
		Sigmoid	96.5%	86.0%	92.0%	72.3%
8 (4)	0.63	ODE	95.9%	98.9%	92.0%	92.6%
		Sigmoid	85.8%	86.8%	91.7%	96.6%
9 (4)	0.71	ODE	99.6%	96.1%	97.1%	87.5%
		Sigmoid	89.4%	80.3%	79.2%	69.5%
10 (5)	1.04	ODE	83.4%	81.2%	98.7%	85.5%
		Sigmoid	88.3%	78.4%	74.4%	80.1%
11 (6)	1.04	ODE	98.2%	99.8%	86.5%	85.1%
		Sigmoid	75.7%	76.3%	67.6%	72.6 %
12 (4)	0.40	ODE	94.6%	91.3%	96.5%	91.7%
		Sigmoid	89.7%	81.5%	88.9%	75.1%
13 (6)	0.88	ODE	97.0%	92.8%	96.1%	98.8%
		Sigmoid	97.4%	85.4%	85.3%	84.3%
14 (4)	0.75	ODE	98.4%	99.1%	99.1%	87.1%
		Sigmoid	90.9%	79.7%	88.6%	79.8%
15 (4)	0.55	ODE	99.6%	96.8%	90.9%	81.5%
		Sigmoid	93.8%	95.1%	81.5%	59.2%
Average	0.78 ±0.64	ODE	96.3% ±4.9%	94.6% ±5.6%	95.2% ±4.4%	90.3% ±4.8%
		Sigmoid	86.8% ±7.6%	81.4% ±7.4%	82.6% ±8.2%	76.3% ±10.1%

PseudoIDs (n)	DPS Diff	Model	Accuracy CSF Abeta42	CSF tTau	HIPPv	ADAS13
16 (5)	0.70	ODE	99.0%	87.8%	99.1%	88.2%
		Sigmoid	91.2%	87.4%	74.5%	70.1%
17 (6)	0.51	ODE	97.8%	86.2%	92.2%	90.7%
		Sigmoid	94.1%	83.3%	81.6%	88.7%
18 (5)	0.50	ODE	97.3%	95.9%	98.4%	95.1%
		Sigmoid	95.5%	90.8%	95.7%	63.4%
19 (4)	0.55	ODE	88.9%	96.3%	82.8%	96.6%
		Sigmoid	79.1%	87.7%	94.3%	85.1%
20 (5)	0.30	ODE	96.8%	98.7%	97.0%	84.9%
		Sigmoid	80.3%	89.1%	80.1%	82.1%
21 (6)	0.75	ODE	95.3%	85.8%	98.3%	84.4%
		Sigmoid	86.3%	72.9%	90.5%	79.8%
22 (4)	0.82	ODE	97.0%	95.7%	89.8%	84.6%
		Sigmoid	72.5%	68.5%	96.6%	81.6%
23 (4)	0.43	ODE	99.2%	91.0%	94.2%	80.0%
		Sigmoid	88.8%	61.3%	80.8%	62.1%
24 (6)	0.49	ODE	98.2%	98.9%	99.9%	96.9%
		Sigmoid	83.5%	89.9%	88.8%	81.1%
25 (6)	0.70	ODE	97.0%	94.4%	93.3%	89.3%
		Sigmoid	93.2%	80.2%	74.1%	70.7%
26 (4)	0.98	ODE	97.7%	93.5%	96.7%	93.8%
		Sigmoid	85.2%	89.1%	91.8%	81.3%
27 (4)	0.38	ODE	96.4%	97.4%	94.4%	96.3%
		Sigmoid	97.9%	79.5%	92.1%	72.6%
28 (5)	0.52	ODE	97.0%	90.6%	87.7%	96.6%
		Sigmoid	85.1%	78.4%	84.6%	83.1%
29 (4)	0.50	ODE	94.2%	91.4%	91.3%	95.0%
		Sigmoid	76.3%	82.7%	75.1%	72.1%
30 (4)	0.57	ODE	94.7%	90.0%	93.6%	83.1%
		Sigmoid	73.1%	80.6%	86.3%	85.8%
Average	0.58 ±0.17	ODE	97.1% ±1.5%	92.3% ±4.1%	94.8% ±3.6%	89.9% ±5.4%
		Sigmoid	85.5% ±7.4%	81.4% ±8.2%	85.8% ±7.5%	77.3% ±7.9%

Poster Number 15

A Predictive and Interpretable ML Framework to Guide LNP Design for Nucleic Acid Delivery

Gaurav Kumar

Mechanical Engineering, Purdue University

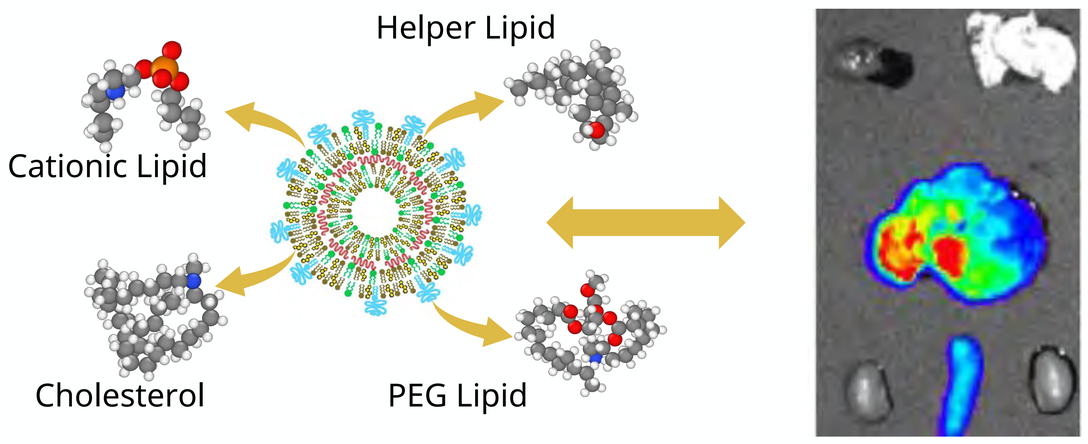
Lipid nanoparticles (LNPs) are highly effective carriers for gene therapies, including mRNA and siRNA delivery, due to their ability to transport nucleic acids across biological membranes, low cytotoxicity, improved pharmacokinetics, and scalability. A typical approach to formulate LNPs is to establish a quantitative structure-activity relationship (QSAR) between their compositions and in vitro/in vivo activities which allows for the prediction of activity based on molecular structure. However, developing QSAR for LNPs can be challenging due to the complexity of multi-component formulations, interactions with biological membranes, stability in physiological environments, and diverse physicochemical properties. To address these challenges, we developed a machine learning framework to predict the activity and cell viability of LNPs for nucleic acid delivery. We curated data from 6,454 LNP formulations reported across 21 independent studies and implemented eleven different molecular featurization techniques, ranging from descriptors and fingerprints to graph-based representations, alongside six machine learning algorithms for binary and multiclass classification. Using scaffold-based five-fold cross-validation, our models achieved classification accuracies exceeding 90% for both activity and cell viability prediction tasks. Among all model-feature combinations, descriptor-based features combined with ensemble models such as balanced random forest and extra trees yielded the highest performance. Through SHAP-based feature attribution and interaction analysis, we identified key physicochemical properties and compositional features driving LNP performance, highlighting the importance of synergistic effects among multiple molecular features. Furthermore, we developed a transfer learning strategy to bridge in vitro-to-in vivo prediction gaps by incorporating base model predictions along with additional biological attributes such as particle size, polydispersity index, and zeta potential. Despite the smaller size and inherent class imbalance of the in vivo dataset, the transfer learning models demonstrated promising predictive performance, with accuracies exceeding 82%. Our findings underscore the potential of interpretable ML frameworks to guide rational LNP design and provide a scalable approach for SAR modeling in complex nanomedicine systems.

A Predictive and Interpretable ML Framework to Guide LNP Design for Nucleic Acid Delivery

Gaurav Kumar, Arezoo Ardekani
School of Mechanical Engineering, Purdue University, West Lafayette

Kumar and Ardekani, ACS Appl. Bio Mater., 2025
DOI: 10.1021/acsabm.4c01716

Motivation

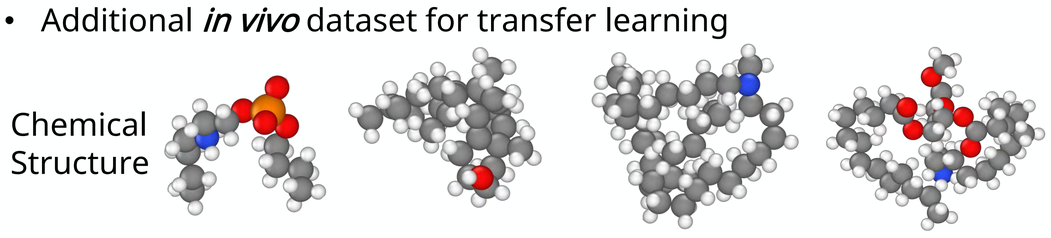


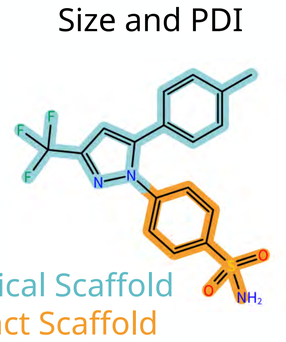
Predict the performance of Lipid Nanoparticles solely from the chemical structure/composition of LNPs

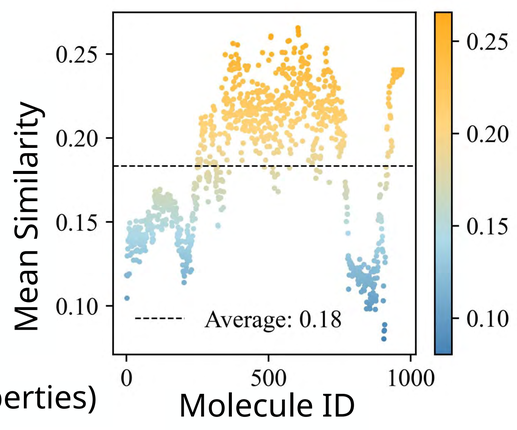
- Identify key physicochemical and structural features that drive LNP performance
- Uncover and leverage interactions between molecular and formulation features to enhance predictive accuracy
- Improve model generalizability by enforcing structural dissimilarity between training and testing LNPs
- Evaluate the effectiveness of graph-based machine learning models for LNP prediction
- Apply transfer learning to bridge in vitro predictions with in vivo performance outcomes

Methods

- **Data Preparation:** 6,454 LNPs from 21 independent *in vitro* studies
  - Data Curation, Data Standardization, Data Splitting
  - Additional *in vivo* dataset for transfer learning



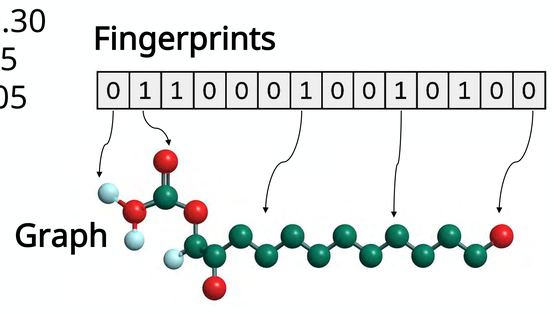




- **Feature Engineering**

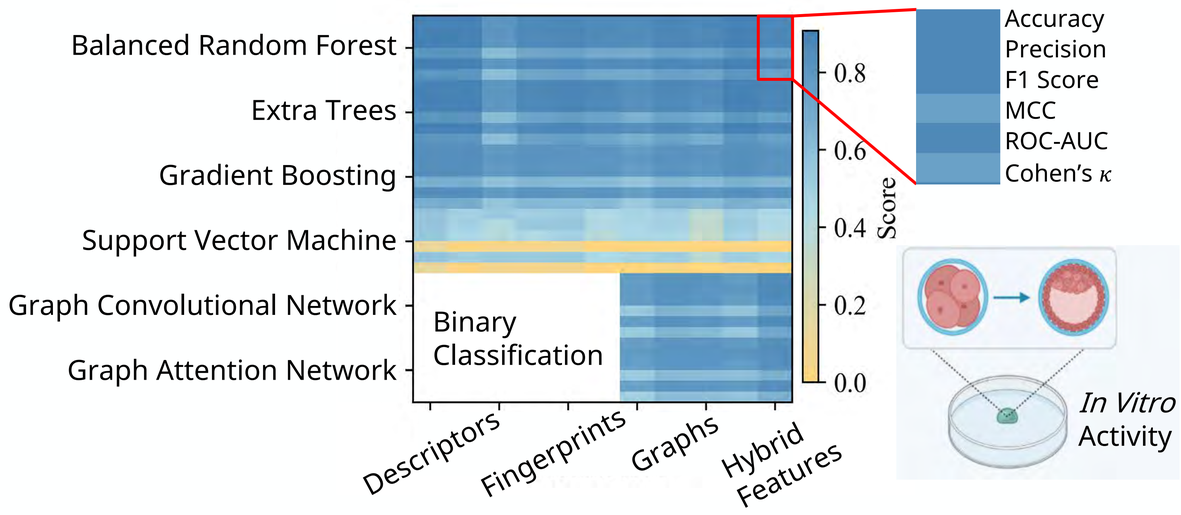
Descriptors (set of 210 physicochemical properties)

Partial Charge	-0.30
Radius of gyration	8.5
Hydrophobic surface area	405

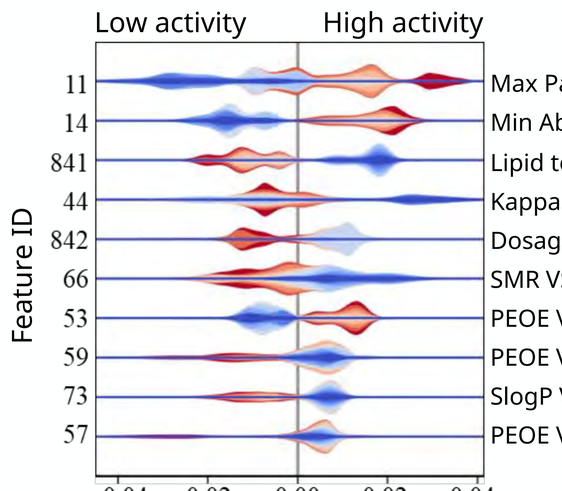


- **Model Training and Evaluation**
- **Feature Analysis**
- **Transfer Learning**

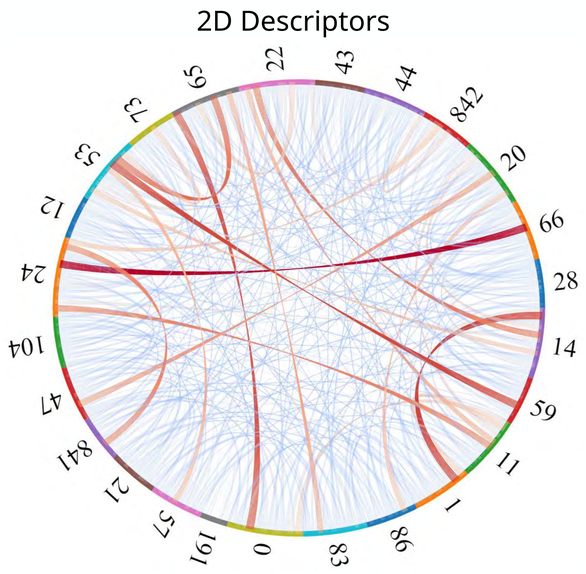
Results



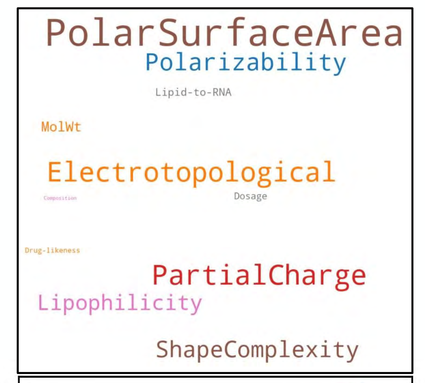
Accurate & precise *in vitro* predictions

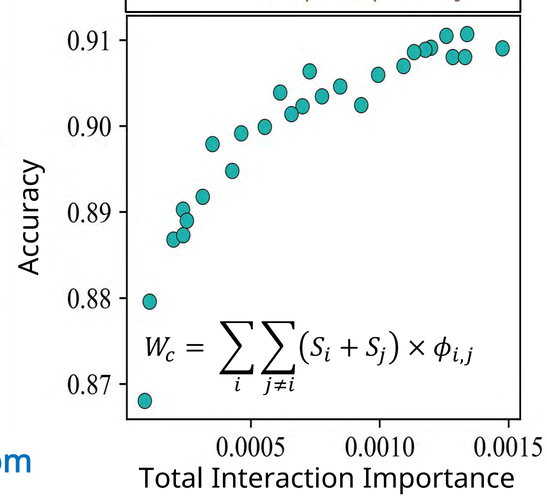


Properties related to surface charge distribution and molecular topology, in addition to LNP composition, were most important



Model performance emerged from synergistic effects of multiple interacting descriptors





Clusters comprising complementary and diverse physicochemical properties performed best: polar surface area, partial charge, molecular shape, lipophilicity

- Over 90% accuracy for *in vitro* predictions, and 82% accuracy for *in vivo* predictions
- Tree-based ensemble models paired with physicochemical descriptors were the best predictors
- Partial atomic charge, topological indices with electrostatic surface area and molecular branching, lipophilicity, lipid to RNA ratio, LNP dosage were key contributors
- Complex feature interactions played a critical role, cluster of features must be considered
- Total interaction importance was a strong indicator of predictive power

Poster Number 16

Implementing Molecular Design Principles towards a New Class of Anti-Apoptotic Mitochondrial Therapeutics

Baylen Ravenscraft

Neurological Surgery, IUSM

Traumatic spinal cord injury (SCI) is a serious condition that causes high mortality and long-term disability. Neuroprotective strategies targeting mitochondrial apoptosis have emerged as promising treatments. Early excitotoxicity in SCI leads to calcium buffering by mitochondria, which enhances oxidative phosphorylation and reactive oxygen species (ROS) production. This triggers a cascade of events, including the translocation of cytochrome c (Cyt c) from the mitochondrial membrane, leading to apoptosis. Altered interactions between Cyt c and the phospholipid cardiolipin (CL) are essential for this process, making them key targets for therapeutic intervention.

Various mitochondrial-targeted therapeutics, such as antioxidants conjugated with mitochondrial localization sidechains like triphenylphosphonium (TPP), have been developed. Another notable compound is L254614 (L25), an imidazole-containing small molecule designed to stabilize the heme of Cyt c. L25 has shown promise using in silico and in liposomal models but has not been tested in primary neuron cultures or animal models. Another therapeutic, SS-31, a mitochondria-targeted tetrapeptide, has shown benefits in various pathology models but has been underexplored in neurotrauma, with mixed results in functional recovery after SCI using 5mg/kg.

In this study, L25 was tested in vitro on primary spinal cord neuronal cultures, revealing neurotoxic effects. In contrast, SS-31 provided neuroprotection against excitotoxicity, measured by assays for cell death, viability, neuronal apoptosis, and neurite morphology. SS-31 was further tested in an in vivo mouse model of SCI, where it showed neuroprotective effects when administered at higher concentrations (100-200ÂµM in vitro and 10mg/kg in vivo). Long-term benefits were observed through locomotor assays, including kinematic analysis with DeepLabCut, and histological analysis, with SS-31 demonstrating improved functional recovery.

However, L25 showed neurotoxic effects, unlike its precursor compounds, which lacked disulfide bonds. To build on these findings, novel libraries of molecules were designed in silico, including mitochondrial localization sidechains (PACPs) and next-generation imidazole derivatives (TPP-L25s, PACP-L25s). Computational simulations indicated that these molecules may reduce oxidative phosphorylation uncoupling and enhance antioxidant effects.

The study also explored a "bait-and-push" mechanism of action for a new class of therapeutics, where electrostatic modulation of Cyt c's interaction with CL stabilizes heme coordination and promotes its displacement from the mitochondrial membrane. This effort employed a three-pronged (human-based, semi-randomized-based, and machine-learning-based) design approach and it is supported by in silico peptide/protein docking analysis.

Overall, this work advances mitochondrial therapeutics for neurotoxicant and neuroprotective applications, integrating in-vitro, in-vivo, and computational approaches to develop novel treatment strategies for SCI and other forms of neurotrauma.

Implementing Molecular Design Principles towards a New Class of Anti-Apoptotic Mitochondrial Therapeutics

Baylen Ravenscraft[1,2], Do-Hun Lee[2], Heqiao Dai[2], Ling-Xiao Deng[2], Abbie Lea Watson[2], Xiao-Ming Xu[2#], & Nai-Kui Liu[2*]

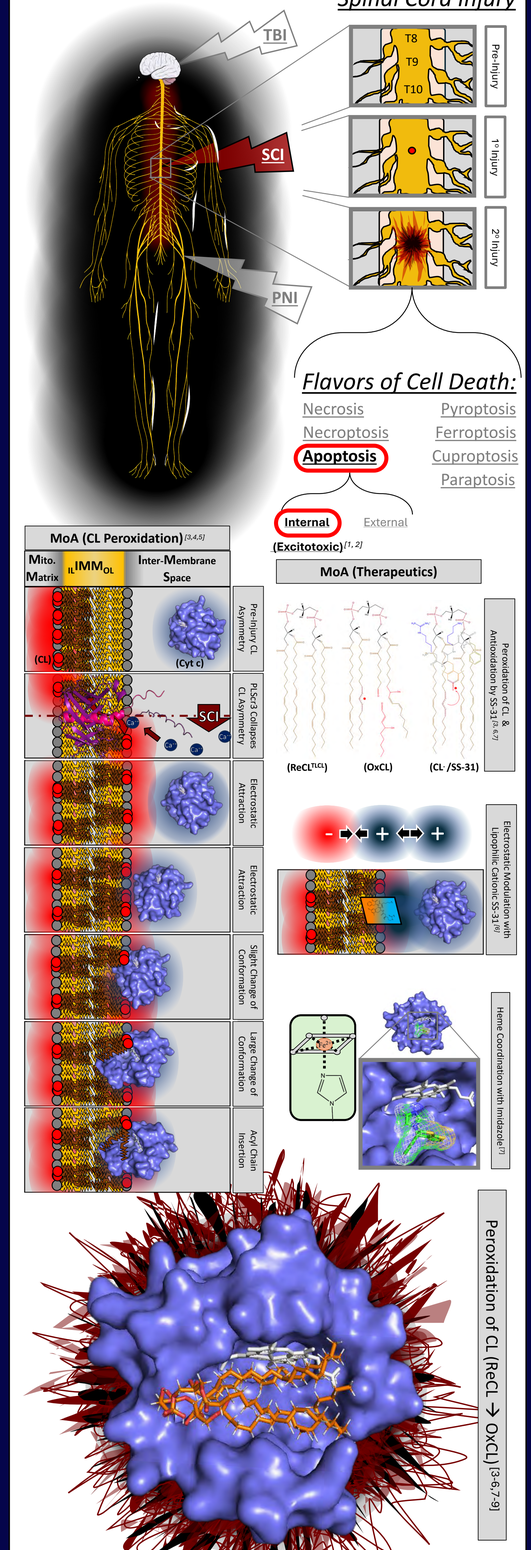
[1]Medical Neuroscience Graduate Program, [2]Neurological Surgery Department, Stark Neuroscience Research Institute (SNRI) of Indiana University School of Medicine, Indianapolis, IN, USA

INDIANA UNIVERSITY
SCHOOL OF MEDICINE
Department of Neurological Surgery
Stark Neurosciences Research Institute

Background:

Neurotrauma

Spinal Cord Injury

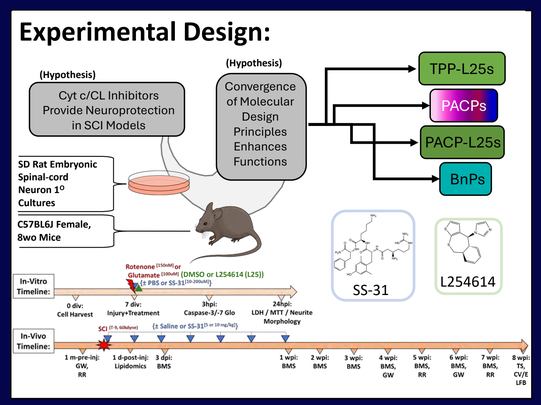


Flavors of Cell Death:

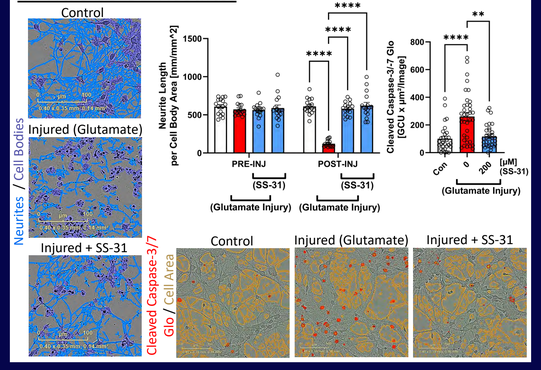
Necrosis
Necroptosis
Apoptosis
Pyroptosis
Ferroptosis
Cuproptosis
Paraptosis

Internal (Excitotoxic)[1, 2] External

Experimental Design:

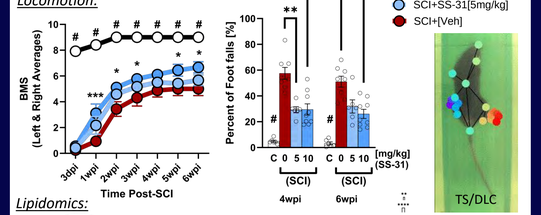


In-vitro SS-31 Results:

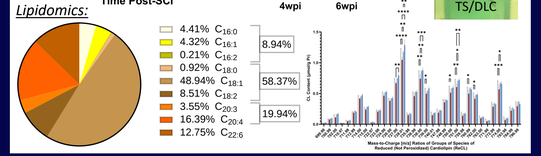


In-vivo SS-31 Results:

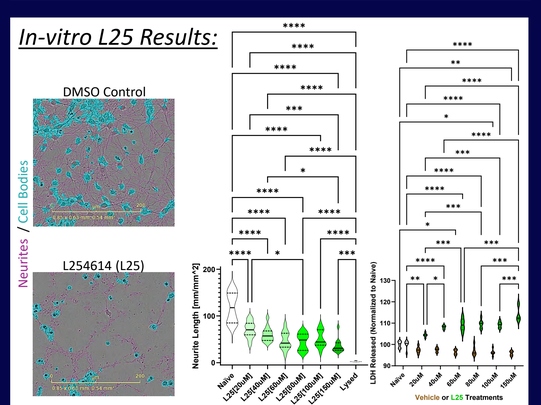
Locomotion:



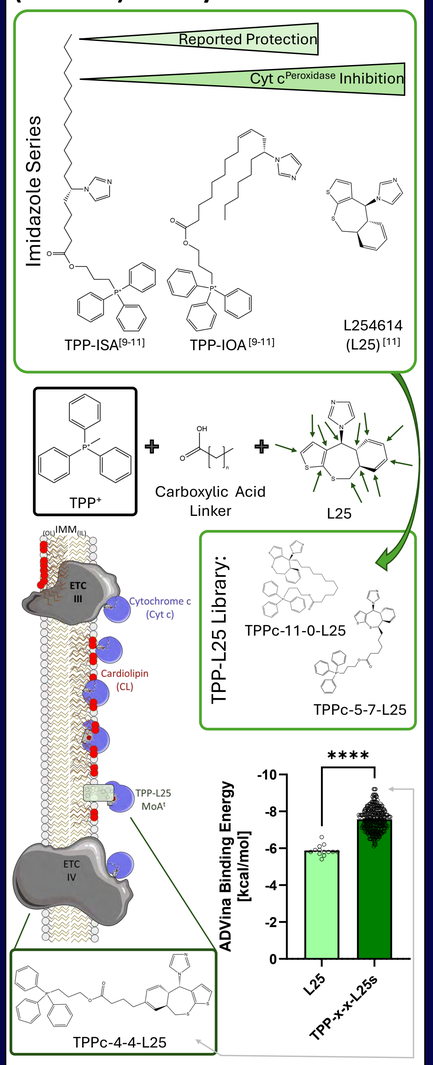
Lipidomics:



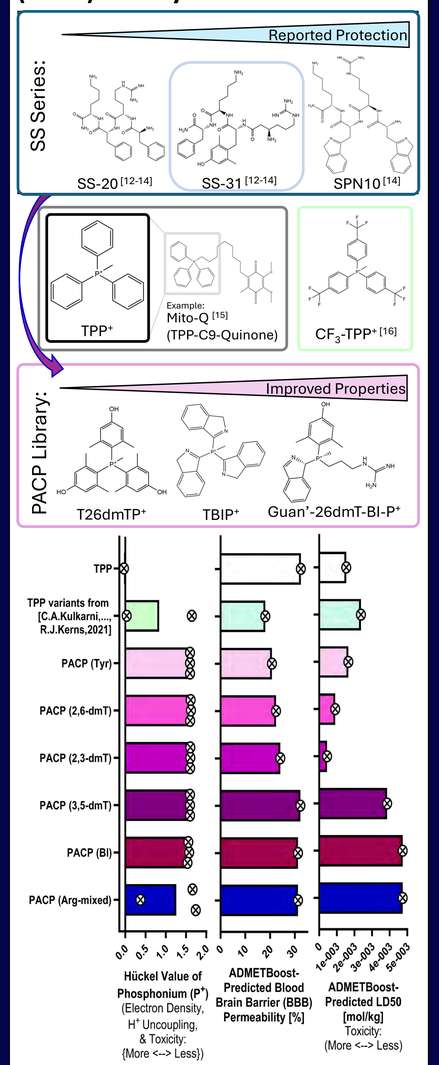
In-vitro L25 Results:



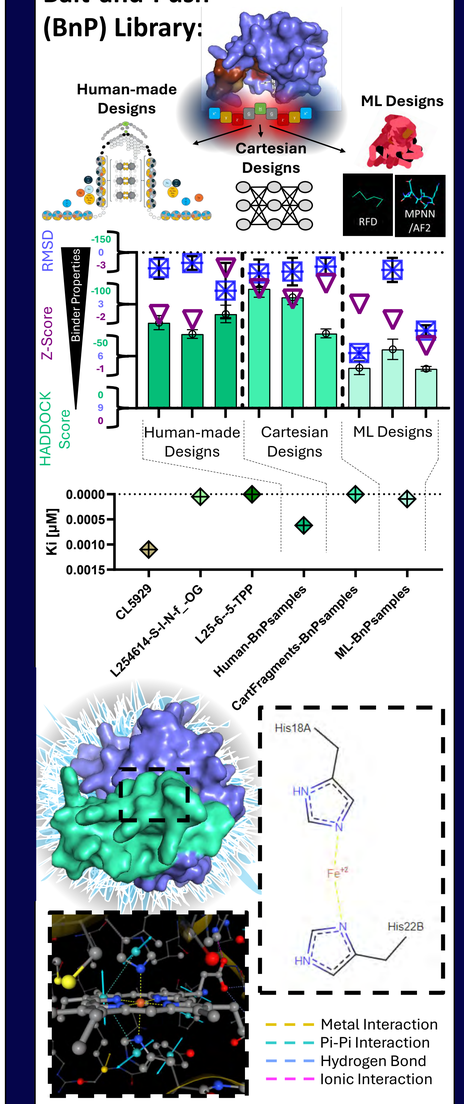
Tri-Phenyl-Phosphonium-L25 (TPP-L25) Library:



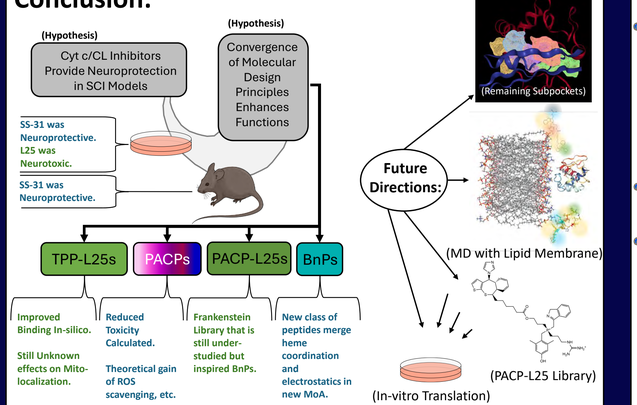
Poly-Aryl-Cationic-Phosphonium (PACP) Library:



Bait-and-Push (BnP) Library:



Conclusion:



Acknowledgements:

This work was supported by Merit Review Award I01 BX002356, I01 BX003705, I01 RX002687 from the U.S. Department of Veterans Affairs, Mari Hulman George Endowment Funds (XMX) and Indiana Spinal Cord & Brain Injury Research Fund from ISDH (NKL).

We sincerely thank the rodents whose lives were sacrificed for this study.

Additional thanks to Gabriel Inés Aparicio, Paula Monje, Jeff Recchia-Rife, Xiangbing Wu, Rene Baugh, Dom Acri, Sam Colin, Jade Harkin, and Bill Truitt for their support, teachings, and patience. Additional thanks to Xianlin Han and colleagues for the collaborative lipidomics service.

References & Software:

1. Ankarcrona, M., (1995). doi.org/10.1016/0896-6273(95)90186-8
2. Liu, N. K., & Xu, X. M. (2012). doi.org/10.3969/j.issn.1673-5374.2012.26.007
3. Nakagawa, Y. (2004) doi.org/10.1007/978-3-662-41088-2_18
4. Kagan, V. E., (2005). doi.org/10.1038/nchembio727
5. Li, M., (2019). doi.org/10.1016/j.str.2019.02.007
6. Mitchell W, (2020) doi.org/10.1074/jbc.RA119.012094.
7. Kagan, V. E., (2009). doi.org/10.1002/mnfr.200700402
8. Kagan, V. E., (2009). doi.org/10.1016/j.freeradbiomed.2009.03.004
9. Jiang, J., (2014). doi.org/10.1016/j.freeradbiomed.2014.02.029
10. Atkinson, J., (2011). doi.org/10.1038/ncomms1499
11. Bakan, A., (2015). doi.org/10.1124/mol.115.097816
12. Zhao, doi.org/10.1074/jbc.M402999200
13. Szeto, (2014). doi.org/10.1038/clpt.2014.174
14. Mitchell, W., (2022). doi.org/10.7554/eLife.75531
15. Taylor K, Smith R, *(2006), US abandoned 20060229278,*
16. Kulkarni, C. A., (2021). doi.org/10.1021/acs.jmedchem.0c01671
17. Mathis, A., (2018). DeepLabCut: ..., Nature neuroscience, 21(9), 1281-1289.
18. ChemDraw Professional, version 22.2, PerkinElmer Informatics
19. Trott, O., & Olson, A. J. (2010). AutoDock Vina: ..., 31(2), 455-461.
20. Watson, J. L., ... & Baker, D. (2023). De novo design of protein structure and function with RFdiffusion. *Nature*, 620(7976), 1089-1100. https://doi.org/10.1038/s41586-023-06415-8
21. Van Zundert, ... & Bonvin, A. M. (2016). The HADDOCK2.2 web server..., Journal of molecular biology, 428(4), 720-725.
22. Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 2.0. 2015.
23. Protein.Plus: Interactive web-based modeling of protein-protein complexes.
24. Pettersen, E. F., ... & Ferrin, T. E. (2004). UCSF Chimera--..., Journal of computational chemistry, 25(13), 1605-1612.
25. Smart.Servier.com (Servier Medical Art), 2024

Poster Number 17

Development and Evaluation of AI Models for Surgical Phase Recognition

Zheyuan Zhang

Weldon School of Biomedical Engineering, Purdue University

Surgical video analysis has emerged as a critical tool for improving operative performance, education, and patient outcomes. However, manual annotation of surgical videos remains a time-consuming and labor-intensive task, hindering the scalability of video-based evaluations in clinical and academic settings. To address this challenge, we present the development and evaluation of artificial intelligence (AI) models for automatic surgical phase recognition across multiple minimally invasive procedures, including laparoscopic cholecystectomy, Roux-en-Y gastric bypass, and colorectal surgeries.

We trained and compared four deep learning modelsâ€”ResNet-based CNN, LSTM, Temporal Convolutional Network (TCN), and Surgformerâ€”on two public datasets: Cholec80 and MultiBypass140. Models with temporal awareness clearly outperformed the static CNN. On Cholec80, TCN and Surgformer achieved accuracies of 86.12% and 89.38%, respectively. LSTM reached 80.35%, while the CNN baseline achieved 77.82%. On the more diverse MultiBypass140 dataset, both TCN and Surgformer surpassed 83% accuracy. These results show that temporal and transformer-based models are better suited for complex surgical workflows. Patient-level phase recognition remains a challenge. Frame-level predictions across long videos can reduce accuracy due to label noise and phase imbalance. We also found that recognition accuracy differs by phase. Longer phases tend to yield better performance, likely due to more consistent visual patterns and greater training data volume.

In parallel, we developed a cloud-based web interface through collaboration with the Surgical Data Science Collective (SDSC). The tool enables clinicians to upload operative videos, visualize AI-generated surgical phase boundaries, annotate timeline segments, and attach relevant notesâ€”facilitating efficient review and analysis. The interface was designed to align with clinical workflows and minimize technical barriers, supporting user interaction via standard web browsers without the need for specialized hardware.

Development and Evaluation of AI Models for Surgical Phase Recognition

Zheyuan Zhang, Fiona R. Kolbinger
Weldon School of Biomedical Engineering, West Lafayette, IN, USA

PURDUE UNIVERSITY
Weldon School of Biomedical Engineering

Background and Introduction

• AI-driven surgical phase recognition can significantly enhance intraoperative workflow understanding and has the advantage of efficiency.

• Current surgical video analyses lack effective clinical integration due to technical barriers and limited user-friendly interfaces.

Methods

Figure 1. Web-based Tool and AI Model Training Pipeline: (a) Surgeons upload laparoscopic videos via the web interface, which automatically detects key surgical phases and timestamps. (b) AI training workflow: video frames are processed using ResNet50 feature extraction and classified with deep learning models..

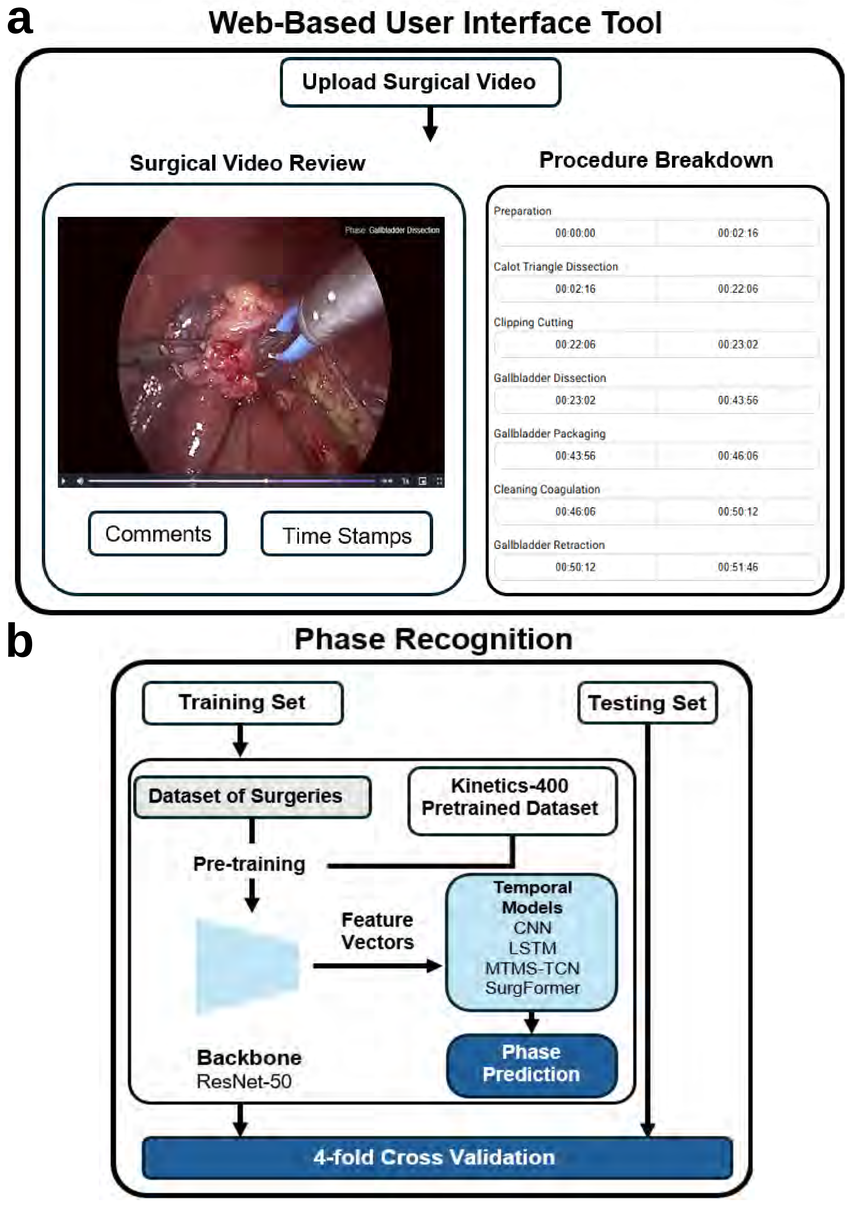
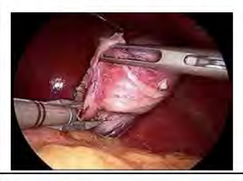


Table 1: Overview of Selected Public Datasets for Laparoscopic Cholecystectomy [1-2], Roux-en-Y Gastric Bypass [3], and Colorectal Surgery [4, 5].

Procedures	Laparoscopic cholecystectomy			Roux-en-Y gastric bypass	Colorectal surgery	
Sample image						
Datasets	M2CAI16 Cholec80 HeiChole			MultiBypass140	HeiCo LapSig300	
Centers	2	1	3	2	1	19
Patients	41	80	33	140	30	300
Phases	7 Phases			10 Phases	9 Phases	

Results

Model	Accuracy of MultiBypass140 (%)	Accuracy of Cholec80 (%)
CNN	78.15 ± 11.92	77.82 ± 5.33
LSTM	83.63 ± 11.36	80.35 ± 4.56
TCN	85.26 ± 10.12	86.12 ± 5.83
Surgformer	84.81 ± 10.23	89.38 ± 6.49

Table 2: Frame-level Model Performance on Two Surgical Datasets.

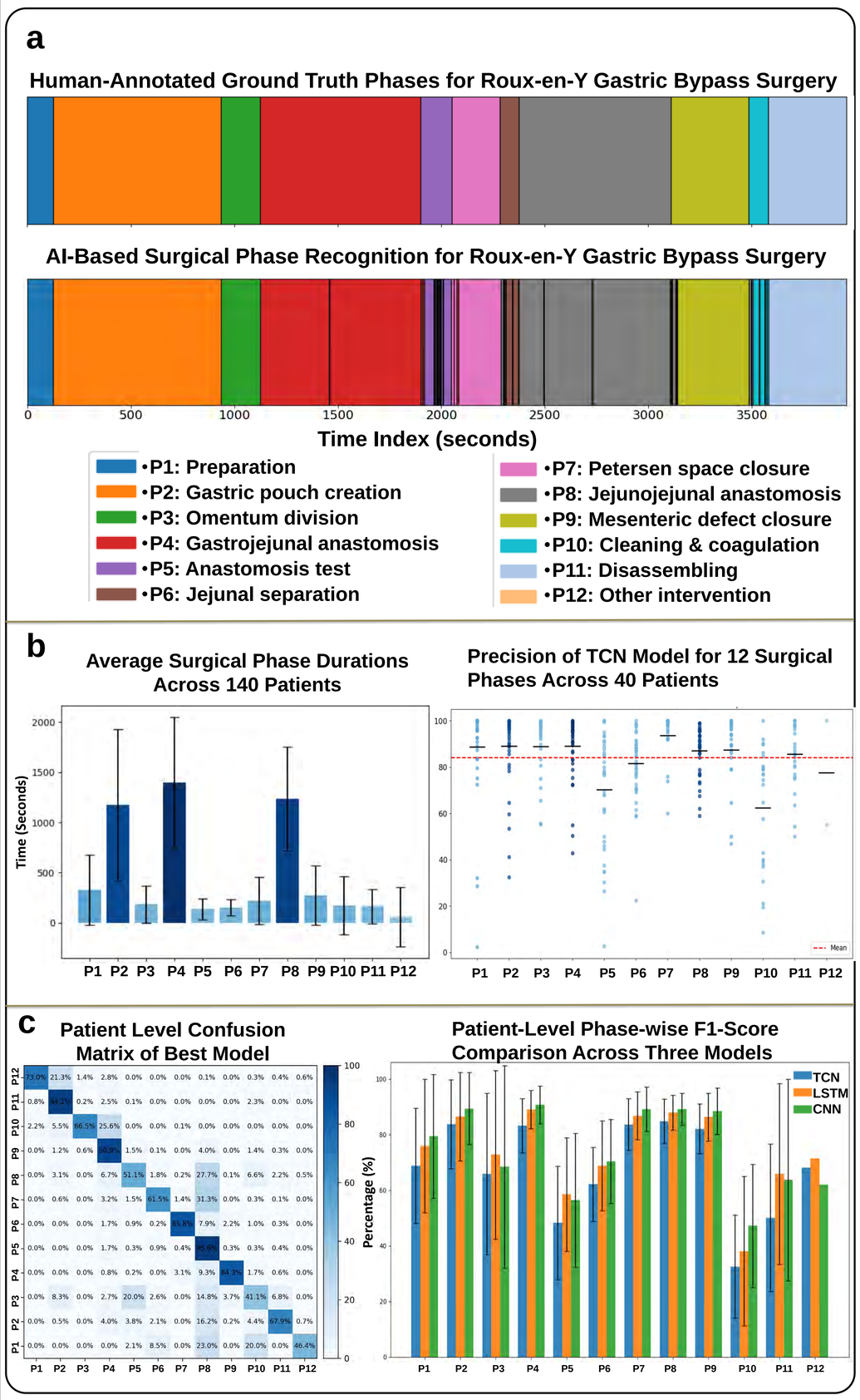


Figure 2. AI Surgical Phase Recognition of Gastric Bypass Surgery: (a) Ground Truth vs. AI-Predicted Surgical Phases, (b) Phase Duration and Model Performance Distribution, (c) Confusion Matrix and Model Comparison

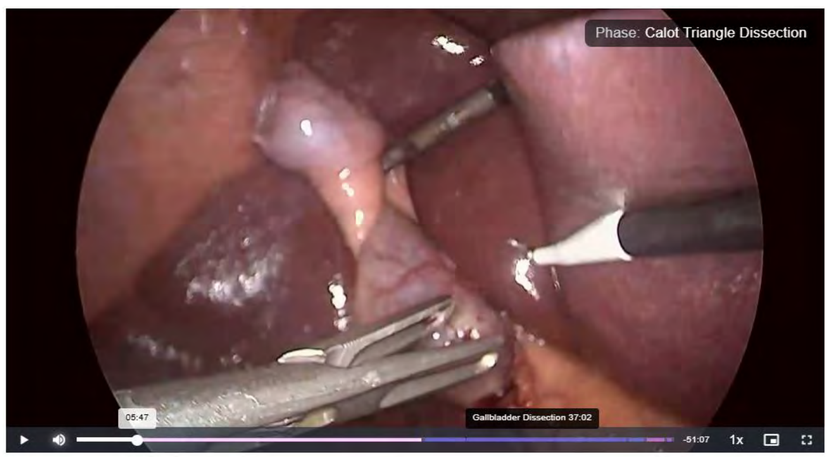


Figure 3. Visualization of Current and Upcoming Surgical Phases via the SDSC Web Interface.

Discussion

• Patient-level phase recognition is crucial, as frame-level analysis of lengthy surgical videos can negatively impact overall accuracy.

• Recognition accuracy varies significantly across phases: longer-duration phases generally achieving higher AI accuracy.

• Model choice should match the surgery type: TCN works best for Gastric Bypass, while Surgformer is better for Cholecystectomy.

Future Work

• Implement all models across three types of surgical procedures and conduct cross-dataset validation.

• Enhance the user interface for improved usability and performance.

References

[1] R. Stauder *et al.*, "The TUM LapChole dataset for the M2CAI 2016 workflow challenge," 2016. [Online]. Available: https://www.synapse.org/ (or relevant link/reference if any)

[2] M. Wagner et al., "Comparative validation of machine learning algorithms for surgical workflow and skill analysis with the HeiChole benchmark," Med. Image Anal., vol. 86, p. 102770, 2023.

[3] J. L. Lavanchy et al., "Challenges in multi-centric generalization: phase and step recognition in Roux-en-Y gastric bypass surgery," Int. J. Comput. Assist. Radiol. Surg., vol. 19, pp. 2249–2257, 2024.

[4] L. Maier-Hein et al., "Heidelberg colorectal data set for surgical data science in the sensor operating room," Sci. Data, vol. 8, pp. 1–11, 2021.

[5] D. Kitaguchi et al., "Automated laparoscopic colorectal surgery workflow recognition using artificial intelligence: Experimental research," Int. J. Surg., vol. 79, pp. 88–94, 2020.

[6] A. P. Twinanda et al., "EndoNet: A Deep Architecture for Recognition Tasks on Laparoscopic Videos," IEEE Trans. Med. Imaging, vol. 36, pp. 86–97, 2016.

[7] A. P. Twinanda et al., "EndoNet: A deep architecture for recognition tasks on laparoscopic videos," IEEE Trans. Med. Imaging, vol. 36, pp. 86–97, 2017.

[8] S. Hochreiter and J. Schmidhuber, "LSTM can solve hard long time lag problems," in Advances in Neural Information Processing Systems, M. Mozer, M. Jordan, and T. Petsche, Eds. Cambridge, MA, USA: MIT Press, 1996.

[9] S. Ramesh et al., "Multi-task temporal convolutional networks for joint recognition of surgical phases and steps in gastric bypass procedures," Int. J. Comput. Assist. Radiol. Surg., 2021.

[10] S. Yang et al., "Surgformer: Surgical Transformer with Hierarchical Temporal Attention for Surgical Phase Recognition," in Proc. Med. Image Comput. Comput. Assist. Interv. (MICCAI 2024), vol. 15006, Springer, 2024.

Machine Learning in Medicine

Poster Number 18

Atrial Fibrillation classification using Machine Learning on ECG signals.

Heiner Castro

Polytechnic Institute, Purdue University

This study explores the use of machine learning to detect atrial fibrillation (AF) from ECG signals by analyzing P wave characteristics. We developed synthetic ECGs in Simulink and validated our approach using the MIT-BIH Arrhythmia database. Feature extraction was performed with NeuroKit2, and classification was achieved using a Support Vector Machine (SVM). Results show high accuracy in distinguishing AF, supporting future integration into clinical decision-making systems.

Purdue University, Polytechnic Institute

Atrial Fibrillation classification using Machine Learning on ECG signals.

Integrated Sensing & Smart Solutions Lab

ABSTRACT

This study explores the use of machine learning to detect atrial fibrillation (AF) from ECG signals by analyzing P wave characteristics. We developed synthetic ECGs in Simulink and validated our approach using the MIT-BIH Arrhythmia database. Feature extraction was performed with NeuroKit2, and classification was achieved using a Support Vector Machine (SVM). Results show high accuracy in distinguishing AF, supporting future integration into clinical decision-making systems.

Authors: Heiner Castro Gutierrez (hcastrog@purdue.edu) *
Dr. Suranjan Panigrahi (spanigr@purdue.edu) **

*PhD student at the School of Engineering Technology.
** Professor at the School of Engineering Technology.

Background

- Our purpose is to assess predictive models in healthcare using modern machine learning techniques and big healthcare datasets to improve the detection of patient deterioration or subacute patient illness in critical care settings.
- By predicting catastrophic clinical events, such as arrhythmias, these models can improve the management of hospitalized patients and help healthcare providers make better decisions.
- We aim to evaluate the performance of machine learning (ML) algorithms for arrhythmia detection.
- We developed an ECG simulation in Simulink as a probe of concept (see Fig. 1). Then, we extended this concept to the MIT-BIH dataset to detect Atrial Fibrillation (AF)
- The Purdue IRB office determined (IRB-2023-1944) that this research does not qualify as Human Subjects Research under federal regulations.

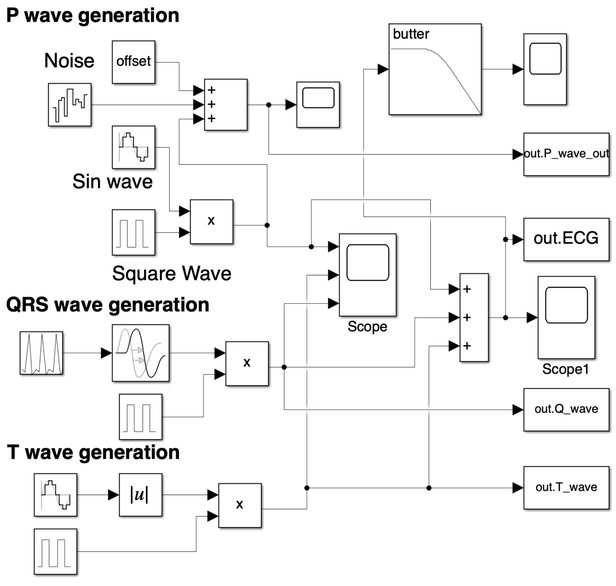


Fig. 1. Simulink Block Diagram for the simulation of an ECG signal.

Methodology, Results, and Discussion

- A Simulink block diagram was created to generate synthetic ECG waves (P, QRS, and T) using sinusoidal and triangular signals, using inputs such as amplitudes, widths, noise, and offsets (Fig. 1).
- A dataset containing low and high P waves (Fig. 2) were created and classified using a Support Vector Machine (SVM) with a linear kernel.
- The model achieved high accuracy (up to 0.99), demonstrating machine learning's effectiveness in classifying P wave amplitudes.
- We processed ECGs from the MIT-BIH Arrhythmia database (see Fig. 3). We used the NeuroKit2 Python library to analyze ECG signals and extract features like P wave peak and width.
- Initial analysis showed that general abnormal and normal ECGs were not distinguishable by P wave amplitude.
- When focusing on atrial fibrillation (AF) ECGs—known to lack P waves— distinct differences emerged.
- Histograms showed (Fig. 4) that low P wave amplitudes correlated with AF ECGs, while high amplitudes indicated normal ECGs.
- Then, we build an SVM model containing 4 predictors such as mean and std deviation of amplitude and width of the P waves each 10s. The results are shown in Table 1.

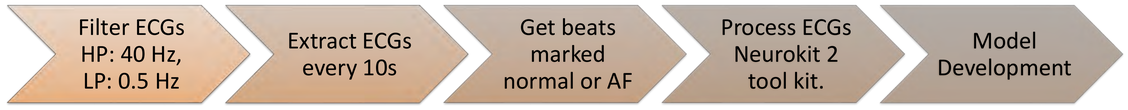


Fig. 3. Block Diagram for the preprocessing of the MIT-BIH ECG signals.

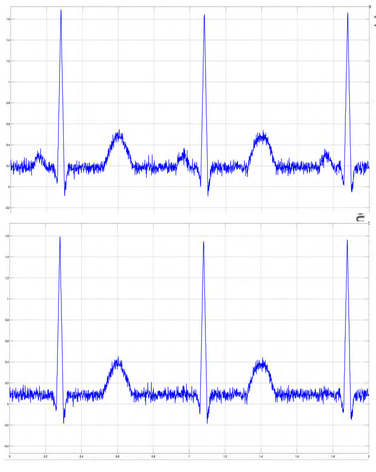


Fig. 2.Sumulated ECG signals by MATLAB: Normal ECG (u), no P wave (bottom).

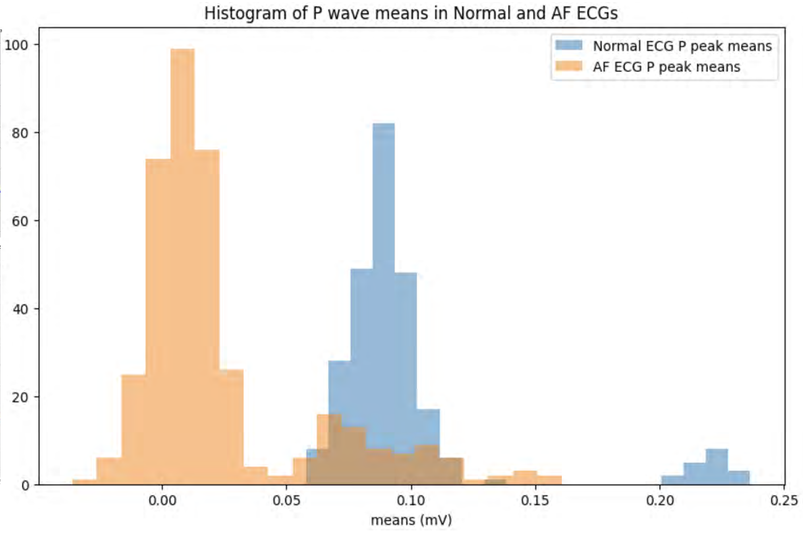


Fig. 4. Histogram for the Normal and AF ECGs in the MIT-BIH database.

SVM Performance Evaluation

Table 1. Linear SVM Performance Metrics.

Validation Set Performance metrics	
Accuracy	96.89 %
Sensitivity	100 %
Specificity	91.55 %
Precision (PPV)	95.31 %
F1-Score	97.60 %
NPV	100 %

Conclusions and Future Work

- This study demonstrates that machine learning techniques, particularly Support Vector Machines (SVM), can effectively classify ECG signals based on P wave characteristics, aiding in the detection of atrial fibrillation (AF).
- By combining simulated ECG signals with real data from the MIT-BIH Arrhythmia database, we validated the approach's potential.
- Our results show that although general abnormal ECGs are not easily distinguishable from normal ECGs using P wave amplitude alone, AF ECGs can be reliably identified.
- The developed SVM model achieved high accuracy by leveraging statistical features (mean and standard deviation) of P wave characteristics over 10-second intervals.
- In future work, we aim to expand this methodology to classify other types of arrhythmias beyond AF using additional ECG features such as QRS complex and T wave morphology.
- We also plan to incorporate deep learning techniques and explore time-series models to improve performance on noisier, real-world clinical datasets.

Poster Number 19

Development of digital biomarkers for inflammation: Exploring ECG morphology from a chest â€“ worn wearable device

Darpit Dave

Biomedical Engineering, Purdue University

Background: Current approaches requires collection blood-based serum biomarkers to track vaccine induced inflammation response indicative of reactogenicity. This can be costly, cumbersome for the individuals and also do not provide a continuous measure of response. Digital health wearables with continuous tracking of physiology present a great alternative. Studies so far have relied on heart rate (HR) and heart rate variability (HRV) to track this response. Despite the success, potential of wearable technology remains under-utilized.
Objective: This study explores the role of ECG morphology in detecting inflammation through a chest-worn wearable ECG devices. Studies based on the 12-lead clinical ECG has lead to great results. However, wearable ECG often produce noisy data due to artifacts. Therefore, this study investigates the potential of reliably capturing information related to ECG morphology from a wearable device.
Methods: Data was collected from 45 patients aged 18 to 69, who received either the Moderna or Pfizer vaccine. The ECG signals were recorded using the Vitalconnect patch at a sampling rate of 125 Hz. We focused on the detection of fiducial points (P, Q, R, S, T peaks) and analyzed ECG beat morphology before and after vaccination. The study aimed to answer two main research questions: (1) Does ECG morphology exhibit signs of vaccine-induced inflammatory responses? and (2) Can we leverage morphology-based features to detect these changes?
Results: The detection of inflammatory responses was consistently observed across all vaccine doses, with significant changes observed in 100% of patients. The median detection time for inflammatory responses was 5 hours from vaccine.
Conclusion: This study demonstrates that combining ECG morphology and machine learning techniques can effectively capture vaccine-induced inflammatory changes through a wearable device. Our findings suggest that future work should explore other morphological components beyond the QT-interval, with potential applications for managing other chronic conditions and early detection of diseases.

Development of digital biomarkers for inflammation: Exploring ECG morphology from a chest – worn wearable device

Darpit Dave[1], Sarwat Amin[1], Damen Wilson[1,2], Matthew P. Ward[1,2], Steven R. Steinhubl[1]

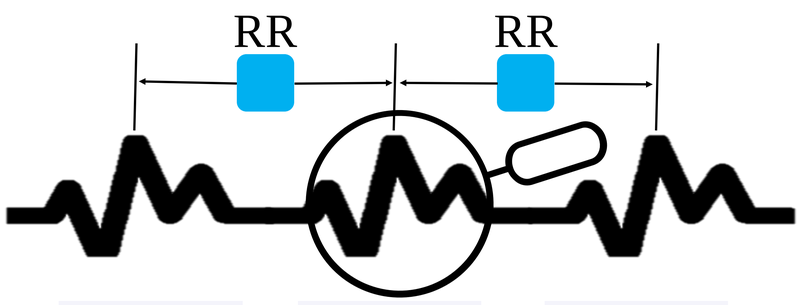
[1]Weldon School of Biomedical Engineering

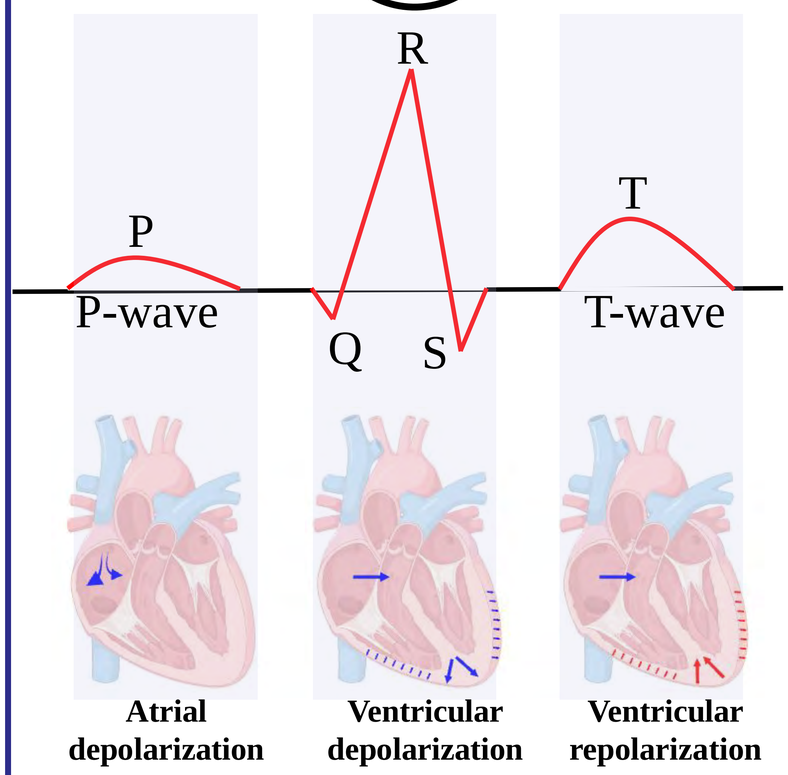
[2]IU School of Medicine

Background

- Studies have shown heart rate and heart rate variability (HRV) to correlate with vaccine-induced inflammation [1]
- However, ECG morphology remains largely unexplored
- Manual feature extraction for morphology can be unreliable given the presence of artifacts with wearable devices

Why morphology matters?



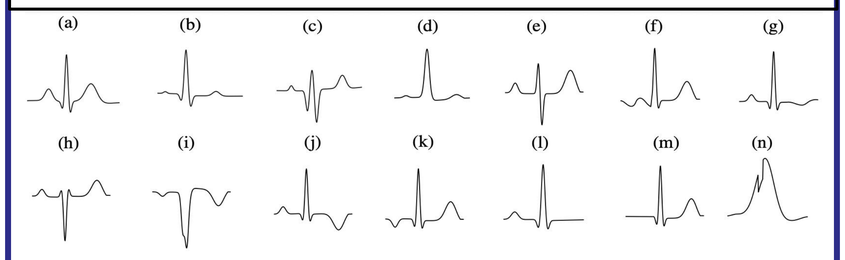


Atrial depolarization | **Ventricular depolarization** | **Ventricular repolarization**

Research Questions

- Does beat morphology show signs of vaccine induced inflammatory response?
- How can we detect leverage morphology-based information for vaccine-induced inflammatory changes?

Wearable ECG data

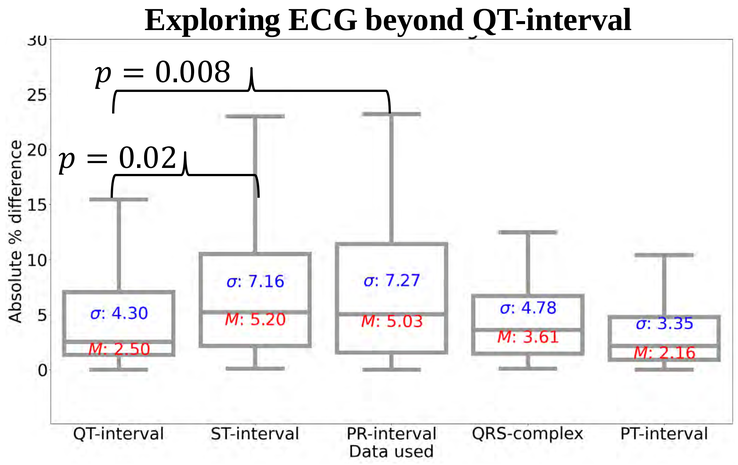


- Wearable ECG is noisy compared to 12-lead ECG used in clinical settings
- Individual beats can exist in various forms making detection of fiducial points (P, Q, R, S, T peaks) challenging [2]

Future Work

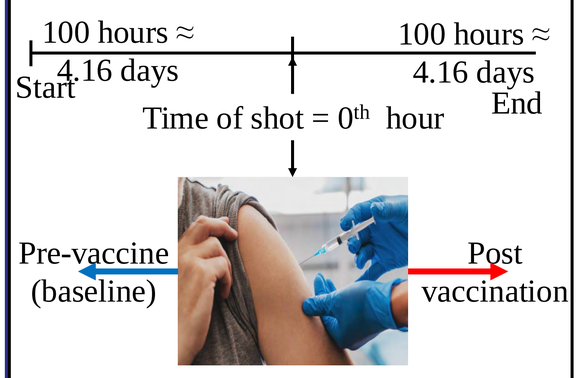
- Combining ECG morphology and RR-interval data to effectively capture vaccine induced inflammatory responses
- Explore morphological components beyond QT-interval and correlation with CAE based output

Exploring ECG beyond QT-interval



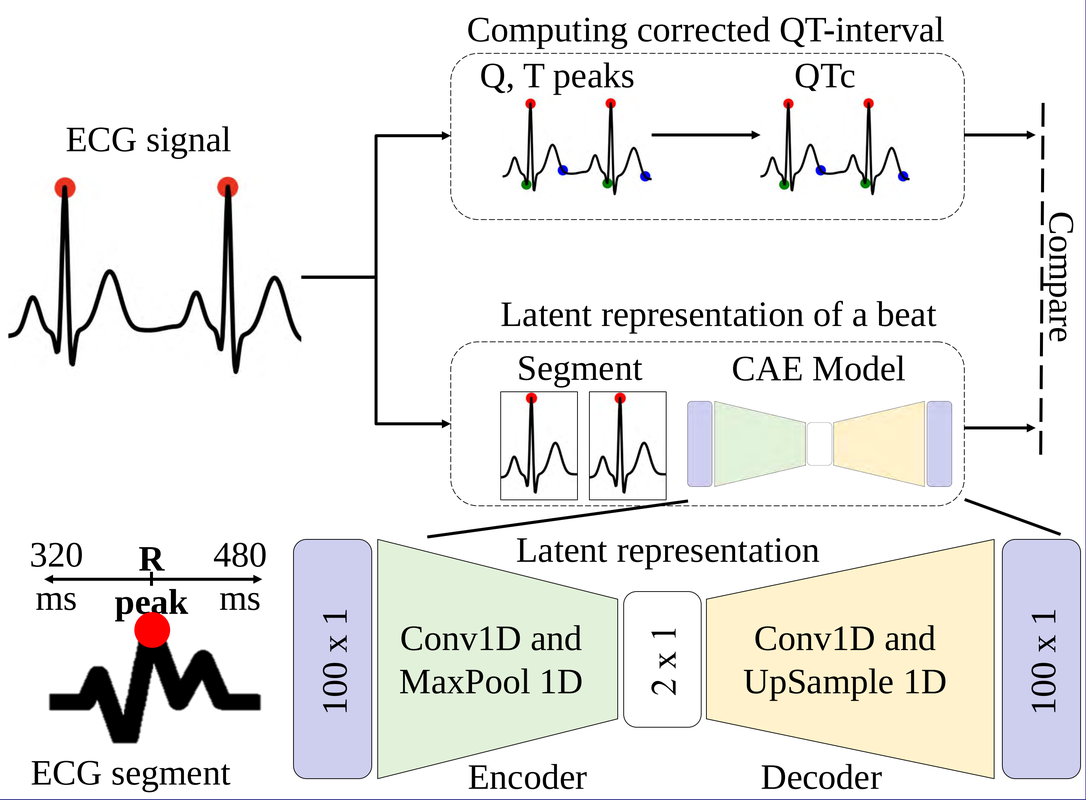
Methods

Data collection



Vaccine type	Moderna and Pfizer
Patient age range	18 – 69 years old
ECG data collected	Vitalconnect patch
Sampling rate	125 Hz

Modeling approach



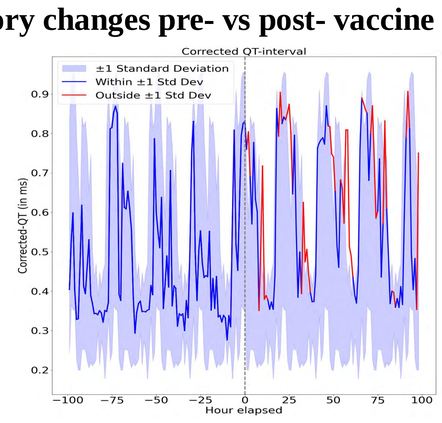
Results

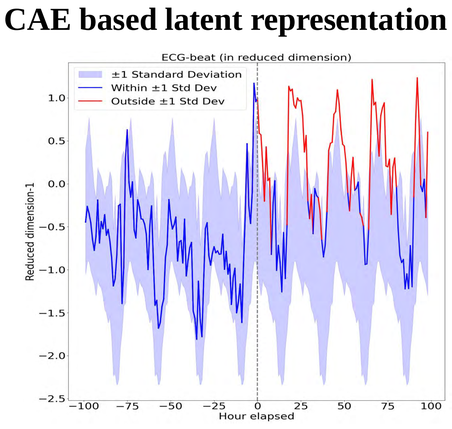
Statistical test: Wilcoxon signed-rank $H_o: m_{baseline} - m_{post-vaccine} = 0, H_a: m_{baseline} - m_{post-vaccine} \neq 0$

Category	Overall (n=45)	V1 (n=10)	V2 (n=26)	Booster (n=9)
# of patients with detectable response	45 (100%)	45 (100%)	45 (100%)	45 (100%)
Difference in absolute % change (Median, SD)	3.3 (5.5)***	1.6 (2.9)**	4.8 (9.7)***	2.1 (5.0)**
Detection time in hours (Median, SD)	5.0 (1)	5.0 (9)	5.0 (1)	3.0 (3)

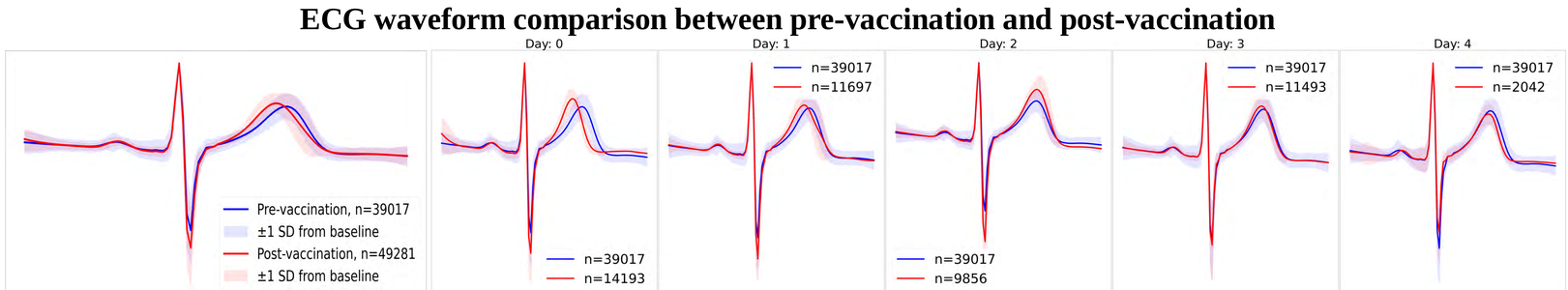
$* \; p-value < 0.05, ** \; p-value < 0.01, *** \; p-value < 0.001$

Inflammatory changes pre- vs post- vaccine : (a) CAE based latent representation of beat, (b) QTc-interval





ECG waveform comparison between pre-vaccination and post-vaccination



Patient and vaccine dosewise Cosinor analysis showing presence of circadian rhythms

References

1. Steinhubl, S. R., Sekaric, J., Gendy, M., Guo, H., Ward, M. P., Goergen, C. J., ... & Wegerich, S. (2025). Development of a personalized digital biomarker of vaccine-associated reactogenicity using wearable sensors and digital twin technology. *Communications Medicine*, *5*(1), 115.
2. Martínez, Arturo, Raúl Alcaraz, and José Joaquín Rieta. "Application of the phasor transform for automatic delineation of single-lead ECG fiducial points." *Physiological measurement* 31, no. 11 (2010): 1467.

Poster Number 20

Deep Learning-Driven Hyperspectral Imaging for Label-Free Classification of Nanoscale Systems

Kaeul Lim

Mechanical Engineering, Purdue University

Hyperspectral imaging (HSI) has emerged as a transformative modality widely applied in remote sensing, precision agriculture, environmental monitoring, and pharmaceutical research. While these applications leverage the spectral and spatial data captured by HSI, its potential for characterizing nanoscale systems, such as therapeutic nanoparticles, remains underexplored. This study introduces a novel HSI framework that integrates 3D convolutional neural networks (3D CNNs) with the synthetic minority oversampling technique (SMOTE) to address class imbalance, high data dimensionality, and spectral noise. Using liposomes as a representative case, our method achieves an average classification accuracy of 99.16%, enabling the precise, label-free characterization of complex systems. By advancing label-free hyperspectral classification through deep learning, this work highlights the transformative potential of spectral imaging in nanoparticle analysis. Beyond therapeutic applications, the scalability and precision of the proposed framework position it as a powerful tool for label-free characterization across diverse scientific and industrial domains.

The 2025 Machine Learning in Medicine, May 9, 2025

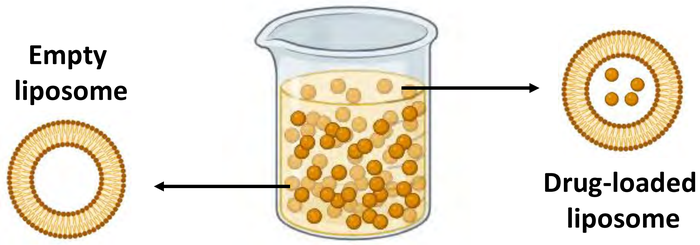
Deep Learning-Driven Hyperspectral Imaging for Label-Free Classification of Nanoscale Systems

Kaeul Lim, Arezoo Ardekani, Mechanical Engineering, Purdue University

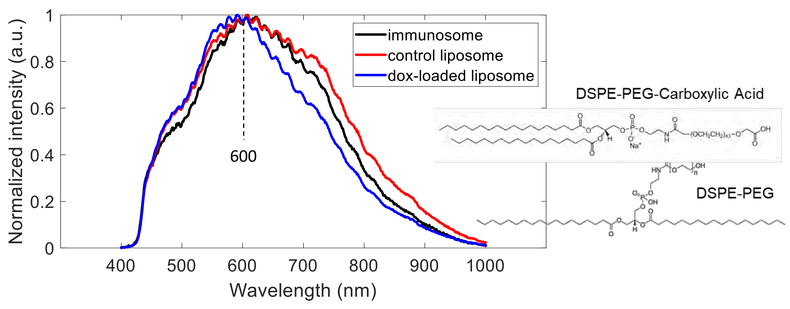
PURDUE UNIVERSITY

CBM Center for Bioanalytic Metrology

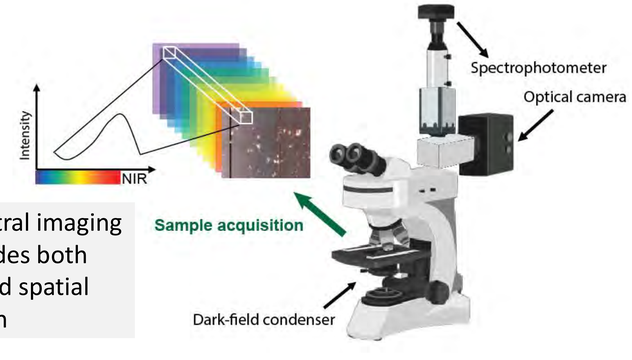
Milestones & Deliverables



- A representative spectral profile of each nanomaterial is distinctive, which can be used to map drug loading and classify different nanoparticles.



Hyperspectral Imaging (HSI) System



Hyperspectral imaging (HSI) provides both spectral and spatial information

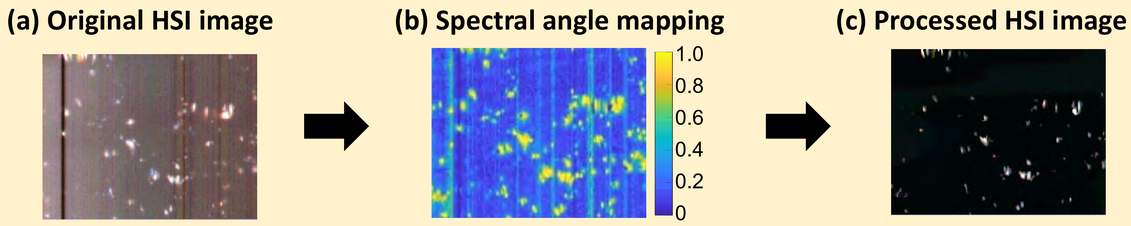
- HSI techniques can be used for the identification of materials of biological particles. Many machine learning algorithms are used for hyperspectral image classification.

Acknowledgments

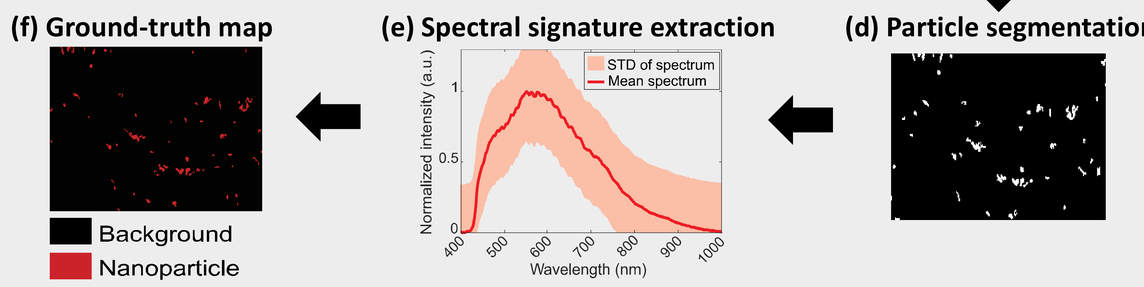
- We are grateful to the NSF Center for Bioanalytic Metrology for providing funding for this project, and to industry members of the CBM for valuable discussions.

Classification of Hyperspectral Images using Machine Learning

- Image preprocessing step using spectral angle matching (SAM) method

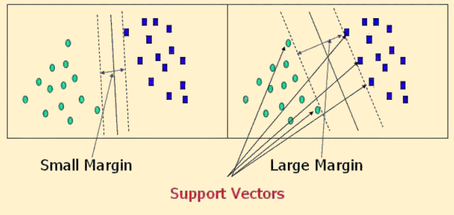


- From label-free hyperspectral images, unique reference spectral profiles for each bio-nanoparticle are estimated.

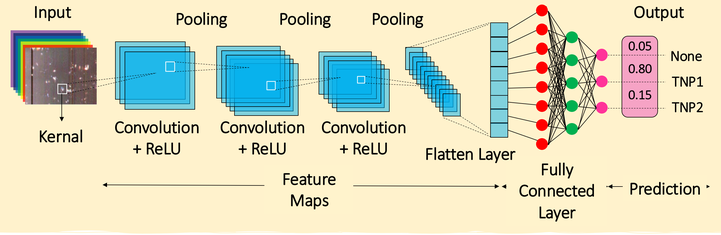


- Machine learning model selection: SVM vs 1D CNN vs 3D CNN

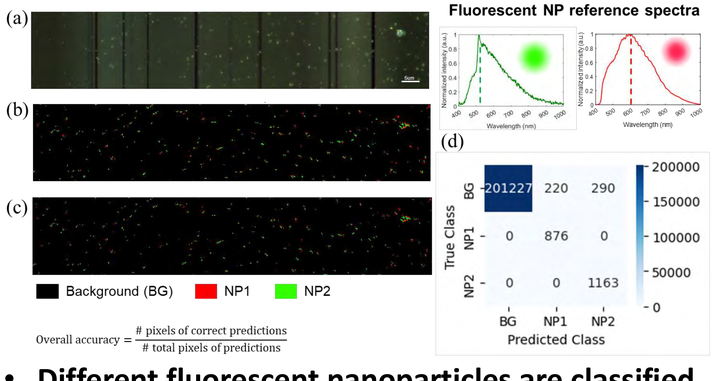
Support Vector Machine (SVM)



Convolutional Neural Network (CNN)



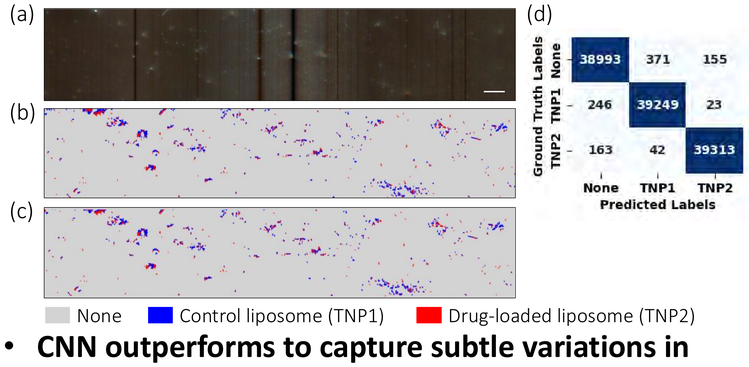
- SVM-based classification result



$$\text{Overall accuracy} = \frac{\#\text{ pixels of correct predictions}}{\#\text{ total pixels of predictions}}$$

- Different fluorescent nanoparticles are classified through the SVM method.

- 3D CNN-based classification result



- CNN outperforms to capture subtle variations in reference spectra. This capability enables the effective classification of diverse biological particles.

Poster Number 21

Error informed heritability vs environment measures in functional connectomes

Tanu Raghav

Edwardson School of Industrial Engineering, Purdue University

The brainâ€™s complex structural and functional organization varies significantly between individuals, influenced by both genetic and environmental factors. In the last decade, functional magnetic resonance imaging (fMRI) studies with repeated sessions have shown the presence of a fingerprint in functional connectomes. More recently, twin studies have provided insights into the extent to which heritability and environment contribute to differences in brain structure and function. However, the contributions of these factors to functional networks are scarce and have been mostly focused on resting-state. How heritability and environmental factors can be assessed and understood across fMRI conditions remain unknown. To address this, we extend the classical ACE and ADE models to include a repeated measurement error term and obtain heritability and environment estimations for each fMRI condition and functional coupling separately. We hypothesize the existence of functional circuits that show similar levels of differentiated heritability and/or environmental factors across all fMRI conditions (including tasks and rest).

Error informed heritability vs environment measures in functional connectomes

Tanu Raghav[1,2], Julie Sara Benny[1,2], Uttara Tipnis[3], Daniel Guerrero[1,2], Mintao Liu[1,2], Jaroslaw Harezlak[4], Joaquín Goñi[1,2,5]

[1]Edwardson School of Industrial Engineering, Purdue University, West Lafayette, IN, USA
[2]Purdue Institute for Integrative Neuroscience, Purdue University, West Lafayette, IN, USA
[3]Lawrence Livermore National Laboratory, Livermore, CA, USA
[4]Department of Epidemiology and Biostatistics, Indiana University, Bloomington, IN, USA
[5]Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA

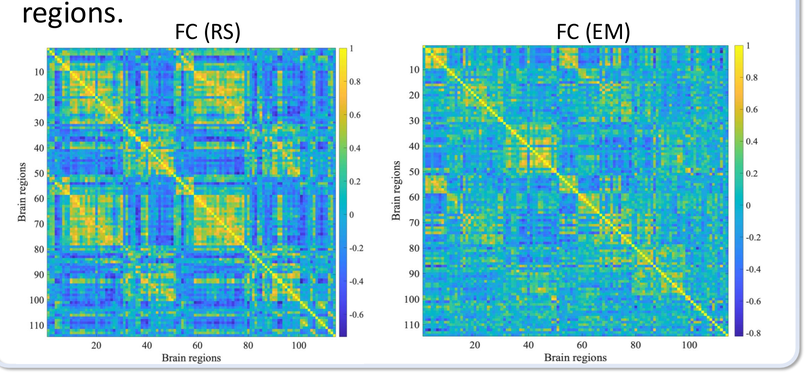
CONN*plexity* Lab

Introduction

The brain's complex structural and functional organization varies significantly between individuals, influenced by both genetic and environmental factors. In the last decade, functional magnetic resonance imaging (fMRI) studies with repeated sessions have shown the presence of a fingerprint in functional connectomes. More recently, twin studies have provided insights into the extent to which heritability and environment contribute to differences in brain structure and function. However, the contributions of these factors to functional networks are scarce and have been mostly focused on resting-state. How heritability and environmental factors can be assessed and understood across fMRI conditions remain unknown. To address this, we extend the classical ACE and ADE models to include a repeated measurement error term and obtain heritability and environment estimations for each fMRI condition and functional coupling separately. We hypothesize the existence of functional circuits that show similar levels of differentiated heritability and/or environmental factors across all fMRI conditions (including tasks and rest).

Functional MRI Data

We used data from the HCP 1200 participants release which includes 426 unrelated participants, 116 pairs of Monozygotic (MZ) twins and 63 pairs of Dizygotic (DZ) twins. fMRI data includes both resting state (RS) and tasks: emotion processing (EM), gambling (GAM), language (LAN), motor (MOT), relational processing (REL), social cognition (SOC), and working memory (WM). For each condition, participants underwent two sessions corresponding to two different acquisitions (left-to-right or LR, and right-to-left or RL. Whole-brain FC is computed for each participant for both sessions (test-retest), each fMRI condition using Pearson's correlation, yielding an m × m symmetric matrix, where m is the number of brain regions.

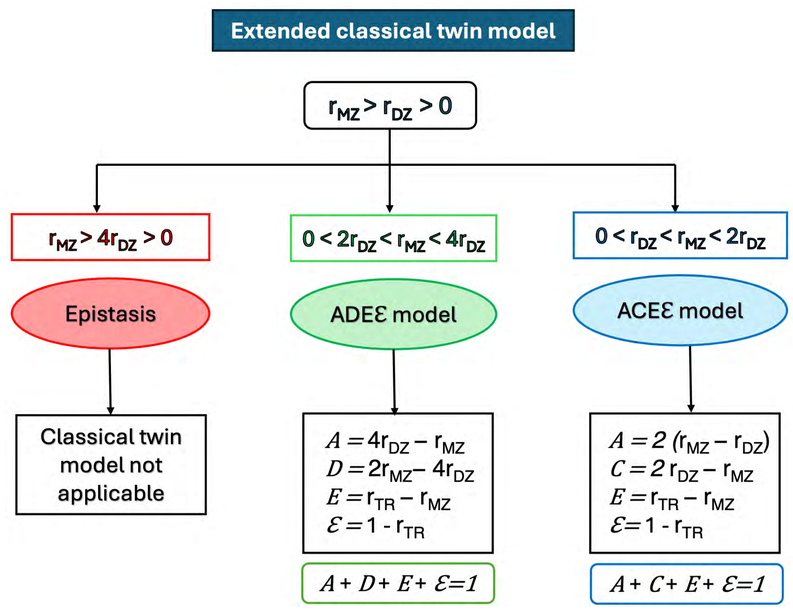


Extended Classical Twin Model

We use classical twin model to separate trait variance into proportions attributable to genetic and environmental factors, most commonly using the ACE and ADE models.

The ACE model separates variance into **additive genetic (A)**, **shared environmental (C)**, and **unshared environmental (E)** components by comparing trait similarity between monozygotic (MZ) twins (~100% genetic similarity) and dizygotic (DZ) twins (~50%). Using correlations between twins (r_{MZ}, r_{DZ}), the model computes A, C and E ensuring A + C + E = 1.

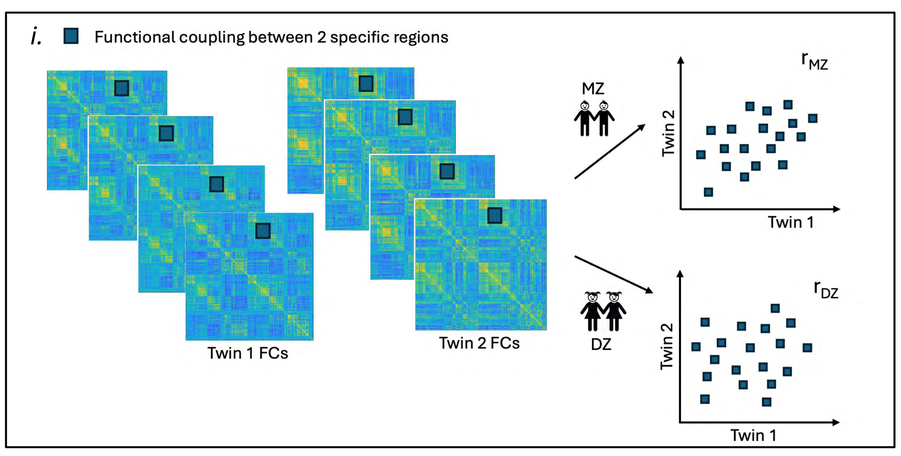
The ADE model replaces the shared environment (C) With **dominant genetic effects (D)** to capture non-additive genetic contributions.



In our study, the trait of interest is functional coupling (FC) between brain regions. Since FC estimation may include measurement noise, we **extend the ACE/ADE model** by introducing a **test-retest correlation** (r_{TR}) to separate E into a **true unshared environmental component E** = $r_{TR} - r_{MZ}$ and a **measurement error component** ($\mathcal{E} = 1 - r_{TR}$), such that A + C/D + E + $\mathcal{E}$ = 1.

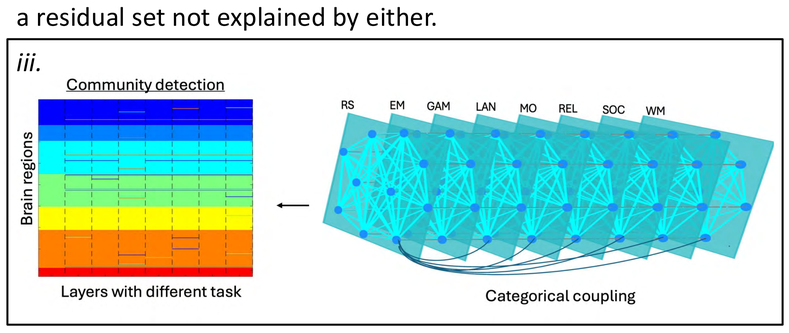
Methodological Framework

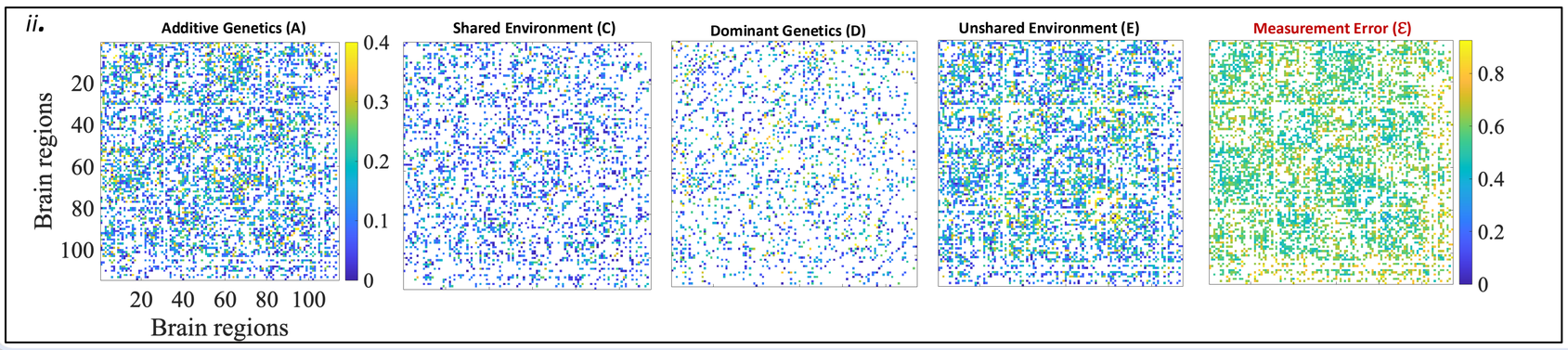
We used a multi-layer community detection framework, with each layer representing one condition, to uncover functional circuits that display a differentiated heritability and/or shared environment across all fMRI conditions. In a multiplex network, each layer contains the same set of nodes but different types of interactions, with inter-layer links connecting corresponding mirror nodes across layers. We constructed 8-layer multiplex networks (one per fMRI condition) for each source of variance (A, C, E, and $\mathcal{E}$), where intra-layer connections represent the heritability (in case of A) within each task, and inter-layer connections link the same node across different fMRI conditions.



These networks are categorical, with all-to-all interlayer structure limited to mirror node connections.

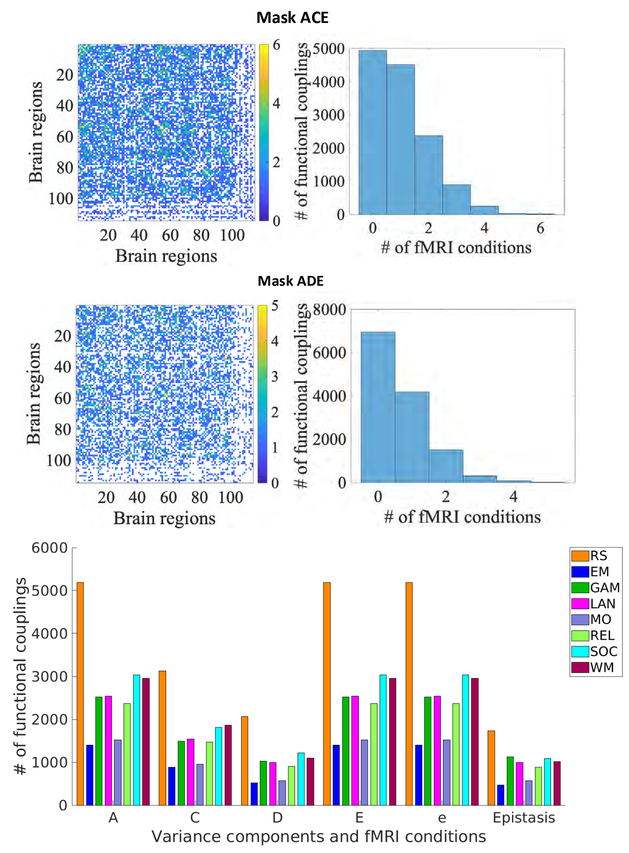
Functional couplings were partitioned across models, with a subset explained by the ACE model, a lesser subset by the ADE model, and a residual set not explained by either.





ACEε and ADEε results on fMRI conditions

Different brain regions and tasks favor distinct genetic models, reflecting condition-specific heritability structures, highlighting heterogeneous genetic influences across brain circuits. This suggests that heritability and environmental factors effects jointly shape brain network connectivity.



References :

Blokland, G. et. al (2013). Twin studies and behavior genetics. In The Oxford Handbook of Quantitative Methods in Psychology, 2, 190-218.

Mucha, P. et. Al (2010). Community structure in time-dependent, multiscale, and multiplex networks. Science, 328, 876-878.

PURDUE INSTITUTE FOR INTEGRATIVE NEUROSCIENCE
Where science meets engineering

Edwardson School of Industrial Engineering

rethink

Poster Number 22

Sparse functional networks reveal backbone structures associated with individual fingerprinting

Daniel Guerrero

Industrial Engineering, Purdue University

Fingerprint in functional connectomes (FC) may enable the development of personalized treatments for neurodegenerative disorders and advance our understanding on how the functional architecture of brain networks relates to cognitive function. Most studies have assessed fingerprinting by measuring identification rates based on fully connected (i.e. unfiltered) functional brain networks. In this work, we assessed how representations of functional connectomes at different density levels (from fully connected to very sparse) uncover cohort-level and subject-specific backbones that produce higher fingerprinting. To do so, we investigated the impact of coupling thresholding, random filtering, and disparity filter on identification rates at a wide range of sparsity. We analyzed 426 unrelated subjects from the Young Adult Human Connectome Project. Our results show how very sparse network regimes (density around 8%; Î±=0.1) increased fingerprinting precision (92%) by a magnitude of a 7% with respect to baseline (unfiltered; Î±=1).

Sparse functional networks reveal backbone structures associated with individual fingerprinting

Daniel Guerrero[1], Mintao Liu[1], Mario Dzemidzic[2], Mahdi Moghadam[1], Yujing Zhang[1], David Kareken[2], Jaroslaw Harezlak[3], Joaquín Goñi[1]

[1]School of Industrial Engineering, Purdue Institute for Integrative Neuroscience, Purdue University
[2]School of Medicine, Indiana University.
[3]School of Public Health, Indiana University

CONN*plexity* Lab

Summary and Hypothesis

Fingerprint in functional connectomes (FC) may enable the development of personalized treatments for neurodegenerative disorders and advance our understanding on how the functional architecture of brain networks relates to cognitive function. Most studies have assessed fingerprinting by measuring identification rates based on fully connected (i.e. unfiltered) functional brain networks. In this work, we assessed how representations of functional connectomes at different density levels (from fully connected to very sparse) uncover cohort-level and subject-specific backbones that produce higher fingerprinting. To do so, we investigated the impact of coupling thresholding, random filtering, and disparity filter on identification rates at a wide range of sparsity. We analyzed 426 unrelated subjects from the Young Adult Human Connectome Project. Our results show how very sparse network regimes (density around 8%; α=0.1) increased fingerprinting precision (92%) by a magnitude of a 7% with respect to baseline (unfiltered; α=1).

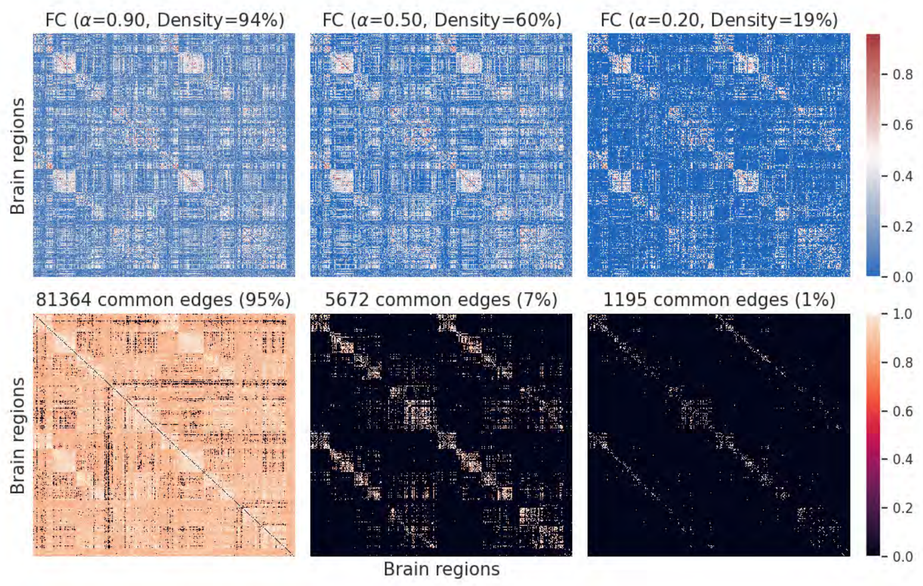
Identification Rate

We use mean squared error (MSE) to measure the pair-wise distance between functional connectomes to distinguish an individual from a population. A process referred to as fingerprinting or subject identification. As FCs capture intrinsic subject-level connectivity profiles, they should be reliable at differentiating individuals.

$$ID\ rate = \frac{Number\ of\ correctly\ identified\ participants}{Total\ number\ of\ participants}$$

Disparity filter (DF)

Disparity Filter is an edge filtering method that uses the weight distribution to infer relevant connections in multiscale networks. The filtering process is informed by the heterogeneous connectivity pattern of individual nodes to produce a network backbone that captures the relevant interactions at different scales.

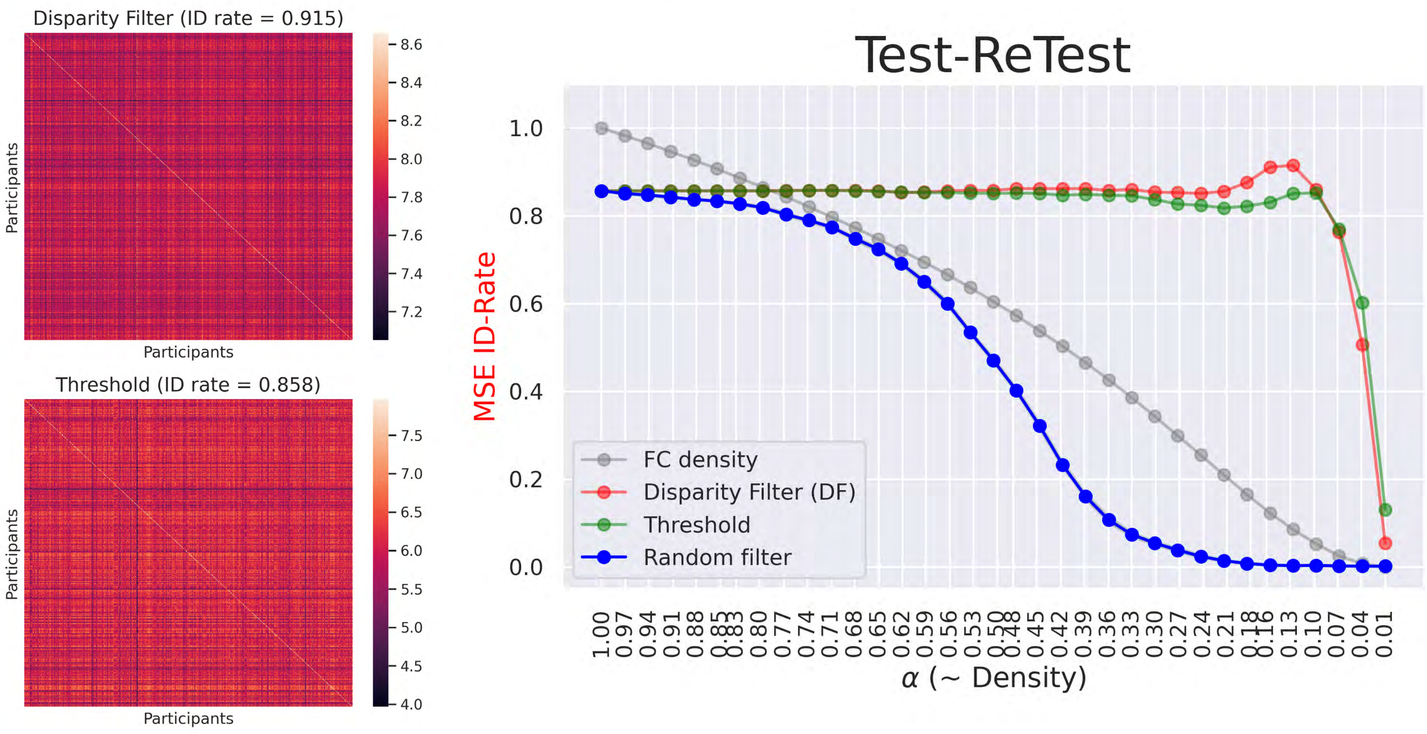


Edges with high finger printing at each density level

Limitations

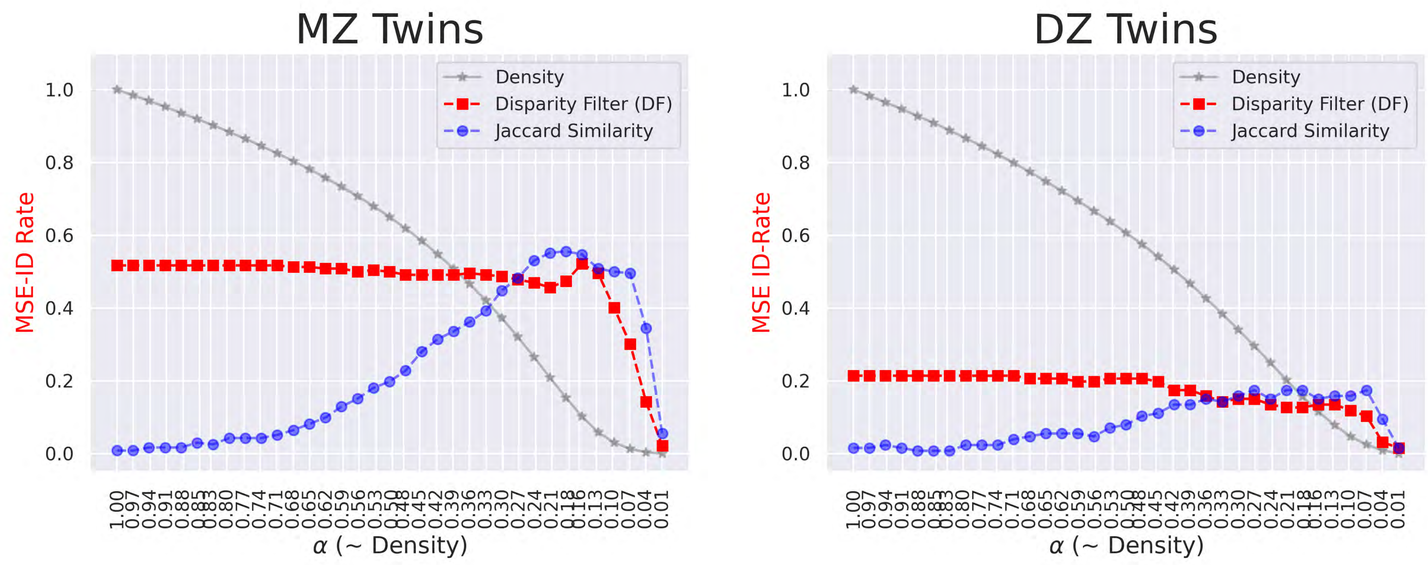
- Theoretically
  - Networks with very different topologies (degree distribution) can be challenging to filter using DF
  - The filtering procedure does not consider the type of information being carried by the edges only their relative magnitudes
- Experimentally
  - In datasets with only one session (scan) available, it is not possible to use ID rate to determine the optimal density level to analyze the connectome

Network density and identification rate



Sparse functional connectomes (sFC) result in higher identification rates than dense connectomes (FC) across resting state sessions. We compare several sparsification methods (disparity filter DS, simple thresholding and random filter) for different density levels. The disparity filter better identifies topological structures with systematically higher identification rates for low density FCs.

Twin-based backbone analysis



Comparison of ID rates when using disparity filter (red) and Jaccard similarity of the filtered FCs backbones at any density level. Our result shows that there are participant-characteristic functional architectures of the brain (comprising around 10% of the functional edges) that lead to high fingerprinting. Such subject-level set of backbones accurately differentiate individuals, and to some extent MZ twins, in large cohorts. This advances our understanding of the functional organization of the brain, its replicability across repeated scans, and its heterogeneity across participants.

NIH CTSI CTR EPAR2169, NIH R21 AA029614, NIH R01 AA029607

Poster Number 23

Predicting Th2 Cytokine Activation in Asthma; A Single-Cell and Bulk RNA-Seq Validation Study

Kianmehr Aalipour

Department of Epidemiology and Biostatistics, School of Public Health, IU - IUB - IUI

Introduction: Asthma is a heterogeneous lung disease with diverse clinical presentations and underlying biological mechanisms. Overactive T helper type 2 (Th2) inflammation is known as a major driver of asthma, especially in mild to moderate and most severe cases. Single-cell RNA sequencing (scRNA-seq) offers high-resolution views of cytokine-producing cells, but it remains difficult to truly identify cytokine activities. Transcript dropout and cellular variability make it difficult to identify cytokine activities in healthy and disease states.
Objective: To develop a predictive model of Th2 cytokine activity at the single-cell level using human allergen-activated CD4â□° T cells and validate these signatures in a bulk RNA-seq dataset from adults with varying asthma severity.
Methods: We analyzed scRNA-seq data (GSE146170) from allergen-stimulated CD4â□° T cells of eight adult patients with allergic asthma. Principal component analysis (PCA) was used to assess transcriptomic variation among Th2 cells. A Lasso-regularized logistic regression model was trained to classify Th2 cells expressing IL5 or IL13 above a scaled threshold (>0.5). IL4 was excluded due to excessive dropout. Model performance was evaluated using 5-fold cross-validation. In this regard, we applied a Th2 Activity Scoreâ€”defined as the average normalized expression of canonical Th2 genes (IL4, IL5, IL13, GATA3, etc.)â€”to bulk RNA-seq data (GSE130499) from adult subjects categorized by asthma severity (mild, moderate, or severe) and healthy controls.
Results: Among 1,755 Th2 cells, 22.0% expressed IL5 only, 10.3% expressed IL13 only, 26.6% co-expressed both, and 41.2% expressed neither cytokine, indicating high transcriptional heterogeneity. These findings motivated a transcriptome-wide modeling approach. In the bulk data, the Th2 Activity Score increased with asthma severity and differed significantly across groups. CLCA1 was the only Th2-pathway gene with significantly different expression across severity levels (adjusted p = 0.0013). Although IL4 was not robustly detected in single cells, it contributed to the bulk-level activity score.
Conclusion: Single-cell gene expression profiles can be used to accurately model cytokine activation, and these signatures remain robust when validated in independent bulk RNA-seq data. This integrative strategy highlights the biological and translational relevance of Th2-driven inflammation in severe asthma.
Keywords: Asthma, Cytokine Activity, Th2 Cells, Machine Learning

Predicting Th2 Cytokine Activation in Asthma; A Single-Cell and Bulk RNA-Seq Validation Study

Kianmehr Aalipour[1], Kedir N Turi[1]

[1]Department of Epidemiology and Biostatistics, School of Public Health, Indiana University, Bloomington
Email Address: ezaali@iu.edu

Introduction

Asthma is a heterogeneous lung disease with diverse clinical presentations and underlying biological mechanisms. Overactive T helper type 2 (Th2) inflammation is known as a major driver of asthma, especially in mild to moderate and most severe cases. Single-cell RNA sequencing (scRNA-seq) offers high-resolution views of cytokine-producing cells, but it remains difficult to truly identify cytokine activities. Transcript dropout and cellular variability make it difficult to identify cytokine activities in healthy and disease states.

Objective: To develop a predictive model of Th2 cytokine activity at the single-cell level using human allergen-activated CD4⁺ T cells and validate these signatures in a bulk RNA-seq dataset from adults with varying asthma severity.

Methods

We analyzed scRNA-seq data (GSE146170) from allergen-stimulated CD4⁺ T cells of eight adult patients with allergic asthma. Principal component analysis (PCA) was used to assess transcriptomic variation among Th2 cells. A Lasso-regularized logistic regression model was trained to classify Th2 cells expressing IL5 or IL13 above a scaled threshold (>0.5). IL4 was excluded due to excessive dropout. Model performance was evaluated using 5-fold cross-validation. In this regard, we applied a Th2 Activity Score—defined as the average normalized expression of canonical Th2 genes (IL4, IL5, IL13, GATA3, etc.)—to bulk RNA-seq data (GSE130499) from adult subjects categorized by asthma severity (mild, moderate, or severe) and healthy controls.

Results

Among 1,755 Th2 cells, 22.0% expressed IL5 only, 10.3% expressed IL13 only, 26.6% co-expressed both, and 41.2% expressed neither cytokine, indicating high transcriptional heterogeneity. These findings motivated a transcriptome-wide modeling approach. In the bulk data, the Th2 Activity Score increased with asthma severity and differed significantly across groups. CLCA1 was the only Th2-pathway gene with significantly different expression across severity levels (adjusted p = 0.0013). Although IL4 was not robustly detected in single cells, it contributed to the bulk-level activity score.

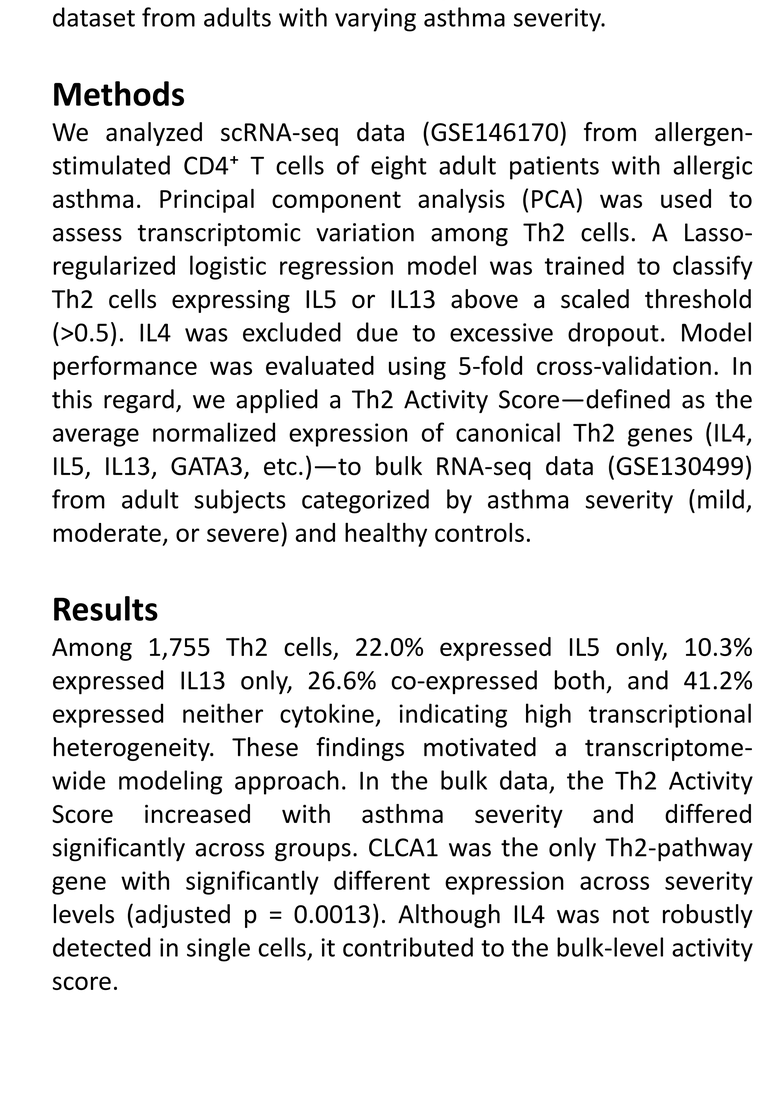


Figure 1: Violon Plots of Key Quality Metrics

Figure 1 shows violin plots of key quality metrics for the Th2 cells, including the number of genes detected per cell, total UMI counts, and percentage of mitochondrial transcripts. These metrics confirm high cell quality prior to modeling. Most cells expressed between 500–1500 genes and had total UMI counts ranging from 1000–6000. Importantly, mitochondrial transcript percentage remained below 10% for most cells, indicating high cell quality and successful removal of stressed or dying cells. These quality control results support the reliability of downstream modeling.

Table 1: Principal Component Analysis

	Top Positive Loadings	Top Negative Loadings
PC1	IL5, IL13, IL4, CSF2, TMSB4X, DUSP2, GADD45G	NCL, DDX21, HSP90AB1, METAP2, RPL36A, EIF2S2
PC2	IL5, IL9, IL1RL1, IL17RB, PRDM1, JAK1	IL2, IL4, CD40LG, TNF, NR4A2, PMAIP1
PC3	ACTB, FABP5, HSPA8, CD2, TXN, PRDX1	IL13, IL5, REL, IL7R, NEAT1, CSF2, INSIG1

Table 1 shows that Th2 cytokines are major drivers of transcriptional variability in this dataset — especially in early PCs (1–2), which supports their biological importance in clustering or modeling

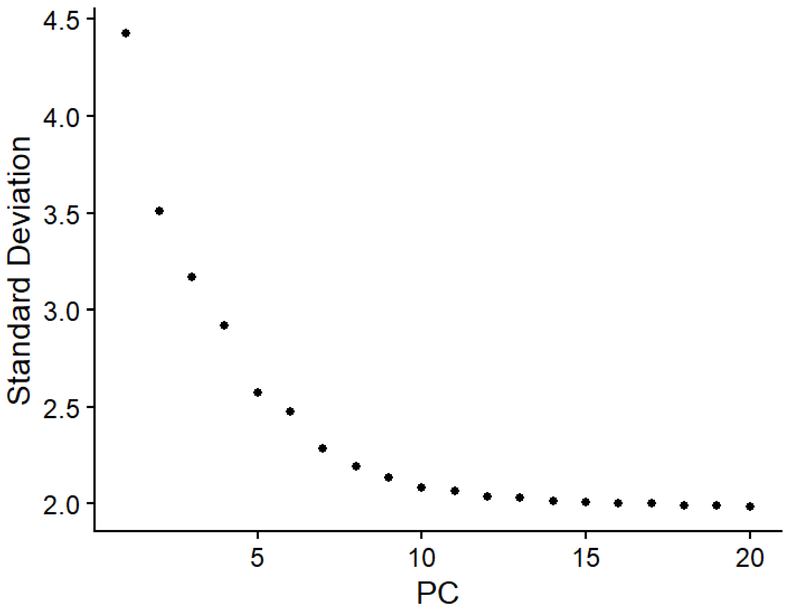


Figure 2. PCA Scree Plot of Variance Explained by Principal Components

The first few principal components capture the majority of variance in the gene expression matrix. PCs 1 through 5 explain the most variation, with a clear elbow observed around PC6–7, after which additional components contribute marginally.

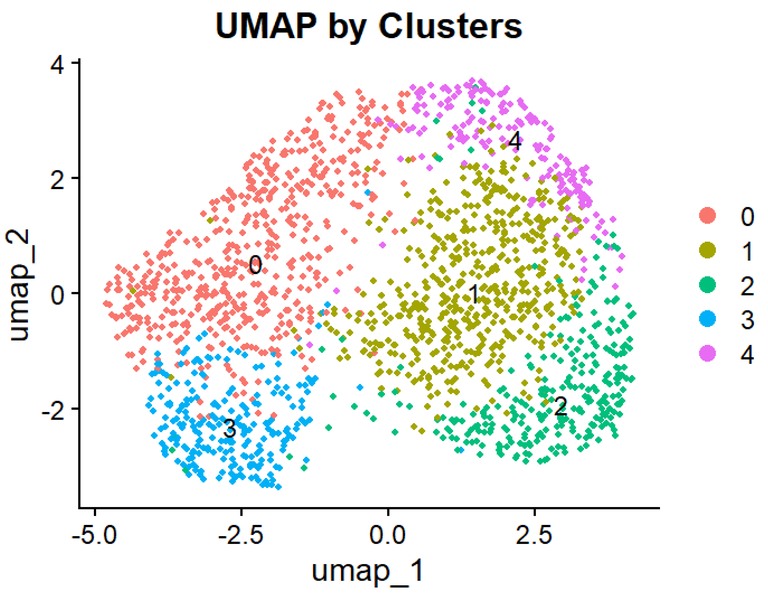


Figure 3: UMAP Visualization of Th2 Cells Clustered by Gene Expression

Figure 3 demonstrates that Uniform Manifold Approximation and Projection (UMAP) was used to embed 1,766 Th2 cells into two dimensions based on PCA-reduced expression of the top 500 variable genes. Cells are colored by unsupervised cluster identity (0–4). The presence of multiple distinct clusters suggests functional heterogeneity within the Th2 population.

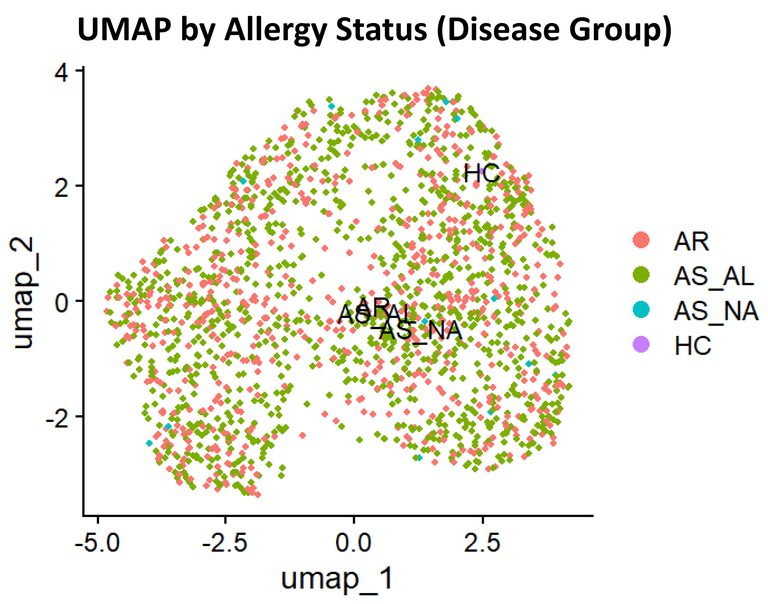


Figure 4: UMAP Projection of Th2 Cells Colored by Clinical Allergy Status

Figure 4 indicates that Cells from patients with asthma and/or allergy (AS_AL, AS_NA, AR) and healthy controls (HC) are colored according to the disease group annotation. While all groups are transcriptionally mixed, slight enrichment of certain clusters by disease status is observed.

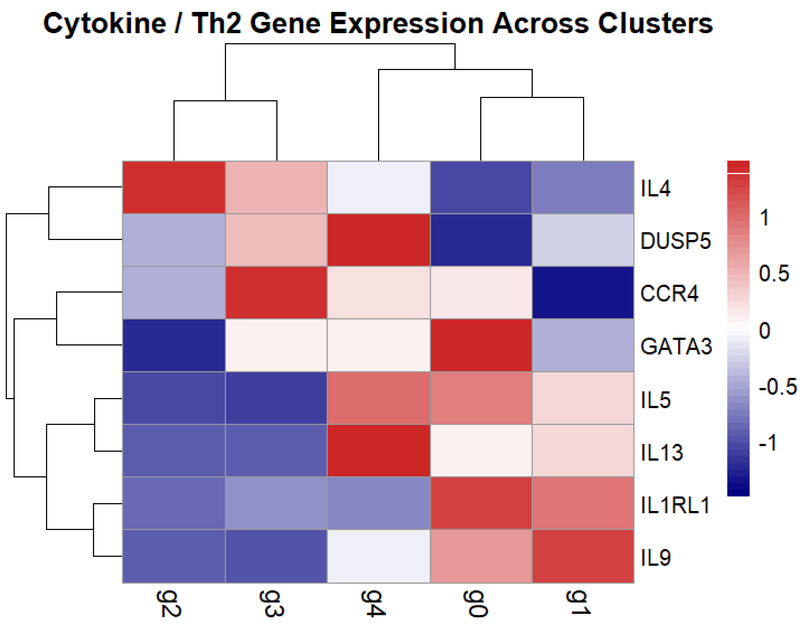


Figure 4: Heatmap of Th2-Related Gene Expression Across Clusters

Average expression of eight Th2 signature genes was computed for each unsupervised cluster (g0–g4) and scaled (Z-score) across rows. High expression (red) and low expression (blue) highlight functionally distinct Th2 subsets. For example, cluster g0 shows elevated IL5 and IL13, while cluster g1 is enriched for IL9 and IL1RL1.

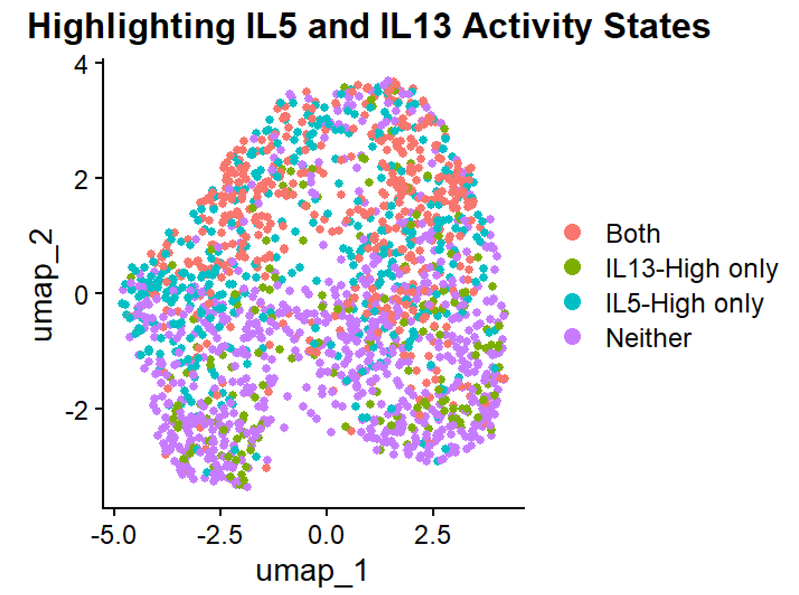


Figure 5: UMAP of Th2 Cells Colored by IL5 and IL13 Activity States

Average expression of eight Th2 signature genes was computed for each unsupervised cluster (g0–g4) and scaled (Z-score) across rows. High expression (red) and low expression (blue) highlight functionally distinct Th2 subsets. For example, cluster g0 shows elevated IL5 and IL13, while cluster g1 is enriched for IL9 and IL1RL1.

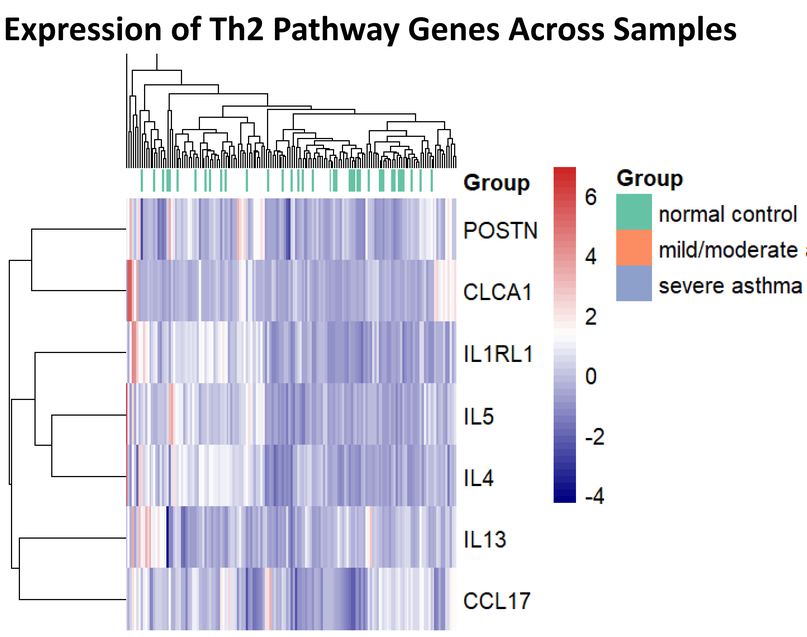


Figure 6: Expression of Th2 Pathway Genes Across Samples by Asthma Severity (GSE130499).

Hierarchical clustering heatmap of normalized gene expression for seven Th2-related genes in BAL cell bulk RNA-seq samples. Samples from severe asthmatics exhibit distinct upregulation patterns (blue bar), while normal controls show lower expression levels (green bar), highlighting the transcriptional activation of Th2 immune pathways in more severe disease states.

Conclusion

Single-cell gene expression profiles can be used to accurately model cytokine activation, and these signatures remain robust when validated in independent bulk RNA-seq data. This integrative strategy highlights the biological and translational relevance of Th2-driven inflammation in severe asthma.