***AI & Health* Symposium**
Friday, September 12
9:00 am - 4:00 pm
North Campus Research Complex
Building 18, Dining Hall

12 - 1:15p

Poster Session (with Lunch)

Posters are presented in alphabetical order according to the first author's last name

Poster 1

Patient Interpretations and Preferences for the Disclosure of Artificial Intelligence Generated Draft Replies to Electronic Messages

Philip Barrison[1] Jodyn Platt[1] Mark Ackerman[2] Charles P. Friedman[1] Alexandra Vinson[1]

[1] Department of Learning Health Sciences, Michigan Medicine
[2] School of Information, University of Michigan (UMSI)

Background: Since their initial public availability, artificial intelligence-generated draft replies (GDRs) to patient portal messages have undergone rapid implementation and evaluation at healthcare systems across the United States. Amid their adoption, experts have debated the extent to which the technology should be disclosed to patients. In these ongoing conversations, patient expectations for disclosure practices are not well understood. This study examines patients' reactions to and preferences for written AI disclosure statements presented at the time of message delivery. Methods: Semi-structured interviews were conducted with 30 adults who had no prior experience with GDRs. Participants shared their initial interpretations and reactions to portal messages with written disclosure statements provided through four case-based vignettes. Following this, participants were informed about how the portal messages were drafted and allowed to ask questions about GDRs. Their preferences and expectations for disclosure methods were elicited at this time. Results: In reacting to GDRs with written disclosure statements, some

participants expressed concerns about error, authorship, and depersonalization. Participants who desired disclosure practices based their reasoning on a perceived right to information, agency, and autonomy in their healthcare. While some participants expressed concerns elicited by written disclosure statements, they argued for preemptive disclosure methods that facilitate education and trust building. Conclusion: As healthcare systems establish best practices for transparency and disclosure of GDRs, accounting for evolving patient perspectives on AI will be crucial to maintaining patient trust and confidence. The reactions and preferences for disclosure methods outlined by participants in this study suggest that careful deliberation should be given not to whether to disclose, but how to disclose, as healthcare systems balance tensions between AI transparency and patient trust.

Poster 2

Implementing a Standardized Squat Classification System: A Clinical and Educational Innovation

Batoul Baydoun[1] Manar Chebli[1] Omar Abdalla[1]
Alexa Jaume[2] MD Ameen Suhrawardy[2] MD Rahul Vaidya[2] MD

[1] Wayne State University
[2] Department of Orthopaedic Surgery - Detroit Medical Center

Introduction: Current orthopaedic exams lack a standardized method to assess global lower extremity (LE) function. The squat, a multi-joint, compound movement, reflects mobility, proprioception, muscle strength, balance, and pain, making it a potential benchmark for comprehensive LE evaluation. This study introduces and validates a novel squat classification system. *Methods:* The system consists of two components: squat depth (Grades 1–4), reflecting eccentric strength, range of motion (ROM), and pain; and ability to rise (Grades A–C), reflecting concentric strength of the extensor mechanism. Twenty-five patients were photographed at maximal squat depth during both assisted and unassisted squats. Hip, knee, and ankle flexion angles were measured using DetroitBoneSetter.com. Four cohorts—orthopaedic residents (n=2), medical students (n=2), undergraduate students (n=3), and ChatGPT—classified each patient. The senior resident's ratings served as the gold standard. Inter-user agreement was calculated with Fleiss' Kappa, and accuracy was defined as percent agreement with the gold standard. *Results:* Medical students achieved near-perfect agreement ($\kappa = 0.83$), followed by residents ($\kappa = 0.77$).

Undergraduates (κ = 0.45) and ChatGPT (κ = 0.49) showed moderate agreement. Accuracy for depth grading was 83.3% for residents and medical students, compared to 59.7% for undergraduates and 41.7% for ChatGPT ($p<0.05$). All groups achieved 100% accuracy for ability-to-rise grading. *Conclusion:* This standardized squat classification system demonstrated high reliability, teachability, and reproducibility, particularly among medical students and residents. Its simplicity and strong performance suggest utility as a practical LE physical exam tool and rehabilitative benchmark. Future work should explore correlations with patient-reported outcomes and its role in tracking functional recovery over time.

Poster 3

Supporting collaborative health management of children with Type 1 Diabetes and their parents

Yoon Jeong Cha[1] Sun Young Park[1] Mark Newman[1]

[1] School of Information, University of Michigan (UMSI)

Type 1 Diabetes (T1D), commonly diagnosed in early childhood, requires lifelong monitoring and treatment to maintain safe glucose levels. Effective management in children often depends on collaboration between parents, children, and other caregivers. Our multi-year research, presented at CHI 2022–2024, investigates how families co-manage T1D and how children can gradually develop independence in their self-care. In CHI 2022, we conducted interviews with 20 child-parent pairs to understand how children build knowledge and motivation for managing their condition. We found that children's understanding of their illness and their confidence in self-care were critical for transitioning to more independent roles. In CHI 2023, we identified four parental strategies—educated guessing, contingency planning, experimentation, and seeking help—categorized along two dimensions: whether the cause of risk is known or unknown, and whether the risk is predictable or unpredictable. In CHI 2024, a three-week probe study with 22 families revealed that shared responsibility in tracking diabetes data—what we call collaborative tracking—can support both child learning and parental oversight. From this, we outlined four tracking approaches: child-independent, child-led, parent-led, and parent-independent. Building on these insights, we recently conducted a series of co-design workshops with children with T1D and their parents to ideate and prototype AI technologies that support collaborative tracking. Through storytelling, drawing activities, and

discussion, we gathered rich input on how children and parents envision technology that helps them manage diabetes together—focusing on themes such as independence, safety, privacy, and decision-making support. These workshops informed the design of an AI agent that facilitates data sharing and joint reflection between parents and children. We are now preparing for a field deployment study to evaluate how this tool functions in families' everyday lives and to assess its impact on engagement, learning, and communication in T1D management.

Poster 4

Toward "AI Scientists" in Practice: A Human–LLM Workflow for Clinical Temperature Phenotyping

Douglas B. Craig[1]

[1] Department of Emergency Medicine, Michigan Medicine

We present a streamlined, reproducible workflow in which a researcher partners with a large language model (GPT-5) and an LLM-enabled coding environment (Cursor) to accelerate the classic cycle of observation → hypothesis → experimental design → interpretation for clinical time-series research. The workflow makes the LLM a methodological co-investigator: (i) formalizing observational prompts into falsifiable hypotheses; (ii) proposing alternative analytic designs; (iii) generating and refactoring executable code for feature engineering and modeling; and (iv) critiquing findings by connecting patterns to mechanistic and clinical literature for interpretation and next-step planning. Human oversight remains central at each gate, approving study assumptions, validating code and outputs, and selecting analyses to advance, while the system captures provenance to support reproducibility and collaboration. As a case study, we examine whether distinct high-resolution temperature dynamics constitute meaningful and predictive phenotypes. Starting from the observation that survival appears to vary with temperature variance, we use LLM-guided design to implement wavelet-based feature extraction and unsupervised clustering of temperature time series, followed by LLM-assisted interpretation that maps clusters to plausible physiological, pharmacologic, and care-process explanations and frames testable follow-up analyses. The same pattern, co-design, co-coding, and co-interpretation, generalizes to scientific research more broadly. This contribution emphasizes method over findings: a practical

recipe for human-LLM co-discovery that is transparent, auditable, and readily integrated with resources for data, compute, and clinical workflow integration.

Poster 5

Towards Personalized Multimodal Efficient Detection of Human Circadian States

Kapotaksha Das[1]

[1] University of Michigan–Dearborn

Circadian rhythms govern essential processes such as sleep, appetite, and hormone levels, with disruptions linked to numerous health issues, particularly among college students. This paper introduces a novel, non-invasive, multimodal approach for reliably detecting aspects of circadian rhythms. We present a synchronized dataset from 16 subjects, utilizing thermal and visual imagery, physiological channels, and survey data. We analyze how different window sizes and overlaps affect segmentation and model performance. Additionally, we enhance prediction accuracy through personalized models using subject-specific training data and one-time survey responses, achieving up to 77.97% mean F1 score for time-based labels and 69.12% for self-reported sleepiness scores. Our findings underscore the importance of personalized modeling. They show that incorporating chronotype-focused survey data and optimizing window sizes significantly enhances model accuracy. This approach sets the stage for practical health monitoring and intervention personalization applications. Future work will aim to incorporate additional modalities, such as audio data, and explore advanced machine learning techniques to improve circadian state detection further.

Poster 6

Large Language Models for the Prediction of Hospital Discharge in Inpatients

Karan Desai[1] Natalie Guzman[1] Anthea Bell[1] Veronica Gilbert[1] Samuel Lehn[1]
Jonathan Spagnoli[2] MD Andrew Wong[2] MD

[1] University of Michigan Medical School
[2] Department of Internal Medicine, University of Michigan,

Introduction: Accurate discharge prediction is essential for hospital operations, impacting patient throughput, bed utilization, staffing, and overall care coordination. Prior research has demonstrated that large language models (LLMs) can be effective with operational tasks in healthcare, including predicting length of stay, hospital readmissions, insurance denial, and morbidity and mortality.1 While LLMs show promise for these use cases, their performance can vary significantly across specialties and case complexity.2 We analyzed LLM discharge prediction performance among inpatients at a tertiary center to identify specific barriers and develop targeted interventions for optimizing model accuracy. *Methods:* We conducted a qualitative analysis of LLM-generated same-day discharge predictions on 1000 inpatients at a large, tertiary academic medical center from January-December 2024. Using Llama 3.3-70B, predictions were generated through two-step prompting: binary prediction (1=discharge, 0=no discharge) followed by reasoning. The model received clinical notes from the previous day through 6am on prediction day. To maximize specificity, the model was instructed to assign "1" only if 100% certain of discharge. We examined 202 prediction errors (24 false positives, 178 false negatives). Four trained team members reviewed clinical notes to assess algorithm validity, identify reasons for prediction-outcome discordance, and extract representative quotes. Errors were categorized by clinical service and type to identify opportunities for targeted model improvements. *Results:* False positives often occurred when the model lacked knowledge of standard clinical workflows, when patients could theoretically be discharged but remained for care planning or shared decision-making, or when documentation was sparse or outdated (Table 1). False negatives occurred when the model was overly conservative, overemphasized early timeline events, failed to recognize outpatient manageable needs, or overemphasis on medical complexity rather than current clinical stability. *Conclusions:* To accurately predict inpatient discharge, LLMs require integration with speciality and workflow specific clinical knowledge for optimal performance. Our analysis revealed service specific intervention opportunities including targeted prompting for recovery protocols, enhanced temporal reasoning, and improved distinction between inpatient versus outpatient management. Addressing these limitations could enable better resource allocation, improved efficiency, and safer discharge decisions, ultimately enhancing patient outcomes and healthcare operations.

Poster 7

AI vs Human Expertise: Measuring Squat Joint Angles Across Clinical Training Levels

Abdurrahman Ehsan[1] Zarrar Chaudhri[1] Omar Abdalla[1]

Alexa Jaume[2] MD Ameen Suhrawardy[2] MD Rahul Vaidya[2] MD

[1] Wayne State University
[2] Department of Orthopaedic Surgery - Detroit Medical Center

Objective: To evaluate the efficacy of ChatGPT and individuals with varying levels of clinical training in measuring joint angles from squat images, assessing accuracy and variability to determine how training influences interpretation. *Methods:* Images from 25 patients performing unassisted and assisted squats were analyzed at maximum squat depth. Four cohorts—Orthopaedic Resident (1), Medical Students (2), Undergraduates (3), and ChatGPT—measured hip, knee, and ankle angles using DetroitBoneSetter.com. A Wilcoxon signed-rank test assessed statistical differences between cohorts; Coefficient of Variation (CV) measured within-group variability. *Results:* Significant differences in hip, knee, and ankle angles were found between undergraduates/medical students and the resident ($p<0.001$). Medical students and undergraduates also differed in knee angles ($p<0.001$), and ChatGPT differed from the resident in ankle angles ($p<0.001$). ChatGPT showed the highest hip CV (33.97%), while residents had the highest knee and ankle CVs (31.49%, 8.2%). Undergraduates had the lowest knee CV (22.28%), and ChatGPT had the lowest ankle CV (5.93%). *Conclusions:* Undergraduates and medical students consistently deviated from resident measurements, suggesting limited clinical exposure affects accuracy. ChatGPT's hip and knee measurements did not significantly differ from the resident's, highlighting its potential in image interpretation. However, its high hip CV suggests occasional large errors, reinforcing the need for physician oversight. While students showed lower variability, their significant deviations from the resident's "gold standard" indicate reliance on oversimplified measurement strategies that overlook patient-specific factors. Findings underscore the role of advanced clinical training in accurate image assessment and suggest AI's complementary role in clinical settings when supervised.

Poster 8

Temporal Harmonization: Improved Detection of Mild Cognitive Impairment from Temporal Language Markers using Subject-invariant Learning

Bao Hoang[1] Siqi Liang[2] Yijiang Pang[1] Hiroko Dodge[3] Jiayu Zhou[2]

[1] Michigan State University
[2] University of Michigan
[3] Harvard Medical School

Mild Cognitive Impairment (MCI) is an early stage of dementia characterized by cognitive decline and behavioral changes. Early detection is crucial for timely interventions, improved clinical trial cohort selection, and the development of targeted therapies. Linguistic markers have recently emerged as a non-invasive, cost-effective method for MCI detection. This study analyzes linguistic markers from conversations between participants and healthcare professionals to distinguish MCI from cognitively normal (NL) individuals. The dynamics of multiple conversations of a subject carry fine-granular linguistic change over time and expect to greatly enhance detection accuracy. However, individual variations in speaking styles pose challenges for learning cognitive characteristics from temporal sequences of conversations. To address this, we propose a temporal harmonization method to mitigate distributional differences in linguistic features across subjects, improving model generalization. Our results show that machine learning models leveraging subject-invariant harmonized temporal features greatly improve the prediction performance of MCI detection from multiple conversations.

Poster 9

Segmentation of early-stage reticular pseudodrusen with limited annotations

Anna Kay[1] Jason Miller[1] Stella Yu[1]

[1] College of Engineering, University of Michigan (CoE)

Age-related macular degeneration (AMD) is the leading cause of irreversible blindness in the United States, affecting over one third of individuals above age eighty. Reticular pseudodrusen (RPD) are a pathologic finding in AMD that are particularly difficult to segment on routine clinical imaging since they are small and indistinct. We developed a model for segmenting RPD from optical coherence tomography (OCT) images. Our model was trained with minimal annotations, using largely self-supervised techniques, and it identifies over thirty times more RPD than the leading prior model. With automated segmentation, we will be able to study AMD progression on a more granular level, and the methods we developed may be applied to other research questions in medical imaging and especially retinal analysis.

Poster 10

Chain-of-Generation: Progressive Latent Diffusion for Text-Guided Molecular Design

Lingxiao Li[1]

[1] School of Information, University of Michigan (UMSI)

Text-conditioned molecular generation seeks to translate natural-language descriptions into novel chemical structures, opening an intuitive channel for scientists to specify functional groups, scaffolds, and physicochemical constraints without handcrafted rules. Diffusion-based models—particularly latent diffusion models (LDMs)—have recently shown promise by performing stochastic search in a continuous latent space that compactly captures molecular semantics. Yet current LDM approaches suffer from weak cross-modal alignment, coarse one-shot conditioning, and evaluation metrics that overlook semantic fidelity. We propose Chain-of-Generation (CoG), a multi-stage latent diffusion framework that decomposes an input prompt into a sequence of semantic sub-prompts. CoG progressively incorporates these segments to define intermediate goals, guiding the denoising trajectory toward molecules that satisfy increasingly rich linguistic constraints. A post-alignment learning phase strengthens the correspondence between textual and molecular latent spaces, while an enhanced denoising network exploits this alignment during generation. To assess fidelity and interpretability, we introduce deep qualitative trajectory analysis, tracing how individual prompt components manifest as structural motifs across diffusion steps. Extensive experiments on benchmark and real-world tasks demonstrate that CoG yields higher semantic alignment, diversity, and controllability than one-shot baselines, producing molecules that more faithfully reflect complex, compositional prompts while offering transparent insight into the generation process.

Poster 11

AI-Supported Spanish Counseling Training: Enhancing Health Behavior Counseling through NLP and Educational Feedback Systems

Do June Min[1,3] Longju Bai[1] Ken Resnicow[2] John D. Piette[2]

Veronica Perez-Rosas[1,4] Rada Mihalcea[1]

[1] Computer Science & Engineering (CSE), University of Michigan
[2] School of Public Health (SPH), University of Michigan
[3] Amazon
[4] Texas State University

Motivational Interviewing (MI) is a key counseling skill for improving health outcomes, yet structured training is often inaccessible in Spanish-speaking, low-resource settings. We developed an AI-supported, web-based system that aims to train and evaluate MI skills, especially the reflection, among health professionals in Honduras and Mexico. Using a Spanish-adapted NLP model based on the PAIR framework, the system aims to score reflection responses and provide automated, example-based feedback via a mobile-optimized interface. Ongoing work focuses on extending the system to other foundational MI skills, calibrating prompt difficulty, and exploring subtitled video modules to further enhance accessibility and scalability.

Poster 12

MedCompReviser: Revising Patient Education Materials for Improved Comprehension and Actionability

Joan Nwatu[1] Michael Rubyan[2] Rada Mihalcea[1]

[1] Computer Science & Engineering (CSE), University of Michigan
[2] School of Public Health (SPH), University of Michigan

Patient education materials (PEMs) play a crucial role in informing patients about their illnesses, treatment options, and self-management practices. Yet, prior studies consistently show that many PEMs are written at reading levels that exceed the abilities of a large portion of the public, with the average adult in the United States reading at or below a sixth-grade level. This gap in readability can limit comprehension, reduce engagement, and ultimately undermine health outcomes. To address this challenge, we propose MedCompReviser, a system that leverages large language models (LLMs) to automatically rewrite PEMs at a sixth-grade reading level while preserving accuracy and clinical intent. The system integrates readability constraints with domain-specific safeguards to reduce medical jargon and enhance accessibility. Ongoing work focuses on systematically evaluating the rewritten materials, including a planned user study to assess comprehension gains, patient engagement, and the perceived trustworthiness of the texts. Anticipated outcomes include improved alignment of

PEM readability with patient literacy levels and practical insights into the role of LLM-based rewriting in healthcare communication.

Poster 13

Agentic Multimodal Health Data Analytics through SQL and GenAI

Yueqi Ren[1] Zhanhao Liu[1] Lin Ma[1] Karthik Ramani[1]

[1] College of Engineering, University of Michigan (CoE)

Healthcare, public health, and medical research face increasing demands for systems that can intelligently collect, integrate, and analyze diverse data sources to improve health outcomes. A major challenge lies in processing unstructured data, such as clinical notes and imaging results, which traditional database methods handle poorly. This project explores the integration of relational database techniques (SQL) with generative AI (GenAI) in an agentic framework to address this gap. Our proposed system consists of three components: (1) a centralized planner that orchestrates and coordinates the execution of agent functions, (2) a retrieval agent that extracts structured information from unstructured medical data, and (3) an analytics agent that translates natural language into queries for structured data analysis. Incoming medical data are first processed by the retrieval agent. When a user submits a query, the analytics agent interprets the natural language request, executes the corresponding SQL query on the structured data generated by the retrieval agent, and returns results that directly address the user's question. By combining SQL's strengths in structured data management with GenAI's semantic reasoning capabilities, the system enables more effective multimodal health data analytics. This work demonstrates a pathway toward leveraging agentic AI to enhance data-driven decision-making in healthcare. Keywords: multimodal health data analytics, SQL, data retrieval, generative AI, agentic systems

Poster 14

Community Perspectives on Health AI: Trust, Transparency, and the Role of Labeling

Morgan Sielaff[1] Jodyn Platt[2,1,3] Reema Hamasha[1]

Sean Tan[2,1] Kerry Ryan[2,4]

[1] Department of Learning Health Sciences, Michigan Medicine
[2] Medical School, University of Michigan
[3] TIERRA
[4] Center for Humanities, Arts, Social Sciences, and Ethics in Medicine

Despite the exponential growth of artificial intelligence (AI) tools in healthcare, patients often lack clear, evidence-based information on how AI tools impact their care. We aimed to identify key information for a health AI tool label and explore community perspectives on ethical and trustworthy AI adoption. In 2024, we conducted five virtual community deliberations across Michigan, engaging 159 participants. Participants completed 20-minute pre- and post-deliberation surveys to assess changes in trust, knowledge, and attitudes about AI in healthcare. The 6-hour deliberation sessions featured a series of expert presentations and two small-group discussions. Participants were recruited from community organizations across Michigan including Friends of Parkside and ACCESS. In the first small-group discussion, participants were cautiously optimistic about AI's potential but voiced concerns about safety, bias, and lack of transparency. There was strong emphasis placed on the need for transparency, oversight, and further educational opportunities. Additionally, participants were interested in governance that would allow for community involvement in AI development and healthcare practice policies. When participants completed the labeling activity, we found that participants deemed privacy and security, health equity, and safety and effectiveness of AI tools as priorities for inclusion on a health AI tool label. The post-deliberation survey results showed that 94% of participants agreed that "health systems should tell patients how they use AI tools in their healthcare." These findings underscore the importance of patient-informed solutions that promote transparency, oversight, and engagement. Community members acknowledge both the potential benefits and risks of AI, emphasizing the need for clear, accessible information about how AI tools are used. There is strong public demand for transparent communication, meaningful engagement, and oversight to ensure that adoption of AI is both trustworthy and patient-centered. A labeling system could help address ethical concerns and support trustworthy and equitable integration of AI into healthcare. These insights may be useful in guiding policy and practice to advance safe, effective and equitable integration of AI in healthcare systems.

Poster 15

Assessing Alcohol Use Disorder: Insights from Lifestyle, Background, and Family History with Machine Learning Techniques

Chenlan Wang[1] Gaojian Huang[2] Yue Luo[2]

[1] Electrical and Computer Engineering (ECE), University of Michigan
[2] San José State University

This study explored how lifestyle, personal background, and family history contribute to the risk of developing Alcohol Use Disorder (AUD). Survey data from the All of Us Program was utilized to extract information on AUD status, lifestyle, personal background, and family history for 6,016 participants. Key determinants of AUD were identified using decision trees including annual income, recreational drug use, length of residence, sex/gender, marital status, education level, and family history of AUD. Data visualization and Chi-Square Tests of Independence were then used to assess associations between identified factors and AUD. Afterwards, machine learning techniques including decision trees, random forests, and Naive Bayes were applied to predict an individual's likelihood of developing AUD.

Poster 16

Augmented Risk Prediction for the Onset of Alzheimer's Disease from Electronic Health Records with Large Language Models

Jiankun Wang[1] Sumyeong Ahn[2] Taykhoom Dalal[3] Xiaodan Zhang[4] Weishen Pan[3] Qiannan Zhang[3] Junyuan Hong[5] Bin Chen[4] Hiroko H. Dodge[6] Fei Wang[3] Jiayu Zhou[1]

[1] School of Information, University of Michigan (UMSI)
[2] Department of Computer Science & Engineering, Michigan State University
[3] Weill Cornell Medicine, Cornell University
[4] Department of Pediatrics & Human Development, Michigan State University
[5] Department of Electrical & Computer Engineering, University of Texas at Austin
[6] Harvard Medical School, Harvard University

Alzheimer's disease (AD) is the fifth-leading cause of death among Americans aged 65 and older. Screening and early detection of AD and related dementias

(ADRD) are crucial for timely intervention and for identifying clinical trial participants. The widespread adoption of electronic health records (EHRs) offers an important resource for developing ADRD screening tools such as machine learning based predictive models. Recent advancements in large language models (LLMs) demonstrate their exceptional capability of encoding knowledge and performing reasoning, which offers them strong potential for enhancing risk prediction. This paper proposes a novel pipeline that augments risk prediction by leveraging LLMs' few-shot inference power to make predictions in cases where supervised learning methods (SLs) may not excel. Specifically, we develop a collaborative pipeline that combines SLs and LLMs via a confidence-driven decision-making mechanism, leveraging the strengths of SLs in clear-cut cases and LLMs in more complex scenarios. We evaluate this pipeline using a real-world EHR data warehouse from Oregon Health & Science University (OHSU) Hospital, including EHRs from over 2.5 million patients and more than 20 million patient encounters. Our results show that the proposed approach effectively combines the power of SLs and LLMs, offering significant improvements in risk prediction of ADRD. Moreover, our findings highlight that removing noisy samples in ICL can significantly boost performance and that increasing the LLM size and fine-tuning on medical data does not necessarily improve performance, which challenges common scaling laws and emphasizes the need for careful consideration in practice.

Poster 17

A Cohort-Grounded Multimodal Biomedical Knowledge Graph for Alzheimer's Disease

Haobo Zhang[1] Jiayu Zhou[1]

[1] School of Information, University of Michigan (UMSI)

Alzheimer's disease (AD) is a complex neurodegenerative disorder with data scattered across diverse research cohorts, modalities, and literature, posing a major barrier to biomarker discovery and translational progress. This proposal introduces the Multimodal Biomedical Knowledge Graph (MMBKG) — a first-of-its-kind resource that integrates heterogeneous AD data, including neuroimaging, genomics, clinical phenotypes, and literature-derived knowledge, into a unified, analyzable framework. Our approach addresses the fragmentation of existing AD datasets (e.g., ADNI, NACC, UK Biobank, All of Us) by harmonizing them using a hybrid OMOP+FHIR model and embedding their

patient-level features into a graph-based structure. Through a novel “cohort grounding” strategy, we dynamically link abstract biomedical knowledge to empirical data across cohorts, enabling hypothesis validation and cross-cohort modeling. We will evaluate MMBKG’s effectiveness in enhancing early AD detection, cognitive outcome prediction, and biomarker identification through deployment on large-scale EHR datasets. The proposed work represents a paradigm shift in AD research, enabling rich, generalizable insights through integrated modeling of multi-modal, real-world and research data. The MMBKG framework will serve as a reusable foundation for knowledge-driven precision medicine in neurodegenerative disease and beyond.

Poster 18

Offspring Health Trajectories After Hypertensive Disorders in Pregnancy: A Retrospective EHR-Based Analysis

Zicheng Jin[1] Yi Liu[1] Ashley Hesson[1] Lana Garmire[1]

[1] Computational Medicine & Bioinformatics, University of Michigan

Introduction: Hypertensive disorders of pregnancy (HDP), including gestational hypertension, preeclampsia, affect approximately 5–10% of pregnancies worldwide and remain a major cause of maternal and perinatal morbidity. While some studies have suggested associations between HDP and offspring health, comprehensive longitudinal analyses across multiple pediatric conditions are limited. Therefore, this study leveraged large-scale EHR data to assess the risks and timing of pediatric conditions in offspring exposed to HDP.

Methods: We analyzed electronic health records from Michigan Medicine, identifying 4,306 mother–child pairs affected by preeclampsia, 5,365 with gestational hypertension, and 32410 controls. Propensity score matching was performed to control for sex, race and gestational age. Kaplan–Meier survival curves and Cox proportional hazards models were used to assess the incidence and timing of 20 pediatric chronic conditions. We also perform trajectory analyses to examine disease pairs and better understand patterns of disease onset and progression.

Results: Based on the survival curves and statistical tests, the Case group shows a faster decline in disease-free probability compared to the Control group across multiple conditions. From the forest plot, some diseases such as sleep disorders and obesity have hazard ratios greater than 1 with statistical significance,

indicating that maternal preeclampsia is significantly associated with an increased risk of these pediatric conditions. Additionally, children exposed to gestational hypertension also exhibit elevated risks for several conditions.

Discussion and Conclusions: This study shows that offspring exposed to hypertensive disorders of pregnancy (especially preeclampsia) have an increased risk of various pediatric diseases, with some conditions occurring at earlier ages. This highlights the need for enhanced early monitoring and intervention for high-risk children. Future work will further refine trajectory analysis and integrate multi-source data to explore potential mechanisms in depth, supporting precise prevention strategies.