

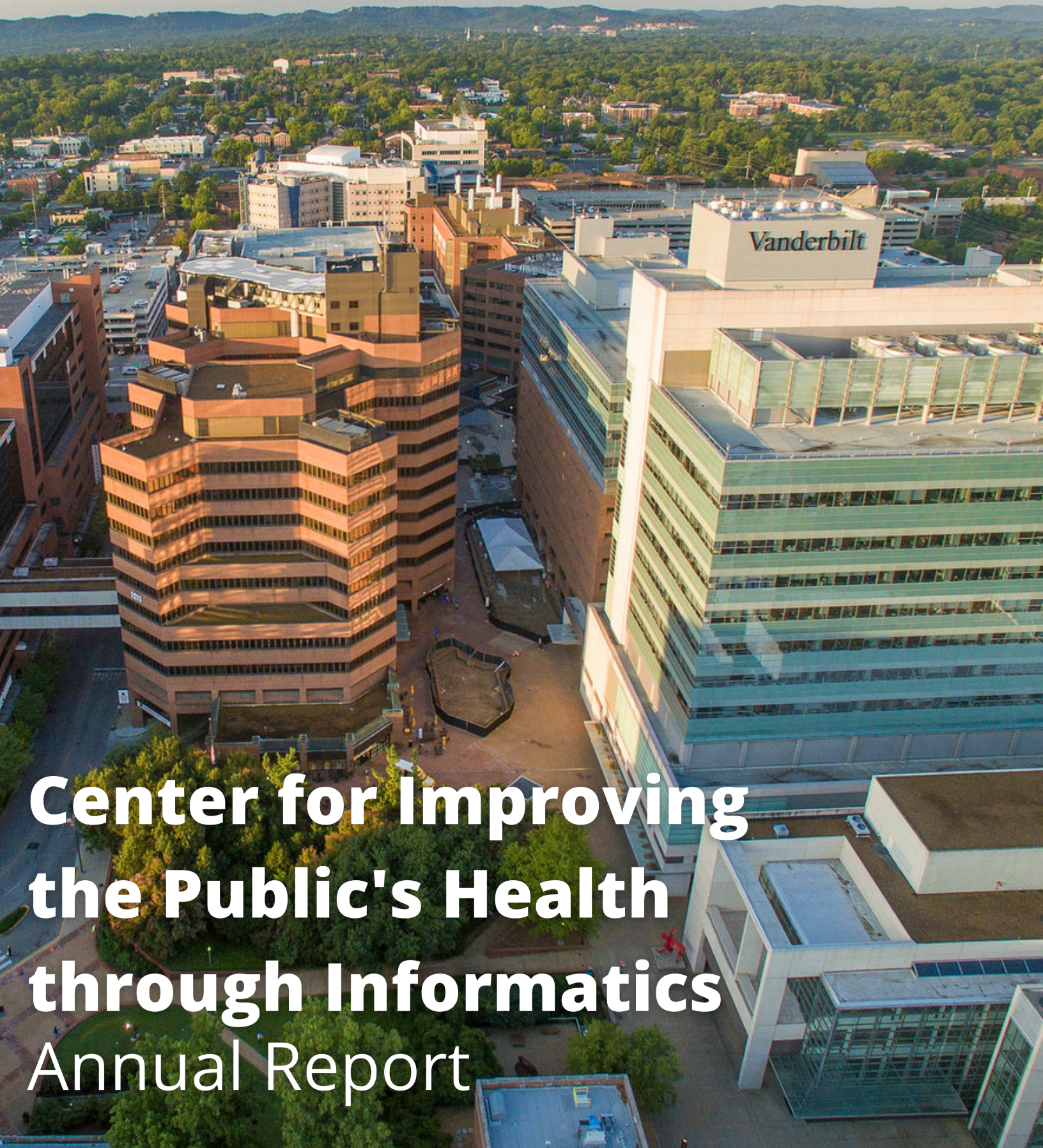


2022-2023

VANDERBILT UNIVERSITY

MEDICAL CENTER

Department of
Biomedical Informatics



Center for Improving the Public's Health through Informatics Annual Report

TABLE OF CONTENTS

Mission.....	02
Director Message.....	03
By the Numbers & Funding by Year.....	04
CIPHI Team.....	05
New Hires & Promotions.....	06
CIPHI Innovations	
IMPROVE Trial.....	07
FDA Sentinel Project.....	08
OMOP.....	10
People of CIPHI	
Spotlight: Alvin Jeffery.....	11
Spotlight: Dax Westerman & NLP Core.....	12
Faculty Highlights.....	13
Contact.....	14

OUR MISSION

To improve the health of populations by developing innovative methods to transform diverse data into actionable knowledge for research, healthcare delivery, and public health.



Pictured (left to right): Jessica Deere, Thomas Reese, Steven Brown, Elliot Fieldstein, Sharon Davis, Josh Smith, Julie Kim, Kim Kondratieff, Taylor Thurston, Michael Matheny, Josh Osmanski, Mohammed Al-Garadi, Mia Garchitorena, Alvin Jeffery, Tina French, Bhavnisha Patel, Michael McLemore

DIRECTOR MESSAGE

**MICHAEL
MATHENY**

MD, MS, MPH,
FACMI

CIPHI DIRECTOR
PROFESSOR, DEPARTMENTS OF
BIOMEDICAL INFORMATICS,
MEDICINE & BIOSTATISTICS



"We've had a wonderful year at CIPHI, and I'm excited to share some of the accomplishments and highlights from our talented faculty, staff, and students in 2022 with you! Since our inception as the Center for Population Health Informatics in 2015, and subsequent re-branding and expansion into Center focusing in both population and public health in 2018 with our former Co-Director, **Melissa McPheeters, PhD, MPH**, we have had steady growth and progress in this domain.

"Our mission has been to help support our investigators, collaborators, and the research community at Vanderbilt University Medical Center and Vanderbilt University in pursuit of innovative ways to leverage healthcare and publicly available data to support the health of patients and patient populations at scale across a number of clinical domains and in collaboration with other academic institutions, state, and federal regulatory agencies.

"None of this would be possible without both the inspiration and perseverance of our community of researchers and collaborators, and I hope that this report can give you a flavor for both the breadth and depth of some of the ongoing work. Our members have been very productive with grants, papers, presentations, panels, but they've also shown an increasing outreach to population and public health agencies that help support patient health.

"I'm very proud of the team's work, and we highly value collaborative and consortia building and participating activities, much of our work is in a larger community. If elements within interest you in your research, or your clinical and operational work, please do reach out to me and the teams leading the efforts to engage in a conversation. It is truly the most fun when we tackle hard problems together!" — **Michael Matheny**

BY THE NUMBERS



**5 VUMC DEPTS
REPRESENTED**

- BIOMEDICAL INFORMATICS
- EMERGENCY MEDICINE
- MEDICINE
- PSYCHIATRY
- SURGERY

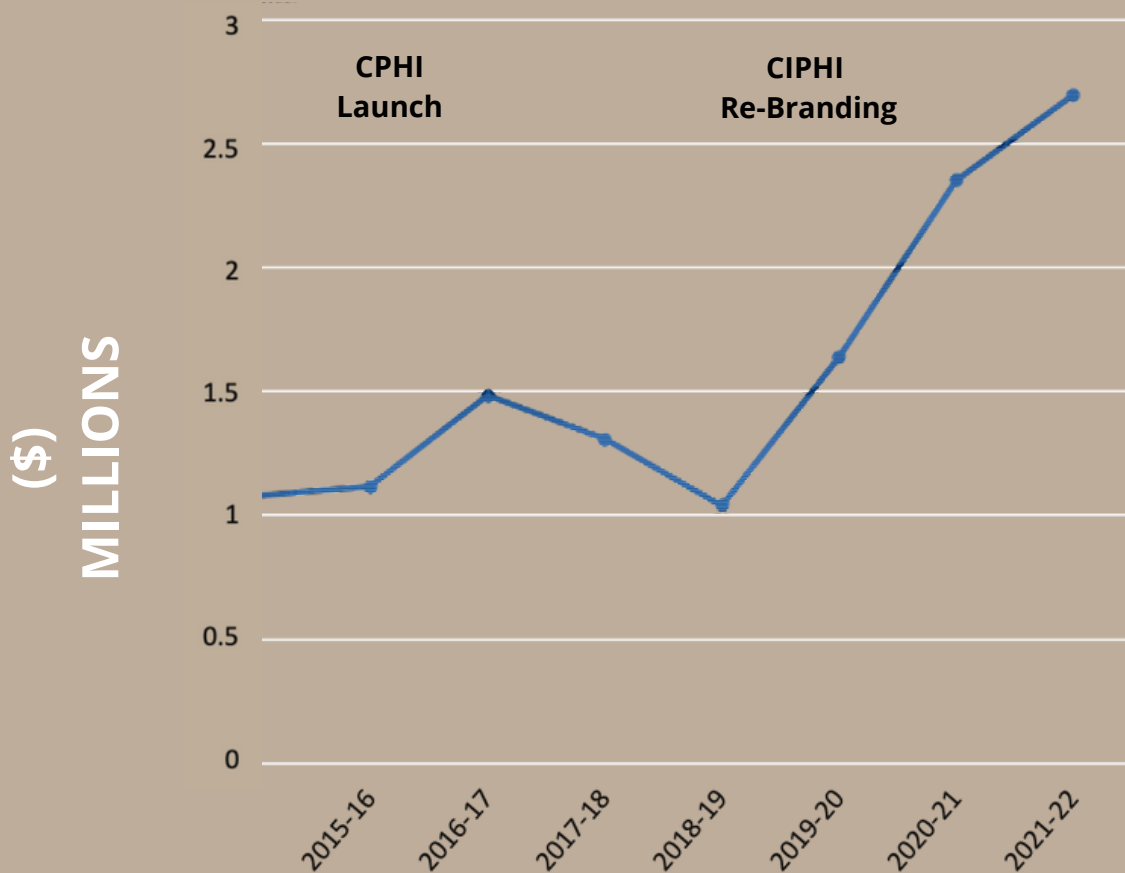


**5+
PRESENTATIONS
AMIA 2022**



**40+
MEMBERS IN
CIPHI**

FUNDING BY YEAR



CIPHI TEAM

FACULTY

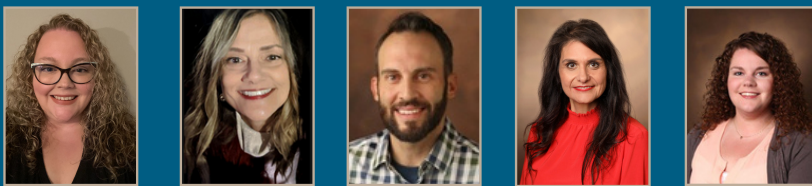


ADMIN STAFF

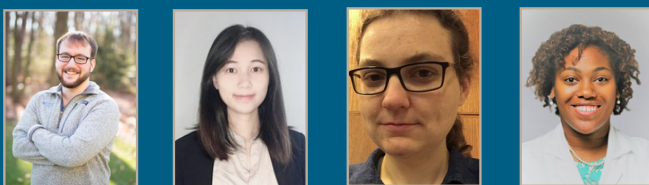
PROGRAMMING STAFF



CLINICAL INFORMATICISTS



TRAINEES & STUDENTS



NEW HIRES & PROMOTIONS

Mohammed Al-Garadi
Research Assistant
Professor
(April 2022)

Jessica Deere
Nurse
Informatics
Data Analyst
(May 2022)

Michele Lenoue-Newton
Research Program
Manager
(Jan 2023)

Matt Martino
Associate
Application
Developer
(Feb 2023)

Josh Osmanski
IT Project
Manager
(Oct 2021)

Bhavnisha Patel
Application
Developer
(Nov 2021)

Thomas Reese
Assistant
Professor
(July 2021)

Lisa Roddy
Nurse
Informatics Data
Analyst
(May 2022)

Taylor Thurston
Senior Project
Manager
(Dec 2021)

Chad Dorn
Promoted to:
Senior
Application
Developer
(Nov 2022)

Daniel Park
Promoted to:
Senior
Application
Developer
(Nov 2022)

Amy Perkins
Promoted to:
Lead
Biostatistician
(Jan 2023)

Katherine Simon
Promoted to:
Senior
Application
Developer
(Aug 2022)

Dax Westerman
Promoted to:
Senior Data
Scientist
(May 2023)

Robert Winter Named Recipient of the 2022 DBMI Admin Excellence Award



Pictured: Elise Russo, 2021 DBMI Admin Excellence Award recipient, Robert Winter, and DBMI Chair Peter Embí

Since joining CIPHI in 2021, **Robert Winter**, Senior Project Manager, has helped lead a large initiative to support data modeling and data transformation across the national VA and to improve the scale, accuracy, and efficiency of the project. He has implemented ticketing systems, documentation of process and workflow of the operational and research projects, and has worked well with technical development staff to pursue the objectives of the projects.

"He consistently provides excellent customer service and support to faculty, staff and students. He goes above and beyond to consider the brand of the Center, and how personnel interact with each other, in order to generate a consistent, efficient set of tools, datasets, processes, and analyses," said **Michael Matheny**.



IMPROVE-AKI TRIAL



Acute kidney injury (AKI) is a common complication of cardiac catheterization, posing higher costs, longer hospital stays and increased short-term and long-term mortality risk.

In a large, randomized trial conducted over a period of 18 months at 19 medical centers of the Veterans Health Administration, a 2x2 trial of intensive quality improvement coaching and an informatics risk-adjusted performance and data visualization tool found a **46% reduction in the odds of AKI after cardiac catheterization.**

Led by Dr. Matheny at Vanderbilt University Medical Center, Jeremiah Brown at Dartmouth Geisel School of Medicine, in Hanover, New Hampshire, and Richard Solomon at Larner College of Medicine at the University of Vermont, in Burlington, the trial was reported Jan. 13, 2023 in the Clinical Journal of the American Society of Nephrology.

"Lots of time in informatics, we build things to support other important work, and the impact is uncertain," **Michael Matheny, MD, MS, MPH**, said. "It's incredibly gratifying to see a set of AI-driven data visualizations to support provider practice have a direct impact on patient health by reducing the incidence of AKI."

AI Algorithm Surveillance and Maintenance

One of most important facets for the informatics intervention in the study was a suite of AI/ML algorithm surveillance and maintenance tools deployed in the trial, created by **Sharon Davis, PhD, MS.**

"The algorithm allowed the trial to update and maintain the risk-adjustment model through the pandemic, which allowed the successful completion of the trial," said **Dr. Matheny.**

"One of the most exciting pieces of the Improve-AKI trial was that we proactively addressed dataset and performance drift right from the start," **Dr. Davis** said.

"Drifting calibration can completely change the insights gleaned from quality benchmarking tools like the one we created for the trial—even going so wrong as to lead care sites to believe they are exceeding expectations when in reality, their patients are experiencing worse outcomes than we expect.

"How can dashboards effectively motivate quality improvement if you can't rely on the information? We made sure our study sites could trust and act on insights from the dashboard by incorporating ongoing updating behind the scenes, maintaining calibration at the national level to ensure local trends were accurate and informative.

"Unfortunately, model monitoring and maintenance is not yet broadly implemented, so it was wonderful to show that we can operationalize our maintenance framework and see its impact supporting clinicians working to improve care quality and patient outcomes."

"Reliable, robust prediction tools are on the way!" – Sharon Davis

To learn more, read this article in the VUMC Reporter:

<https://news.vumc.org/2023/02/02/machine-learning-aids-injury-prevention-in-cardiac-cath-labs/>





PROJECT

Vanderbilt University Medical Center assumed an expanded role in U.S. Food and Drug Administration (FDA) medical product safety monitoring through a program called Sentinel, developed and operated for the FDA by Boston-based Harvard Pilgrim Health Care Institute. This program supports national electronic safety surveillance of regulated medical products after they've reached market — drugs, vaccines, blood products, and other biologicals, etc. According to the FDA, Sentinel is the **largest multisite distributed database in the world dedicated to medical product safety**.

VUMC has collaborated in the Sentinel Operations Center since it was launched in 2009, and currently the VUMC's OC effort is led by Margaret Adgent. Following the FDA's announcement on five years' additional funding for Sentinel, in 2019, the FDA established a new "Innovation Center" designed to advance the use of electronic health records and publicly available data sources for medical product safety surveillance, led by Sebastian Schneeweiss at Mass General Brigham. **Michael Matheny** and **Kevin Johnson** (until Dr. Johnson's departure from Vanderbilt) have served as the leads for VUMC, one of the four innovation lead sites along with Duke and Kaiser Permanente Washington.

"While Sentinel has had marked success relying primarily on structured administrative and billing data from health care payers, the goal for the Innovation Center has been to expand safety surveillance into the realm of electronic health records, where some of the richest information is stored as text," **Dr. Matheny** said.

To learn more, read this article in the VUMC Reporter:

<https://news.vumc.org/2019/10/16/vumc%E2%80%88assumes-new-role-in-fda%E2%80%88safety-monitoring/>



Below are descriptions of a few important Sentinel projects!

Project: Augmenting Date of Death and Cause of Death Ascertainment in Sentinel

"This project is supporting the FDA's need to determine date and contributing causes of death as close to real-time as possible by exploring the use of publicly available data and probabilistic patient linkage to electronic health records (EHR) to extract this information. We are using the state and national vital statistics records as a reference, and deploying novel machine learning natural language processing (NLP) and modeling algorithms in order to accomplish this," said **Mohammed Al-Garadi, PhD** (Co-Investigator).

"In collaboration with Rishi Desai at Mass General Brigham, we hope to innovate by integrating multiple sources of data in supporting a probabilistic framework to determine the likelihood of contribution to mortality for common causes of death, and integrate that into the common data model for wide reuse by Sentinel site participants. Ruth Reeves and Mohammed Al-Garadi are leading the NLP and analytics to push this forward," said **Michael Matheny, MD, MS, MPH** (Project Co-PI). Pictured: **Ruth Reeves & Mohammed Al-Garadi**



SENTINEL PROJECTS CONT.



Project: *Advancing Scalable Natural Language Processing Approaches for Unstructured Electronic Health Record Data*

This project, in collaboration with David Carrell at Kaiser Permanente Washington, is exploring incorporation of structured and unstructured EHR data into Sentinel. We developed automated phenotyping approaches, using NLP and machine learning, to identify patients with Symptomatic COVID-19. From that success, the FDA has extended our project to use the same methods to identify anaphylaxis, a condition of great interest to the FDA. **Joshua Smith, PhD** (Project Co-PI)

Project: *Evaluation of Existing Approaches to EHR-Based Signal Identification*

This project conducted a literature review of methods for adverse event signal identification using EHR data. Despite broad interest in utilizing EHRs for this, we found that published efforts fail to leverage the full breadth and depth of available data. The development of best practices and reference standards would promote the expansion of EHR-based pharmacovigilance. **Joshua Smith, PhD** (Project PI), **Sharon Davis, PhD, MS** (Co-Investigator). Publication: <https://link.springer.com/article/10.1007/s40264-023-01325-0>

Project: *Development and Evaluation of EHR Information Extraction Pipeline and Tree Based Scan Statistic (TBSS) Methods for EHR-Based Signal Detection*

Tree-based scan statistics (TBSS) are data mining methods used to prioritize statistical alerts when assessing a drug of interest for potential adverse events. This method has been used successfully in claims data, however, current TBSS implementations do not leverage the rich information available in EHRs. In collaboration with Shirley Wang at Mass General Brigham, this project is working to develop approaches for combining structured and unstructured EHR data for use with TBSS. **Joshua Smith, PhD** (Project Co-PI)



Project: *Improving Probabilistic Phenotyping of Incident Outcomes through Enhanced Ascertainment with Natural Language Processing*

"Our project developed and validated scalable phenotyping algorithms using NLP to identify new clinical events that we might like to prevent or survey for in post-marketing safety surveillance. We completed the project successfully this year and have a manuscript undergoing FDA public affairs clearance right now," said **Colin Walsh, MD, MA** (Project PI)



Project: *Developing Chart Review Tools to Identify Adverse Events*

"The project aims to demonstrate the use and portability of a chart review tool to assist and streamline adverse event extraction from electronic medical records. Recent gains in artificial intelligence and machine learning are still limited by the lack of high-quality labeled clinical data sets. This project will help streamline the chart review extraction process, reducing the time between hypothesis generation and the evaluating the potential clinical benefits," said **Daniel Fabbri, PhD** (Project PI).



Project: *Representation of Unstructured Data Across Common Data Models*

This project seeks to develop best practices around integration of NLP derived data into the Sentinel common data model. Led by Keith Marsolo at Duke, this multi-site investigation and two clinical use cases, we have developed a series of recommendations and guidelines for normalizing various types of NLP outputs using two clinical use cases as exemplars, and are collaborating with Sentinel to implement and pilot these recommendations across the Innovation Center. **Ruth Reeves, PhD** (Project Site Lead)



Project: *Using Unsupervised Learning to Generate Code Mapping Algorithms to Harmonize Data Across Data Systems*

Led by Xu Shi at the University of Michigan, this project developed and evaluated scalable, privacy-preserving, data-driven statistical methods to describe and mitigate data heterogeneity due to coding differences between healthcare systems. "We need to tackle the failure of clinical prediction model generalization from all angles, including differences in how health systems capture data. This project faces that challenge head on with novel methods to sort out and standardize coding practice patterns," said **Sharon Davis, PhD, MS** (Site PI)

OMOP INITIATIVE

The importance of the rise of healthcare related common data models cannot be overstated. Multiple large mature data models have arisen in our field, including Sentinel, I2B2, PCORNet, and OMOP, and translating healthcare data into a more standardized representation has allowed the acceleration of large observational analyses and insight to come from real-world data.

In particular, the **Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM)** is the data model for the Observational Health Data Sciences and Informatics (OHDSI) community, which seeks to continue to develop tools and systems to support real-work evidence generation using observational data.

Since 2015, as part of a national research infrastructure initiative at the Department of Veteran Affairs called VINCI, led by Scott DuVall, we have been responsible for transforming and ensuring the quality of the national VA healthcare data that has seen steady growth and is in use by over 500 projects across the country to serve research and operations.

"Because medical data are generated differently, whether from pharmacies or labs or EHR tools, the data requires translation. OMOP reduces time and cost by having developers and scientists clean up the front end so researchers don't lose time on their research,"

Robert Winter said. "Our work is doing the backend magic so the data models are cleaned and structured so it's understandable internationally."

He continues: "Understanding where the value is means you have to go through a big data project. A lot of this seems trivial, but some of these projects take multiple years. We have to be sure they're precise enough to move the needle of treatments and workflows of care."

Our Center has also developed a set of data transformation tools to allow rapid development of analytic data sets for observational analyses that are reusable across research project using OMOP, and accelerate research projects' ability to get to the fun work of discovery and inference generation.

For investigators that are interested in executing code and research in both the VA and VUMC in parallel, this represents a tremendous opportunity.

Benefits of the VA OMOP CDM

- Extensive data cleaning
- Robust integration of disparate data
- Back-reference to source data
- Integrates VA specific vocabularies
- Leverages a national community to crowdsource data quality improvements

Current Audience

- Veterans Affairs researchers
- Audiences can use a standard data model and apply that to each location to their local dataset

*Fun Fact: The VINCI OMOP initiative delivers the national VA EHR data in a normalized OMOP CDM—including about **25 million veterans!***



To learn more, contact Robert Winter at robert.winter@vumc.org.



FACULTY SPOTLIGHT

ALVIN JEFFERY

PHD, RN-BC, CCRN-K,
FNP-BC

ASSISTANT PROFESSOR VANDERBILT
UNIVERSITY SCHOOL OF NURSING VUMC
DEPARTMENT OF BIOMEDICAL
INFORMATICS

Dr. Alvin Jeffery has made leaps and bounds this year in the areas of substance use disorder (SUD) research. In July 2022, he received the **Moore Foundation grant** to focus on electronic health record (EHR) customization.

In September 2022, he received a **\$2.4 million grant** from the NIH's National Institute on Drug Abuse (NIDA) that will advance identification of SUDs in order to accelerate genetics studies.

"I'm thrilled my lab is growing in personnel and productivity thanks to a Moore Foundation grant supporting EHR usability efforts and an NIH grant focused on the genetics of substance use disorder," **Dr. Jeffery** said.

RECOGNITIONS & AWARDS

- Provost's Faculty Grant for Culminating Projects in Immersion, Vanderbilt University
- Dean's Award for Recognition of Faculty Achievement in Research Endeavors, Vanderbilt University School of Nursing
- Funding: Sept 2022 – July 2027
- He is also a co-Investigator on an NCI U01 as well as an NIAAA R03
- Accepted into Babson College's L-SPRINT Program (sponsored by the National Institute on Drug Abuse)



CORE SPOTLIGHT

DAX WESTERMAN, MS

SENIOR DATA SCIENTIST & OPERATIONS DIRECTOR

NATURAL LANGUAGE PROCESSING SUPPORT SERVICES CORE (NLPSSC)

The **Natural Language Processing Support Services Core (NLPSSC)** launched in February 2022. It was developed as an important resource for data science, software development and infrastructure support in projects leveraging NLP in and outside of VUMC. Since its launch, the NLPSSC has worked with **8 PIs** over **9 projects** in the last year. A total of **2 projects** have been completed.

If you are in need of NLPSSC's services, the Core offers the following:

- Selection, extraction, cleaning, and processing of large documents sets
- Text extraction from imaged documents (OCR to NLP)
- Implementation of novel methods or modification of existing tools
- Data adapter creation to novel source acquisition
- Method containerization for cross-site portability
- Cross-site concept mapping
- Corpus- and cohort-level analytics and reporting
- Evaluation of algorithms against publicly available datasets

- Algorithm integration with existing systems (e.g., method implementation within EHR)
- Database design and query optimization
- User Interface design and development
- Collaboration on novel method design or model development
- Method execution across multiple platforms (e.g., VUMC Azure, ACCRE, VA-VINCI, etc.)

Contact Info:

- Submit requests to:
https://redcap.link/nlpssc_services
- NLPSSC Information: Email
nlpssupportservicescore@vumc.org

NLPSSC Members:

- **Josh Smith**, Faculty Director
- **Ruth Reeves**, NLP Faculty Consultant

FACULTY HIGHLIGHTS



VA Cardiac Rehab Study (2021-2023)

In partnership with CIPHI, **Justin Bachmann, MD, MPH**, and his team have been evaluating national data in an effort to better understand eligibility for and participation of veterans in VA cardiac rehab programs. While cardiac rehab is strongly recommended after "MI," "PCI" or coronary artery bypass grafting, rehab services are historically underutilized. This has led to higher mortality rates in certain populations, an outcome this study hopes to further shine a light on. This study builds on previous work and aims to provide further insights in this area of veteran care, and Dr. Bachmann is submitting a HSR&D IIR (R-01 Type) grant further extending this work.



Glenn Gobbel, DVM, PhD, MS, has had a tremendous year of proposals being funded. He is working with Saket Girotra (UT Southwestern) as a site PI for a follow-on R-01 looking at the long-term survival and outcomes among patients with peripheral artery disease, where his work focuses on natural language processing information extraction. He is working with Dennis Oh (San Francisco VA) as a site PI for a VA IIR study that seeks to create a NLP system to enable a comparison of outcomes between in-person and telehealth-based dermatology, since its meteoric rise in recent years. He has been refunded for a VA operational partnership with Elliot Fielstein, also one of our Center faculty, to leverage NLP for mental health process and outcome information extraction. He is working with Jian Gui (Dartmouth) as a site PI for a new study seeking to predict suicide risk in veterans using VA clinical notes and NLP in the study. In addition, he is working with Ashley Griffin (Palo Alto VA) as a site PI in a new study to identify incidence and prevalence of use of non-pharmacologic therapies in VA patients. "It's exciting to see the use of NLP expanding into so many different areas of medicine and to be a part of that," **Dr. Gobbel** said.



Stephen Deppen, PhD, has two large projects that began in the Fall 2022. The first was a competitive renewal for a U01 from the National Cancer Institute's Early Detection Research Network for VUMC to be the national clinical validation center for imaging and biological biomarker development and validation. This large multi-year effort will engage experts in machine learning for processing imaging and clinical data as well as implementation science to examine drivers of change in the clinical management of lung cancer diagnosis when biomarkers are added to the bedside. The second project is a VA CSR&D Merit Award also with Eric Grogan to create a lung cancer risk screening algorithm specific to veterans that expands the pool of veterans who may benefit from lung cancer screening with low-dose CT scans. The proposal incorporates veterans' service related exposures and current risk factors, and will integrate the Department of Defense and VA's EHR data to creating a longitudinal dataset covering more than 16 million veterans with detailed healthcare data covering, on average, 13 years to define environmental, epidemiological and clinical data. This effort links the extensive depth of VA data knowledge from CIPHI with Vanderbilt's NLP and Phenotyping cores. If successful, it will be one of the largest implementations of machine learning across the VAs health record and could impact care among six million veterans.



In the past year, **Thomas Reese, PharmD, PhD**, has settled nicely into his new home at DBMI. He's particularly proud of the research, networking, and teaching he's done since joining the department. Briefly, he will be the site PI on an R18 (AHRQ). He was invited as key personnel on a substance use disorder P grant as well as to the NIA-funded Junior Investigator Intensive on deprescribing. As an instructor of BMIF 6332, he received good feedback about restructuring the course. Finally, he has five first- or last-author papers pending publication. "My success was only possible because of the guidance from my mentorship team and support from DBMI as well as CIPHI and the Vanderbilt Clinical Informatics Center!" said **Dr. Reese**.

CONTACT US



Michael Matheny
michael.matheny@vumc.org



<https://www.vumc.org/cphi/welcome>

VANDERBILT  UNIVERSITY
MEDICAL CENTER®

Department of
Biomedical Informatics