THE 2025 ANNUAL JOINT BIOSTATISTICS SYMPOSIUM

Case Western Reserve University ~ Cleveland Clinic ~ Ohio State University

CASE WESTERN RESERVE UNIVERSITY

Cleveland Clinic

THE OHIO STATE UNIVERSITY

APRIL 7, 2025
CASE WESTERN RESERVE UNIVERSITY
TINKHAM VEALE UNIVERSITY CENTER

Case Western Reserve University ~ Cleveland Clinic ~ Ohio State University

2025 ANNUAL JOINT BIOSTATISTICS SYMPOSIUM

April 7, Case Western Reserve University
Tinkham Veale University Center – Kelvin & Eleanor Smith Foundation Grand Ballroom
11038 Bellflower Road, Cleveland, OH 44106

Directions to the Lot S-29
11173 East Boulevard, Cleveland, Ohio 44106

AGENDA

10:00 am	**Registration and Poster Set up**
10:30 am-10:50 am	**Opening remarks and introduction: J. Michael Oakes, Senior VP of Research; Gary Schwartz, Director of Case CCC**
10:50 am – 11:30 pm	**Christopher Boyer, Ph.D. from Cleveland Clinic Lerner College of Medicine**
	Title: Counterfactual prediction: A framework for estimation, validation, and transportability of models under hypothetical interventions
11:30 pm – 1:00 pm	**Lunch and Poster Session**
1:00 pm – 1:40 pm	**Andy Ni, Ph.D. from Ohio State university**
	Title: Contrast Weighted Learning for Individualized Treatment Rule Estimation
1:40 pm – 1:50 pm	**Break**
1:50 pm – 2:30 pm	**Holly Hartman, Ph.D. from Case Western Reserve University**
	Title: Sequential, Multiple Assignment, Randomized Trials (SMARTs) with continuous outcomes and tailoring functions
2:30 pm – 2:40 pm	**Break**
2:40 pm – 2:50 pm	**Poster Awards and Remarks: Jonthan Haines, PQHS Chair**
2:50 pm -3:00 pm	**Remarks and Keynote Speaker Introduction:**
	Dean Stan Gerson
3:00 pm – 4:00 pm	**Jeff Leek, Ph.D. Chief Data Officer, Vice President, and J Orin Edson Foundation Chair of Biostatistics at the Fred Hutchinson Cancer Center**
	Title: Building an oncology AI engine at the Fred Hutch and beyond
4:00pm	**Adjourn**

Keynote Speaker:
Jeff Leek, Chief Data Officer, Vice President, and J Orin Edson Foundation Chair of Biostatistics at the Fred Hutchinson Cancer Center.

"Building an oncology AI engine at the Fred Hutch and beyond"

Artificial intelligence and machine learning are having a transformative impact across industries; but real impact in the biomedical sciences and other regulated industries has been slower. In this talk, I will discuss the challenges and opportunities in each component of an AI system in biomedicine – the data, the models, the interfaces, and the people and discuss our efforts at the Fred Hutch to build an "artificial intelligence engine" that connects the most up to date clinical information with collaborative groups of AI researchers and oncologists to drive practical AI impact. I will also discuss our efforts to extend this model nationally with the Cancer AI Alliance (CAIA). I will discuss how and why I believe biostatisticians, working with diverse other technologists, medical scientists, and ethical and legal experts, can have a major impact on biomedicine for the next decade.

Biography:

Dr. Jeff Leek is Chief Data Officer, Vice President, and J. Orin Edson Foundation Chair of Biostatistics at the Fred Hutchinson Cancer Center. He is a prominent biostatistician whose work focuses on developing statistical methods, software, and data resources for analyzing large-scale genomic and biomedical data. His research has contributed significantly to understanding gene expression and improving reproducibility in science. Prior to joining Fred Hutch, Dr. Leek was a professor at the Johns Hopkins Bloomberg School of Public Health and co-director of the Johns Hopkins Data Science Lab. There, he co-developed one of the earliest and most widely adopted online data science programs, enrolling over 8 million learners globally, and partnered with community-based organizations to promote data science education as a tool for economic and public health advancement. His contributions span methodology, education, and advocacy for open and reproducible science. Dr. Leek is a Fellow of the American Statistical Association and a recipient of both the Mortimer Spiegelman Award and the COPSS Presidents' Award, one of the highest honors in the field of statistics.

Speaker Abstracts

Christopher Boyer, Ph.D.

Department of Quantitative Health Sciences, Cleveland Clinic.

Title: Counterfactual prediction: A framework for estimation, validation, and transportability of models under hypothetical interventions

Abstract: Counterfactual prediction methods are required when a model will be deployed in a setting where treatment policies differ from the setting where the model was developed, or when a model provides predictions under hypothetical interventions to support decision-making. However, estimating and evaluating counterfactual prediction models is challenging because, unlike traditional (factual) prediction, one does not observe the full set of potential outcomes for all individuals. Here, we discuss how to fit or tailor a model to target a counterfactual estimand, how to assess the model's performance, and how to perform model and tuning parameter selection when (1) data are available from a randomized trial in a source population that differs from the target population in which the model will be deployed, (2) data are available from an observational study conducted in the target population, or (3) both are available. We provide identifiability and estimation results for building a counterfactual prediction model and for multiple measures of counterfactual model performance including loss-based measures, the area under the receiver operating characteristics curve, and calibration. Importantly, our results allow valid estimates of model performance under counterfactual intervention even if the candidate model is misspecified, permitting a wider array of use cases. We illustrate these methods using simulation and apply them to representative counterfactual prediction tasks in cardiovascular disease.

Andy Ni

Division of Biostatistics, College of Public Health, The Ohio State University.

Title: Contrast Weighted Learning for Individualized Treatment Rule Estimation

Abstract: Precision medicine aims to tailor medical decisions based on patient-specific characteristics. An individualized treatment rule (ITR) assigns an optimal treatment to a patient based on their personal characteristics. An archetypal ITR estimation approach is outcome-weighted learning (OWL) based on a weighted classification framework with clinical outcomes as the weights. Existing OWL methods are susceptible to irregularities of outcome distributions, such as outliers and heavy tails. Moreover, it is unclear how these methods can be used for multivariate survival outcomes that are commonly encountered in cancer and cardiovascular research. In this study, we propose contrast-weighted learning (CWL) for ITR estimation. CWL exploits the flexibility and robustness of contrast functions between pairs of patients to enable robust ITR estimation for a wide range of clinical outcomes. By introducing win status into the contrast function, CWL naturally handles multivariate survival outcomes while incorporating the clinical importance of different events into ITR estimation. We established the theoretical properties of the estimated ITR. We conducted simulations to evaluate the finite sample performance of CWL in comparison to several alternative ITR estimation methods. Finally, we apply the proposed CWL method to three real datasets to demonstrate its real-world application.

Holly Hartman

Department of Population & Quantitative Health Sciences, Case Western Reserve University.

Title: Sequential, Multiple Assignment, Randomized Trials (SMARTs) with continuous outcomes and tailoring functions

Abstract: Multi-stage trials like sequential, multiple assignment, randomized trials (SMARTs) have typically relied on a binary variable to define response which is used in assigning the next stage treatment assignment. Instead, we develop a function of a continuous outcome to assign a probability of staying on the same treatment and then randomly assign the next treatment using a multinomial distribution. First, we develop a new trial design for small sample SMARTs (snSMARTs). The overall goal of the trial is to determine the optimal first stage treatment. We use a function, called the mapping function, to map the first stage outcome to a probability of staying on the same treatment and Bayesian regression methods to analyze data from both stages. Re-randomization based on a mapping function of a continuous outcome allows for snSMARTs to be conducted without requiring a binary outcome. Next, we apply similar concepts to a standard size SMART with continuous outcomes where the goal is to determine the optimal dynamic treatment regimen (DTR). We present a new trial design for SMARTs that use a tailoring function instead of a binary tailoring variable. In this trial design, we simultaneously develop a tailoring variable and estimate the DTR. We perform simulation studies to compare the proposed design with continuous outcomes to standard designs with binary outcomes. The proposed designs results in more efficient treatment effect estimates and similar or better outcomes for trial patients.

Posters

Clinical Applications

1. Syed Hassaan Ahmed Bukhari (with Hassaan A. Bukhari, Shivangi Kewalramani, Luke Witzigreuter, Rajkumar Dhar, Jafar Pourbemany, Natalia Amadio, Larisa G. Tereshchenko), Cleveland Clinic
   Automated Detection and Statistical Analysis of Non-Physiological Artifacts and Electrode Misplacement in ECG Signals

2. Gi-Ming Wang (with Curtis Tatsuoka), Case Comprehensive Cancer Center
   Bayesian Ordered Lattice Design for Phase I Clinical Trials

3. Joshua A. Gerlick (with Stefan Agamanolis, Peter J. Whitehouse, Philip A. Cola), Case Western Reserve University
   Next-Generation Eldercare: Validating Robotic Technologies to Support Individuals with ADRD

4. Shivangi Kewalramani (with Shivangi Kewalramani, Larisa G. Tereshchenko), Cleveland Clinic
   Anatomic Location of Heart Vector Origin and Its Association with Vectorcardiogram Morphology

5. Rajkumar Dhar (with Larisa Tereshchenko), Cleveland Clinic
   Automated QT Detection in ECG Signals Using Deep Learning

6. Krishna Mridha (with Ming Wang, and Lijun Zhang), Case Western Reserve University
   Predicting CKD Progression with Machine Learning: Insights from CRIC Data and XAI Techniques

Epidemiology

7. Kathryn Kohler Menta (with Dr. Holly Hartman, Dr. Maureen Curley), Case Western Reserve University
   Assessing the Association Between Pregnancy Outcomes and Suicide Risk Using Bayesian Network Meta-Analysis

Genomics and Bioinformatics

8. Grant Konkel (with Jiasen Zhang, Weihong Guo, Liangliang Zhang), Case Western Reserve University
   Unrolling Regularized Non-negative Matrix Factorization for Spatial Transcriptomics Deconvolution

9. Jiasen Zhang (with Weihong Guo, Liangliang Zhang, Xi Qiao), Case Western Reserve University
   Bayesian Spatial Transcriptomic Data Deconvolution using Graph Laplacian Prior

10. Yiqian Zhang (with Jonas Schluter, Lijun Zhang, Xuan Cao, Robert R Jenq, Hao Feng, Jonathan Haines, Liangliang Zhang), University of Illinois Urbana-Champaign
    Review and revamp of compositional data transformation: A new framework combining proportion conversion and contrast transformation

11. Claudia Cabrera (with Lijun Zhang PhD), Case Western Reserve University
    Machine Learning for Prognostication in HPV-Associated vs. HPV-Negative Oropharyngeal Cancer

12. Fode Tounkara, The Ohio State University
    A Flexible Copula-Based Approach for the Analysis of Secondary Phenotypes in Ascertained Samples

Statistical Methods

13. Ruoyuan Qian (with Bo Lu), The Ohio State University
    Issues and Remedies in External Control Data Borrowing for Clinical Trials Through Matching: A Statistical Perspective

14. Lingpeng Shan (with Michelle J Naughton, Electra D Paskett, and Michael L. Pennell), The Ohio State University
    Bayesian Variable Selection for Joint Models of Heterogeneous Longitudinal Variables and a Binary outcome

15. Fandi Chang, The Ohio State University

Treatment effect estimation in the presence of cluster size dependent treatment heterogeneity in stepped wedge designs

16. Chenyu Liu (with Xi Qiao, Zihan Zhu, Liangliang Zhang), Case Western Reserve University
    Suppressing Odds Ratio Inflation: Detection and Correction of Perfect Separation in Logistic Regression

17. Junhui Mi (with Emily C. Zabor, Sujata Patil, Rahul D. Tendulkar, Sarah M. C. Sittenfeld), Cleveland Clinic
    Combining missing data imputation and internal validation in clinical risk prediction models

18. Raghav Awasthi (with Gayan Samuditha Mend Mend Arachchig, Xiaofeng Zhu), Case Western Reserve University
    Unsupervised evaluation of pre-trained DNA language model embeddings

19. Zihan Zhu (with Yiqian Zhang, Wenhao Li, Liangliang Zhang), Case Western Reserve University
    Enhanced Pattern Recognition and Biomarker Identification through Beta Diversity Analysis in Microbiome Data

20. Congli Ma (with Abdul-Nasah Soale∗ , Congli Ma, Siyu Chen, Obed Koomson), Case Western Reserve University
    On metric choice in dimension reduction for Fréchet regression

21. Kari O'Donnell (with Hovmand, P.S., Gunzler, D.D.), Center for Community Health Integration
    A generative framework for comparative causal modeling of nonlinear feedback systems

22. Yihe Yang (with Noah Lorincz-Comi, Xiaofeng Zhu), Case Western Reserve University
    Selecting associated variants as instrument variables can bias multivariable Mendelian randomization

23. Wenjing Liu, The Ohio State University
    Impact of ignoring correlation in variable selection: A simulation-based study using LASSO

24. Mengxuan Li (with Yihe Yang, Xiaofeng Zhu), Case Western Reserve University
    Cell Type-Specific contributions of PRS of 71 complex traits

25. Xuerong Wang (with Meghna Kalra, Kiryung Lee, Yoonkyung Lee), The Ohio State University
    Graph-regularization approach to single-cell RNA sequencing data analysis with applications to cell clustering

26. Yan Zou (with Penglei Gao, Xiaofeng Wang), Case Western Reserve University
    Deep Survival Analysis for Competing Risk Modeling with Functional Covariates and Missing Data Imputation

Date	Host	Keynote Speaker	CCF Speaker	OSU Speaker	CWRU Speaker
04/18/1991	OSU	Gary Koch	M. Schluchter	J. Klein	NA
05/16/1995	CCF	Martin Tanner	J.S. Rao	D. Pearl	NA
05/17/1996	OSU	Bruce Weir	N. Obuchowski	M. Irwin	NA
05/2/1997	CCF	Nan Laird	T. Greene	J. Hsu?	J. Witte
05/28/1998	OSU	Kathryn Roeder	J. Gassman	H. Doss	R. Elston
05/13/1999	CWRU	Bruce Lindsay	N. Obuchowski	S. Lin	J. Sun A. Ivanova
05/11/2000	OSU	Scott Zeger	H. Ishwaran	F. Wright	J. Albert
05/10/2001	CCF	Butch Tsiatis	R. O'Brien	S. Lemeshow	N. Gordon
05/9/2002	OSU	Steven Piantadosi	C. Apperson-Hansen	L. Shen	NA
05/2/2003	CWRU	Nancy Geller	G. Beck	D. Burr	M. Diaz-Insua
05/13/2004	OSU	Mike West	X. Peng	S. Lin	T. Radivoyevitch
05/13/2005	CCF	Donald Rubin	L. Li	Y. Lee	J Albert
5/18/2006	OSU	C.R. Rao	M. Kattan	M.Lee	R. Elston
5/17/2007	CWRU	Cyrus Mehta	D. Babineau	B. Lu	J. Sun
5/15/2008	OSU	Frank Harrell	P. Imrey	H. Nagaraja	T. Radivoyevitch
5/14/2009	CCF	Colin Begg	X. Wang	M. Pennell	P. Fu
5/13/2010	OSU	Betz Halloran	Y. Xu	L. Kubatko	R. O'Brien
5/5/2011	CWRU	Rafael Irizarry	B. Hu	A. Shoben	A. Sattar
5/17/2012	OSU	Weng Kee Wang	J. Schold	H. Zhu	N. Morris
4/11/2013	CCF	Tom Louis	L. Li	G. Rempala	D. Gunzler
4/10/2014	OSU	Donald Berry	S. Griffith	V. Vu	J. Sun
4/20/2015	CWRU	Marie Davidian	J. Dalton	S. Kurtek	J. Albert
4/14/2016	OSU	Rebecca Betensky	E. Mascha	G. Brock	C. Chun
4/20/2017	CCF	Francesca Dominici	T. Dey	C. Song	A, Sattar
4/19/2018	OSU	Michael Boehnke	J. Rajeswaran	J. Huling	C. Tatsuoka
4/19/2019	CWRU	Kathryn Roeder	B. Hobbs	K. F Au	W. Bush
4/21/2021	Virtual - OSU	Rebecca Hubbard	T. H. Hwang	E Kenah	H. Feang
4/13/2023	CCF	Anastasia Ivanova	E. Zabor	X. Cai	L. Zhang
4/19/2024	OSU	Mei-Cheng Wang	L. Tereshchenko	F. Tounkara	M. Wang
4/7/2025	CWRU	Jeff Leek	C. Boyer	A. Ni	H. Hartman

Acknowledgement

Thanks to the Joint Biostatistics Symposium Committee and Sponsors

Local Committee:

Liangliang Zhang, Ming Wang, Holly Hartman, Hao Feng, Jeffery M. Albert, Lijun Zhang, Pingfu Fu, Abdus Sattar, Yilun Sun, Thomas E. Love, Yeunjoo E. Song, Xiaofeng Zhu, Mireya Diaz-Insua, Sudha K. Iyengar, Jonathan Haines from Department of Population & Quantitative Health Sciences, Weihong Guo from Department of Mathematics, Applied Mathematics, and Statistics, Case Western Reserve University.

Steering Committee:

Xiaofeng Zhu, Department of Population & Quantitative Health Sciences, Case Western Reserve University
Gerald Beck, Department of Quantitative Health Sciences, The Cleveland Clinic
Michael Pennell, Division of Biostatistics, College of Public Health, The Ohio State University
Shili Lin, Department of Statistics, The Ohio State University
Qing Wu, Department of Biomedical Informatics, The Ohio State University

Leadership support

J.Michael Oakes, Senior Vice President for Research, Office of Research and Technology Management
Gary Schwartz, Director of Case Comprehensive Cancer Center
Jonathan Haines, Chair of Department of Population and Quantitative Health Sciences
Stan Gerson, Dean, School of Medicine

Administrative Support

Tuesday Gibson, Department of Population and Quantitative Health Sciences
ASA Student Chapter at Case Western Reserve University

Financial Support

Jack H. Berman, MD Lecture Fund, School of Medicine
Case Comprehensive Cancer Center
Department of Population and Quantitative Health Sciences