

---

---

# BOOK OF ABSTRACTS

---

## EUROPEAN CAUSAL INFERENCE MEETING EUROCIM 2018

Florence, 11-13 April 2018

Presentations



# Contents

Essential concepts of causal inference – a remarkable history ( <i>Donald B. Rubin</i> )	1
Model averaging in instrumental variable analysis for consistent and efficient estimation when a plurality of candidate instruments are valid ( <i>Stephen Burgess, Verena Zuber, Apostolos Gkatzionis, Christopher Foley</i> )	1
Oracle-optimal confidence intervals for causal effects in Mendelian randomization via two stage hard thresholding with voting ( <i>Guo Z., Kang, H., Cai T.T., Small D.S.</i> )	2
Selection bias in mendelian randomization ( <i>Apostolos Gkatzionis and Stephen Burgess</i> )	2
Stochastic intervention on continuous instrument: Estimating the effects of visitation redicivism ( <i>Jacqueline A. Mauro, Edward H. Kennedy, Daniel Nagin</i> )	3
Data driven tuning parameters for matching with stronger instruments ( <i>Samuel D. Pimentel, Rachel R. Kelz, Luke Keele</i> )	3
Generalized Optimal Matching for Inference and Policy Learning ( <i>Nathan Kallus</i> )	4
High-Dimensional Doubly Robust Inference For Regression Parameters ( <i>Oliver Dukes, Vahe Avagyan and Stijn Vansteelandt</i> )	4
Data-adaptive doubly robust instrumental variable methods for treatment effect heterogeneity ( <i>K. DiazOrdaz, R. Daniel, N. Kreif</i> )	5
Model Misspecification And Bias For Inverse Probability Weighting And Doubly Robust Estimators ( <i>Ingeborg Waernbaum and Laura Pazzagli</i> )	6
Nonparametric Bayes models for doubly robust causal inference for high-dimensional data ( <i>Joseph Antonelli and Francesca Dominici</i> )	6
Model-assisted design of experiments in the presence of social network information ( <i>Edoardo Airoldi</i> )	7
Handling Time-Dependent Confounding Using a Structural Nested Cumulative Survival Time ( <i>Shaun Seaman, Stijn Vansteelandt and Oliver Dukes</i> )	7
Optimal Balancing of Time - Dependent Confounders for Marginal Structural Models ( <i>Michele Santacatterina and Nathan Kallus</i> )	8
Estimating the Causal Effect of State Transitions Using Structural Mean Models ( <i>Paul Clarke and Yanchun Bao</i> )	9
Mediation Analysis Synthetic Control ( <i>Giovanni Mellace and Alessandra Pasquini</i> )	9
A note on causal mediation analysis with a survival outcome ( <i>Vanessa Didelez</i> )	11
Insights on Variance Estimation for Blocked and Matched Pairs Designs ( <i>Nicole Pashley and Luke Miratrix</i> )	11
Penalized Spline of Propensity Methods for Treatment Comparisons in Longitudinal Treatment Settings ( <i>Tingting Zhou, Michael Elliott, Roderick Little</i> )	12

Bayesian Inference about Optimal Treatment Regimens Using Observational Data ( <i>Thomas Klausch, Peter van de Ven, Johannes Berkhof</i> ) . . . . .	12
Choosing Optimal Dynamic Treatment Regime Methods when Analyzing Observational Data ( <i>Beth Ann Griffin, Denis Agniel, Daniel Almirall</i> ) . . . . .	13
Time-Varying Survivor Average Causal Effects with Semicompeting Risks ( <i>Leah Comment, Fabrizia Mealli, Corwin Zigler</i> ) . . . . .	14
Local Confounding Adjustment in Exposure Response Modeling ( <i>Francesca Dominici, Georgia Papadogeorgou</i> ) . . . . .	14
Overlap in Observational Studies with High-Dimensional Covariates ( <i>Alexander D'Amour, Peng Ding, Avi Feller, Lihua Lei, and Jasjeet Sekhon</i> ) . . . . .	15
Augmented Minimax Linear Estimation: Generalizing Double-Robust Balancing Estimators ( <i>David A. Hirshberg and Stefan Wager</i> ) . . . . .	15
Calibrated Inverse Probability of Treatment Weights By Covariate Balancing for Marginal Structural Models ( <i>Sean Yiu and Li Su</i> ) . . . . .	16
Estimating Population Average Causal Effects in the Presence of Non-Overlap: A Bayesian approach ( <i>Rachel C. Nethery, Fabrizia Mealli, Francesca Dominici</i> ) . . . . .	17
Instrument-based estimation with binarized treatments: Issues and tests for the exclusion restriction ( <i>Martin Eckhoff Andresen and Martin Huber</i> ) . . . . .	17
Causal Interpretation in Survival Analysis ( <i>Kjetil Røysland</i> ) . . . . .	18
Marginal Structural Models for Survival Analysis in Continuous Time: Theoretically Appealing and Practically Feasible ( <i>Mats J. Stensrud</i> ) . . . . .	18
Rethinking meta-analysis: addressing problems of non-transportability when combining treatment effects across patient populations ( <i>Vo Tat Thang, Raphael Porcher, Stijn Vansteelandt</i> ) . . . . .	19
Discovering effect modification and randomization inference in air pollution studies ( <i>Kwonsang Lee, Dylan Small, and Francesca Dominici</i> ) . . . . .	20
Causal Inference in the Presence of Interference ( <i>Michael Hudgens</i> ) . . . . .	20
Average treatment effects in the presence of unknown interference ( <i>Fredrik Sävje, Peter M. Aronow, Michael G. Hudgens</i> ) . . . . .	21
Causal inference for interfering units with cluster and population level treatment allocation programs ( <i>Georgia Papadogeorgou, Fabrizia Mealli, Corwin Zilger</i> ) . . . . .	21
Propensity score methods for causal inference in observational studies with network interference ( <i>Bruno Arpino, Luca De Benedictis and Alessandra Mattei</i> ) . . . . .	22
On the perils of stacking thresholds on RD designs ( <i>Margherita Fort, Andrea Ichino, Enrico Rettore, Giulio Zanella</i> ) . . . . .	22
Disease subtypes in aetiological research ( <i>Lorenzo Richiardi</i> ) . . . . .	23
How to obtain valid tests and confidence intervals for treatment effects after confounder selection? ( <i>Stijn Vansteelandt and Oliver Dukes</i> ) . . . . .	23
Data-Adaptive Variable Importance with CV-TMLE ( <i>Chris J. Kennedy and Alan E. Hubbard</i> ) . . . . .	24
Challenges in Estimating the Causal Effects of a Complex Multi-Dimensional Exposure ( <i>Sarasvati Bahadursingh, Sarah Lewington, Sofia Massa, Iona Millwood, Naomi Allen</i> ) . . . . .	24

## Essential concepts of causal inference – a remarkable history

11<sup>th</sup> April  
09:30 - 10:30

Donald B. Rubin  
Harvard University

I believe that a deep understanding of cause and effect, and how to estimate causal effects from data, complete with the associated mathematical notation and expressions, only evolved in the twentieth century. The crucial idea of randomized experiments was apparently first proposed in 1925 in the context of agricultural field trails but quickly moved to be applied also in studies of animal breeding and then in industrial manufacturing. The conceptual understanding seemed to be tied to ideas that were developing in quantum mechanics. The key ideas of randomized experiments evidently were not applied to studies of human beings until the 1950s, when such experiments began to be used in controlled medical trials, and then in social science - in education and economics. Humans are more complex than plants and animals, however, and with such trials came the attendant complexities of non-compliance with assigned treatment and the occurrence of “Hawthorne” and placebo effects. The formal application of the insights from earlier simpler experimental settings to more complex ones dealing with people, started in the 1970s and continue to this day, and include the bridging of classical mathematical ideas of experimentation, including fractional replication and geometrical formulations from the early twentieth century, with modern ideas that rely on powerful computing to implement aspects of design and analysis.

---

## Model averaging in instrumental variable analysis for consistent and efficient estimation when a plurality of candidate instruments are valid

11<sup>th</sup> April  
11:00 - 11:30

Stephen Burgess<sup>1,2</sup>, Verena Zuber<sup>1</sup>, Apostolos Gkatzionis<sup>1</sup>, Christopher Foley<sup>1</sup>

<sup>1</sup>MRC Biostatistics Unit, University of Cambridge, Cambridge, UK, CB2 0SR

<sup>2</sup>Cardiovascular Epidemiology Unit, University of Cambridge, Cambridge, UK, CB1 8RN

Several robust methods for instrumental variable analysis have been developed that do not require all candidate instruments to be valid to give consistent estimates of a causal parameter. We here develop a model averaging method that gives consistent estimates under a ‘plurality of valid instruments’ assumption. This means that the largest subset of candidate instruments with the same causal estimate (in a large sample size) comprises the valid instruments. The method considers a mixture distribution of estimates derived from each subset of candidate instruments. The estimates are weighted such that subsets with more candidate instruments receive more weight, unless candidate instruments in the subset have heterogeneous causal estimates, in which case that subset is severely downweighted. The mode of this mixture distribution is the causal estimate.

The method is asymptotically efficient, and has good robustness properties. We present an example analysis in the context of Mendelian randomization in which the mixture distribution is bimodal. This proposed method suggests two distinct mechanisms by which inflammation affects coronary heart disease risk, with subsets of genetic variants suggesting both positive and negative causal effects.

11<sup>th</sup> April  
11:30 - 11:45

## Oracle-optimal confidence intervals for causal effects in Mendelian randomization via two stage hard thresholding with voting

Guo Z.<sup>1</sup>, Kang, H.<sup>2</sup>, Cai T.T.<sup>3</sup>, Small D.S.<sup>3</sup>

<sup>1</sup>Rutgers University (Statistics),

<sup>2</sup>University of Wisconsin-Madison (Statistics),

<sup>3</sup>The Wharton School of Business, University of Pennsylvania (Statistics)

A major challenge in Mendelian randomization/instrumental variables (IV) analysis is to find instruments that are valid, or have no direct effect on the outcome and are ignorable. Typically one is unsure whether all of the putative IVs are in fact valid. We propose a general inference procedure in the presence of invalid IVs, called Two-Stage Hard Thresholding (TSHT) with voting. TSHT uses two hard thresholding steps to select strong instruments and generate candidate sets of valid IVs. Voting takes the candidate sets and uses majority and plurality rules to determine the true set of valid IVs. In low dimensions with invalid instruments, our proposal (i) correctly selects valid IVs, (ii) consistently estimates the causal effect, (iii) produces valid confidence intervals for the causal effect, and (iv) has oracle-optimal width in many settings, even if the so-called 50% rule or the majority rule is violated. In high dimensions, we establish nearly identical results without oracle-optimality. We also propose extensions of TSHT to two-sample summary data, which is popular in Mendelian randomization settings. Finally, we apply our method to simulated and real data and find that our proposal outperforms traditional and recent methods in MR for inferring causal effects.

---

11<sup>th</sup> April  
11:45 - 12:00

## Selection bias in mendelian randomization

Apostolos Gkatzionis and Stephen Burgess

MRC Biostatistics Unit, University of Cambridge

Mendelian randomization is the use of genetic information to assess whether a causal relationship exists between a risk factor and an outcome. It is the application of instrumental variables analysis in epidemiology, using genetic variants as instruments. Like most epidemiological studies, Mendelian randomization analyses can be affected by selection bias. This typically occurs when selection of participants in the study is influenced by either the risk factor or the outcome. In this case selection becomes a collider between the genetic instrument and confounders of the risk factor-outcome association. Restricting attention to the selected individuals induces dependence between the genetic instrument and confounders, violating one of the core Mendelian randomization assumptions and introducing bias.

We investigate by simulation the impact of selection bias in Mendelian randomization analyses. Our results suggest that selection bias is often weaker than other sources of bias, but can be a real concern if the effect of the risk factor and/or outcome on selection is strong. We discuss how the magnitude of bias is affected by selection frequency, instrument strength, the strength of the confounder-risk factor and confounder-outcome associations and the causal DAG structure. We also investigate the performance of inverse probability weighting as a method to correct selection bias. As an application, we

consider a recently reported finding that lipoprotein(a) is a causal risk factor for cardiovascular disease risk but not for cardiovascular mortality. We explore whether this rather counterintuitive finding could be explained by selection bias.

---

**Stochastic intervention on continuous instrument: Estimating the effects of visitation recidivism**

11<sup>th</sup> April  
12:00 - 12:15

Jacqueline A. Mauro, Edward H. Kennedy, Daniel Nagin  
Carnegie Mellon University, Departments of Statistics and Public Policy  
Carnegie Mellon University, Department of Statistics  
Carnegie Mellon University, Heinz College

Robust causal inference with Instrumental Variables (IVs) is a powerful tool in observational settings, but is often limited to estimating the effect of intervening on all subjects at a specific level vs. on none. In many cases, this is not the effect that is most interesting to subject matter experts, or it is inappropriate given the data. In this paper, we develop nonparametric instrumental variables methods which extend modern influence function-based doubly robust IV methods to more general settings by investigating the effects of arbitrary interventions on continuous instruments. These do not require the strong parametric modeling assumptions of other methods. This new methodology was motivated by the need to investigate the effects of visitation on recidivism. We apply one version of these new estimators to the question of whether visitation will reduce the chances of recidivism for prisoners released from Pennsylvania prisons in 2008. We find that policies which facilitate visitation may lead to a decrease in recidivism.

---

**Data driven tuning parameters for matching with stronger instruments**

11<sup>th</sup> April  
12:15 - 12:30

Samuel D. Pimentel<sup>1</sup>, Rachel R. Kelz<sup>2</sup>, Luke Keele<sup>3</sup>

<sup>1</sup>Department of Statistics, University of California, Berkeley

<sup>2</sup>Department of Surgery, Hospital of the University of Pennsylvania

<sup>3</sup>McCourt School of Public Policy, Georgetown University

Instrumental variable methods are widely used in observational studies. In observational studies of this type, investigators use some naturally occurring variation as a haphazard nudge for exposure to a treatment. Given that the instrument is not randomly assigned in these contexts, possible instrument-outcome confounding is a source of bias. However, stronger instruments reduce the risk of bias from confounders of this type. Specialized matching algorithms can be used to both reduce bias from observed confounders and strengthen the instrument. However, these matching algorithms have tuning parameters that users must select with little guidance. We propose a sample splitting procedure for the selection of these tuning parameters. The match we propose relies on an optimal tradeoff between sample size and the strength of the instrument. We use robustness to bias from unobserved confounders to derive data driven tuning parameters. We apply our methods to a data on the effects of c-sections. Specifically,

we estimate the effect c-sections on in-hospital complications. In this application, we use variability in a physicians' tendency to deliver via c-sections as an instrumental variable for birth by c-section.

---

11<sup>th</sup> April  
14:00 - 14:30

## **Generalized Optimal Matching for Inference and Policy Learning**

Nathan Kallus  
Cornell University

We develop an encompassing framework for matching, covariate balancing, and doubly-robust methods for causal inference from observational data called generalized optimal matching (GOM). The framework is given by generalizing a new functional-analytical formulation of optimal matching, giving rise to the class of GOM methods, for which we provide a single unified theory to analyze tractability, consistency, and efficiency. Many commonly used existing methods are included in GOM and, using their GOM interpretation, can be extended to optimally and automatically trade off balance for variance and outperform their standard counterparts. As a subclass, GOM gives rise to kernel optimal matching (KOM), which, as supported by new theoretical and empirical results, is notable for combining many of the positive properties of other methods in one. KOM, which is solved as a linearly-constrained convex-quadratic optimization problem, inherits both the interpretability and model-free consistency of matching but can also achieve the  $\sqrt{n}$ -consistency of well-specified regression and the efficiency and robustness of doubly robust methods. In settings of limited overlap, KOM enables a very transparent method for interval estimation for partial identification and robust coverage. We demonstrate these benefits in examples with both synthetic and real data.

---

11<sup>th</sup> April  
14:30 - 14:45

## **High-Dimensional Doubly Robust Inference For Regression Parameters**

Oliver Dukes<sup>1</sup>, Vahe Avagyan<sup>1</sup> and Stijn Vansteelandt<sup>1,2</sup>

<sup>1</sup>Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Belgium.

<sup>2</sup>Department of Medical Statistics, London School of Hygiene and Tropical Medicine, UK

It is known that after variable selection, standard inferential procedures for regression parameters may not be uniformly valid; there is no finite sample size at which a procedure is guaranteed to attain its nominal coverage/size (within pre-specified error margins). This problem is exacerbated in high-dimensional settings where variable selection becomes unavoidable. This has prompted a flurry of activity in developing uniformly valid confidence intervals and hypothesis tests for a low-dimensional regression parameter (e.g. the effect of a treatment  $A$  on an outcome  $Y$ ) in high-dimensional models. Recent proposals aim to 'debias' or 'desparsify' either the estimating equations for the treatment effect or the estimates themselves. So far there has been limited focus on model misspecification, although this is inevitable in high-dimensional settings.

We attempt to unify the recent proposals and obtain inferential procedures for a treatment effect parameter in a high-dimensional generalized linear model that are doubly robust. This means that they are uniformly valid if either a working model for the exposure or the outcome is correctly specified. They moreover deliver uniformly consistent standard error estimates even when both working models are misspecified. When all models are correct, then our confidence intervals will continue to attain their nominal coverage under sparsity conditions weaker than those typically invoked in the literature. The proposal is illustrated in a simulation study and an analysis of data from the Ghent University Intensive Care Unit, where we consider the effect of a change in glycemia level on mortality in critically ill patients.

---

## **Data-adaptive doubly robust instrumental variable methods for treatment effect heterogeneity**

11<sup>th</sup> April  
14:45 - 15:00

K. DiazOrdaz<sup>1</sup>, R. Daniel<sup>2</sup>, N. Kreif<sup>3</sup>

<sup>1</sup>Department of Medical Statistics, LSHTM

<sup>2</sup>Division of Population Medicine, Cardiff University

<sup>3</sup>Centre for Health Economics, University of York

We consider the estimation of the average treatment effect in the treated as a function of baseline covariates, where there is a valid (conditional) instrument.

We describe two doubly robust (DR) estimators: a locally efficient g-estimator, and a targeted minimum loss-based estimator (TMLE). These two DR estimators can be viewed as generalisations of the two-stage least squares (TSLS) method to semi-parametric models that make weaker assumptions. We exploit recent theoretical results that extend to the g-estimator the use of data-adaptive fits for the nuisance parameters.

A simulation study is used to compare standard TSLS with the two DR estimators' finite-sample performance, (1) when fitted using parametric nuisance models, and (2) using data-adaptive nuisance fits, obtained from the Super Learner, an ensemble machine learning method.

Data-adaptive DR estimators have lower bias and improved coverage, when compared to incorrectly specified parametric DR estimators and TSLS. When the parametric model for the treatment effect curve is correctly specified, the g-estimator outperforms all others, but when this model is misspecified, TMLE performs best, while TSLS can result in huge biases and zero coverage.

Finally, we illustrate the methods by reanalysing the COPERS (COping with persistent Pain, Effectiveness Research in Self-management) trial to make inference about the causal effect of treatment actually received, and the extent to which this is modified by depression at baseline.

---

11<sup>th</sup> April  
15:00 - 15:15

## **Model Misspecification And Bias For Inverse Probability Weighting And Doubly Robust Estimators**

Ingeborg Waernbaum<sup>1</sup> and Laura Pazzagli<sup>2</sup>

<sup>1</sup>Umeå University,  
<sup>2</sup>University of Perugia

In the causal inference literature a class of semi-parametric estimators is called robust if the estimator has desirable properties under the assumption that at least one of the working models is correctly specified. A standard example is a doubly robust estimator that specifies parametric models both for the propensity score and the outcome regression. When estimating a causal parameter in an observational study the role of parametric models is often not to be true representations of the data generating process, instead the motivation for their use is to facilitate the adjustment for confounding, for example by reducing the dimension of the covariate vector, making the assumption of at least one true model unlikely to hold.

In this paper we propose a crude analytical approach to study the large sample bias of estimators when all models are assumed to be approximations of the true data generating process, i.e., all models are misspecified. We apply our approach to three prototypical estimators, two inverse probability weighting (IPW) estimators, using a misspecified propensity score model, and a doubly robust (DR) estimator, using misspecified models for the outcome regression and the propensity score. To analyze the question of when the use of two misspecified models are better than one we derive necessary and sufficient conditions for when the DR estimator has a smaller bias than the simple IPW estimator and when it has a smaller bias than the IPW estimator with normalized weights. If the misspecification of the outcome model is moderate the comparisons of the biases of the IPW and DR estimators suggest that the DR estimator has a smaller bias than the IPW estimators. However, all biases include the PS-model error and we suggest that a researcher is careful when modeling the PS whenever such a model is involved. For numerical and finite sample illustrations we also include a simulation study.

---

11<sup>th</sup> April  
15:15 - 15:30

## **Nonparametric Bayes models for doubly robust causal inference for high-dimensional data**

Joseph Antonelli and Francesca Dominici  
Harvard T.H Chan School of Public Health

We introduce causal inference methods for high-dimensional scenarios that utilize non-parametric Bayesian methods to alleviate modeling assumptions, and sparsity inducing priors to reduce the dimension of the covariate space. By leveraging the Bayesian framework, we can acquire the posterior distribution of any estimator that can be defined as a function of the treatment and outcome model (e.g. standard doubly robust estimator). We illustrate how the proposed framework allows us to extend many existing estimators to highly nonlinear or high-dimensional scenarios while maintaining desirable small sample properties. We account for all of the uncertainty inherent in the data by combining our posterior samples with an efficient resampling procedure that will only slightly increase computation time, while giving credible intervals that obtain good

frequentist properties in finite samples and achieve nominal coverage rates. We show via simulation the ability of the proposed estimator to flexibly estimate causal effects in high-dimensions. Finally, we apply our proposed procedure to estimate the effect of environmental exposures on health outcomes.

---

## **Model-assisted design of experiments in the presence of social network information**

11<sup>th</sup> April  
16:00 - 16:30

Edoardo Airoidi  
Harvard University

Classical approaches to causal inference largely rely on the assumption of “lack of interference”, according to which the outcome of an individual does not depend on the treatment assigned to others, as well as on many other simplifying assumptions, including the absence of strategic behavior. In many applications, however, such as evaluating the effectiveness of healthcare interventions that leverage social structure, assessing the impact of product innovations and ad campaigns on social media platforms, or experimentation at scale in large IT companies, assuming lack of interference and other simplifying assumptions is untenable. Moreover, the effect of interference itself is often an inferential target of interest, rather than a nuisance. In this talk, we will formalize technical issues that arise in estimating causal effects when interference can be attributed to a network among the units of analysis, within the potential outcomes framework. We will introduce and discuss several strategies for experimental design in this context centered around a judicious use statistical models, which we refer to as “model-assisted” design of experiments. In particular, we wish for certain finite-sample properties of the estimator to hold even if the model catastrophically fails, while we would like to gain efficiency if certain aspects of the model are correct. We will then contrast design-based, model-based and model-assisted approaches to experimental design from a decision theoretic perspective.

---

## **Handling Time-Dependent Confounding Using a Structural Nested Cumulative Survival Time**

11<sup>th</sup> April  
16:30 - 16:45

Shaun Seaman<sup>1</sup>, Stijn Vansteelandt<sup>2,3</sup> and Oliver Dukes<sup>2</sup>

<sup>1</sup>MRC Biostatistics Unit, University of Cambridge, UK

<sup>2</sup>Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Belgium

<sup>3</sup>Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, UK

Time-varying confounding is a commonly encountered problem when estimating the total effect of an exposure on a time-to-event outcome using data from an observational study. Failing to account for confounding leads to bias in the estimated exposure effects, but standard regression adjustment for confounding eliminates the indirect effect of the exposures. Marginal structural models, fitted by inverse probability weighting, can be

used, but these can be subject to the problem of highly variable weights when the confounders are strongly predictive of the exposure or when the exposure is continuous. Structural nested accelerated failure time models can be fitted by g-estimation, but this requires artificial recensoring, which causes loss of information.

We have developed an alternative method using a structural nested cumulative survival time model. This method has the advantage that it avoids inverse probability weighting and artificial recensoring. The assumed model is closely related to the structural nested cumulative failure time model of Picciotto et al. (2011), but our method can accommodate irregular measurement times and exposure effect estimates can be calculated in closed form using g-estimation.

The effects of exposures at different times can be estimated in an unrestricted way or, to gain efficiency and to reduce the number of parameters that need to be interpreted, restrictions can be imposed. We shall describe the structural nested cumulative survival time model, discuss how to use standard software to estimate parameters, present results from simulation studies, and consider possible restrictions on parameters.

---

## Optimal Balancing of Time - Dependent Confounders for Marginal Structural Models

11<sup>th</sup> April  
16:45 - 17:00

Michele Santacatterina<sup>1</sup> and Nathan Kallus<sup>2</sup>

<sup>1</sup>Unit of Biostatistics, Karolinska Institutet, Stockholm, Sweden, 17177

<sup>2</sup>School of Operations Research and Information Engineering and Cornell Tech, Cornell University, New York, New York 10011

Marginal structural models (MSM) have been used to estimate the causal effect of a time-varying treatment on an outcome of interest with longitudinal data in observational studies. Despite their theoretical appeal, these methods are highly sensitive to misspecifications of the treatment assignment model and can lead high-variance estimates due to extreme weights whenever true propensities may be close to 0 or 1. In this talk, we present Kernel Optimal Weighting (KOW) which provides weights that optimally balance time-invariant and time-dependent confounders while controlling for the precision of the resulting MSM estimate. The method is built upon optimal balance to directly address the error objective, using reproducing kernel Hilbert spaces to describe the potential imbalances due to time-dependent confounders based on a new interpretation of the g-computation formula. We evaluate the performance of KOW in a simulation study, comparing it with inverse probability weighting, stable inverse probability weighting, and the recently proposed covariate balancing propensity scores. We further extended KOW to control for informative censoring. We apply KOW to study the effect of treatment initiation on time to death among people who live with HIV and to study the effect of negative advertisements on vote shares in the United States.

## Estimating the Causal Effect of State Transitions Using Structural Mean Models

11<sup>th</sup> April  
17:00 - 17:15

Paul Clarke and Yanchun Bao  
University of Essex, UK

Structural mean models (SMMs) are a class of semiparametric models originally designed to estimate causal effects of treatment regimes from observational clinical trials in which there is patient noncompliance. Despite being an appropriate modelling strategy, SMMs have not been widely used for modelling longitudinal panel data (Vansteelandt and Joffe 2014). In this presentation, we use SMMs to estimate the causal effect of employment-status transitions on mental health using the British Household Panel Study (BHPS). We explore the issues raised by focussing on transitions rather than states, and propose an extended g-estimator which accounts for the problems this creates.

G-estimation is carried out using the generalized method of moments (Clarke et al. 2015). We contrast the assumptions required for g-estimator consistency with those needed for the maximum likelihood estimation of multiprocess models like those used by Steele et al. (2013), and compare the robustness of these approaches in a simulation study.

### References:

- Clarke PS, Palmer TM, Windmeijer F (2015) Estimating structural mean models with multiple instrumental variables using the generalised method of moments, *Statistical Science* 30, 96-117.
- Steele F, French R, Bartley M (2013) Adjusting for selection bias in longitudinal analyses using simultaneous equations modeling: the relationship between employment transitions and mental health, *Epidemiology* 24, 703-711.
- Vansteelandt S, Joffe M (2014) Structural nested models and g-estimation: the partially realized promise, *Statistical Science* 29, 707-731

---

## Mediation Analysis Synthetic Control

11<sup>th</sup> April  
17:15 - 17:30

Giovanni Mellace<sup>1</sup> and Alessandra Pasquini<sup>2</sup>

<sup>1</sup>University of Southern Denmark

<sup>2</sup>Sapienza University of Rome

The Synthetic Control Method (SCM) allows estimating the casual effect of an intervention even when data are available only on one treated and few control units. This is possible by using information on pre-intervention period to build a “synthetic control” mimicking what would have happened to the treated unit, in post-intervention period, in absence of the intervention. In mediation analysis, the total effect of a treatment is decomposed into the direct effect of the treatment and its indirect effects which go through intermediate variables (mediators) lying on the causal pathway between the treatment and the outcome of interest. We introduce the Mediation Analysis Synthetic Control (MASC), a generalization of SCM that allows decomposing the total effect of the intervention into its indirect effects, which go through observed mediators, and its

direct effect. To give an example, MASC re-weights the control unit post-intervention outcomes, choosing weights that minimize the distance between treated and control in pre-intervention observable characteristics (including pre-intervention values of the outcome and the mediator), as well as post-intervention values of the mediator, to mimic what would have happened to the treated, in absence of the intervention, if her mediator was set to her potential mediator under treatment. Finally, we illustrate MASC with a simple factor model with interactive fixed effects and show convergence of the direct and indirect effects estimator as the number of pre-intervention periods goes to infinity. An extensive simulation study and an empirical application are currently in progress.

---

## **A note on causal mediation analysis with a survival outcome**

12<sup>th</sup> April  
9:30 - 10:30

Vanessa Didelez

Leibniz Institute, University of Bremen

Various notions of (in)direct effects are based on nested counterfactuals. These may have problems of interpretation and identifiability when the outcome of interest is a survival or time-to-event and the mediator a set of longitudinal measurements. We propose an alternative definition that does not suffer from those problems, following the extended graphical approach of Robins and Richardson (2011). We give assumptions allowing identification of such alternative mediated effects resulting in the familiar mediation g-formula. Relations to path-specific effects and edge-interventions will be discussed. The approach is an interesting alternative motivation for any causal mediation setting, but especially for survival outcomes.

---

## **Insights on Variance Estimation for Blocked and Matched Pairs Designs**

12<sup>th</sup> April  
11:00 - 11:30

Nicole Pashley and Luke Miratrix

Harvard Graduate School of Education

The current literature on blocking has two main branches. The first focuses on larger blocks, with multiple treatment and control units in each block. The second focuses on matched pairs, with a single treatment and control unit in each block. For larger blocks, variance estimation is relatively straightforward. For matched pairs, however, because one cannot directly estimate variance within a block we have to use estimators that look at variation across the blocks. These alternative estimators have been evaluated under different assumptions than found in the large block literature. Because of this, these two literatures do not handle cases with blocks of varying sizes, but which contain singleton treatment or control units. This has also created some confusion regarding the benefits of blocking in general. In this talk, we reconcile the literatures by carefully examining the performance of different estimators of treatment effect and of variance under several different frameworks. We also provide variance estimators for experiments with many small blocks of different sizes and for experiments with mixtures of large and small blocks. We finally discuss in which situations blocking is or is not guaranteed to reduce the variance of our estimator.

---

12<sup>th</sup> April  
11:30 - 11:45

## **Penalized Spline of Propensity Methods for Treatment Comparisons in Longitudinal Treatment Settings**

Tingting Zhou, Michael Elliott, Roderick Little  
University of Michigan Department of Biostatistics

Observational studies lack randomized treatment assignment, and as a result valid inference about causal effects can only be drawn by controlling for confounders. When time dependent confounders are present that serve both as mediators of treatment effects and affect future treatment assignment, standard regression methods for controlling for confounders fail. We propose a robust Bayesian approach to causal inference in this setting we term Penalized Spline of Propensity Methods for Treatment Comparison (PENCOMP), which builds on the Penalized Spline of Propensity Prediction method for missing data problems. PENCOMP estimates causal effects by imputing missing potential outcomes with flexible spline models, and draws inference based on imputed and observed outcomes. We demonstrate that PENCOMP has a double robustness property for causal effects - that is, unbiased estimates can be obtained when either the treatment assignment or mean model is correctly specified, and simulations suggest that it tends to outperform doubly-robust marginal structural modeling when the relationship between propensity score and outcome is nonlinear or when the weights are highly variable. We further consider the issue of “overlap” - that is, restricting inference to the subset of the population in which the sampled data allows estimation of all of the treatments of interest without extrapolation. This is particularly important in the setting of longitudinal treatments for which PENCOMP is designed, where the set of estimable treatment patterns may be quite limited. We apply our method to evaluate the effects of antiretroviral treatment on CD4 counts in HIV infected patients.

12<sup>th</sup> April  
11:45 - 12:00

---

## **Bayesian Inference about Optimal Treatment Regimens Using Observational Data**

Thomas Klausch, Peter van de Ven, Johannes Berkhof  
Department of Epidemiology and Biostatistics, VU University Medical Centre Amsterdam, The Netherlands

Using patient characteristics for personalizing treatment decisions is an increasingly important objective in evidence-based medicine and statistics. Optimal treatment regimens enable personalization of treatment decisions in such a way that the expected benefit from treatment is maximized.

Based on the Rubin Causal Model, we suggest a Bayesian inference framework that allows estimating and updating optimal treatment regimens for dichotomous treatment outcomes using observational data under the assumption of unconfoundedness. Optimal treatment regimens are estimated using predictions of the response probabilities to all treatments and a loss function that penalizes sub-optimal decisions. As the true functional relationship between the potential outcomes and patient characteristics is unknown, we apply flexible Bayesian modelling strategies.

An important feature of our approach is the possibility to quantify uncertainty concerning the expected benefit under the optimal treatment regimen and the optimality

of the patient-specific treatment decisions. We also include a Bayesian sensitivity analysis to account for uncertainty regarding the partial correlation between the potential outcomes which is always unobserved in practice. Earlier approaches cannot quantify the aforementioned uncertainties which are, however, pivotal for evaluating treatment regimens.

We demonstrate the usefulness of our approach for a selection of modelling techniques in nonlinear multivariate settings by a simulation study and an application to data from head and neck cancer patients collected within the European “BD2DECIDE” consortium. We evaluate Bayesian additive regression trees, Bayesian polynomial regression, and Bayesian B-splines using regularizing and sparse priors, suggested when both variable selection and function approximation are required.

---

## Choosing Optimal Dynamic Treatment Regime Methods when Analyzing Observational Data

Beth Ann Griffin<sup>1</sup>, Denis Agniel<sup>2</sup>, Daniel Almirall<sup>3</sup>

<sup>1</sup>RAND Corporation; Co-Director RAND Center for Causal Inference

<sup>2</sup>RAND Corporation, Santa Monica

<sup>3</sup>University of Michigan, Institute for Social Research

12<sup>th</sup> April  
12:00 - 12:15

Novel statistical methods to estimate decision rules to help inform personalized treatment decisions are actively being developed, including methods like Q-learning, outcome weighted learning, decision lists, and extensions thereof. The area of dynamic treatment regimens (DTR) is an active area of statistical research providing a wide range of applied areas with a valuable opportunity to generate new hypotheses about how best to optimize treatments and interventions for individuals. New DTR methodologies are needed to help researchers appropriately analyze large longitudinal, observational databases commonly found in healthcare. In this talk/poster, the speaker will share new work utilizing DTR methodology together with observational data, including detailed discussion on how to select between DTR methods by evaluating their performance using hold-out data. The methodology will be applied within the context of selecting optimal treatment regimens for adolescent substance users from a large ( $n > 14,000$ ), observational cohort of youth undergoing substance abuse treatment collected as part of the Global Appraisal of Individual Needs (GAIN) assessment system. Information was collected on setting of care (inpatient/residential, intensive outpatient, outpatient, or none) at approximately 3-month intervals for 12 months. The speaker will also describe novel incorporation of stakeholder feedback as inputs in formulating decision rules which was done as part of the ongoing study.

12<sup>th</sup> April  
12:15 - 12:30

## **Time-Varying Survivor Average Causal Effects with Semicompeting Risks**

Leah Comment<sup>1</sup>, Fabrizia Mealli<sup>2</sup>, Corwin Zigler<sup>1</sup>

<sup>1</sup>Department of Biostatistics, Harvard T.H. Chan School of Public Health

<sup>2</sup>Department of Statistics, Informatics, Applications “G. Parenti,” University of Florence

In semicompeting risks problems, non-terminal time-to-event outcomes such as time to hospital readmission are subject to truncation by death. Such settings are often evaluated with parameters from illness-death models, but evaluating causal treatment effects with such models is problematic due to the evolution of incompatible risk sets over time. As an alternative, the survivor average causal effect (SACE) is a principal stratum causal effect of a treatment on the non-terminal event among units that would survive regardless of the assigned treatment. Traditional SACE formulations specify a single time point past which an individual is deemed to have always survived.

We propose a new causal estimand, the time-varying SACE (TV-SACE), for non-terminal events in the semicompeting risks setting. We adopt a Bayesian estimation procedure that is anchored to parameterization of illness-death models for both treatment arms but maintains causal interpretability. We outline several frailty specifications and highlight their connection to assumptions from the principal stratification literature. The method is demonstrated with data on hospital readmission for pancreatic cancer patients.

---

12<sup>th</sup> April  
14:00 - 14:30

## **Local Confounding Adjustment in Exposure Response Modeling**

Francesca Dominici, Georgia Papadogeorgou

Harvard T.H. Chan School of Public Health

In the last two decades, environmental regulations have substantially reduced ambient exposure to air pollution. However, an important and politically charged question remains: is long term exposure to low air pollution levels, well below the safety standards, still harmful? Several parametric and semi-parametric regression modeling approaches have been developed for flexibly estimating the exposure-response (ER) curve. However, these approaches: 1) are not formulated in the context of a potential outcome framework for causal inference; 2) adjust for the same set of potential confounders across all levels of exposure; and 3) do not account for model uncertainty regarding covariate selection. We propose an innovative Bayesian causal inference framework for ER estimation that allows for: a) different confounders and different strength of confounding at the different exposure levels; and b) model uncertainty regarding confounders' selection and the shape of the ER. As a consequence, it provides a principled way of assessing the observed covariates' confounding importance at different exposure levels. We apply our methods to estimate the whole causal ER and the causal effects of changes in PM2.5 at low levels on all-cause cardiovascular hospitalization rates.

## Overlap in Observational Studies with High-Dimensional Covariates

12<sup>th</sup> April  
14:30 - 14:45

Alexander D'Amour, Peng Ding, Avi Feller, Lihua Lei, and Jasjeet Sekhon

Department of Statistics, UC Berkeley

Causal inference in observational settings typically rests on a pair of identifying assumptions: (1) unconfoundedness and (2) covariate overlap, also known as positivity or common support. Investigators often argue that unconfoundedness is more plausible when many covariates are included in the analysis. Less discussed is the fact that covariate overlap is more difficult to satisfy in this setting. In this paper, we explore the implications of overlap in high-dimensional observational studies, arguing that this assumption is stronger than investigators likely realize. Our main innovation is to frame (strict) overlap in terms of bounds on a likelihood ratio, which allows us to leverage and expand on existing results from information theory. In particular, we show that strict overlap bounds discriminating information (e.g., Kullback-Leibler divergence) between the covariate distributions in the treated and control populations. We use these results to derive explicit bounds on the average imbalance in covariate means under strict overlap and a range of assumptions on the covariate distributions. Importantly, these bounds grow tighter as the dimension grows large, and converge to zero in some cases. We examine how restrictions on the treatment assignment and outcome processes can weaken the implications of certain overlap assumptions, but at the cost of stronger requirements for unconfoundedness. Taken together, our results suggest that adjusting for high-dimensional covariates does not necessarily make causal identification more plausible.

---

## Augmented Minimax Linear Estimation: Generalizing Double-Robust Balancing Estimators

12<sup>th</sup> April  
14:45 - 15:00

David A. Hirshberg<sup>1</sup> and Stefan Wager<sup>2</sup>

<sup>1</sup>Department of Statistics, Columbia University

<sup>2</sup>Stanford Graduate School of Business

For estimation of average treatment effects, balancing weights are an increasingly popular alternative to inverse probability weights. Their use has strong empirical and theoretical justification. In particular, several popular balancing estimators fall into the class of minimax linear estimators, which have been shown to be nearly optimal in fixed design among all estimators. This property holds for minimax linear estimators for a large class of estimands: the class of linear functionals of regression functions.

Balancing weights for binary treatments are chosen to ensure the weighted average of an unknown regression function over one (e.g. control) subsample approximates the expectation of that function on another (e.g. treatment) subsample. And for any linear functional, the minimax-optimal weights ensure a generalization of this balance property: the weighted average of an unknown regression function approximates the linear functional evaluated at that function. The existence of such weights follows from the Reisz representation theorem under a generalization of the familiar ‘overlap’ condition. Like balancing weights for more familiar problems, they can be estimated by taking

this approximate equality, for a large set of possible regression functions, as estimating equations for the weights.

In this talk, I will propose a general approach to estimating linear functionals satisfying this generalized overlap condition. The proposed estimators incorporate the minimax-optimal weights into a double-robust augmented weighting estimator. I will give general conditions under which these estimators are semiparametrically efficient and demonstrate via simulation the performance of these estimators. As examples, I will discuss several types of average treatment effect for continuous-valued treatments.

---

## **Calibrated Inverse Probability of Treatment Weights By Covariate Balancing for Marginal Structural Models**

12<sup>th</sup> April  
15:00 - 15:15

Sean Yiu and Li Su

MRC Biostatistics Unit, School of Clinical Medicine, University of Cambridge, UK

Marginal structural models (MSMs) estimated by inverse probability weighting (IPW) offer an approach to quantifying the causal effect of treatment sequences on repeated outcome measures in the presence of time-varying confounding. However, if weights are from maximum likelihood estimation (MLE), the treatment process after weighting will in general not be statistically exogenous, i.e., after weighting treatment assignment still depends on measured confounders conditional on treatment history, even if the treatment model is correctly specified. This is because, with correct specification IPW achieves sequential randomization, but this will only remove systematic, not chance, imbalances of empirical confounder distributions. For a particular sample, these chance imbalances lead to estimation error. Therefore, when the IPW estimator is applied to different samples, its estimation error will result in increased variance and mean squared error after taking expectations even if it is unbiased.

In this work, we propose to improve the IPW estimator of the MSMs by calibrating initial weights (e.g., from MLE) such that the calibrated weights satisfy covariate balancing restrictions implying that the treatment process is statistically exogenous in the current sample after weighting, as characterized by a chosen treatment model. Simulation shows that our method improves the efficiency of the IPW estimator when the initial weights are consistently estimated, and can be more robust under model misspecification as it directly optimizes covariate balance. We apply our method to a natural history study of HIV for estimating the cumulative effect of highly active antiretroviral therapy on CD4 counts over time.

## **Estimating Population Average Causal Effects in the Presence of Non-Overlap: A Bayesian approach**

12<sup>th</sup> April  
15:15 - 15:30

Rachel C. Nethery<sup>1</sup>, Fabrizia Mealli<sup>2</sup>, Francesca Dominici<sup>1</sup>

<sup>1</sup>Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA

<sup>2</sup>Department of Statistics, Informatics, Applications, University of Florence, Florence, Italy

The majority of causal inference studies rely on the assumption of positivity, or overlap, to identify population or sample average causal effects. While the overlap assumption has been shown to be untenable in a substantial proportion of the analyses presented in the causal inference literature, existing methods to handle non-overlap, such as trimming or down-weighting data in the non-overlap region, all suffer from the limitation of changing the estimand to a less interpretable and impactful one. This consequence is often unacceptable in practice, and researchers may be willing to make additional, minimal assumptions in order to preserve the ability to estimate population average causal effects. We seek to make two contributions on this topic. First, we introduce principled definitions of the regions of propensity score overlap and non-overlap. Second, we develop a Bayesian approach for estimating population average causal effects in the presence of non-overlap and in the context of non-linear causal effects. The promising performance of our method is demonstrated in simulations and an application to real data.

---

## **Instrument-based estimation with binarized treatments: Issues and tests for the exclusion restriction**

12<sup>th</sup> April  
16:00 - 16:30

Martin Eckhoff Andresen<sup>1</sup> and Martin Huber<sup>2</sup>

<sup>1</sup>Statistics Norway

<sup>2</sup>University of Fribourg, Dept. of Economics

When estimating local average and marginal treatment effects using instrumental variables (IV), multivalued endogenous treatments are frequently binarized based on a specific threshold in treatment support. However, such binarization introduces a violation of the IV exclusion if (i) the IV affects the multivalued treatment within support areas below and/or above the threshold and (ii) such IV-induced changes in the multivalued treatment affect the outcome. We discuss assumptions that satisfy the IV exclusion restriction with the binarized treatment and permit identifying the average effect of (i) the binarized treatment and (ii) unit-level increases in the original multivalued treatment among specific compliers. We derive testable implications of these assumptions and propose tests, which we apply to the estimation of the returns to (binary) college graduation instrumented by college proximity.

12<sup>th</sup> April  
16:30 - 16:45

## Causal Interpretation in Survival Analysis

Kjetil Røysland  
University of Oslo

It has recently been emphasized, see (Hernan 2010), that the common interpretation of hazards in survival analysis, as risk of death during an interval  $[t, t + \Delta)$  for an individual alive at  $t$ , is often not true. Hazards have a built-in bias since when conditioning on recent survival, we actually condition on a collider that is likely to open a non-causal pathway from the exposure to the event of interest. It does not even matter if the underlying model is causal. A simple example in (Aalen, Cook, RÅyysland 2015) shows that this can also be a problem in RCTs. There are potentially dramatic consequences from this because modern survival analysis has very much been centered around the concept of hazards. As matter of fact, clinicians have been advised to focus on hazards, instead of other measures, since they where thought to contain more dynamical information.

In order to to deal with this situation, we suggest a flexible method to estimate other parameters than the hazard from causal models in survival analysis. These parameters do not have the built the built-in bias as the hazards have, and will often allow interpretations that are more intuitive to clinicians. Our method combines the non-parametric additive hazard model with stability theorems for stochastic differential equations to define a class of estimators that are easily implemented on a computer. We also present a proof that these estimators are consistent, given fairly weak assumptions.

---

## Marginal Structural Models for Survival Analysis in Continuous Time: Theoretically Appealing and Practically Feasible

12<sup>th</sup> April  
16:45 - 17:00

Mats J. Stensrud

Department of Biostatistics, University of Oslo, Domus Medica Gaustad, Sognsvannsveien 9,  
0372 Oslo, Norway

Marginal structural models (MSMs) allow us to estimate causal effects, even in the presence of time-dependent confounders. The strategy is intuitively appealing and offers a sharp distinction between confounder adjustment and model selection. In particular, MSMs are suitable for time to event studies: Survival analysis typically involves processes that develop over time, which may include time-dependent confounders.

The standard MSMs for survival analysis treat time as a discrete process, usually by fitting logistic regressions to approximate a proportional hazards model. Using counting process theory, we have provided an alternative formulation of MSMs. This approach has desirable properties: 1) We can obtain consistent effect estimates not only when the treatment weights are known, but also when these weights are estimated from the data. 2) In particular, we can estimate the treatment weights consistently by flexible additive hazard models. 3) Based on theory for differential equations, we have shown that effect estimates on the hazard scale can be transformed to other parameters that are easier to interpret causally. 4) Time is treated as it is naturally perceived: As a continuous process.

We have also applied continuous-time MSMs to approach a controversial topic in oncology. It is debated whether radical surgery or radiation therapy is the ideal treatment for localized prostate cancer. We used data from the Norwegian Cancer Registry

to compare these treatment regimens. In contrast to a naive analysis, we found that the marginal cumulative incidence of treatment failure is similar between the regimens, accounting for the competing risk of death.

---

## **Rethinking meta-analysis: addressing problems of non-transportability when combining treatment effects across patient populations**

Vo Tat Thang<sup>1,2</sup>, Raphael Porcher<sup>2</sup>, Stijn Vansteelandt<sup>1,3</sup>

12<sup>th</sup> April  
17:00 - 17:15

<sup>1</sup>Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Ghent, Belgium

<sup>2</sup>INSERM, U1153 Epidemiology and Biostatistics Sorbonne Paris Cité Research Center (CRESS), Methods of therapeutic evaluation of chronic diseases Team (METHODS), Paris, F-75014 France; Paris Descartes University, Sorbonne Paris Cité, France

<sup>3</sup>Department of Medical Statistics, Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, UK

Standard methodology for meta-analysis often focuses too much on deriving a summary treatment effect but remains silent as how to interpret it - as a consequence of being implicit about the patient population for which this summary statistics is described. Further, some common treatment effect measures – i.e. odd ratios (OR) and hazard ratios (HR) are vulnerable to the problem of non-collapsibility, which makes them difficult to pool. In this study, we develop novel meta-analysis approaches for randomized clinical trials, which use individual patient data (IPD) from all trials to infer the treatment effect for the patient population in a given trial, based on direct standardization using either outcome regression (OCR) or inverse probability weighting (IPW). Accompanying random-effect meta-analyses are developed, which enable disentangling heterogeneity in treatment effects due to differential covariate distribution from that due to differential outcome generating mechanism.

Across four settings, simulation experiments are conducted to evaluate the two proposed approaches with respect to (i) consistency and (ii) the performance in heterogeneity assessment. Both OCR-based and IPW-based approaches are effective for transforming the evidence across different populations. However, the new meta-analysis approach based on OCR is more successful at correctly capturing treatment effect heterogeneity. Meanwhile, the IPW-based approach is better at preventing the risk of invalid findings due to positivity violation, via examining the weight distribution involving in IPW. The two approaches are then applied to reanalyze a published IPD meta-analysis assessing the effect of vitamin D on the risk of respiratory tract infection.

In conclusion, the novel framework is promising to overcome the important drawbacks of standard meta-analysis methodology. This is because it yields treatment effect estimates that express the effectiveness for well-defined patient populations. Future frameworks should focus on (i) improving the performance of the IPW approach and (ii) generalizing both OCR and IPW to aggregated data, which makes them more applicable in practice.

12<sup>th</sup> April  
17:15 - 17:30

## **Discovering effect modification and randomization inference in air pollution studies**

Kwonsang Lee<sup>1</sup>, Dylan Small<sup>2</sup>, and Francesca Dominici<sup>1</sup>

<sup>1</sup>Department of Biostatistics, Harvard School of Public Health

<sup>2</sup>Department of Statistics, University of Pennsylvania

Studies have shown that exposure to air pollution, even at low levels, significantly increases mortality. As regulatory actions are becoming prohibitively expensive, robust evidence to guide the development of targeted interventions to reduce air pollution exposure is needed. In this paper, we introduce a novel statistical method that splits the data into two subsamples: (a) Using the first subsample, we consider a data-driven search for de novo discovery of subgroups that could have exposure effects that differ from the population mean; and then (b) using the second subsample, we quantify evidence of effect modification among the subgroups with nonparametric randomization-based tests. We also develop a sensitivity analysis method to assess the robustness of the conclusions to unmeasured confounding bias. Via simulation studies and theoretical arguments, we demonstrate that since we discover the subgroups in the first subsample, hypothesis testing on the second subsample can focus on these subgroups only, thus substantially increasing the statistical power of the test. We apply our method to the data of 1,612,414 Medicare beneficiaries in New England region in the United States for the period 2000 to 2006. We find that seniors aged between 81-85 with low income and seniors aged above 85 have statistically significantly higher causal effects of exposure to PM<sub>2.5</sub> on 5-year mortality rate compared to the population mean.

---

13<sup>th</sup> April  
9:00 - 10:00

## **Causal Inference in the Presence of Interference**

Michael Hudgens

University of North Carolina at Chapel Hill

A fundamental assumption usually made in causal inference is that of no interference between individuals (or units), i.e., the potential outcomes of one individual are assumed to be unaffected by the treatment assignment of other individuals. However, in many settings, this assumption obviously does not hold. For example, in infectious diseases, whether one person becomes infected depends on who else in the population is vaccinated. In this talk we will discuss recent approaches to assessing treatment effects in the presence of interference.

---

## **Average treatment effects in the presence of unknown interference**

13<sup>th</sup> April  
10:30 - 11:00

Fredrik Sävje<sup>1</sup>, Peter M. Aronow<sup>2</sup>, Michael G. Hudgens<sup>3</sup>

<sup>1</sup>Department of Political Science, Yale University.

<sup>2</sup>Department of Political Science & Department of Biostatistics, Yale University.

<sup>3</sup>Department of Biostatistics, University of North Carolina, Chapel Hill

We investigate large-sample properties of treatment effect estimators under unknown interference in randomized experiments. The inferential target is a generalization of the average treatment effect estimand that marginalizes over potential spillover effects. We show that estimators commonly used to estimate treatment effects under no-interference are consistent for the generalized estimand for most experimental designs under limited but otherwise arbitrary and unknown interference. The rates of convergence depend on the rate at which the amount of interference grows and the degree to which it aligns with dependencies in treatment assignment. Importantly for practitioners, the results imply that if one erroneously assumes that units do not interfere in a setting with limited, or even moderate, interference, standard estimators are nevertheless likely to be close to an average treatment effect if the sample is sufficiently large.

---

## **Causal inference for interfering units with cluster and population level treatment allocation programs**

13<sup>th</sup> April  
11:00 - 11:15

Georgia Papadogeorgou<sup>1</sup>, Fabrizia Mealli<sup>2</sup>, Corwin Zilger<sup>1</sup>

<sup>1</sup>Harvard T.H. Chan School of Public Health

<sup>1</sup>University of Florence

Interference arises when an individual's potential outcome depends on the individual treatment level, but also on the treatment level of others. A common assumption in the causal inference literature in the presence of interference is partial interference, implying that the population can be partitioned in clusters of individuals whose potential outcomes only depend on the treatment of units within the same cluster. Previous literature has defined average potential outcomes under counterfactual scenarios where treatments are randomly allocated to units within a cluster. However, within clusters there may be units that are more or less likely to receive treatment based on covariates or neighbors' treatment. We define estimands that describe average potential outcomes for realistic counterfactual treatment allocation programs taking into consideration the units' covariates, as well as dependence between units' treatment assignment. We discuss these estimands, propose unbiased estimators and derive asymptotic results as the number of clusters grows. Finally, we estimate effects in a comparative effectiveness study of power plant emission reduction technologies on ambient ozone pollution.

---

## Propensity score methods for causal inference in observational studies with network interference

13<sup>th</sup> April  
11:15 - 11:30

Bruno Arpino<sup>1</sup>, Luca De Benedictis<sup>2</sup> and Alessandra Mattei<sup>3</sup>

<sup>1</sup>Universitat Pompeu Fabra, Barcelona, Spain

<sup>2</sup>University of Macerata, Italy, <sup>3</sup>University of Florence, Italy.

Propensity score estimators have recently been extended to non-standard data structures including network data. Nevertheless with network data the Stable Unit Treatment Value Assumption (SUTVA), and specifically its no interference component, may be questionable as potential outcomes may depend not only on the unit's treatment assignment but also on those of other units in the network. Recently, propensity score methods have been used to deal with network data allowing for interference. We contribute to this literature by investigating the role of the network structure in drawing inference on average individual and global causal effects and on spillover effects in the presence of network interference. Our analysis is motivated by the evaluation of the causal effect of the General Agreement on Tariffs and Trade (GATT) on bilateral international trade flows. We focus on causal effects of an individual treatment, GATT membership, on international trade accounting for interference generated by the trade network. We use a realistic simulation design inspired by the ComTrade UN database used by many studies on the GATT membership effect on international trade. We consider different scenarios by varying the way treatment is assigned and the characteristics of the network. We compare estimators that ignore interference with parametric and semi-parametric estimators that accounts for the network structure by using different network summaries. We conclude by applying the considered techniques on the real data to assess the effect of GATT on trade.

13<sup>th</sup> April  
12:00 - 12:15

## On the perils of stacking thresholds on RD designs

Margherita Fort, Andrea Ichino, Enrico Rettore, Giulio Zanella

University of Trento, FBK-IRVAPP and IZA

In RD designs with multiple thresholds, one for each program, it is common to normalize the score so that all of the cutoffs are centered at zero. We show that if a researcher adopts such normalization and also uses the score-distance of each observation from the respective threshold as the running variable (a practice we refer to as "stacking"), then standard diagnostic tools such as continuity of the score distance and continuity of pre-treatment covariates may reject the stacked design even if the separate RD designs around each threshold are perfectly valid. Moreover, the outcome of interest may appear discontinuous at the stacked threshold even if the treatment effect is zero in each separate RD design. This situation arises in the presence of three conditions: (a) each threshold is the value of the running variable for a marginal subject exposed to the treatment, so that there is one observation located exactly at each threshold; (b) programs have different probabilities of being observed in a neighborhood of zero-distance from their threshold; (c) there is sorting of heterogeneous subjects across programs. When these conditions hold and stacking is necessary, dropping observations at zero distance from each threshold is advisable.

## **Disease subtypes in aetiological research**

13<sup>th</sup> April  
12:00 - 12:30

Lorenzo Richiardi

Cancer Epidemiology Unit, Department of Medical Sciences, University of Turin, Italy

It is becoming increasingly common to classify diseases into subtypes defined by distinct molecular profiles. This activity is linked to the concept of precision medicine and often aims at improving the prediction of response to treatment and prognosis. Although for clinical purposes it might be less relevant to understand if the different subtypes represent distinct pathogenesis mechanisms, this is a key aspect for aetiological research aiming at answering the question whether the disease subtypes have different risk factors. This presentation will discuss a simplified causal framework to reason about potential biases in analyses of risk factors of disease subtypes. It will discuss the cases of weak effect and lack of causality and provide numerical and real examples. These sources of biases should be considered in aetiological research of disease subtypes, which cannot ignore the underlying pathogenic mechanisms.

---

## **How to obtain valid tests and confidence intervals for treatment effects after confounder selection?**

13<sup>th</sup> April  
12:30 - 12:45

Stijn Vansteelandt<sup>1,2</sup> and Oliver Dukes<sup>2</sup>

<sup>1</sup>Ghent University and the <sup>2</sup>London School of Hygiene and Tropical Medicine

The problem of how to best select variables for confounding adjustment forms one of key challenges in the evaluation of exposure or treatment effects in observational studies. Routine practice is often based on stepwise selection procedures that use hypothesis testing, change-in-estimate assessments or the lasso, which have all been criticised for - amongst other things - not giving sufficient priority to the selection of confounders. This has prompted vigorous recent activity in developing procedures that prioritise the selection of confounders, while preventing the selection of so-called instrumental variables that are associated with exposure, but not outcome (after adjustment for the exposure). A major drawback of all these procedures is that there is no finite sample size at which they are guaranteed to deliver treatment effect estimators and associated confidence intervals with adequate performance. This is the result of the estimator jumping back and forth between different selected models, and standard confidence intervals ignoring the resulting model selection uncertainty. In this talk, I will develop insight into this by evaluating the finite-sample distribution of the exposure effect estimator in linear regression, under a number of the aforementioned confounder selection procedures. I will then make a simple but generic proposal for generalised linear models, which overcomes this concern (under weaker conditions than competing proposals).

---

13<sup>th</sup> April  
12:45 - 13:00

## **Data-Adaptive Variable Importance with CV-TMLE**

Chris J. Kennedy and Alan E. Hubbard

Division of Biostatistics, University of California, Berkeley, CA, USA

Big data holds the promise of discovering relevant relationships using mining algorithms. We propose an approach that uses cross-validation to define the target estimand, estimate the parameter and provide robust inference. We present an estimator (and corresponding central limit theorem) of this data adaptive target parameter: cross-validated targeted maximum likelihood estimation (CV-TMLE). The algorithm uses the training sample to define levels of an explanatory variable and uses the validation sample to estimate a variable importance measure comparing these levels. Super learning and causal-inference-inspired estimands are both used to define the target levels and to estimate the adjusted association within the validation sample. We borrow power across the entire sample by a global bias-reduction step (CV-TMLE), which also results in asymptotically-linearly estimators with inference based on the influence curve. Within the proposed algorithm (varimpact), data cleaning, accounting for missing data, model fitting, etc. is automated and the algorithm is scalable for use with large samples and many variables. We apply varimpact to both clinical and epidemiological data.

---

13<sup>th</sup> April  
13:00 - 13:15

## **Challenges in Estimating the Causal Effects of a Complex Multi-Dimensional Exposure**

Sarasvati Bahadursingh, Sarah Lewington, Sofia Massa, Iona Millwood, Naomi Allen  
Nuffield Department of Population Health, University of Oxford

In epidemiology, a complex multi-dimensional exposure is often reduced to more simple summary measures for analysis, and public health recommendations are based on estimates of the causal effects of the latter. This approach, however, may result in biased estimates when potential confounding by different dimensions of the exposure is ignored. It may also ignore the potential causal effects of individual dimensions of the complex exposure and their interaction - knowledge which may inform more detailed and accurate public health recommendations on the most beneficial or risky patterns of exposure. Conversely, incorporating multiple dimensions of an exposure into analysis can quickly diminish power and make the findings less easy to interpret. Here, we will use a study of the effects of alcohol consumption on cardiovascular risk in a cohort of 500,000 participants from the UK Biobank to illustrate the importance and challenges of incorporating several dimensions of a complex exposure into analysis of causal effects. We will also discuss how such an approach has the potential to improve public health recommendations

---