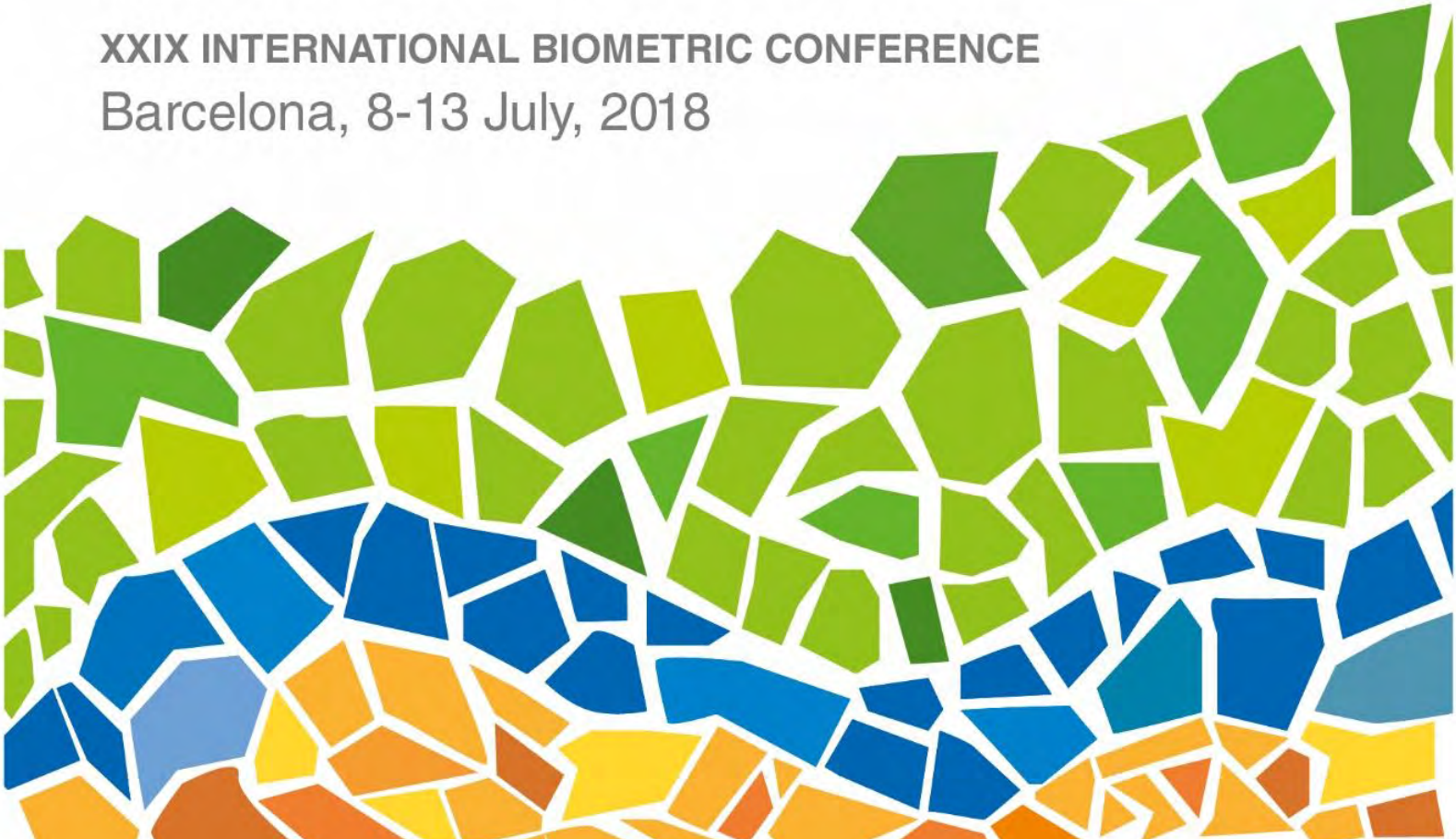
PROCEEDINGS' BOOK

BARCELONA

IBC 2018

XXIX INTERNATIONAL BIOMETRIC CONFERENCE

Barcelona, 8-13 July, 2018



Index

37

A Four-Way Decomposition of Socioeconomic Status and Mortality After First Myocardial Infarction

Ronnie Pingel
Uppsala University, Uppsala, Sweden

This study use Swedish registry data (n=162 000) to analyze how education and income are related to 1 year mortality among MI survivors. The excess relative risk of death 1 year after admission among those having low education is 50% higher compared to those having high education. We then decompose this excess relative risk of education into the component that
is due to only the mediation of income, due to only interaction between education and income, due to both mediation and interaction of income and education, and the controlled direct effect of education. The results show that 30% of the excess relative risk is due to a controlled direct and 50% is due to the pure indirect effect. We find no evidence of interaction between education and income. Thus, by unifying mediation and interaction, we are able to gain greater insight how socioeconomic variables are related to each other.

38

Dynamic Modeling of Multivariate Latent Processes and Their Causal Relationships: Application to Alzheimer's Disease

Bachirou O. Taddé, Hélène Jaqmin-Gadda, Daniel Commenges, Jean François Dartigues, Cécile PROUST-LIMA
U1219, INSERM, Bordeaux, Please Select, France

Alzheimer's disease gradually affects several dimensions including the cerebral anatomy with brain atrophies, the cognitive functioning with a decline in various functions and the functional dependency with impairments in the daily living activities. Hypothetical schemes have been proposed to describe the multiple dimensions involved in AD and define expected dynamic relationships between dimensions. However because of their complexity, they have not yet been translated into statistical models that simultaneously combine the dynamic and multidimensional aspects, and explore the causal relationships.
We propose a new dynamic model that accounts for all these features. The model defines dimensions as latent processes and combines a multivariate linear mixed model and a system of difference equations to model trajectories and causal relationships of the system of latent processes in finely discrete time. Parameters are estimated in the maximum likelihood framework enjoying a closed form for the likelihood. The estimation procedure and the impact of the time discretization on the causal interpretations are evaluated in simulations. The model is illustrated on data from cognitive aging studies to assess dynamic causal relationships in years preceding dementia diagnosis. Are considered dimensions such as cerebral anatomy, global cognitive functioning and functional dependency, or different subtypes of cognitive functions such as episodic memory, speed, verbal fluency and executive functioning.

Keywords: causality, mixed models, difference equations, latent process, longitudinal data

39

Statistical mediation analysis in cardiovascular epidemiology – challenges and case studies

Josef Fritz, Hanno Ulmer
Department of Medical Statistics, Informatics and Health Economics, Medical University of Innsbruck, Innsbruck, Austria

Statistical mediation analysis, that is to investigate whether and how much of the effect of a variable of interest goes through pre-specified intermediate variables, has experienced a substantial upturn over the last decade, particularly through the adoption of ideas from the field of causal inference, and herein specifically the counterfactual framework. However, practical applications of these new methods are still scarce. After explaining challenges and some of the pitfalls of mediation when applied to observational data and ways to circumvent them, we will illustrate the application of novel causal inference mediation analysis methods on three case studies from the field of cardiovascular epidemiology. We will describe our impression and the challenges of applying some of

these methods to large cohort data.
The case studies use data of two prospective, population-based cohorts, namely the Vorarlberg Health Monitoring and Promotion Programme and the Malmö Diet and Cancer Study, consisting of about 180,000 and 23,000 participants, respectively. The three methods for mediation which we applied are (i) natural effect models (proposed by T. Lange), (ii) a regression-based approach allowing also for interactions (proposed by T.J. VanderWeele), and (iii) the inverse odds ratio-weighted approach (proposed by E.J. Tchetgen Tchetgen). All statistical routines needed for analyses were programmed from scratch in R software; example R code is provided.
In the first case study we assessed if sex/gender differences regarding mortality of coronary heart disease (CHD) are mediated by traditional cardiovascular risk factors. Secondly, we investigated age dependencies in metabolic mediation of body mass index on CHD mortality. Lastly, we asked if CHD risk conferred by family history/genetics measured as a genetic risk score is mediated by traditional metabolic risk factors. In all three applications, we observed that substantial parts of the total effect can be explained by the mediators under consideration.
We demonstrate the feasibility of novel mediation methods from the field of causal inference for the analysis of epidemiological studies and the consistency of results among different methods.

40

On the use of electronic health records for confounder control in quality of care measures

Els Goetghebeur
Applied Mathematics, Computer Science and Statistics, Ghent University, Gent, Belgium

Today's evaluation of center performance in patient care draws heavily on outcome indicators of quality-of-care. To allow for an unbiased assessment and fair comparisons of healthcare centers in those terms, patient mix adjustment should involve enough important confounders, measured reliably on the relevant patient population Electronic health records can in principle provide such a rich source of confounder data. They come with two problems however. 1) More covariates thus added introduce more records with missing data, especially among high-risk patients. 2) For reasons of confidentiality acces may be limited to aggregate patient data. The first problem will yield complete case analyses that embellish centers with incomplete registration. It is unclear how much of this can be alleviated through MAR methods and raises the question of whether such additional confounder control can avoid undue bias. We take advantage of complete mortality outcome to examine this selection versus confounder bias trade-off on standardised hospital risks.
In the Swedish Riksstroke register, well-known for generally quite high coverage, we find that under missingness at random, complete case analysis produces underestimated center-specific risks and misclassifies more hospitals as outlying centers. Unfortunately, centers with the better registration coverage tend to be penalised by the appearance of their quality indicator. Here, relatively small differences in missingness across hospitals have non-negligable impact in ways that matter. We examine when an added covariate risks introducing more bias through the level of missingness it brings and how the analysis could avoid stimulating incorrect or incomplete registration for better apparent results.
Secondly, we study how standardised center risks can be obtained from aggregate covariate data per center. For linear models, this follows immediately from the estimating equations and in a derived way for generalised linear models with iteratively re-weighted least squares algorithms (see Lee, Brown and Ryan, 2017). Theoretically, Cox regression models could also be fit if specifically weighted averages were available per center each time a patient dies. In practice more feasible updating schemes will be needed and we study their impact on the obtained standardised measures.

This paper presents joint work with Arnout Van Messem and Marie Eriksson

41

Global test for high-dimensional mediation: testing groups of potential mediators

Vera Djordjilovic[1], Christian M. Page[2, 3], Jon Michael Gran[1, 2], Therese H. Nøst[4], Torkjel M. Sandanger[4], Marit B. Veierød[1], Magne Thoresen[1]

MEDIZINISCHE UNIVERSITÄT
INNSBRUCK

Statistical mediation analysis in cardiovascular epidemiology – challenges and case studies

Josef Fritz[1], Hanno Ulmer[1]

[1] Department of Medical Statistics, Informatics and Health Economics, Medical University of Innsbruck, Innsbruck, Austria

BACKGROUND

Statistical mediation analysis, that is to investigate whether and how much of the effect of a variable of interest goes through pre-specified intermediate variables, has experienced a substantial upturn over the last decade, particularly through the adoption of ideas from the field of causal inference, and herein specifically the counterfactual framework. However, practical applications of these new methods are still scarce. We illustrate the application of two novel causal inference mediation analysis methods on three case studies from the field of cardiovascular epidemiology.

DEFINITIONS & CONCEPTS

- A generic definition of direct and indirect effects needs the concept of counterfactuals (Robins & Greenland, 1992)
- The natural direct effect (comparing exposure level $X = x$ with $X = x^*$) is defined as

$$NDE(Y)_{x,x^*} = E\left[Y_{x,M_{x^*}}\right] - E\left[Y_{x^*,M_{x^*}}\right]$$

- The natural indirect effect is defined as

$$NIE(Y)_{x,x^*} = E\left[Y_{x,M_x}\right] - E\left[Y_{x,M_{x^*}}\right]$$

- It can be shown that NDE and NIE sum up to the total effect which is defined as

$$TE(Y)_{x,x^*} = E[Y_x] - E[Y_{x^*}]$$

- It is not possible to estimate these effects directly from the observed data via standard regression techniques

METHOD 1: 2-STAGE REGRESSION-BASED APPROACH FOR MULTIPLE MEDIATORS

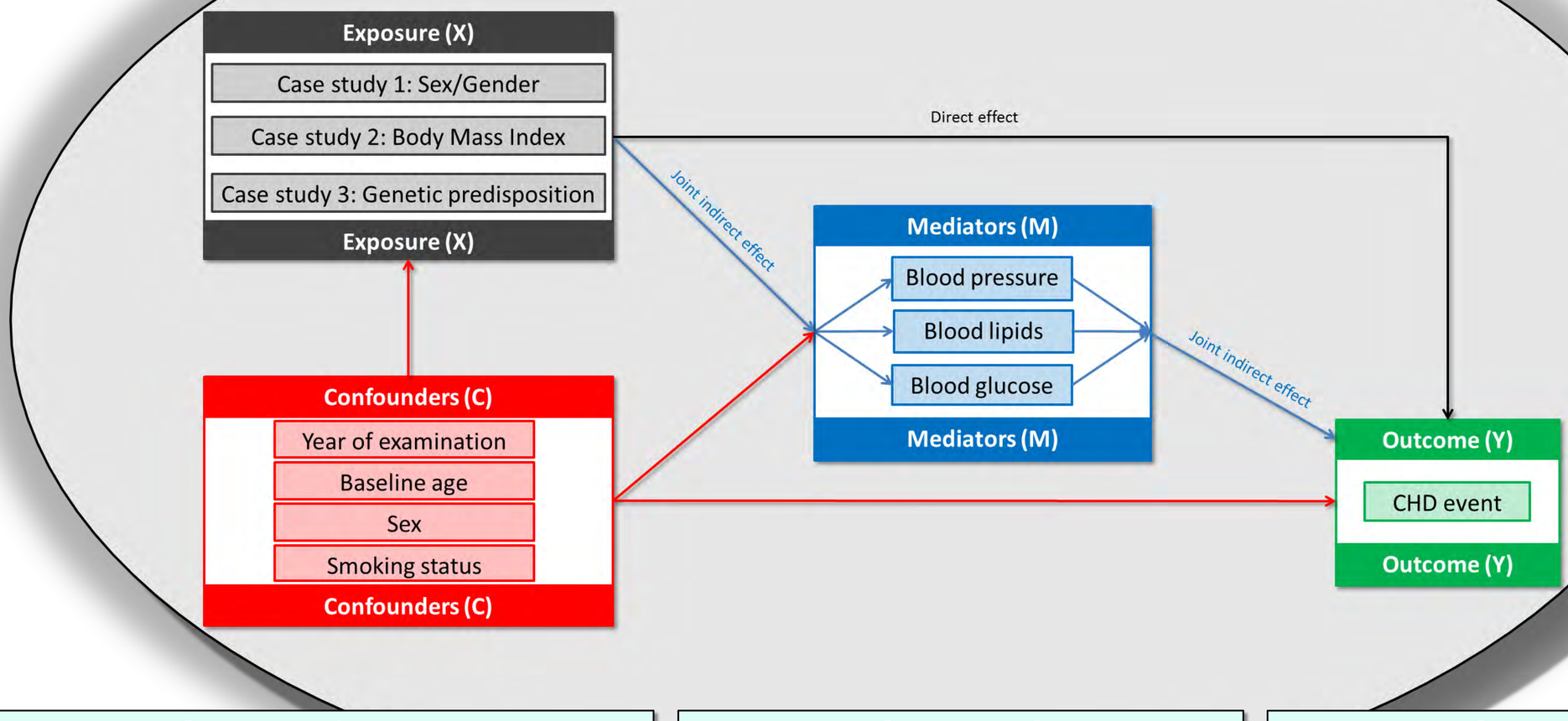
- Extension of the well known “Baron & Kenny” approach from 1986
- Encompassing exposure-mediator interaction
- Mediator model conditioning on exposure (X) and confounder (C): $E(M|x,c) = \alpha + \beta x + \gamma c$
- Outcome model conditioning on exposure (X), mediator (M), and confounder (C): $E(Y|x,m,c) = \delta + \zeta x + \eta m + \vartheta xm + \kappa c$
- Then (for binary exposure) $NDE = \zeta + \vartheta(\alpha + \gamma c)$ and $NIE = \beta(\eta + \vartheta)$
- Can be extended to the case of multiple mediators
- For details see VanderWeele, 2012, 2014

METHOD 2: NATURAL EFFECTS MODELS

- Idea: modify/expand the original dataset by imputing and weighting techniques so that NDE and NIE can be estimated from this modified dataset via one standard regression model
- Specifically, a counterfactual outcome is modelled:

$$E\left[Y_{x,M_{x^*}}\right] = \alpha + \beta x + \gamma x^*$$

- β captures the NDE and γ captures the NIE
- Very flexible approach; e.g., implementation of exposure-covariate interaction is straightforward
- For details see Lange, 2012
- The idea can be generalized to cases of distinct multiple pathways (Lange, 2014)



CASE STUDY 1

- Data: Vorarlberg Health Monitoring and Promotion Programme (VHM&PP), Austria, including 172,262 individuals with 3,892 CHD death events
- Exposure: sex/gender
- Method: natural effects model
- The CHD mortality difference between sexes decreased with age
  <50 years: HR = 4.7 (95% CI: 3.5-6.1)
  ≥50 years: HR = 1.9 (95% CI: 1.7-2.1)
- The extent to which metabolic factors and smoking contributed to these total effects varied with age
  <50 years: 41% (95% CI: 27-54%) explained indirectly
  ≥50 years: 8% (95% CI: 4-12%) explained indirectly
- Fritz J, Edlinger M, Kelleher C, Strohmaier S, Nagel G, Concin H, Ruttmann E, Hochleitner M, Ulmer H. Mediation analysis of the relationship between sex, cardiovascular risk factors and mortality from coronary heart disease: findings from the population-based VHM&PP cohort. *Atherosclerosis*. 2015;243(1):86–92.

CASE STUDY 2

- Data: VHM&PP including 111,303 individuals with 2,127 CHD death events
- Exposure: body mass index (BMI) in WHO categories (normal weight, overweight, obese)
- Method: two-stage regression-based approach
- Interaction with age
- While in younger (<65 years) individuals about half of the total BMI effect was mediated by metabolic factors, it was only 20% in elderly (≥65 years) obese individuals
- Fritz J, Strohmaier S, Nagel G, Concin H, Ulmer H. Re: Mediators of the Effect of Body Mass Index on Coronary Heart Disease / A strong interaction between age and overweight/obesity on the risk of coronary heart disease in the context of metabolic mediation. *Epidemiology*. 2016;27(3):e13–14.

CASE STUDY 3

- Data: Malmö Diet and Cancer (MDC) study including 23,595 individuals with 2,213 CHD events
- Exposure: (i) family history of CHD; (ii) genetic risk score based on 50 CHD-related SNPs (GRS50)
- Method: natural effects model
- A fraction (~20%) of the CHD risk associated with family history (HR=1.5) or with GRS50 (HR = 1.5 (highest vs. other quintiles)) was mediated through dyslipidaemia and hypertension, but not through diabetes
- Fritz J, Shiffman D, Melander O, Tada H, Ulmer H. Metabolic Mediators of the Effects of Family History and Genetic Risk Score on Coronary Heart Disease—Findings From the Malmö Diet and Cancer Study. *J Am Heart Assoc*. 2017; 6(3):e005254.

Contact information:

Josef Fritz, PhD
Department of MSIG
Medical University of Innsbruck
Schoepfstraße 41 / 1
6020 Innsbruck, Austria
P: +43 512 9003 70915
E: josef.fritz@i-med.ac.at
I: http://www.i-med.ac.at/msig/

MEDICAL STATISTICS INFORMATICS HEALTH ECONOMICS

General references:

[1] Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology*. 1992;3(2):143-55.
[2] VanderWeele TJ. Causal mediation analysis with survival data. *Epidemiology*. 2012;22(4):582-585.
[3] VanderWeele TJ, Vansteelandt S. Mediation Analysis with Multiple Mediators. *Epidemiol Method*. 2014;2(1):95-115.
[4] Lange T, Vansteelandt S, Bekaert M. A simple unified approach for estimating natural direct and indirect effects. *Am J Epidemiol*. 2012;176(3):190-5.
[5] Lange T, Rasmussen M, Thygesen L. Assessing natural direct and indirect effects through multiple pathways. *Am J Epidemiol*. 2014;179(4):513-8.