Mendelian randomization: In the age of large-scale accessible genomics data

An international conference organised by the Medical Research Council Integrative Epidemiology Unit at the University of Bristol

11th to 13th July 2017
Colston Hall, Bristol, BS1 5AR, UK
General Information

Local organisation committee

Professor George Davey Smith - University of Bristol
Dr Jack Bowden - University of Bristol
Dr Tom Gaunt - University of Bristol
Dr Sarah Lewis - University of Bristol
Dr Lindsey Gaunt - University of Bristol
Dr Doretta Caramaschi - University of Bristol
Dr Andrew Crawford - University of Bristol
Dr Neil Davies - University of Bristol
Dr Philip Haycock - University of Bristol
Dr Diana Santos Ferreira - University of Bristol
Dr Kaitlin Wade - University of Bristol
Dr Michelle Taylor - University of Bristol
Sally John - Biogen Idec, USA
Dr Lon Cardon - GlaxoSmithKline, USA
Dr Brian Ference - Wayne State University, USA
Professor Vanessa Didelez - Leibniz Institute for Prevention Research and Epidemiology, Bremen
Dr Daniel Benjamin - University of Southern California, USA
Dr Sekar Kathiresan - Broad Institute, USA
Dr Paul Brennan - International Agency for Research on Cancer
Ryan Langdon - University of Bristol
James Yarmolinsky - University of Bristol
Audrey Hayes - University of Bristol

Registration

Registration desk opens at 8:00am on Tuesday 11th July, and will remain open until 13:30. The registration desk will be open between 8:30am and 9:30am on Wednesday 12th and Thursday 13th July.

Scientific Session

Plenary sessions are held in Hall 1. Parallel sessions are held in the Hall 1, The Lantern and The Terrace Bar, as advertised in the schedule. All talks are 10 minutes long, followed by questions.

Posters

Posters can be displayed and viewed in the Stalls Bar during the scheduled poster sessions from 17:00 to 18:00 Tuesday 11th and Wednesday 12th July. Posters presented by early career researchers will be judged by the judging panel during both sessions. Prize winners will be announced at the closing address on Thursday 13th July.

Lunch and Refreshments

During morning and afternoon breaks, refreshments are served in both the Stalls Bar and Foyer. Lunch will be served in the Stalls Bar. No food or drink to be taken into the Auditorium or The Lantern. Lunches and refreshments are included in the registration fee.

Conference Dinner

The Conference Dinner is held at Bambalan, located adjacent to the conference venue.
To attend the conference dinner you should have reserved your place during registration. You will be issued with an event ticket during registration. Guests are advised to bring a warm layer of clothing, so they can enjoy the open-air roof terrace, which provided stunning views of Bristol City Centre.

Seating is also provided indoors. Guests are advised to arrive by 18:30.

Internet Access

Internet access at Colston Hall is available via either the ‘Colston Hall’ or ‘delegate’ wireless provision. For the ‘delegate’ service please use ‘chdelegate’ as the password.

Twitter

Throughout the conference we will be using Twitter to connect delegates through their experience using @mrc_ieu #mrconf17.

Left Luggage

There is limited space available at the conference venue for left luggage. Please enquire at the Registration Desk. The organisers cannot accept liability for left luggage.

Parking

No parking is available at the conference venue, and delegates are encouraged to use public transport wherever possible.

Sponsors

Our thanks go to our generous sponsor: Metabolon.

---

M**ETABOLON**

---
Short course on Mendelian randomization
Monday 10th July, Canynge Hall, Room LG-09, Whatley Road, Bristol, BS8 2PS
Course tutors
Sarah Lewis (course organiser), Neil Davies, Rebecca Richmond, Michelle Taylor, Philip Haycock, Jack Bowden

Course aims and objectives
Mendelian randomization is a study that uses genetic variants as instrumental variables (proxies) to test the causal effect of a (non-genetic) risk factor. Since its first proposal in 2003 it has been increasingly used to determine causal effects in observational epidemiology and is used in a large amount of the applied research in the MRC Integrative Epidemiology Unit and throughout the School of Social and Community Medicine at the University of Bristol.

This course aims to provide an introduction to the conduct, assumptions, strengths and limitations of Mendelian randomization.

Course outline
By the end of the course students will:
- Understand the principles and assumptions of instrumental variable analyses
- Understand the properties of genetic variants that make them suitable to be used as instrumental variables
- Understand the strengths and limitations of Mendelian randomization for making causal inference in epidemiology
- Be able to complete a straightforward instrumental variable analysis and correctly interpret the results from that analysis.
- Understand that the two core parts of a Mendelian randomization analysis (the association of genetic variants with the risk factor of interest and the association of genetic variants with the outcome) can be done on separate sets of studies; so called ‘two-sample Mendelian randomization’
- Know the basics of how to undertake a two-sample Mendelian randomization study

Future availability
A Mendelian randomization short course will run as a two-day short course on 21st and 22nd February 2017. Further details of this and other short courses can be found at: www.bristol.ac.uk/social-community-medicine/shortcourse/.
MR-Base user workshop
Thursday 13th July, Colston Hall, Terrace bar, 1:30pm-5pm

Workshop tutors
Gibran Hemani, Philip Haycock, Jie Zheng, Tom Gaunt

Workshop aims and objectives
The aim of this workshop is to show users how to get the most out of MR-Base for their research. Users will learn about:
1. Recommended workflows
2. How to use the website and R package
3. The different types of analyses that can be implemented, including focused and hypothesis-free approaches
4. Interpretation of outputs, including consideration of sensitivity analyses and potential violations of MR assumptions.

Course outline
The workshop will last three and a half hours. The first half will be spent covering the functionality of MR-Base, including a MR study of lipids and human health as an applied example. In the second half users will be given the opportunity to implement their own analyses of interest. Experienced users and developers of MR-Base will be on hand to give detailed support and feedback. The course will not cover the theoretical underpinnings or limitations of MR in detail.
By the end of the course students will be able to:
• use MR-Base to perform focused and hypothesis-free two-sample MR analyses
• interpret the robustness of their results to violations of MR assumptions.

Note for attendees
Those registered for the workshop are kindly asked to bring a laptop in order to follow along with the presentation. Further details can be found on the MR-Base website: www.mrbase.org.
About the Medical Research Council Integrative Epidemiology Unit (MRC IEU)

Background

The Medical Research Council Integrative Epidemiology Unit at the University of Bristol (MRC IEU) is one of the MRC’s flagship University Units. It was established on 1st June 2013 with Professor George Davey Smith as Director. Substantial investment from the Medical Research Council and the University of Bristol provides resources for staff, research support and infrastructure in order to execute world-leading multidisciplinary research in a vibrant and forward-looking environment.

Vision and objectives

The IEU brings together an innovative collection of research programmes and cross-cutting themes, underpinned by core research activities that will collectively fuel a step-change in causal analysis and their application.

The aim of the MRC IEU is to apply the novel causal methods developed in the IEU to key research questions related to causes of bone, cardiometabolic, reproductive, mental and other aspects of ill-health; and to ensure that the results from these studies are appropriately translated into clinical/public health practice and industrial partnerships. A major focus of IEU activities is the integration of omic measures (genomic, epigenomic, transcriptomic and metabolomic) into epidemiological investigations. There are six scientific programmes (Mendelian randomization; epigenetics; recall by genotype studies; genetics of bone-related phenotypes; women’s reproductive and cardiometabolic health; and health behaviours) and three cross-cutting themes (data capture; data mining and bioinformatics; and statistical and econometric methodology). Training, public engagement and dissemination are cornerstones of the IEU activity. The academic and societal impact of IEU research is set to grow, with an increasing focus on industrial partnerships as methodological advances become more relevant in a translational context.

IEU PhD Programme

The aim of the PhD programme is to train graduates from a wide range of disciplines (e.g. biochemistry, biology, physiology, mathematics, statistics, physics, engineering and computational science) to integrate molecular, cellular, clinical and population data to identify the causal effects of potentially modifiable exposures on health-related outcomes. PhD studentships are funded for four years, on a 1+3 model.

Information and Contact:

Web: www.bris.ac.uk/ieu
E-mail: ieu@bris.ac.uk
Follow us on Twitter @mrc_ieu
Programme
Tuesday 11th July

8:00am Registration

9:30am Session 1 Welcome address (Hall 1)

Professor John Iredale, Pro Vice Chancellor for Health, University of Bristol

Keynote lecture: Professor Martijn Katan, Emeritus Professor of Nutrition, VU University Amsterdam
The challenges of nutrition science, and how Mendelian randomization can help

10:00am Session 2 Large scale biobanks (Hall 1)

Chair: Professor George Davey Smith, MRC Integrative Epidemiology Unit, University of Bristol, UK

Professor Sir Rory Collins, Principal Investigator UK Biobank, University of Oxford, UK
UK Biobank: a resource for genetic-epidemiological research
Dr Zhengming Chen, University of Oxford, UK
China Kadoorie Biobanks: opportunities for personalized medicine and pharmacogenomics

10:45am-11:15am Break

11:15am Session 2 Large scale biobanks (Hall 1)

Dr Catherine Schaefer, Kaiser Permanente Division of Research in Oakland, USA
The Kaiser Permanente Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort
Dr Christopher O’Donnell, Center for Population Genomics & Boston Veterans Administration Healthcare, USA
Million Veterans cohort
Professor Bjørn Asvold, K.G. Jebsen Center for Genetic Epidemiology, Norway
Mendelian randomization in a large-scale population study linked with national health registries: the HUNT Study in Norway

12:30pm-1:30pm Lunch

1:30pm-3:00pm Parallel sessions

1:30pm Session 3 MR in the Social Sciences (Hall 1)

Chair: Dr Dan Benjamin, University of Southern California, USA

Professor Dalton Conley (Sociology Department, Princeton University, USA) & Professor George Davey Smith (MRC Integrative Epidemiology Unit, University of Bristol, UK)
“Does MR have any place in the social sciences?”

1:30pm Session 4 Applications in pregnancy and maternal exposures (The Lantern)
Chair: Professor Aroon Hingorani, *Institute of Cardiovascular Science*, and Farr Institute, University College London, UK

**Professor Debbie Lawlor**, MRC Integrative Epidemiology Unit, University of Bristol, UK

Mendelian randomization and testing causal effects of maternal (intrauterine) exposures on offspring outcomes - where are we now?

**Dr Rachel Freathy**, University of Exeter, UK

Causal associations between maternal characteristics and offspring birth weight

**William David Thompson**, University of Exeter Medical School, UK

Mendelian Randomisation shows circulating maternal vitamin D and calcium may be causally associated with birth weight

**Dr Michelle Taylor**, MRC Integrative Epidemiology Unit, University of Bristol, UK

Assessing causality in associations between maternal adiposity and perinatal birth outcomes in mothers and offspring: A Mendelian randomization approach

**Dr Ge Zhang**, Human Genetics Division, Cincinnati Children’s Hospital Medical Center, USA

Dissecting the causal mechanisms of maternal blood pressure and blood glucose level on birth weight: A genetic score analysis

**1:30pm**

**Session 5: Applications (Terrace Bar)**

Chair: Professor Nicole Soranzo, Wellcome Trust Sanger Institute, UK

**Professor David Evans**, MRC Integrative Epidemiology Unit, University of Bristol, UK and University of Queensland Diamantina Institute, Brisbane, Australia

Using structural equation modelling in Mendelian randomization studies

**Dr Andrew Crawford**, MRC Integrative Epidemiology Unit, University of Bristol, UK

Identifying and utilising genetic variants associated with morning plasma cortisol: a CORtisol NETwork (CORNET) analysis

**Gunnhild Aaberge Vie**, Department of Public Health and Nursing, NTNU, Norway

Body mass index and risk of musculoskeletal pain – a Mendelian randomisation analysis in the HUNT Study

**Alice Carter**, MRC Integrative Epidemiology Unit, University of Bristol, UK

Investigating the combined association of BMI and alcohol consumption on liver disease biomarkers: a Mendelian randomization study of over 90,000 adults from the Copenhagen General Population cohort

**Professor Mary Schooling**, CUNY School of Public Health, New York, USA

Could HTR7, TAC3 or AXL be new potential genetic targets for ischemic heart disease?

**Dr Claudia Langenberg**, MRC Epidemiology Unit, UK

TBC

**3:00pm-3:30pm Break**

**3:30pm-5:00pm Parallel sessions**

**3:30pm**

**Session 6 Applications in social science and health behaviours (Hall 1)**

Chair: Dr Neil Davies, MRC Integrative Epidemiology Unit, UK
Dr Dan Benjamin*, University of Southern California, USA
GWAS of educational attainment in over 700,000 Individuals
Applying Mendelian randomization to epigenetics: practicalities and limitations*

Professor Philipp Koellinger*, Center for Neurogenomics and Cognitive Research, VU
University Amsterdam, The Netherlands
Large-scale genetic study of risk tolerance and risky behaviors identifies novel loci and reveals shared genetic influences

Professor Melinda Mills*, University of Oxford
GWAS results of human reproductive behaviour

3:30pm Session 7 Methodology (The Lantern)
Chair: Dr Steve Burgess, University of Cambridge, UK

Dr Dylan Small*, University of Pennsylvania, USA
Sensitivity Analysis and Power for Instrumental Variables Studies with Application to Mendelian Randomization

Helmut Farbmacher, Max Planck Society
On the Use of the Lasso for Instrumental Variables Estimation with Some Invalid Instruments

Lai Jiang, Lady Davis Institute, McGill University, USA
Constrained Instrument Variable Method and its Application to Mendelian Randomization with Observed Pleiotropy

Chia-Yen Chen, Massachusetts General Hospital
Detection and correction of pleiotropic bias in multivariable Mendelian randomization

Professor Frank Windmeijer, MRC Integrative Epidemiology Unit, University of Bristol, UK
Instrumental Variables Estimation of causal effects in the presence of invalid instruments

3:30pm Session 8 Applications in neuropsychiatric and immune diseases (Terrace Bar)
Chair: Professor Dave Evans, University of Queensland Diamantina Institute, Brisbane, Australia

Dr Beate St Pourcain*, Max Planck Institute for Psycholinguistics, The Netherlands
Modelling changes in genetic variances during development – The re-birth of twin modelling strategies in the GWAS era

Dr Dylan Williams, Department of Medical Epidemiology & Biostatistics, Karolinska Institutet, Sweden
Circulating Insulin-like Growth Factors and Alzheimer’s disease: a Mendelian randomization study

Dr Alastair Noyce, University College London Institute of Neurology, UK
Estimating the causal influence of BMI on risk of Parkinson’s disease: a Mendelian randomization study

Dr Jessica Tyrrell, University of Exeter, UK
Genetic analyses to test for a causal relationship between obesity and depression

Tea Skaaby, Research Centre for Prevention and Health, Denmark
Estimating the causal effect of body mass index on hay fever, asthma, and lung function: a Mendelian randomization analysis

5:00pm Poster session 1 (Stalls Bar)

7:00pm Conference dinner, Bambalan

Wednesday 12th July

9:00am Session 9 Methodology (Hall 1)

Chair: Dr Gib Hemani, MRC Integrative Epidemiology Unit, UK
Assistant Professor Benjamin Neale*, Harvard Medical School and Massachusetts General Hospital, USA

Selection: another mode of action by phenotype
Dr Cosetta Minelli*, Population Health and Occupational Disease, Imperial College, London, UK

A Bayesian approach to pleiotropy in Mendelian randomization
Dr Stephen Burgess*, MRC Biostatistics Unit/Cardiovascular Epidemiology Unit, University of Cambridge, UK

Robust methods for Mendelian randomization: towards more reliable causal inferences

10:30am-11:00am Break

11:00am-12:30pm Parallel sessions

11:00am Session 10 Applications in cardiovascular disease and diabetes (Hall 1)

Chair: Professor Nishi Chaturvedi, MRC Unit for Lifelong Health and Ageing, UCL, UK
Dr Christina Ellervik*, University of Copenhagen, Denmark/Harvard Medical School, USA

Lactase persistence, milk intake, and ischemic heart disease, type 2 diabetes, hip fracture, and total mortality in the Danish general population
Linda Vissers, Julius Center for Health Sciences and Primary Care, UMC Utrecht

Dr Raymond Noordam, Leiden University Medical Center, The Netherlands

High-sensitivity c-reactive protein, low-grade systemic inflammation and type 2 diabetes mellitus: A two-sample Mendelian randomization study
Dr Kaitlin Wade, MRC Integrative Epidemiology Unit, University of Bristol, UK

Assessing the causal role of body mass index on cardiovascular health in young adults: Mendelian randomization and recall-by-genotype analyses
Danielle Rasooly, Harvard Medical School, USA

Helicobacter pylori identified as a modulator of the onset of Type 2 Diabetes

11:00am Session 11 Methodology (The Lantern)

Chair: Dr Jack Bowden, MRC Integrative Epidemiology Unit, UK
Fernando Hartwig*, Federal University of Pelotas, Brazil
Mendelian randomisation via the zero modal pleiotropy assumption: the mode-based estimate

Dr Roelof Smit, Leiden University Medical Center, The Netherlands
Simulating survival bias in Mendelian randomization

Dr Rachael Hughes, University of Bristol, UK
Selection bias in instrumental variable analyses

Dr Mats Julius Stensrud, University of Oslo, Norway
Survival bias in Mendelian Randomisation studies: Does the magnitude matter?

Jeremy Labrecque, Erasmus Medical Center, The Netherlands
Potential biases in MR studies: Considerations when exposures vary over the life-course

11:00am Session 12 Applications of lipid profiling (Terrace Bar)
Chair: Professor Debbie Lawlor, MRC Integrative Epidemiology Unit, UK

Dr Peter Würtz *, University of Helsinki & Nightingale Health Ltd
Metabolomic profiling and Mendelian randomisation of lipid-lowering targets: comparison of statins, ezetimibe and PCSK9-inhibition in 72,000 individuals

Dr Michael Holmes, University of Oxford, UK
PCSK9 genetic variants, metabolomics and cardiovascular risk factors and disease in the China Kadoorie Biobank

Dr Fotios Drenos, MRC Integrative Epidemiology Unit, University of Bristol, UK
Exploring the causal relationships between metabolic measures from a targeted NMR approach

Chen Li, MRC Epidemiology Unit, University of Cambridge
Characterizing gene-specific associations of LDL cholesterol with type 2 diabetes using untargeted metabolomics

Benjamin Sun, University of Cambridge, UK
Identifying genetic determinants of the human plasma proteome to facilitate two-sample Mendelian randomisation

12:30pm-1:30pm Lunch

1:30pm-3:00pm Parallel sessions

1:30pm Session 13 Applications to biomarker and target identification (Hall 1)
Chair: Professor Caroline Relton, MRC Integrative Epidemiology Unit, UK

Professor Brian Ference*, Wayne State University, USA
Using Mendelian randomization to inform the design and anticipate the results of randomized trials

Dr Karol Estrada, Computational Biology & Genomics, Biogen, Boston, USA
Use of Mendelian Randomization to Validate Human Genetic Targets for Alzheimer’s Disease

Dr Alisa Kjaergaard, Department of Clinical Biochemistry, Aarhus University Hospital, Denmark
Improving biomarker prediction by circumventing the component of variance explained by genotype: Biomarker De-Mendelization: proof of principle

Dr Martin Hornshaw, Metabolon,
Beyond the genome. How the metabolome can help uncover gene function

Dr Lavinia Paternoster, MRC Integrative Epidemiology Unit, University of Bristol, UK
Mendelian Randomization for disease progression: opportunities and methodological challenges

1:30pm

Session 14 Application to social science and health behaviours (The Lantern)

Chair: Dr Luisa Zuccolo, MRC Integrative Epidemiology Unit, UK

Dr Suzi Gage*, Department of Psychological Sciences, University of Liverpool, UK
Using 2-sample MR to investigate associations between substance use and schizophrenia risk

Dr Taavi Tillmann, Department of Epidemiology & Public Health, University College London, UK
Causal association between Education and Coronary Heart Disease: a two-sample Mendelian randomization study

Xuejie Ding, Department of Sociology, University of Oxford, UK
Educational attainment and allostatic load in later life: Evidence using genetic markers

Dr Neil Davies, MRC Integrative Epidemiology Unit, University of Bristol, UK
The causal effects of education on health: evidence from Mendelian randomization and the raising of the school leaving age 14.

Dr Iona Millwood, Clinical Trial Service Unit & Epidemiological Studies Unit, University of Oxford, UK
Mendelian randomisation of alcohol intake and cardiovascular diseases in Chinese adults

Professor Philipp Koellinger, Center for Neurogenomics and Cognitive Research, VU University Amsterdam, The Netherlands
Genetic Instrumental Variable (GIV) regression: Explaining socioeconomic and health outcomes in non-experimental data

1:30pm

Session 15 Methodology (Terrace bar)

Chair: Dr Frank Dudbridge, Department of Health Sciences, University of Leicester, UK

James Staley, MRC Integrative Epidemiology Unit, University of Bristol, UK
Semiparametric methods for estimation of a non-linear exposure-outcome relationship using instrumental variables in Mendelian randomization

Shirbi Ish-Shalom, Biomedical Informatics, Stanford University School of Medicine, USA
A survey of causal inferences via Mendelian randomization across the landscape of available GWAS

Dr Tom Michoel, The University of Edinburgh, UK
Efficient And Accurate Causal Inference With Hidden Confounders From Genome-Transcriptome Variation Data
Dr Michel Nivard, Department of biological psychology, VU University, The Netherlands
Power asymmetry should give us pause when interpreting results from summary statistic based Mendelian-randomisation.

Dr Dipender Gill, Imperial College London, UK
A test for violation of the InSIDE assumption

Dr Jack Bowden, MRC Integrative Epidemiology Unit, University of Bristol, UK
Improving the accuracy of two-sample summary data Mendelian randomization: moving beyond the NOME assumption

3:00pm-3:30pm Break

3:30pm-5:00pm Parallel sessions

3:30pm  Session 16 Pleiotropy (Hall 1)
Chair: Dr Cosetta Minelli, Population Health and Occupational Disease, Imperial College, London, UK

Dr Hans van Kippersluis*, Erasmus School of Economics, Erasmus University Rotterdam, The Netherlands
Pleiotropy-robust Mendelian Randomization

Dr Jack Bowden, MRC Integrative Epidemiology Unit, University of Bristol, UK
Improving the analysis and interpretation of two-sample summary data Mendelian Randomization via Radial plot regression

Amand Floriaan Schmidt, Institute of Cardiovascular Science, University College London, UK
Mendelian randomization with Egger pleiotropy correction and weakly informative Bayesian priors

Dr Christopher Foley, MRC Biostatistics Unit, University of Cambridge, UK
Demystifying causal effect heterogeneity of composite risk-factors in multi-instrument Mendelian Randomisation studies using a novel Bayesian feature selection algorithm

Dr Eleanor Sanderson, MRC Integrative Epidemiology Unit, University of Bristol, UK
An examination of multivariable Mendelian randomization: translating and extending econometric theory from the single sample to the two sample summary data setting

3:30pm  Session 17 Applications in cancer (The Lantern)
Chair: Professor Richard Martin, MRC Integrative Epidemiology Unit, UK

Dr Paul Brennan*, International Agency for Research on Cancer, France
Elucidating the role of diabetes in cancer using an MR approach

Associate Professor Stuart MacGregor*, QIMR Berghofer Medical Research Institute, Australia
Applications of MR in cancer

Robert Carreras-Torres, International Agency for Research on Cancer, France
Solving complex relationships between metabolic parameters and renal cell carcinoma risk: from Mendelian randomization to matrix analysis.

3:30pm  **Session 18 Applications (Terrace bar)**

Chair: Dr Beate St Pourcain, *Max Planck Institute for Psycholinguistics, The Netherlands*

**Dr Gibran Hemani, MRC Integrative Epidemiology Unit, University of Bristol, UK**
The causal map of 150 complex traits and diseases: a first draft

**Ruth Boxall, University of Oxford, UK**
Association of a LDL-cholesterol genetic instrument with cardiovascular risk factors and events in the China Kadoorie Biobank

**Shan Luo, School of Public Health, The University of Hong Kong**
Mendelian randomization study compared to randomized controlled trial: a systematic review and meta-analysis

**Dr Jie Zheng, MRC Integrative Epidemiology Unit, University of Bristol, UK**
Do blood lipid levels influence bone mineral density? Findings from a Mendelian randomization study

**Dr Ashley Budu-Aggrey, MRC Integrative Epidemiology Unit, University of Bristol, UK**
Body mass index and inflammatory skin disease: is there a causal relationship?

**Hanieh Yaghootkar, Institute of Biomedical and Clinical Science, University of Exeter Medical School, UK**
Uncoupling “healthy obesity” from “unhealthy obesity” genetic variants to provide a tool for Mendelian randomization studies

5:00pm  **Poster session 2 (Stalls Bar)**
Thursday 13th July

9:00am  Session 19 Invalid instruments (The Lantern)

Chair: Dr Nic Timpson, MRC Integrative Epidemiology Unit, UK

Professor Eric Tchetgen Tchetgen*, School of Public Health, Harvard University, USA
Recent developments on robust inference in Mendelian randomization studies

Dr Frank Dudbridge*, Department of Health Sciences, University of Leicester, UK
Mendelian randomisation across a range of exposure

Wes Spiller, MRC Integrative Epidemiology Unit, University of Bristol, UK
Linear Slichter Regression: using gene-environment interactions to correct for pleiotropic bias in Mendelian randomization analyses

10:30am-11:00am Break

11:00am  Session 20 Molecular mediation (The Lantern)

Chair: Dr Sarah Lewis, MRC Integrative Epidemiology Unit, UK

Professor Caroline Relton*, MRC Integrative Epidemiology Unit, University of Bristol, UK
Mendelian randomisation for molecular mediation: Applications in epigenetics

Dr Tom Richardson, MRC Integrative Epidemiology Unit, University of Bristol, UK
Causal epigenome-wide association study identifies CpG sites with evidence of a mediatory role in cardiovascular disease risk

Dr Nicole Warrington*, The University of Queensland Diamantina Institute, The University of Queensland, Australia
Can MR using structural equation modellng inform us about the fetal origins hypothesis?

12:15pm  Prize giving and close

12:30pm  Lunch
Session 2: Large scale biobanks
Chair: Professor George Davey Smith, MRC Integrative Epidemiology Unit, University of Bristol, UK
UK Biobank: a resource for genetic-epidemiological research

Presenter: Sir Rory Collins
Oxford. Email: audrey.hayes@bristol.ac.uk

Co-authors: Rory Collins
Principal Investigator UK Biobank, University of Oxford, UK

Abstract
UK Biobank has been set up to allow reliable assessment of the relevance of a very wide range of different types of exposure to a very wide range of different diseases, not just those that kill or are cancers (which have been extensively studied) but also many other conditions that cause much disability (such as dementia). Enrolment of 500,000 men and women aged 40 to 69 was achieved during 2006-2010. Detailed questionnaire, interview and measurement data were obtained, and multiple samples of blood, urine and saliva stored for analysis. Activity is now focused on further enhancing the phenotyping of the participants (including by genotyping and biochemical assays for all 500,000 participants, and multi-modal imaging of 100,000 of them) and on carefully characterising the disease outcomes being detected during follow-up by linkage to primary and secondary care health-related records. The UK Biobank Resource is available (via an internet-based access management system at www.ukbiobank.ac.uk) to all researchers (academic or commercial) anywhere in the world (without preferential or exclusive access) for any health-related research that is in the public interest.

Keywords:
biobank
China Kadoorie Biobank: opportunities for personalized medicine and pharmacogenomics

Presenter: Prof. Zhengming Chen
University of Oxford. Email: zhengming.chen@ctsu.ox.ac.uk

Abstract
Chronic diseases, such as stroke, IHD, and cancer are the leading causes of disability and death worldwide. Despite recent advances, our ability to prevent and treat these conditions is still limited. Understanding what causes these diseases in diverse populations with different lifestyles, environments and genetic architectures can lead to improved disease prevention and risk prediction, and the development of personalised medicine. Unique opportunities to fulfill these goals are offered by prospective “biobank” studies.

China Kadoorie Biobank (CKB) is one of the world’s largest prospective biobanks ever established. It involves 512,891 adults recruited during 2004-08 from 10 diverse areas in China, with extensive data collected at baseline and periodic resurveys, on lifestyle, environmental, physiological (e.g. blood pressure, adiposity, lung function) and psychological factors (e.g. depression, sleeping disorder, personality trait), and with long-term storage of biological samples (DNA, plasma, urine). To date, >0.5 million fatal and non-fatal disease events of >1000 different types (e.g. stroke, heart disease, cancer, diabetes, fracture, depression, suicide, cataract and rheumatoid arthritis) have been recorded. These exposure and health outcome data are now being complemented by blood assays of genetic (e.g. 800K SNPs, including 80K functional variants, using custom designed array for Chinese), metabolomic (e.g. ~250 metabolites), proteomic (e.g. up to 400 biomarkers) and infective (~20 pathogens) biomarkers, with many novel findings starting to emerge, including phenome-wide analyses of the impact of modulating levels and/or activity of major potential or actual drug targets (e.g. lipoprotein-related PLA2G7, CETP, PCSK9, HMGCR, and NPC1L1 proteins). The uniquely powerful and increasingly rich resources in CKB will enable many important discoveries relevant to risk prediction, disease prevention and treatment of many common conditions.

Keywords:
biobank, china, genetics, cohorts
Mendelian randomization in a large-scale population study linked with national health registries: the HUNT Study in Norway

Presenter: Prof Bjørn O Åsvold
K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway. Email: bjorn.o.asvold@ntnu.no

Co-authors: Bjørn O Åsvold [1]
1 K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway

Abstract
In the HUNT Study, the entire adult population of Nord-Trøndelag county, Norway, has been invited to health surveys since the 1980s. The surveys cover a wide range of health-related topics and include questionnaires, clinical examination and blood and urine sampling. Approximately 70,000 participants from the HUNT2 (1995-97) and HUNT3 (2006-08) Surveys have now been genome-wide genotyped using a customized Illumina HumanCoreExome genotyping array. Further genotypes have been imputed using a merged reference panel of samples from HUNT and the Haplotype Reference Consortium. The national identity number of all Norwegian citizens enables linkage of HUNT study data with local and national health registries, such as hospital records, information from general practitioners, national registries of births, cancer, and causes of death, and the national prescription database. I will describe opportunities for Mendelian randomization studies using these data, exemplified by a study of PCSK9 and HMGCR variants.

Keywords:
mendelian randomization, health registries, serum lipids
Session 4: Applications in pregnancy and maternal exposures
Chair: Professor Aroon Hingorani, Institute of Cardiovascular Science, and Farr Institute, University College London, UK
Mendelian randomization and testing causal effects of maternal (intrauterine) exposures on offspring outcomes - where are we now?

Presenter: Deborah Lawlor
MRC Integrative Epidemiology Unit at the University of Bristol. Email: d.a.lawlor@bristol.ac.uk

Abstract
Few of the methodological developments in MR have considered the specific situation of using genetic IVs to test the causal effect of exposures in pregnant women on postnatal offspring outcomes. In this presentation, I will describe specific ways in which the IV assumptions might be violated when MR is used to test such intrauterine effects. I will highlight the importance of considering the extent to which there is overlap between genetic variants in offspring that influence their outcome with genetic variants used as IVs in their mothers. Where there is overlap, and particularly if it generates a strong association of maternal genetic IVs with offspring outcome via the offspring genotype, the exclusion restriction assumption of IV analyses will be violated. I will suggest a set of analyses that ought to be considered when MR is used to address research questions concerned with intrauterine effects on post-natal offspring outcomes, and provide details of how these can be undertaken and interpreted. These additional analyses include the use of genetic data from offspring and fathers, examining associations using maternal non-transmitted alleles, and using simulated data in sensitivity analyses. I will briefly discuss the extent to which new methods that have been developed for exploring violation of the exclusion restriction assumption in the two-sample setting (MR-Egger and median based methods) might be used when exploring intrauterine effects in one-sample MR. Real examples will be used to illustrate the key issues and recommended analyses.

Keywords:
developmental origins, mendelian randomization, iv assumption violation
Understanding the causal nature of associations between maternal characteristics and offspring birth weight

Presenter: Dr. Rachel Freathy
University of Exeter. Email: r.freathy@ex.ac.uk

Co-authors: Rachel M. Freathy [1,2]
1 University of Exeter Medical School, University of Exeter, Exeter EX2 5DW
2 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN

Abstract
Women who are overweight or who have diabetes in pregnancy tend to have bigger babies with higher rates of associated complications. However, it is unclear which obesity-related maternal characteristics are causally related to birth weight. Recently, we published a study of 30,487 mothers and babies, which showed evidence of causal associations between higher maternal body mass index (BMI) or fasting glucose and higher birth weight. Our results also supported a causal relationship between higher maternal blood pressure and lower offspring birth weight.

We have used data from the first release of UK Biobank genetic data (n=48,632 women reporting birth weight of first child) to replicate these associations. We have additionally performed a meta-analysis of genome-wide association studies (GWAS) in 86,577 European women from 25 studies (n=37,945 from Early Growth Genetics (EGG) Consortium studies; n=48,632 from UK Biobank) to aim to identify novel pathways by which the maternal genotype might influence birth weight through the intrauterine environment.

Analysis in UK Biobank confirmed associations between higher maternal glucose and higher birth weight, and between higher maternal blood pressure and lower offspring birth weight. However, there was little evidence to support the overall causal association between maternal BMI and birth weight. This may in part reflect the opposing nature of the glucose and blood pressure associations, both of which are positively associated with BMI. The GWAS identified 10 loci associated with birth weight, which implicate a number of biological pathways and highlight the importance of accounting for fetal genotype.

Follow-up is planned in the upcoming release of the full UK Biobank genetic dataset, which will enhance power. Ultimately, these analyses will prioritise modifiable maternal risk factors for follow-up in intervention studies designed to reduce rates of very high (or very low) birth weight and associated complications.

Keywords:
birth weight, maternal genotype, fetal genotype, mendelian randomisation, gwas
Mendelian Randomisation shows circulating maternal vitamin D and calcium may be causally associated with birth weight

Presenter: Mr William David Thompson
Institute of Biomedical and Clinical Science, University of Exeter Medical School, University of Exeter.
Email: wdt204@exeter.ac.uk

1 Institute of Biomedical and Clinical Science, University of Exeter Medical School, University of Exeter; 2 MRC Integrative Epidemiology Unit at the University of Bristol; 3 School of Social and Community Medicine, University of Bristol; 4 Avon Longitudinal Study of Parents and Children, University of Bristol; * Joint senior authors who contributed equally

Abstract

Background
Observational studies have shown positive associations between maternal vitamin D or calcium levels and birth weight, though these may be explained by confounding. We investigated whether these traits in the mother were associated with the birth weight of the child.

Methods
We used UKBiobank (n=48,632), ALSPAC (n=7,853) and EFSOCH (n=746) data to perform Mendelian Randomisation using trait-specific SNPs. We made three weighted allele scores (WAS), one for vitamin D synthesis, vitamin D metabolism and for Calcium, using genome-wide significance SNPs previously reported by published GWAS. Weights from those GWAS were used in our independent data sets. Causative estimates were calculated using the ratio estimator, which is equivalent to the inverse-variance weighted method. We used MR-Egger to check for pleiotropy and “leave-one-out” analysis to check for single SNP effects. We also used ALSPAC and EFSOCH data to adjust for fetal genotypes, and compared WAS-own birth weight (i.e. fetal) with maternal WAS-first child birth weight associations in UKBiobank.

Results
We calculated that a 10% higher maternal 25[OH]D level (synthesis score) was causally associated with a 15g (95% CI: 5g to 26g) higher birth weight, and a 1SD higher calcium level was causally associated with a 45g (95% CI: -5g to 96g) higher birth weight. The equivalent result for 25[OH]D level metabolism was 1g (95% CI: -7g to 9g). Vitamin D results were similar with the inverse-variance weighted and MR-Egger methods. For calcium, there was evidence of pleiotropy, and single SNP influences. There was no evidence (from conditional analyses or from maternal vs fetal WAS-birth weight associations) that the results were driven by fetal genotypes.

Conclusion
Our results suggest that higher maternal vitamin D synthesis causes higher birth weight. Research is required to explore the extent to which supplements of vitamin D and/or calcium might be useful for optimal birth weight.

Keywords:
25[oh]d, calcium, mr, birth weight, ukbiobank
Assessing causality in associations between maternal adiposity and perinatal birth outcomes in mothers and offspring: A Mendelian randomization approach

Presenter: Dr Michelle Taylor  
*MRC Integrative Epidemiology Unit, University of Bristol. Email: michelle.taylor@bristol.ac.uk*

Co-authors: Michelle Taylor [1], Maria-Carolina Borges [1], Rebecca Richmond [1], Jane West [2], Rosie McEachan [2], John Wright [2], Debbie Lawlor [1]  
1 *MRC Integrative Epidemiology Unit and School of Social and Community Medicine, University of Bristol, Bristol, BS8 2BN*  
2 *Bradford Institute for Health Research, Bradford Royal Infirmary, Bradford, BD9 6RJ*

Abstract
Aim: To carry out Mendelian randomization (MR) analysis to determine the causal effect of maternal BMI on a number of perinatal outcomes.

Methods: Data from the Avon Longitudinal Study of Parents and Children and the Born in Bradford cohort (White European individuals only) were used (N=3,554 to 9,480 for different outcomes). Multivariable regression associations between maternal BMI and each outcome was examined. We used 97 genetic variants associated with BMI in a weighted allele score in inverse variance IV analyses to obtain causal estimates of maternal BMI on each outcome. We used MR-Egger and weighted median (WM) IV analysis to explore potential bias due to horizontal pleiotropy. For outcomes where offspring genetic predisposition for BMI was likely to play a role, analyses were adjusted for offspring genotype.

Results: In observational analysis, a 1 standard deviation (SD, ~4.95 kg/m²) increase in maternal BMI was associated with a 0.12 SD (95% CI 0.10-0.14) increase in offspring birthweight, 1.64 (95% CI 1.57-1.75) greater odds of gestational hypertension, 1.34 (95% CI 1.25-1.43) greater odds of induced labour and 1.36 (95% CI 1.28-1.46) greater odds of macrosomia. In MR analyses, a 1 SD increase in maternal BMI was associated with a 0.16 SD (95% CI 0.04-0.28) increase in offspring birthweight, 1.84 (95% CI 1.12-3.03) greater odds of gestational hypertension, 1.55 (95% CI 1.07-2.23) greater odds of induced labour and 2.10 (95% CI 1.20-3.71) greater odds of macrosomia. Results from MR Egger regression and WM were generally consistent. For many other outcomes, MR estimates were statistically consistent, but the 95% confidence intervals were wide and included the null.

Conclusions: Our findings support a causal effect of maternal BMI on several perinatal outcomes with greater BMI resulting in poor perinatal health. MR analyses are underestimated and we are currently working with more studies in a large consortia to better address this research question.

Keywords: bmi, perinatal health, alspac, born in bradford
Dissecting the Causal Mechanisms of Maternal Blood Pressure and Blood Glucose Level on Birth Weight: A Genetic Score Analysis

Presenter: Dr. Ge Zhang

Human Genetics Division, Cincinnati Children’s Hospital Medical Center, Cincinnati, Ohio, United States of America. Email: ge.zhang@cchmc.org

Co-authors: Ge Zhang[1], Amit Srivastava[1], Jonas Bacelis[2], Kari Teramo[4], Mikko Hallman[5], Bo Jacobsson[3], Louis Muglia[1]

1 Human Genetics Division, Cincinnati Children’s Hospital Medical Center, Cincinnati, Ohio, United States of America, Center for Prevention of Preterm Birth, Perinatal Institute, Cincinnati Children’s Hospital Medical Center and March of Dimes Prematurity Research Center Ohio Collaborative, Cincinnati, Ohio, United States of America
2 Department of Obstetrics and Gynecology, Sahlgrenska University Hospital, Gothenburg, Sweden
3 Department of Genes and Environment, Division of Epidemiology, Norwegian Institute of Public Health, Oslo, Norway, Department of Obstetrics and Gynecology, Institute of Clinical Sciences, Sahlgrenska Academy, Gothenburg University, Gothenburg, Sweden
4 Obstetrics and Gynecology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
5 PEDEGO Research Center, University of Oulu and Department of Children and Adolescents, Oulu University Hospital, Oulu, Finland

Abstract

Background

Observational epidemiological studies indicate that maternal blood pressure and blood glucose level are associated with their infant’s birth weight. Generally, high maternal blood pressure is associated with low birth weight and high maternal blood glucose is associated with increased birth weight. Different mechanisms have been postulated to explain these associations. This study aimed to investigate the causal mechanisms underlying the observed associations using a parental-specific genetic score approach.

Methods and Findings

We conducted a genetic score analysis using phenotype and genome-wide SNP data of 9,747 mother/infant pairs from multiple birth cohorts collected from European countries and the United States. We built genetic scores using SNPs known to be associated with blood pressure (66 SNPs) and fasting blood glucose level (65 SNPs) to index these maternal phenotypes respectively. To avoid the interference in causal inference due to the transmission of maternal alleles, we inferred parental transmission of the SNPs and constructed parental-specific haplotype genetic scores. Our haplotype genetic score association analysis showed that the maternal transmitted haplotype score of blood pressure was significantly associated with birth weight (P = 3.91e-05), suggesting genetic variants associated with blood pressure have direct effects in the fetus. In contrast, the maternal non-transmitted haplotype score of blood glucose was significantly associated with birth weight (P = 2.67e-3), indicating direct causal effect of maternal blood glucose concentration on birth weight.

Conclusions

Our results demonstrate that maternal blood pressure and blood glucose level are associated with birth weight through different causal mechanisms. The association between maternal blood pressure and birth weight is mainly defined by fetal genetics, while maternal glucose level is more likely to have a direct causal effect on birth weight.

Keywords: birth weight, maternal blood pressure, maternal blood glucose, genetic score, causal mechanism
Session 5: Applications
Chair: Professor Nicole Soranzo, Wellcome Trust Sanger Institute, UK
Using Structural Equation Modelling in Mendelian Randomization Studies

Presenter: Prof David Evans
University of Queensland Diamantina Institute, Translational Research Institute, University of Queensland. Email: d.evans1@uq.edu.au

Co-authors: David M. Evans [1,2], Nicole M. Warrington [2]
1 University of Queensland Diamantina Institute, Translational Research Institute, University of Queensland
2 MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol

Abstract
Mendelian randomization (MR) is a method of estimating the causal effect of modifiable environmental exposures on medically relevant outcomes, identifying molecular biomarkers that are likely to be causal for disease, and determining the suitability of drug targets for pharmacological intervention. However, MR studies are currently performed using very simple statistical methods based on e.g. two stage least squares and Wald ratios. These approaches potentially lack flexibility to model more complicated causal networks involving many different variables, bidirectional relationships, and horizontal pleiotropy, which in some cases may invalidate analyses and bias estimates of causal effects. Structural Equation Modelling (SEM) is a very flexible statistical tool that allows the modelling of complex linear dependencies between variables and the estimation of causal effects. Despite its potential advantages, SEM has yet to have been employed in MR studies except in the simplest of situations. In this presentation I give an overview of SEM, outline its advantages and limitations, and discuss examples of attempts by our group and others to combine SEM with Mendelian randomization principles to estimate complicated causal effects that may not be easy to estimate using simple ratio or two stage least squares models.

Keywords:
structural equation modeling, mendelian randomization, direction of causation, twin model, dohad
Identifying and utilising genetic variants associated with morning plasma cortisol: a CORtisol NETwork (CORNET) analysis

Presenter: Dr Andrew Crawford
University of Bristol. Email: andrew.crawford@bris.ac.uk

Co-authors: Andrew A Crawford [1, 2]
Brian R Walker [2]
on behalf of the CORNET consortium
1 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN
2 BHF Centre for Cardiovascular Science, Queen’s Medical Research Institute, University of Edinburgh, Edinburgh, UK

Abstract
The latest genome-wide association meta-analysis conducted by the CORNET consortium investigated morning plasma cortisol in 25,314 individuals from 17 cohort studies. Genetic variation in the SERPINA6/A1 locus on chromosome 14 reached a genome-wide level of significance (top SNP rs12589136, p=3.2 x 10^-19). Linkage disequilibrium score regression analyses estimated a SNP heritability of 4% (0.038, se 0.018). Positive genetic correlations were identified between plasma cortisol and LDL cholesterol and total cholesterol, and negative correlations with obesity, BMI, body fat, schizophrenia, chronotype and cigarettes smoked per day (absolute rg effect sizes between 0.23 and 0.63, p<0.05). Pathway analyses were performed using MAGMA and suggested over-representation of genes involved in lipid metabolism and circadian pathways. Two sample Mendelian randomisation analyses were performed using three independent SNPs in the SERPINA6/A1 locus as an instrument for plasma cortisol and publicly available summary GWAS data. These analyses suggested that a 1SD higher plasma cortisol is causally associated with a 35% increase in type 2 diabetes (IVW method, lnOR 0.30, se 0.13, p=0.02, DIAGRAM, cases=26,488; controls=83,964) and 28% increase in risk of coronary artery disease (lnOR 0.25, se 0.11, p=0.03, CARDIoGRAM, cases=22,233; controls=64,762). They also suggest a 1SD higher BMI is causally associated with a 0.09SD reduction in plasma cortisol (beta -0.09, se 0.04, p=0.02, GIANT 2015, n=339,224). Consistent estimates were obtained when using maximum likelihood or weighted median approaches. Our results, combined with similar findings from prospective, observational analyses, suggest approaches which target tissue-specific cortisol action, such as inhibition of the cortisol regenerating enzyme 11βHSD1, should be evaluated thoroughly for their effects on CVD outcomes.

Keywords:
cortisol, cardiovascular, type 2 diabetes, two sample mendelian randomisation
Body mass index and risk of musculoskeletal pain – a Mendelian Randomisation analysis in the HUNT Study

Presenter: Gunnhild Aaberge Vie

Department of Public Health and Nursing, Faculty of Medicine and health sciences, NTNU, Norwegian University of Science and Technology, Trondheim, Norway. Email: gunnhild.vie@ntnu.no

Co-authors: Gunnhild A Vie[1], Ben Brumpton[2,3,4], Egil A Fors[1], Ellen H Modalsli[1,5], Bjørn Olav Åsvold[2,6], Johan Håkon Bjørngaard[1,7]

1 Department of Public Health and Nursing, Faculty of Medicine and health sciences, NTNU, Norwegian University of Science and Technology, Trondheim, Norway
2 K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, Faculty of Medicine and health sciences, NTNU, Norwegian University of Science and Technology, Trondheim, Norway
3 Department of Thoracic Medicine, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway.
4 MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol, UK
5 Department of Dermatology, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway
6 Department of Endocrinology, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway
7 Forensic Department and Research Centre Brøset, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway

Abstract

Objectives: To use a Mendelian Randomisation (MR) approach to examine the robustness of the association between body mass index (BMI) and musculoskeletal pain previously identified in observational studies.

Methods: Among 65,762 participants in the second or third wave of the Nord-Trøndelag Health Study (HUNT), we estimated the cross-sectional association between observed BMI and self-reported muscle or joint pain lasting three consecutive months or more during the previous year, with logistic regression adjusted for age, sex, education and smoking status. To improve causal inference, we performed one-sample and two-sample MR analyses. We created a weighted genetic risk score (GRS) using 74 available single nucleotide polymorphisms (SNPs) and external weights from the GIANT consortium. The association between GRS and outcome was estimated with logistic regression and the association between GRS and BMI with linear regression, adjusted for age and sex. The causal odds ratio (OR) was estimated using the ratio method with bootstrapped confidence intervals in one-sample analyses, and using inverse variance weighting of summarized SNP-exposure data from the GIANT consortium and SNP-outcome data from HUNT in two-sample analyses. MR Egger regression was performed as sensitivity analyses for the latter.

Preliminary results: Measured BMI was associated with increased odds of pain (adjusted OR 1.23 (95% CI 1.20-1.25) per 5 units of BMI). Mendelian randomisation analyses supported an association between BMI and pain (OR 1.29, 95% CI 1.06-1.57) in one-sample MR and using inverse variance weighting in two-sample MR (OR 1.17, 95% CI 1.04-1.31). MR Egger gave a weaker association (OR 1.06, 95% CI 0.80-1.41) with possible pleiotropy (intercept β=0.003, 95%CI -0.005-0.01).

Conclusion: We found associations between genetically higher BMI and pain, supporting associations identified in observational analyses. However, associations might be overestimated due to pleiotropy.

Keywords: body mass index, musculoskeletal pain, mendelian randomisation
Investigating the combined association of BMI and alcohol consumption on liver disease biomarkers: a Mendelian randomization study of over 90,000 adults from the Copenhagen General Population cohort

Presenter: Miss Alice Carter
1MRC Integrative Epidemiology Unit at the University of Bristol, UK; 2 School of Social and Community Medicine. Email: alice.carter@bristol.ac.uk

Co-authors: Alice Carter,1,2 Maria-Carolina Borges,1,2 Marianne Benn,3,4,6 Anne Tybjaerg-Hansen,3,5,6 George Davey Smith,1,2 Børge G Nordestgaard.3,4,5 Debbie A Lawlor.1,2
1MRC Integrative Epidemiology Unit at the University of Bristol, UK
2 School of Social and Community Medicine, University of Bristol, UK
3The Copenhagen General Population Study, Herlev Hospital, Copenhagen, Denmark
4 Copenhagen University Hospital, Faculty of Health Sciences, University of Copenhagen, Denmark
5Department of Clinical Biochemistry, Herlev Hospital, Copenhagen, Denmark
6Department of Clinical Biochemistry, Rigshospitalet, Copenhagen, Denmark

Abstract
Body mass index (BMI) and alcohol consumption are suggested to independently and interactively increase the risk of liver disease. We aimed to assess the combined effect of BMI and alcohol consumption on liver disease biomarkers using factorial Mendelian randomization (MR).

We used multivariable regression and MR analysis to estimate individual and joint associations of BMI and alcohol consumption on the biomarkers, alanine aminotransferase (ALT) and gamma-glutamyl transferase (GGT). We then undertook a factorial MR study splitting participants according to median of BMI allele score and then by ADH1B genotype (AA/ AG and GG), which resulted in four groups named low BMI/low alcohol (-BMI/-alc), low BMI/high alcohol (-BMI/+alc), high BMI/low alcohol (+BMI/-alc) and high BMI/high alcohol (+BMI/+alc).

Observational and MR analysis provided evidence of positive associations of BMI and alcohol with circulating ALT and GGT (P=<0.0001 for all models) In the factorial MR analyses, considering the +BMI/+alc group as the reference, mean circulating ALT and GGT levels were lowest in the –BMI/-alc group (-1.93% (95% CI: -4.03;0.16) and -3.49% (95% CI: -6.05;-0.93) for ALT and GGT respectively). Individuals with -BMI/+alc and +BMI/-alc had lower mean circulating ALT and GGT compared to the reference group (+BMI/+alc). For ALT, the mean difference to reference group was larger for -BMI/+alc group (-1.15% (95% CI: -1.76, -0.55)) than for +BMI/-alc (-0.23% (95% CI: -2.37, 1.90)). For GGT, the mean difference was similar for these two factorial groups (-0.89 (95% CI: -1.62, -0.16) and -0.87 (95% CI: -3.42, -1.74) for -BMI/+alc and +BMI/-alc respectively.)

Both multivariable and MR analyses support greater BMI and alcohol consumption increasing circulating levels of biomarkers related to liver damage. The factorial MR provides some evidence of a joint association, where intervening on both BMI and alcohol might reflect the greatest reduction in liver disease.

Keywords:
bmi, alcohol, liver disease, factorial
Could HTR7, TAC3 or AXL be new potential genetic targets for ischemic heart disease?

Presenter: C Mary Schooling
School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong. Email: cms1@hku.hk

Co-authors: C Mary Schooling [1], Levi Waldron [2], MK Kwok [1], SL Lin [1], SL Au Yeung [1]
1 School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China
2 CUNY School of Public Health, New York, NY 10027, USA

Abstract
Despite successes cardiovascular disease prevention and treatment has recently encountered several late stage drug failures. To address this gap, we used the paradigm of growth and reproduction trading-off against longevity to identify new directions for cardiovascular disease interventions, focusing on drivers of the reproductive axis, specifically gonadotropin releasing hormone (GnRH) and anti-psychotics/anti-depressants targeting serotonin, histamine and gamma-aminobutyric acid because of their potential endocrine effects. Where possible we used two-sample Mendelian Randomization, applied to a large densely genotyped ischemic heart (IHD) case (n=60,801)-control (n=123,504) study, CARDIoGRAMplusC4D 1000 genomes, to estimates effects on ischemic heart disease (IHD). We also used a gene-based test applied to the same dataset to obtain an overall estimate of the association of the relevant target genes to IHD. Serotonin and its precursor, tryptophan were not associated with IHD (odds ratio (OR) 0.99, 95% confidence interval 0.74 to 1.34 and OR 0.63 95% CI 0.27 to 1.50 respectively) but the tryptophan metabolite, kynurenine was associated with IHD (OR 2.28, 95% CI 1.42 to 3.67). Histidine, a precursor of histamine, was unrelated to IHD (OR 0.94, 95% CI 0.84 to 1.06). Of 28 serotonin, histamine or gamma-aminobutyric acid autosomal genes targeted by anti-psychotics/anti-depressants 2 genes (GABRG1 p-value=0.04) and (HTR7 p-value=0.005) were associated with IHD. Of the 14 autosomal genes (GNRHR, GNRH1, KISS1, FGFR1, FGFR8, KISS1R, TAC3, TACR3, PROKR2, PROK2, SEMA3A, SOX10, AXL and SEMA3E) functionally related to GnRH deficiency 2 genes TAC3 (p-value=0.03) and AXL (p-value=0.00003) were associated with IHD. HTR7, TAC3 and AXL are related to existing interventions, some selective serotonin receptor inhibitors (HTR7), tachykinin receptor 3 antagonists (TAC3) and Axl protein inhibitors (AXL), that could perhaps be re-purposed to prevent and treat IHD.

Keywords:
ihd, gnrh, axl, tac3, kynurenine
Session 6: Applications in social science and health behaviours
Chair: Dr Neil Davies, MRC Integrative Epidemiology Unit, UK
Large-scale genetic study of risk tolerance and risky behaviors identifies novel loci and reveals shared genetic influences

Presenter: Philipp Koellinger
Vrije Universiteit Amsterdam. Email: p.d.koellinger@vu.nl

Co-authors: The Social Science Genetic Association Consortium (SSGAC)

Abstract
Risk tolerance is an important variable in the behavioral sciences, but few genetic variants have so far been found to robustly associate with it or with risky behaviors. We conducted large-scale genome-wide association studies (GWAS) of general risk tolerance and four risky behaviors: automobile speeding propensity, alcoholic drinks per week, whether one has ever been a smoker, and the lifetime number of sexual partners. Our GWASs identified several genome-wide significant loci associated with these outcomes or their first principal component. We report evidence of substantial pleiotropy between general risk tolerance and the four risky behaviors. Polygenic scores of risk tolerance predict a range of risky behaviors, personality variables, and ADHD. Our results begin to elucidate the biological mechanisms that underlie variation in risk tolerance and risky behaviors.

Keywords:
gwas, risk tolerance, smoking, alcohol, risky behaviors
GWAS of human reproductive behaviour

Presenter: Prof. Melinda Mills  
*University of Oxford. Email: melinda.mills@nuffield.ox.ac.uk*

Co-authors: Nicola Barban, GWAS human reproductive behaviour consortium  
*University of Oxford*

**Abstract**  
The genetic architecture of human reproductive behavior—age at first birth (AFB) and number of children ever born (NEB)—has a strong relationship with fitness, human development, infertility and risk of neuropsychiatric disorders. However, very few genetic loci have been identified, and the underlying mechanisms of AFB and NEB are poorly understood. In the presentation, we report a large genome-wide association study of both sexes including 251,151 individuals for AFB and 343,072 individuals for NEB. We identified 12 independent loci that are significantly associated with AFB and/or NEB in a SNP-based genome-wide association study and 4 additional loci associated in a gene-based effort. These loci harbor genes that are likely to have a role, either directly or by affecting non-local gene expression, in human reproduction and infertility, thereby increasing understanding of these complex traits. Additional extensions of this study are also discussed.

**Keywords:**  
fertility, behavioural genetics, genome-wide association studies, reproduction
Session 7: Methodology
Chair: Dr Steve Burgess, University of Cambridge, UK
Sensitivity Analysis and Power for Instrumental Variables Studies with Application to Mendelian Randomization

Presenter: Dylan Small
University of Pennsylvania. Email: dsmall@wharton.upenn.edu

Co-authors: Xuran Wang [1], Yang Jiang [1], Nancy R. Zhang [1], Dylan S. Small [1]
1 Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, PA USA 19104

Abstract
In observational studies to estimate treatment effects, unmeasured confounding is often a concern. The instrumental variable (IV) method can control for unmeasured confounding when there is a valid IV. To be a valid IV, a variable needs to be independent of unmeasured confounders and only affect the outcome through affecting the treatment. When applying the IV method, there is often concern that a putative IV is invalid to some degree. We present an approach to sensitivity analysis for the IV method which examines the sensitivity of inferences to violations of IV validity. Our approach is based on extending the Anderson-Rubin test and is valid regardless of instrument. A power formula for this sensitivity analysis is presented. We illustrate its usage via examples about Mendelian randomization studies and its implications via a comparison of using rare vs. common genetic variants as instruments.

Keywords:
anderson-rubin test, instrumental variable, measure of instrument strength
On the Use of the Lasso for Instrumental Variables Estimation with Some Invalid Instruments

Presenter: Helmut Farbmacher
Max Planck Society. Email: farbmacher@mpisoc.mpg.de

Co-authors: Frank Windmeijer [1], Neil Davies [2], George Davey Smith [3]
University of Bristol [1,2,3]

Abstract
We investigate the behaviour of the Lasso for identifying invalid instruments in linear instrumental variables models for estimating causal effects of exposures on outcomes, as proposed recently by Kang, Zhang, Cai and Small (2016, Journal of the American Statistical Association). Invalid instruments are such that they fail the exclusion restriction and enter the model as explanatory variables. We show that for this setup, the Lasso may not select all invalid instruments in large samples if they are relatively strong. Consistent selection also depends on the correlation structure of the instruments. We propose a median estimator that is consistent when less than 50% of the instruments are invalid, but its consistency does not depend on the relative strength of the instruments or their correlation structure. This estimator can therefore be used for adaptive Lasso estimation. The methods are applied to a Mendelian randomisation study to estimate the causal effect of BMI on diastolic blood pressure using data on individuals from the UK Biobank, with 94 single nucleotide polymorphisms as potential instruments for BMI.

Keywords:
causal inference, invalid instruments, lasso, mendelian randomisation
Constrained Instrument Variable Method and its Application to Mendelian Randomization with Observed Pleiotropy

Presenter: Lai Jiang
Lady Davis Institute, McGill University. Email: lai.jiang@mail.mcgill.ca

Co-authors: Lai Jiang [1] [7], Karim Ouakacha [2], J. Brent Richards [3], Vanessa Didelez [4], Pedro Rosa [5] [6], Andrea L. Benedet [6], Sulantha Mathotaarachchi [6], and Celia M.T.Greenwood [1] [7].
1 Lady Davis Institute, McGill University, Montreal, Canada
2 Université du Québec à Montréal (UQAM), Montreal, Canada
3 Department of Medicine, McGill University, Montreal, Canada
4 University of Bremen, Bremen, Germany
5 Department of Neurology & Neurosurgery and Psychiatry, McGill University, Montreal, Canada
6 Translational Neuroimaging Laboratory, McGill University Research Centre for Studies in Aging, Douglas Hospital, McGill University, Montreal, Canada
7 Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada

Abstract
In Mendelian randomization, genetic variants (SNPs) are used to construct instrumental variables to estimate the causal effect of a phenotype of interest on a disease outcome. However, valid MR inference is only possible when the genetic variants have no indirect effect on the outcome other than through the phenotype considered. Pleiotropy occurs when a genetic variant influence the response though multiple phenotypes of interest, and therefore renders it an invalid instrumental variable in MR studies, unless the pleiotropy can appropriately be taken into account.

When potentially pleiotropic phenotypes are observed, we propose a novel Constrained Instrumental Variable (CIV) method to construct valid instrumental variables and perform adjusted causal effect estimation correcting for pleiotropy. A constrained optimization approach using smoothed-L0 norm penalty is introduced to give sparse models and consequently correct for the causal estimate bias due to pleiotropy effects.

Simulations have been conducted under different types of pleiotropy to compare the performance of CIV method with other popular methods. Results show that our method leads to valid instrument selection and causal effect estimation with reduced bias. We also analyzed genetic factors and biomarkers of Alzheimer’s disease data from Alzheimer’s Disease Neuroimaging Initiative (ADNI) study. We identified a list of valid genetic instrument for each of these biomarkers of interest and provided adjusted estimation of their causal effect on AD progression. However, further work and underlying knowledge about the biology of the disease may still be needed to interpret our results.

In conclusion, we propose a new approach to construct valid instruments and adjusted causal effect estimator when pleiotropy is present and observable. Our data analysis illustrates that our CIV method leads to automatic feature selection and improved estimation of causal effects under a variety of types of pleiotropy.

Keywords: mendelian randomization, pleiotropy, constrained instrumental variables, l0 penalization, alzheimer’s disease
Detection and correction of pleiotropic bias in multivariable Mendelian randomization

Presenter: Chia-Yen Chen
Massachusetts General Hospital. Email: cychen@broadinstitute.org

Co-authors: Chia-Yen Chen [1,2,3], Marie Verbanck [4], Ron Do [4], Benjamin Neale [1,2,3]
1 Analytic and Translational Genetics Unit, Massachusetts General Hospital
2 Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital
3 Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard
4 Charles Bronfman Institute for Personalized Medicine, Department of Genetics and Genomics, Icahn School of Medicine at Mount Sinai

Abstract
Inferring the causal relationships between traits and diseases is essential to understanding disease etiology and the development of treatments. Mendelian randomization (MR) is a widely-used method that uses single nucleotide variants (SNVs) as instruments to obtain causal effects estimates. With the increasing number of trait-associated loci identified by genome-wide association studies (GWAS), methods have been developed to aggregate estimates from multiple instrumental variables, such as Multi-phenotype MR ([MPMR]; Do et al. 2013 Nat. Genet.) and MR-Egger regression.

A key assumption of MR analysis is exclusion restriction, which states that the genetic instrument solely acts on the disease outcome through the risk factor. Pleiotropy is a pervasive phenomenon in human genetics where it’s been shown that many trait-associated loci have multiple effects on other traits. Consequently, multi-instrument MR approaches are increasingly subject to violations of the exclusion restriction assumption.

To address this question, we propose a novel method, MR Pleiotropy RESidual Sum and Outlier (MR-PRESSO) test, for detection and correction of the pleiotropic in a multi-instrument MR analysis. To evaluate this approach, we performed a series of simulations with and without pleiotropic effects of the SNVs, and compared MR-PRESSO’s ability to detect pleiotropy bias with existing methods. Our simulations demonstrate that our method can identify presence of SNVs with pleiotropic effects and has higher statistical power than MR-Egger regression and non-inflated type 1 error than the Q-statistics. Our simulations further showed that MR-PRESSO can reduce pleiotropy bias in the causal estimate of MR without loss of statistical power in detecting causality. Finally, we successfully apply our method on GWAS summary data for body mass index and C-reactive protein. Our method improves on existing approaches to multi-instrument MR analysis by minimizing pleiotropic bias.

Keywords:
pleiotropy, multi-phenotype mr, mr-egger, q-statistics, gwas summary statistics
Instrumental Variables Estimation of Causal Effects in the Presence of Invalid Instruments

Presenter: Frank Windmeijer
University of Bristol. Email: f.windmeijer@bristol.ac.uk

Co-authors: Frank Windmeijer [1]
Fernando Hartwig [2]
Jack Bowden [1]
1 MRC Integrative Epidemiology Unit, University of Bristol, UK
2 Federal University of Pelotas, Brazil

Abstract
In Mendelian randomization methods, genetic markers are used as instrumental variables (IVs) for a confounded exposure. Some may not be valid instruments if they have e.g. pleiotropic effects, affecting the outcome directly, invalidating the exclusion restriction. We propose a new method to identify valid instruments, based on the per marker estimates of the causal effect and their associated confidence intervals (CIs). IVs are grouped together when their CIs overlap and the markers in the largest group are selected as the valid instruments. This leads to consistent selection under the assumption that the valid instruments form the largest group. The method can be applied to both one- and two-sample settings.

Keywords:
iv estimation, exclusion restriction, confidence interval
Session 8: Applications in neuropsychiatric and immune diseases
Chair: Professor Dave Evans, University of Queensland Diamantina Institute, Brisbane, Australia
Modelling changes in genetic variances during development - The re-birth of twin modelling strategies in the GWAS era

Presenter: Beate St Pourcain
1) Language and Genetics Department, Max Planck Institute for Psycholinguistics, The Netherlands; 2) MRC Integrative Epidemiology Unit (MRC IEU), University of Bristol, UK; 3) Donders Institute for Brain, Cognition and Behaviour, Radboud University, The Netherlands. Email: Beate.StPourcain@mpi.nl

Co-authors: Beate St Pourcain [1,2,3], Lindon J Eaves [4], Simon E Fisher [1,4], Susan M Ring [2,5], Sarah Medland [6], David M Evans [2,7]. George Davey Smith [2,5]
1 Language and Genetics Department, Max Planck Institute for Psycholinguistics, The Netherlands
2 MRC Integrative Epidemiology Unit (MRC IEU), University of Bristol, UK
3 Donders Institute for Brain, Cognition and Behaviour, Radboud University, The Netherlands
4 Department of Human and Molecular Genetics, Institute for Psychiatric and Behavioral Genetics, Commonwealth University School of Medicine, Richmond, Virginia, USA
5 School of Social and Community Medicine, University of Bristol, UK
6 Psychiatric Genetics, QIMR Berghofer Medical Research Institute, Queensland, Australia
7 University of Queensland Diamantina Institute, Translational Research Institute, University of Queensland, Queensland, Australia

Abstract
Genetic correlation studies exploiting genome-wide genotyping information in unrelated individuals have advanced our understanding of the pleiotropic relationships among many complex traits. Besides a reliable representation of the genetic architecture, the interpretability of such findings crucially depends on the underlying phenotype definition. Psychiatric disorders, for example, can often be considered as diagnostic entities, defined by clinical criteria including the age of onset, while human behaviour changes continuously during development. Thus, the extent to which genetic aetiologies are shared between traits may depend on the genetic architecture of the phenotype during a defined developmental window. Our work aims to assess developmental variations in complex genetic trait architectures by flexibly modelling latent genetic factor structures within a multivariate context.
For this, we have developed a multivariate analysis framework combining whole-genome genotyping information in unrelated individuals with structural equation modelling techniques (Genetic-relationship-matrix structural equation model, GSEM) using a Full Maximum Likelihood approach. Analogous to multivariate twin research, multivariate GSEM can include Cholesky decomposition models, common pathway models or independent pathway models, among others.
Using repeatedly assessed social-communication traits in the ALSPAC study (Ns 5,551, Social Communication Disorders Checklist, 8 to 17 years), we will show how genetic factors can be modelled longitudinally using standard model fitting procedures. We will also show that the identified GSEM for social-communication difficulties is consistent with the previously described developmental trait profiles for genetic overlap with psychiatric disorder.
Our work shows that an understanding of developmental variations in complex genetic trait architectures is important for studies investigating shared genetic aetiologies between phenotypes.

Keywords: structural equation modelling, longitudinal analysis, genetic variance decomposition, genetic-relationship matrix structural equation modelling, genetic relationship matrix
Circulating Insulin-like Growth Factors and Alzheimer’s disease: a Mendelian randomization study

Presenter: Dr Dylan Williams
Department of Medical Epidemiology & Biostatistics, Karolinska Institutet, Stockholm, Sweden.
Email: dylan.williams@ki.se

Co-authors: Ida K Karlsson [1], Nancy L. Pedersen [1, 2], Sara Hägg [1]
1. Department of Medical Epidemiology & Biostatistics, Karolinska Institutet, Stockholm, Sweden
2. Department of Psychology, University of Southern California, Los Angeles, CA, USA

Abstract
Objective: Epidemiological studies have reported associations of lower circulating insulin-like growth factor (IGF) peptides with higher risk of Alzheimer’s disease (AD). We aimed to examine whether genetically predicted variation in IGF1, its main binding protein (IGFBP3), and/or the molar ratio of these peptides are associated with AD risk in Mendelian randomization analyses.

Methods: We first conducted meta-analyses of genotype-AD associations for ten IGF-related single nucleotide polymorphisms (SNPs), using published summary genome-wide association study (GWAS) statistics from the International Genomics of Alzheimer’s Project (IGAP; N= 17008 cases; 37154 controls). We assessed whether any SNP-disease associations replicated in an independent sample derived from the Swedish Twin Registry (N = 984 cases; 10304 controls). We also derived Wald estimators for magnitudes of associations of IGF peptide variation with AD, using a sub-set of seven of the SNPs for which relevant summary statistics were retrievable from published GWAS.

Results: Meta-analyses of SNP-AD associations did not suggest that variation in IGF1, IGFBP3, or the molar ratio of these, affect AD risk. Only one SNP was associated with AD in IGAP data. This variant is located in the gene FOXO3, implicated in human longevity. In a meta-analysis of both IGAP and secondary data, the odds ratio of AD per FOXO3 risk allele was 1.04 (95% confidence interval: 1.01, 1.08; P=0.008).

Conclusions: These findings suggest that circulating IGF1 and IGFBP3 are not important determinants of Alzheimer’s risk. FOXO3 function may influence Alzheimer’s development via pathways that are independent of IGF signaling (i.e. pleiotropic actions).

Keywords:
alzheimer’s disease, insulin-like growth factors
Estimating the causal influence of BMI on risk of Parkinson’s disease: a Mendelian randomization study

Presenter: Dr Alastair Noyce
University College London Institute of Neurology. Email: a.noyce@ucl.ac.uk

Co-authors: Alastair J Noyce [1], Demis A Kia [1], Gibran Hemani [2,3], Aude Nicolas [4], T Ryan Price [4], Eduardo De Pablo-Fernandez [1], Philip C Haycock [2,3], Patrick A Lewis [1,5], Thomas Foltynie [6], George Davey Smith [2,3], IPDGC collaborators, Anette Schrag [7], Andrew J Lees [1], John Hardy [1], Andrew Singleton [4], Mike A Nalls [4,8], Neil Pearce [9], Debbie A Lawlor [2,3], Nicholas W Wood [1]


Abstract

Objectives
To explore the association between higher body mass index (BMI) and Parkinson’s disease (PD) using Mendelian randomization (MR).

Methods
Two-sample MR was undertaken using genome-wide association (GWA) study data. The associations between the genetic instruments and BMI were obtained from the GIANT consortium and consisted of the per-allele difference in mean BMI for 77 independent variants that reached genome-wide significance. The per-allele difference in log-odds of PD for each of these variants was estimated from a recent meta-analysis, which included 13,708 cases of PD and 95,282 controls. The inverse variance weighted method was used to estimate a pooled odds ratio (OR) for the effect of a 5kg/m2 higher BMI on PD. Evidence of directional pleiotropy averaged across all variants was sought using MR-Egger regression. Frailty simulations were used to assess whether causal associations were affected by mortality selection.

Results
A combined genetic instrumental variable expected to confer a lifetime exposure of 5kg/m2 higher BMI was associated with a lower risk of PD (OR 0.82, 95% CI 0.69-0.98). MR-Egger regression gave similar results, suggesting that directional pleiotropy was unlikely to be biasing the result (intercept 0.002, p-value=0.654). However, the apparent protective influence of higher BMI could be at least partially induced by survival bias in the PD GWA study, as demonstrated by frailty simulations. Other important limitations of this application of MR include the inability to analyse non-linear associations, to undertake sub-group analyses, and to gain mechanistic insights.

Conclusions
In this large study using two-sample MR, we found that variants known to influence BMI had effects on Parkinson’s disease in a manner consistent with higher BMI leading to lower risk of PD. The mechanism underlying this apparent protective effect warrants further study.

Keywords: parkinson’s disease, body mass index, mendelian randomisation
Genetic analyses to test for a causal relationship between obesity and depression

Presenter: Dr Jessica Tyrrell
University of Exeter. Email: j.tyrrell@exeter.ac.uk

Co-authors: Jessica Tyrrell1, Andrew R Wood1 Hanieh Yaghoootkar1, Robin Beaumont1, Samuel E Jones1, Marcus A Tuke1, Katherine S Ruth1, Rachel M Freathy1, Anna Murray1, Michael N Weedon1, Timothy M Frayling1
1 Genetics of Complex Traits, University of Exeter Medical School, University of Exeter, UK

Abstract
Background: Obesity and depression are chronic modern conditions. Numerous observational studies have demonstrated associations between obesity and depression but the causal directions are poorly understood. These associations differ by sex, with obese women disproportionately affected by depression. 
Objective: To test for bidirectional causal relationships between obesity and depression in 120,000 UK Biobank participants using Mendelian Randomisation (MR).
Methods: Firstly, we created genetic risk scores for depression and body mass index (BMI) in 120,000 individuals of white British origin in the UK Biobank, using genetic variants identified in the most recent genome-wide association scans of these traits. We tested the BMI GRS for association with several depression variables including single episode major depression, recurrent depression and any depression and the depression GRS was tested for association with BMI and obesity. We also performed MR-Egger and median MR as sensitivity analyses to check for pleiotropy and stratified our analyses by sex.
Results: Our MR analyses suggest that depression causally associates with higher BMI. A 1 unit increase in natural logged OR of depression was causally associated with a 0.19 (95%CI: 0.03, 0.34) SD higher BMI, similar results were noted for males and females. There was nominal evidence of a causal relationship between BMI and depression in all individuals. Sex-stratified MR provided no evidence of a relationship in males, but borderline evidence in females with a one SD higher BMI associating with a 1.24 (95%CI: 1.00, 1.52) higher odds of depression. Results were similar in the sensitivity analyses.
Conclusions: Initial analyses in 120,000 UK Biobank participants provides some evidence for a bidirectional causal relationship between BMI and depression, especially in women. Further refinement of the causality of these associations will provide insight into this complex relationship between obesity and mental health.

Keywords:
depression, obesity, uk biobank, mendelian randomisation
Estimating the causal effect of body mass index on hay fever, asthma, and lung function: a mendelian randomization analysis

Presenter: Dr. Tea Skaaby
Research Centre for Prevention and Health, Capital region of Denmark, Glostrup, Denmark. Email: tea.skaaby.01@regionh.dk

Co-authors: Tea Skaaby [1], Amy E. Taylor [2,3], Betina H. Thuesen [1], Rikke K. Jacobsen [1], Nele Friedrich [1,4], Line Tang Møllehave [1], Susanne Hansen [1], Sofus C. Larsen [5], Uwe Völker [6], Matthias Nauck [4], Henry Völzke [7], Torben Hansen [8], Oluf Pedersen [8], Torben Jørgensen [1,9,10], Lavinia Paternoster [2], Marcus Munafò [2,3], Niels Grarup [8], Allan Linneberg [1,11,12].

1 Research Centre for Prevention and Health, Centre for Health, Capital Region of Denmark, Copenhagen, Denmark. 2 MRC Integrative Epidemiology Unit (IEU) at the University of Bristol, Bristol, UK 3 UK Centre for Tobacco and Alcohol Studies, School of Experimental Psychology, University of Bristol, Bristol 4 Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Germany 5 Research unit for Dietary Studies, the Parker Institute, Frederiksberg and Bispebjerg Hospitals, The Capital Region, Frederiksberg, Denmark 6 Interfaculty Institute for Genetics and Functional Genomics, University Medicine and Ernst-Moritz-Arndt University Greifswald, Germany 7 Institute for Community Medicine, University Medicine Greifswald, Germany 8 The Novo Nordisk Foundation Center for Basic Metabolic Research, Section on Metabolic Genetics, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark 9 Department of Public Health, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark 10 Faculty of Medicine, Aalborg University, Denmark 11 Department of Clinical Experimental Research, Rigshospitalet, Glostrup, Denmark 12 Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

Abstract

Background: Observational studies have shown that body mass index (BMI) is positively associated with asthma. However, observational data are prone to confounding and reverse causation. In Mendelian randomization, genetic variants are used as un-confounded markers of exposures to examine causal effects. We examined the causal effect of BMI on asthma, hay fever, allergic sensitization, serum total immunoglobulin E (IgE), forced expiratory volume in one second (FEV1), and forced vital capacity (FVC).

Methods: We included 490,497 participants in the observational and 162,124 participants in the genetic analyses. A genetic risk score (GRS) was created using 26 BMI-associated single nucleotide polymorphisms (SNPs). Results were pooled in meta-analyses and expressed as odds ratios (ORs) or β-estimates with 95% confidence interval (CI).

Findings: The GRS was significantly associated with asthma (OR=1.009; 95% CI: 1.004, 1.013), but not with hay fever (OR= 0.998; 95% CI: 0.994, 1.002), or allergic sensitization, serum total immunoglobulin E (IgE), forced expiratory volume in one second (FEV1), and forced vital capacity (FVC). Effect sizes estimated by instrumental variable analyses were OR=1.07 (95% CI: 1.03, 1.10) for asthma, a 9 ml decrease in FEV1 (95% CI: 2.0–15 ml decrease), and a 16 ml decrease in FVC (95% CI: 7.0–24 ml decrease) per 1 kg/m2 higher BMI.

Conclusion: The results support the conclusion that increasing BMI is causally related to higher prevalence of asthma and decreased lung function, but not with hay fever or biomarkers of allergy.

Keywords: serum specific ige, hay fever, asthma, allergic sensitization, lung function.
Bayesian approaches to pleiotropy in Mendelian randomization

Presenter: Dr Cosetta Minelli
Population Health and Occupational Disease, NHLI, Imperial College London, London SW3 6LR, UK.
Email: cosetta.minelli1@imperial.ac.uk

Co-authors: Prof John Thompson [1], Dr Cosetta Minelli [2]
1 Department of Health Sciences, University of Leicester, Leicester LE1 7RH, UK
2 Population Health and Occupational Disease, NHLI, Imperial College London, London SW3 6LR, UK

Abstract
In two-sample Mendelian randomization (MR) with multiple instruments, MR-Egger is increasingly used to detect and adjust for pleiotropy. SNP-outcome are regressed against SNP-exposure association estimates, with slope representing the causal effect estimate and intercept the average pleiotropy. Due to uncertainty in the estimation of SNP-specific pleiotropy, MR-Egger can suffer from low statistical power even when using many instruments, and is typically used as a sensitivity analysis. We developed two Bayesian methods to improve detection and adjustment for pleiotropy. The first is a Bayesian Model Averaging (BMA) approach over three possible pleiotropy models, where results are averaged with weighting based on evidence from the data: 1) Bayesian fixed-effect meta-analysis of SNP-specific MR estimates, which assumes no pleiotropy; 2) Bayesian random-effects meta-analysis, which assumes pleiotropy that cancels out; 3) MR-Egger, which assumes directional pleiotropy. Simulation work demonstrated a good trade-off between bias and precision of the causal effect estimate, and we suggested BMA for MR studies where there is no knowledge about pleiotropy. We show how BMA might also be improved by grouping SNPs based on prior knowledge on possible pleiotropic mechanisms. The second approach is a Bayesian shrinkage model with a spike and slab prior for the pleiotropic effects, where the spike centred on 0 is for SNPs with no evidence of pleiotropy from the data. Importantly, this method does not require the InSIDE (Instrument Strength Independent of Direct Effect) assumption, unlike the others. The down-weighting of pleiotropic SNPs, combined with improved estimation of SNP-specific pleiotropy by shrinkage, result in a powerful approach, although results are sensitive to our prior believes on the proportion of pleiotropic SNPs among our instruments. We illustrate and compare these different approaches using an MR study on the effect of age at menarche on adult lung function.

Keywords:
mendelian randomization, pleiotropy, bayesian model averaging, bayesian spike and slab model, mr-egger
Robust methods for Mendelian randomization: towards more reliable causal inferences

Presenter: Dr Stephen Burgess
MRC Biostatistics Unit / Cardiovascular Epidemiology Unit, University of Cambridge. Email: sb452@medschl.cam.ac.uk

Co-authors: Stephen Burgess [1], Jessica Rees [1], Jack Bowden [2]

Abstract
Mendelian randomization is a technique for assessing the causal role of a modifiable risk factor on a disease outcome using genetic data. Recent advances in genome-wide association studies and the increasing availability of publicly available summary data on associations of genetic variants with risk factors and disease outcomes in large sample sizes have enabled powerful Mendelian randomization analyses to be performed relatively quickly and simply. However, these analyses typically use a large number of genetic variants. When at least one of the genetic variants does not satisfy the instrumental variable assumptions, causal estimates will be biased and Type 1 error rates inflated. Novel methods are presented for obtaining causal inferences from summarized data under weaker assumptions that those of a typical Mendelian randomization investigation, including MR-Egger regression, a median-based approach, and approaches for dealing with outlying genetic variants. The talk is illustrated using the example of HDL-cholesterol on coronary artery disease risk: a naive analysis including all genome-wide significant variants suggests a protective effect of HDL-cholesterol that is not supported by the biological evidence, whereas the MR-Egger and weighted median approaches suggest a null causal effect.

Keywords:
robust methods, summarized data, statistical methodology
Session 10: Applications in cardiovascular disease and diabetes
Chair: Professor Nishi Chaturvedi, MRC Unit for Lifelong Health and Ageing, UCL, UK
Lactase persistence, milk intake, and ischemic heart disease, type 2 diabetes, hip fracture, and total mortality in the Danish general population

Presenter: Christina Ellervik
University of Copenhagen, Denmark/Harvard Medical School, USA. Email: christina@ellervik.dk

Co-authors: Helle KM Bergholdt[1,2], Børge G Nordestgaard[2,3,4], Anette Varbo[2,3], Christina Ellervik[2,5,6]

[1]Department of Clinical Biochemistry, Naestved Hospital, Denmark. [2]Faculty of Health and Medical Sciences, University of Copenhagen, Denmark. [3]Department of Clinical Biochemistry and the Copenhagen General Population Study, Herlev and Gentofte Hospital, Copenhagen University Hospital, Denmark. [4]The Copenhagen City Heart Study, Frederiksberg Hospital, Copenhagen University Hospital, Denmark. [5]Department of Research, Nykoebing Falster Hospital, Nykoebing Falster, Denmark. [6]Department of Laboratory Medicine, Boston Children’s Hospital/Harvard Medical School, Boston, MA, USA.

Abstract

The genetic variant LCT-13910 C/T is associated with downregulation of the lactase enzyme activity in adults of European descent, referred to as lactase non-persistence (genotype CC) and lactase persistence (genotypes TC and TT), and affects the ability of adults to digest the lactose in milk. Previous meta-analyses of population based studies have shown a reduced risk of ischemic heart disease and type 2 diabetes in individuals consuming dairy products vs. those who do not. The association of milk intake with fractures have shown contradictory results.

We investigated the association between lactase persistence, milk intake and

1) risk of ischemic heart disease
2) type 2 diabetes
3) hip fracture
4) total mortality

in three general population studies of approximately 100,000 adult Danes. We used the genetic variant LCT-13910 C/T (rs4988235) as proxy/surrogate for long-term milk intake in Mendelian randomization studies to assess indirectly whether there may be a causal association between milk intake and risk of the endpoints.

We found no evidence of association between milk intake and risk of ischemic heart disease, type 2 diabetes, hip fracture or total mortality observationally or genetically in the general Danish population.

We discuss the results from a Mendelian randomization perspective.

Keywords:
lactase persistence, milk intake, ischemic heart disease, type 2 diabetes, hip fracture

Presenter: Linda Vissers
Julius Center for Health Sciences and Primary Care, UMC Utrecht. Email: l.e.t.vissers-13@umcutrecht.nl

Co-authors: Linda E.T. Vissers [1], Ivonne Sluijs [1], Fumiaki Imamura [2], Nita G. Forouhi [2], Yvonne T. van der Schouw [1], on behalf of the InterAct consortium
[1] Julius Center for Health Sciences and Primary Care, UMC Utrecht, the Netherlands
[2] Medical Research Council Epidemiology Unit, Cambridge, United Kingdom

Abstract
Background – High intake of dairy products, particularly yogurt, has been associated with lower risk of diabetes in observational studies, although not consistently. Genetic lactase persistence (LP) enables digestion of dairy sugar (lactose).
Purpose – To investigate if dairy intake is causally related to incident diabetes. Methods – We used data from 21,900 individuals from 8 European countries of the European Prospective Investigation into Cancer and Nutrition (EPIC)-InterAct case-cohort study, including 9,742 incident diabetes cases. We used the CardioMetabochip or Illumina 660W quad chip for genotyping and imputed the SNP of interest, rs4988235. Baseline dietary intakes were assessed with food frequency questionnaires. We assessed the association between genetic LP and dairy intake and confounders of the relationship between dairy and diabetes. IV estimates were obtained through 2SLS, using a Prentice-weighed Cox regression as second stage.
Results – Homozygous LP was present in 36.5% of the cohort, ranging from 5.0% in Florence, Italy to 55.7% in Malmö, Sweden. Every additional LP allele was associated with a 25.5 (95%CI: 16.9, 34.2) g/day higher milk intake, but not with intake of non-milk dairy products. The LP SNP was also associated with intake of fruit (-8.0 g/day, 95%CI: -15.4,-0.64), tea (-13.1 g/day, 95%CI: -26.6, 0.34) and wine (-8.6 g/day, 95%CI: -14.8, -2.5). The IV analysis showed that milk intake was not related to diabetes with an HR of 0.99 per 25 mg/day (95%CI: 0.92,1.06).
Conclusion – This MR analysis suggests that milk intake is not causally related with diabetes, in line with observational evidence. The LP SNP used in this study had a modest magnitude of association with lower intake of fruit, tea and wine, which highlights the challenges in interpretation in nutritional epidemiology, but we consider it unlikely that this has caused the observed null result.

Keywords:
diabetes, milk, dairy products, nutrition
High-sensitivity c-reactive protein, low-grade systemic inflammation and type 2 diabetes mellitus: A two-sample Mendelian randomization study

Presenter: Dr. Raymond Noordam
Leiden University Medical Center. Email: r.noordam@lumc.nl

Co-authors: Raymond Noordam [1], Charlotte H Oudt [1], Maxime M Bos [1], Roelof AJ Smit [1,2], Diana van Heemst [1]
1 Department of Internal Medicine, Section Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands. 2 Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands.

Abstract
The role of inflammation in the development of type 2 diabetes mellitus (T2D) remains unclear. We investigated the associations of high sensitivity c-reactive protein (hsCRP) with T2D and glycemic traits using two-sample Mendelian Randomization. For this, we used publically available summary-statistics data from genome-wide association studies on T2D (DIAGRAM: 12,171 cases; 56,862 controls) and glycemic traits (e.g., glucose, insulin; MAGIC: 46,186 participants without diabetes mellitus). We combined the effects of the genetic instrumental variables through inverse-variance weighting (IVW), and MR-Egger regression and weighted-median estimation as sensitivity analyses which take into account potential violations of the assumptions of instrumental variable analyses. Analyses were conducted using 15 known hsCRP genetic instruments and 6 hsCRP genetic instruments not involved in inflammatory processes beyond hsCRP concentration regulation. With IVW, we found no evidence for a causal association between the combined effect of the genetic instrumental variables for hsCRP and T2D (odds ratio per 1 ln[hsCRP in mg/L]: 1.15; 95% confidence interval: 0.93, 1.42). However, robust associations were observed with MR-Egger regression and weighted-median estimation (odds ratio with 95% confidence interval per 1 ln[hsCRP in mg/L], MR-Egger regression: 1.29; 1.08, 1.49; weighted-median estimation: 1.21; 1.02, 1.39). We found no association with T2D for the combination of hsCRP-specific genetic instruments nor did we found associations with glycemic traits in any of the analyses. Therefore, elevated low-grade inflammation, as reflected by a high hsCRP concentration in blood, but not hsCRP itself, is likely a causal risk factor for T2D, but it remains unclear for glycemic traits.

Keywords:
c-reactive protein, inflammation, type 2 diabetes mellitus, two-sample mendelian randomization
Assessing the causal role of body mass index on cardiovascular health in young adults: Mendelian randomization and recall-by-genotype analyses

Presenter: Kaitlin Wade
*Integrative Epidemiology Unit. Email: kaitlin.wade@bristol.ac.uk*

Co-authors: Kaitlin H. Wade[1,2], Scott T. Chiesa[3], Alun D. Hughes[4], Nish Chaturvedi[4], Marietta Charakida[3], Alicja Rapala[3], Vivek Muthurangu[3], Tauseef Khan[3], Nicholas Finer[3], Naveed Sattar[5], Laura D. Howe[1,2], Abigail Fraser[1,2], Debbie A. Lawlor[1,2], George Davey Smith[1,2], John E. Deanfield[3], Nicholas J. Timpson[1,2]

1 MRC Integrative Epidemiology Unit (IEU) at the University of Bristol, Bristol, UK
2 School of Social and Community Medicine, Faculty of Health Sciences, University of Bristol, Bristol, BS8 2BN
3 Vascular Physiology Unit, Institute of Cardiovascular Science, University College London, 1 St. Martin’s Le Grand, London EC1A 4NP
4 Cardiometabolic Phenotyping Group, Institute of Cardiovascular Science, University College London, London, W1T 7HA
5 Institute of Cardiovascular and Medical Sciences, BHF Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow, G12 8TA

Abstract

Background: Mendelian randomization (MR) studies of body mass index (BMI) and cardiovascular health in late life suggest causal relationships, but the nature of these has not been explored in younger ages.

Objectives: Using MR and recall-by-genotype (RbG) methodologies, we aimed to estimate the causal effect of BMI on detailed measures of cardiovascular health in a population of young adults.

Methods: Data from the Avon Longitudinal Study of Parents and Children were used. For MR analyses, a genetic risk score (GRS) comprising 97 independent genetic variants (constructed using external weights) was used to test the causal effect of each unit increase in BMI (kg/m2) on selected cardiovascular phenotypes measured at age 17 (N=7909). An independent sample from the same cohort participated in a RbG study at age 21, which enabled more detailed cardiovascular phenotyping (N=418; 191/227 from the lower/upper ~30% of a genome-wide GRS distribution).

Results: Difference in mean BMI between RbG groups was 3.58kg/m2 (95% CI: 2.53, 4.63; P=6.09x10^-11). In both MR and RbG analyses, results indicated that higher BMI causes higher blood pressure (BP) and left ventricular mass (indexed to height2.7, LVMI) in young adults (e.g. difference in LVMI per kg/m2 using MR: 1.07g/m2.7; 95% CI: 0.62, 1.52; P=3.87x10^-06 and per 3.58kg/m2 using RbG: 1.65g/m2.7 95% CI: 0.83, 2.47; P=0.0001). Additionally, RbG results indicated a causal role of higher BMI on higher stroke volume (difference per 3.58kg/m2: 1.49ml/m2.04; 95% CI: 0.62, 2.35; P=0.001) and cardiac output (difference per 3.58kg/m2: 0.11l/min/m1.83; 95% CI: 0.03, 0.19; P=0.01). Neither analysis supported a causal role of higher BMI on heart rate.

Conclusions: Complementary MR and RbG causal methodologies, together with a range of appropriate sensitivity analyses, showed that higher BMI is likely to cause worse cardiovascular health even in youth. These consistent results support efforts to prevent or reverse obesity in the young.

Keywords: body mass index, cardiovascular traits, mendelian randomization, recall-by-genotype, causality
Helicobacter pylori identified as a modulator of the onset of Type 2 Diabetes

Presenter: Danielle Rasooly
Department of Biomedical Informatics, Harvard Medical School, Boston, MA 02115. Email: drasooly@g.harvard.edu

Co-authors: Adam S Brown [1], Chirag J Patel [1]
[1] Department of Biomedical Informatics, Harvard Medical School, Boston, MA, 02115

Abstract
Recent studies have implicated a number of infectious diseases in the etiology and onset of Type 2 Diabetes (T2D). However, these studies are limited to candidate infections and lack comprehensive assessment of the full spectrum of infectious diseases. To address this unmet need, we performed an observational investigation, “Infection-Wide Association Study” (IWAS), on a case-control cohort from a proprietary insurance claims dataset to identify infectious diseases that increase an individual’s likelihood of developing T2D. We examined over 600 diagnosis codes for history of infectious disease, and adjusted for patient age, gender, obesity, family history of diabetes, high blood pressure, high cholesterol, Vitamin D deficiency, fasting glucose, and cardiovascular disease. Using this methodology, we identified Helicobacter pylori (H. pylori) to be associated with T2D onset at an odds ratio (OR) of 1.58 (N=46k diabetics and 80k non-diabetics) after correcting for multiple hypothesis testing and replicated the observational findings in an independent electronic medical record dataset (OR = 2.69, N=29.4K diabetics and 28K non-diabetics). An antibiotic prescribed to H. pylori patients was associated with decreased risk of T2D (OR = 0.84, N=7k diabetics and 173k non-diabetics). Of course, all associations are non-causal, and potentially biased and/or false positive. In this talk, we describe use of Mendelian Randomization (MR) to causally associate H. pylori in T2D risk. Specifically, we have obtained H. pylori GWAS summary statistics [1] and are investigating the causal relationship between H. pylori seroprevalence and T2D and risk factors for T2D, including HbA1C, fasting glucose, and fasting insulin. Our study is among the first that combines a data-driven observational study on the association of infectious disease in T2D with a MR estimate of causal effects of H pylori, if any.

Keywords:
helicobacter pylori, type 2 diabetes, infection-wide association study, iwas
Session 11: Methodology
Chair: Dr Jack Bowden, MRC Integrative Epidemiology Unit, UK
Mendelian randomisation via the zero modal pleiotropy assumption: the mode-based estimate

Presenter: MSc. Fernando Pires Hartwig
Postgraduate Programme in Epidemiology, Federal University of Pelotas, Pelotas, 96020220 Brazil. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN, UK. Email: fernandophartwig@gmail.com

Co-authors: George Davey Smith[1,2], Jack Bowden[1,2]
1 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN, UK.
2 School of Social and Community Medicine, University of Bristol, BS8 2BN, UK.

Abstract
Mendelian randomization (MR) is being increasingly used to strengthen causal inference in observational studies. Availability of summary data of genetic associations for a variety of phenotypes from large genome-wide association studies (GWAS) allows straightforward application of MR using summary data methods, typically in a two-sample design. In addition to the conventional inverse variance weighting (IVW) method, recently developed summary data MR methods, such as the MR-Egger and weighted median approaches, allow a relaxation of the instrumental variable assumptions. In this talk, the recently developed method to obtain a single causal effect estimate from multiple genetic instruments, the mode-based estimate (MBE) is presented. The MBE exploits the ZERo Modal Pleiotropy Assumption (ZEMPA), which holds if the largest number of similar (identical in infinite samples) individual-instrument causal effect estimates comes from valid instruments. ZEMPA therefore exploits heterogeneity of horizontal pleiotropic effects to perform consistent causal effect estimation even if the majority of instruments are invalid. The MBE has less bias and better coverage than other established methods under the causal null in many simulation scenarios with different patterns of horizontal pleiotropy. In absence of heterogeneity, the MBE is less powered than the IVW and weighted median methods, but more powered than the MR-Egger regression. The MBE relaxes the instrumental variable assumptions, and should be used in combination with other approaches in a sensitivity analysis framework.

Keywords:
causality, instrumental variables, genetic variation, mendelian randomization, genetic pleiotropy.
Simulating survival bias in Mendelian randomization

Presenter: Drs. Roelof Smit, PhD student
Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands. Email: r.a.j.smit@lumc.nl

Co-authors: Roelof AJ Smit[1,2], Stella Trompet[1,2], J. Wouter Jukema[1], Saskia le Cessie[3]
1. Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands
2. Department of Internal Medicine, section of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands
3. Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands

Abstract
Introduction
The number of studies utilizing genetic variants as instrumental variables is rapidly increasing. While it has previously been theorized that collider stratification bias may distort results from Mendelian randomization (MR) analyses performed in older populations, this has received little attention in the literature. We aimed to quantify the effects of this possible threat to MR-studies using odds ratios typically found in cardiovascular epidemiology.

Methods
Using Monte Carlo simulations, we assessed the extent to which the association between a genetic instrument (G) and both dichotomous and continuous outcomes of interest (Y) are biased when the analysis is restricted to individuals having survived until a certain age. We illustrate the bias for a range of scenarios, incorporating varying degrees of i. age-at-inclusion, ii. phenotypic variance explained by the genetic instrument, and iii. odds ratios for uncorrelated risk factors. For all scenarios, simulations were repeated 1,000 times using 10,000 randomly generated observations.

Results
Preliminary results show that restricting the analyses to survivors leads to biased estimation of the causal effect of G on Y, with increasing age-at-inclusion, explained variance, and greater odds ratios leading to ever greater bias and decreases in statistical power. In addition, when compared to the G-Y association, use of two-stage least-squares (2SLS) models worsens this burden.

Conclusion
Using a simple framework, we were able to show that collider-stratification bias may seriously distort results from MR-studies performed in the oldest old. We describe considerations how to assess the presence of collider stratification bias in MR-studies. Given that these mechanisms should hold true for any sufficiently selected population, our findings support the position that a critical appraisal of possible collider stratification bias is warranted when considering populations for MR-analyses.

Keywords:
survival bias, monte carlo simulations
Selection bias in instrumental variable analyses

Presenter: Dr Rachael A. Hughes
School of Social and Community Medicine, University of Bristol. Email: rachael.hughes@bristol.ac.uk

Co-authors: Neil M. Davies [2], Kate Tilling [1,2]
1 School of Social and Community Medicine, University of Bristol, BS8 2PS
2 MRC Integrative Epidemiology Unit, University of Bristol, BS8 2BN

Abstract
Background: Study participants are rarely a true random sample of the population they are intended to represent, and both known and unknown factors can influence selection of participants. Failure to account for selection in an instrumental variable (IV) analysis, such as a Mendelian randomisation analysis, may lead to bias. The circumstances in which selection bias occurs are not well understood.

Aims: To review the circumstances in which a two stage least squares (2SLS) IV analysis is biased by selection, and to illustrate the effects of selection bias.

Methods: We use directed acyclic graphs (DAGs) to depict assumptions about selection and show how DAGs can be used to determine when selection bias occurs. Using simulations, we assess the magnitude of the selection bias of the 2SLS estimate, and coverage of its 95% confidence interval (CI) for a range of selection scenarios. We repeat the simulation study for (1) weak and strong instruments (partial R2 0.04 and 0.39 respectively), (2) linear and non-linear treatment-instrument associations, and (3) causal and non-causal treatments.

Results: The 2SLS estimate is unbiased and CI coverage is close to 95% when selection does not depend on any factors, and when selection only depends on the instrument. However, the 2SLS estimate is biased with poor CI coverage if selection: 1) depends on treatment and/or outcome, 2) depends on treatment and instrument, 3) depends on instrument and confounder (of the outcome-treatment association), and 4) depends on treatment and confounder. The effects of selection bias are larger for weak instruments than strong instruments (e.g. downward bias of 0.154 and 0.047 for a weak and strong instrument respectively). For some selection scenarios, the 2SLS estimate is unbiased if the treatment does not affect the outcome.

Conclusion: Selection bias can have a major effect on an IV analysis. Statistical methods are needed for estimating causal effects from non-random samples.

Keywords:
instrumental variable analysis, selection bias, simulation study, two stage least squares estimator
Survival bias in Mendelian Randomisation studies: Does the magnitude matter?

Presenter: Dr Mats Julius Stensrud  
*Department of Biostatistics, University of Oslo. Email: m.j.stensrud@medisin.uio.no*

Co-authors: Mats Julius Stensrud [1], Morten Valberg [1], Susanne Strohmaier [2], Kjetil Roysland[1] and Odd Aalen [1]  
1 *Department of Biostatistics, Institute of Basic Medical Sciences, University of Oslo*  
2 *Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, and Harvard Medical School, Boston, USA*

**Abstract**  
Survival bias is a potential problem when subjects are lost to follow-up, and this selection issue may arise in a wide range of biomedical studies. Controlling for the bias is difficult because subjects may be lost due to unmeasured factors. Mendelian Randomisation (MR) studies may be particularly prone to survival bias: MR studies rely on genetic variants that are carried from conception, but subjects are often included into these studies in late adulthood.

We will explore two results that can suffer from survival bias: 1) The association between Low Density Lipoprotein (LDL) cholesterol and all-cause mortality in the elderly, and 2) The association between alcohol consumption and all-cause mortality in the elderly. Then, we will suggest a method to handle the bias. In particular, we will present a novel survival analysis approach based on the Cox Proportional hazards model. To adjust for the survival bias, we will use data on the familial clustering of diseases, which e.g. is available from Scandinavian twin registries.

The suggested approach is explored in a brief simulation study, suggesting that sensible estimates are obtained. Subsequently, the approach is applied to the two data examples. For the effect of LDL variants on all-cause mortality, an estimated hazard ratio of 1.13 [1.06, 1.44] may be adjusted to 1.87 [1.16, 4.15]. For the effect of the Alcohol Dehydrogenase 1B (ADH1B) gene, a reported hazard ratio of 0.68 [0.54,0.87] may be adjusted to 0.52 [0.37, 0.77].

Our analysis suggests that loss to follow-up may bias MR studies towards the null-hypothesis. In fact, we derive that an unadjusted hazard ratio may not only be a biased estimate, but it may sometimes be invalid for hypothesis testing. This result should be communicated to applied researchers interested in MR.

**Keywords:**  
methods, hazard ratio, causality, cox proportional hazards, frailty
Potential biases in MR studies: Considerations when exposures vary over the life-course

Presenter: Jeremy Labrecque
_Department of Epidemiology, Erasmus Medical Center. Email: j.labrecque@erasmusmc.nl_

Co-authors: Jeremy A Labrecque [1], Sonja A Swanson [1]
_1 Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands_

Abstract
Objectives: A core assumption of Mendelian randomization (MR) is the exclusion restriction. When exposures vary over the life-course, the value of the exposure when it is assessed is not usually equal to the value of the exposure at all points in time. When exposure history prior to assessment can also affect the outcome (i.e., violating the exclusion restriction), the MR estimate will be biased for both the effect of exposure at the assessed time and the effect of full exposure history. We explore the former potential bias with simulations.

Methods: We simulated data to mimic an MR study using FTO variants as the instrument, body mass index (BMI) in mid-life as the exposure and a generic outcome. We also simulated BMI in early adulthood. Parameters were informed by the literatures on the relationship between FTO and BMI and the longitudinal auto-correlation of BMI. We simulated two scenarios – (i) a direct effect of prior BMI on the outcome and (ii) selection into the study related to prior BMI – and varied the bias parameters in each of these scenarios. We used a simulated population size of 30,000 and each simulation was run 10,000 times.

Results: When early adulthood BMI had a direct effect of increasing the outcome by 1 case per 100, the MR estimate was biased by 0.30 cases per 100 (interquartile range [IQR]: -0.04, 0.66). In our most extreme selection bias scenario, when 27% of the data were excluded, the MR estimate was biased by 0.44 cases per 100. In both scenarios, 95% coverage probabilities were also reduced.

Conclusion: Violations of the exclusion restriction through the effect of or selection on prior exposure history can be an important source of bias in MR studies. Careful consideration of whether exposure previous to assessment can directly affect the outcome and whether selection on previous exposure has occurred should be included in MR studies.

Keywords: exclusion restriction, time-varying exposures, simulations, biases
Session 12: Applications of lipid profiling
Chair: Professor Debbie Lawlor, MRC Integrative Epidemiology Unit, UK
Metabolomic profiling and Mendelian randomisation of lipid-lowering targets: comparison of statins, ezetimibe and PCSK9-inhibition in 72,000 individuals

Presenter: Peter Würtz
*University of Helsinki & Nightingale Health Ltd. Email: peter.wurtz@helsinki.if*

**Abstract**

Background: Metabolomic profiling of statin therapy has been shown to yield insights into the drug mechanism and potential side-effects (*J Am Coll Cardiol* 2016;67:1200). This concept study demonstrates that combining metabolomics with genetic proxies can provide molecular insights into effects of known and novel drug targets.

Aim: Statins are first-line therapy for lowering cardiovascular risk, but also other LDL lowering drugs are approved, in particular ezetimibe and PCSK9-inhibitors. We assessed the metabolomic profile of these 3 drugs via genetic proxies. The results were compared with the detailed metabolic changes seen in a large statin trial.

Methods: 228 blood biomarkers, including lipoprotein subclasses, fatty acids, and amino acids, were quantified by NMR metabolomics for 72,225 individuals. The metabolite association of PCSK9 and NPC1L1 (target of ezetimibe) were compared to the corresponding associations for genetic variants in HMGCR (target of statins) and statin therapy.

Results: For the same lowering in LDL cholesterol, the overall metabolic associations with variation in PCSK9 and HMGCR were highly reminiscent and matched the randomized effect of statin therapy. Statin therapy was more efficacious at lowering cholesterol in VLDL than the predicted effect of PCSK9 inhibition. Association patterns were similar for alterations in lipoprotein lipid composition and fatty acid balance, and null associations across non-lipid metabolites. Genetic proxies in NPC1L1 displayed more prominent metabolic differences as compared with statin therapy.

Conclusions: Genetic variation in PCSK9 results in similar detailed metabolic effects as statin therapy; however, for a fixed lowering of LDL cholesterol, PCSK9-inhibition is predicted to be less efficacious at lowering triglyceride-rich lipoproteins. The study exemplifies how metabolomic characterization of genetic proxies for drug targets can inform on mechanisms and off-target effects.

**Keywords:**

cholesterol-lowering drugs, metabolomics, phenotypic screen, off-target effects, target validation
PCSK9 genetic variants, metabolomics and cardiovascular risk factors and disease in the China Kadoorie Biobank

Presenter: Michael V Holmes  
*University of Oxford. Email: michael.holmes@ndph.ox.ac.uk*

Co-authors: Michael V Holmes[1,2,3], Ruth Boxall[1], Kuang Lin[2], Derrick A Bennett[2], Iona Y Millwood[1,2], Robert J Clarke[2], Robin G Walters[2] and Zhengming Chen[2] on behalf of CKB working group

1 Medical Research Council Population Health Research Unit at the University of Oxford  
2 Clinical Trial Service Unit & Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, BHF Centre of Research Excellence, Big Data Institute Building, Roosevelt Drive, University of Oxford, UK  
3 National Institute for Health Research Oxford Biomedical Research Centre, Oxford University Hospital, Old Road, Oxford OX3 7LE, UK

Abstract

Objective: PCSK9 inhibition is a new therapeutic approach to lowering LDL-cholesterol (LDL-C) and a recent phase III clinical trial showed efficacy for cardiovascular disease. We assessed the association of PCSK9 genetic variants with metabolomics and risk of cardiovascular risk factors and end-points in the China Kadoorie Biobank (CKB).

Methods: 17,500 individuals with directly-quantified LDL-C were genotyped using a bespoke GWAS array, and associated SNPs at the PCSK9 locus were selected using FINEMAP. A genetic score, with internal weights (with 100x cross-validation) according to their effects on LDL-C, was used in 105,000 individuals with GWAS (including 4,500 with NMR metabolomics and 22,500 with ultrasound-quantified measures of subclinical atherosclerosis).

Results: 3 SNPs within the PCSK9 locus (rs151193009, rs471705 and rs505151) independently associated with LDL-C. In a multivariable model, their respective per-allele (SE) associations with standardized LDL-C were: 0.64 (0.05); 0.10 (0.01) and 0.16 (0.02). The PCSK9 genetic instrument associated with lower particle and cholesterol concentrations of intermediate and low density lipoprotein subclasses. Scaled to a 1-SD lower LDL-C (equivalent to 0.71 mmol/l), the PCSK9 genetic instrument associated with lower cIMT (-0.25 SDs; 95%CI: -0.36, -0.14) and lower risk of plaque (odds ratio [OR] 0.61; 0.46, 0.79). A 1-SD lower LDL-C from the PCSK9 genetic instrument associated with a lower risk of major occlusive vascular events (15,670 cases; OR 0.81; 95%CI: 0.69, 0.95) and, examining the individual components, there was a directionally concordant trend for major coronary events (4949 cases; 0.78; 0.60, 1.01) and ischaemic stroke (11,502 cases; 0.84; 0.70, 1.00). There was no association with haemorrhagic stroke (6,197 cases; 1.04; 0.83, 1.30)

Conclusions: Genetic inhibition of PCSK9 was associated with less subclinical atherosclerosis and a lower risk of cardiovascular disease in Chinese adults.

Keywords: pcsk9, mendelian randomization, ischaemic stroke, haemorrhagic stroke
Exploressing the causal relationships between metabolic measures from a targeted NMR approach

Presenter: Dr Fotios Drenos
MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Oakfield House, Oakfield Grove, BS8 2BN, Bristol, UK. Email: f.drenos@bristol.ac.uk

Abstract
Metabolomics is the study of biological systems at the level of their chemical processes and compounds involved in them. The study of the correlations between the metabolic measures can provide valuable information for the organisation of the system, the flow of information and its relevance to disease. As the application of metabolomics becomes more widespread in epidemiology, our ability to explore these associations increases. Using bidirectional Mendelian randomisation on 230 metabolic measures (with 78 of them ratios) from a targeted NMR metabolomics platform in 5353 children and 4120 mothers from the ALSPAC study, we explore the correlations between the metabolic measures in order to further understand the underlying mechanisms of their relationships. We identified a 20758 phenotypic correlations between the metabolic measures which were mostly also reflected at the genetic level. Mendelian randomisation revealed a 1976 replicated bidirectional associations and 204 single direction causal effects including known and novel pathways. Evidence for pleiotropy were present in a number of the associations. The results, in addition to their biological significance concerning the probable underlying mechanisms explaining the observed associations, provide a clear illustration of the challenges for the use of Mendelian randomisation on metabolomics data. We will also present the application of computational methods that can address some of the limitation evident in this work and provide a way towards computational methods more suitable for use in metabolomics data.

Keywords:
metabolomics, mendelian randomisation, computational methods
Characterizing gene-specific associations of LDL cholesterol with type 2 diabetes using untargeted metabolomics

Presenter: Miss Chen Li  
*MRC Epidemiology Unit, University of Cambridge. Email: cl645@mrc-epid.cam.ac.uk*

Co-authors: Chen S Li [1], Luca A Lotta [2], Isobel D Stewart [3], Robert A Scott [4], Nicholas J Wareham [5], Claudia Langenberg [6]  
1. *MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge, Cambridge, CB2 0QQ, United Kingdom*

Abstract
Low-density lipoprotein cholesterol (LDL-C) levels have been suggested to be causally linked to an increased risk of type 2 diabetes (T2D), with locus-specific effects. We aim to systematically investigate a) differences in associations between established LDL-C lowering genetic variants and T2D risk and b) to characterise their metabolomic and lipidomic profiles. We selected lead SNPs of 58 loci previously associated with lower LDL-C in genome-wide association studies, and investigated their associations with T2D risk using 2-sample inverse variance weighted Mendelian randomization (MR) methods. Metabolomic profiles were measured using Metabolon’s HD4 discovery platform in up to 7,496 individuals in the EPIC-Norfolk study. As a result, LDL-C lowering alleles were jointly associated with T2D risk (OR=1.14 [95%CI, 1.04-1.26]). A total of 14 of the 58 loci showed nominally significant associations with T2D, with high heterogeneity (I^2 = 76.6%), and effect sizes and directions ranging from 0.21 (HNF1A, p-value = 1.06X10^-9) to 3.72 (LOC84931, p-value = 0.002) per 1-SD genetically predicted lower levels of LDL-C. Metabolomic characterisation showed a specific signature for the 14 T2D-associated LDL-C loci, with significant associations across different (lyso)phospholipids. For a given LDL-C lowering effect, the alleles of APOC1 and FADS1 showed opposite effects on both T2D (OR = 1.35 [95% CI, 1.19-1.53] versus 0.53 [0.37-0.76]) and 7 different (lyso)phospholipid species, the majority (6/7) of which carried oleoyl acid (18:1) or linoleoyl acid (18:2) at the sn-2 position. In conclusion, Genetic predisposition to lower LDL-C was associated with a higher risk of T2D, with substantial heterogeneity between loci. Metabolomic analyses highlight lipid species that are associated with genetic predisposition to both, increased T2D risk and lower LDL-C, and potential which may partially mediate the observed heterogeneity of genetic associations.

Keywords:  
LDL cholesterol, type 2 diabetes, metabolomics, lipidomics
Identifying genetic determinants of the human plasma proteome to facilitate two-sample Mendelian randomisation

Presenter: Mr Benjamin Sun
University of Cambridge. Email: bs446@medschl.cam.ac.uk

Co-authors: Benjamin B Sun [1], Joseph C Maranville [2], James E Peters [1], Caroline S Fox [2], Robert M Plenge [2], John Danesh [1], Heiko Runz [2], Adam S Butterworth [1]
[1] MRC/BHF Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK.

Abstract
Mendelian randomisation (MR) has been most successful when applied to soluble proteins. For example, previous studies have used MR approaches to help prioritise some proteins (eg, interleukin-6, lipoprotein[a]) and deprioritise others (eg, C-reactive protein, lipoprotein-associated phospholipase A2) as potential therapeutic targets for cardiovascular disease. However, the generalisation of the approach has been constrained by the lack of suitable genetic instruments indexing a wide variety of proteins.

To address this we conducted genome-wide association studies of 10.6 million variants against levels of 2,994 plasma proteins measured in 3,301 healthy blood donors from the INTERVAL study. We identified 1,927 associations for 1,478 proteins, representing a 4-fold increase on existing knowledge. We describe the genetic architecture of these proteins and the potential for the derivation of allele scores for two-sample MR. We show how this catalogue of genetic associations can be used to inform causal roles for protein biomarkers in disease and to identify the likely mediating pathways underpinning known GWAS loci. Our data provides a valuable toolkit of genetic instruments to allow the wider scientific community to elucidate further causal insights into human disease and to identify therapeutic targets.

Keywords:
genomics, proteomics, mendelian randomisation, protein qualitative trait loci.
Session 13: Applications to biomarker and target identification
Chair: Professor Caroline Relton, MRC Integrative Epidemiology Unit, UK
Use of Mendelian Randomization to Validate Human Genetic Targets for Alzheimer’s Disease

Presenter: Dr Karol Estrada
Biogen. Email: karol.estrada@biogen.com

Co-authors: Karol Estrada[1], KD Nguyen[1,1], Paola Bronson[1], Brian Wipke[2], Ayla Ergun[1], Richard M. Ransohoff[2], Will Chen[1], Michelle Penny[1], Sally John[1]
1Computational Biology & Genomics, Biogen, Boston MA. 2 Neuroinflammation/Acute Neurology, Biogen Boston MA

Abstract
Introduction. Alzheimer’s disease (AD), the most common neurodegenerative condition, has no interventions to slow progression. Genome-wide Association Studies (GWAS) on AD have identified >22 associated loci including a genetic variant in CD33 that is also associated with splicing of CD33 exon 2. We aimed to validate the human genetic evidence of CD33 as a therapeutic target for AD.

Methods. We used publicly available IGAP summary statistics (Lambert, 2013) to evaluate the effect of a CD33 genetic variant on AD risk. To identify potential novel AD variants, we analyzed whole exome sequencing data of 4,186 AD cases and 2,772 controls from ADSP. We obtained eQTL summary stats on 922 RNA-seq blood samples (Battle, 2014). We used protein QTL summary statistics from 1,332 plasma samples (Suhre, 2017) to evaluate exposure. Finally, we used a Wald Ratio test implemented in TwoSampleMR to estimate the potential therapeutic effect of targeting CD33 for AD.

Results. A genetic variant (rs3865444) near CD33 was associated with AD in the IGAP meta-analysis (OR=0.94, P=3x10^{-6}). In the exome data we tested 6 coding variants and did not find significant association with AD. We confirmed the exon 2 exclusion associated with the rs3865444_A variant in the whole-blood samples (P=1x10^{-222}). In the plasma results, rs3865444 was significantly associated with lower CD33 (β= -1.028, P=2x10^{-52}. The Mendelian Randomization analysis using rs3865444 as an instrumental variable to evaluate the impact on AD via modulation of CD33 was significant (P=3.1x10^{-6}), however, it predicted a change of only OR=0.94 [95% CI 0.91-0.96] per each decrease of one standard deviation in CD33 levels.

Conclusion: MR predicts that reducing CD33 levels substantially will not achieve a sufficiently large therapeutic effect in a clinical study on AD patients or prodromal AD patients.

Keywords:
alzheimer’s, genetics, cd33, eQTL, pQTL
Improving biomarker prediction by circumventing the component of variance explained by genotype: Biomarker De-Mendelization: proof of principle

Presenter: Dr Alisa Kjaergaard  
Department of Clinical Biochemistry, Aarhus University Hospital, Denmark. Email: alisa@devedzic.dk


1 Department of Clinical Biochemistry, Aarhus University Hospital, Denmark  
2 Department of Clinical Biochemistry, Herlev and Genterfte Hospital, Denmark; The Copenhagen City Heart Study; The Copenhagen General Population Study; University of Copenhagen, Denmark  
3 Departments of Oncology and Medicine, Herlev and Genterfte Hospital; University of Copenhagen, Denmark  
4 MRC Integrative Epidemiology Unit, University of Bristol, United Kingdom

Abstract

Background: In observational studies, the Mendelian randomization approach can be used in order to circumvent confounding, bias and reverse causation, and to assess a potential causal association between a biomarker and risk of disease. If, on the other hand, a substantial component of variance in a non-causal biomarker is explained by genotype, then genotype could potentially attenuate the observational association and the strength of a predictive biomarker. In order to circumvent the component of variance explained by genotype, an approach that can be seen as the inverse of Mendelian randomization seems plausible. A good candidate for a proposed de-Mendelization proof of principle is a biomarker with strong genotype-phenotype correlation, and where only phenotype (biomarker concentration) and not genotype is associated with risk of disease. Plasma YKL-40 is such a biomarker. A promoter SNP rs4950928 in the gene for YKL-40 explains 14% of the variation, and genotype is associated with doubling and tripling in plasma YKL-40 concentrations.

Methods: Cohort studies in 21,165 individuals form the Danish general population with measurements of both plasma YKL-40 concentration and rs4950928 genotype. Endpoints alcoholic liver cirrhosis and lung cancer were chosen due to their very different magnitude of association with plasma YKL-40 concentrations. Several different methods for biomarker de-Mendelization are explored.

Results: Compared to usual methods, all de-Mendelization methods improved observational associations slightly. However, the areas under the receiver operating characteristics curves hardly differed between the usual and de-Mendelization methods.

Conclusions: Even when genotype explains 14% of the variance in a biomarker, we found no useful empirical improvement in risk prediction by biomarker de-Mendelization.

Keywords: ykl-40, biomarker de-mendelization, risk prediction, general population, proof of principle
Mendelian Randomization for disease progression: opportunities and methodological challenges

Presenter: Dr Lavinia Paternoster
MRC Integrative Epidemiology Unit, University of Bristol. Email: l.paternoster@bristol.ac.uk

Co-authors: Lavinia Paternoster [1], Kate M Tilling [1], George Davey Smith [1]
1. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN

Abstract
Genome-wide association (GWAS) and Mendelian Randomization (MR) studies have been hugely successful in identifying causal mechanisms of disease, with proclaimed potential to inform drug discovery. However, as the majority of these studies focus on disease onset, though they may be informative for prevention of disease, they have unclear utility in informing disease treatment. Though examples of factors causing both disease incidence and disease progression exist (e.g. LDL cholesterol and initial/subsequent coronary artery events), examples also exist where this is not the case (e.g. a Crohn’s disease GWAS observed independent genetic variants for onset and progression). In order to identify effective treatments for disease, there is therefore a major opportunity to study disease progression as the outcome of interest. However, there are several methodological and practical challenges that must be overcome:

1) Collider Bias. A fundamental issue in progression studies, whereby selection of participants according to a characteristic (such as having a disease) will introduce inverse associations between all independent positively-coded risk factors for that characteristic, including between genetic variants and confounding factors.
2) Confounding with disease stage at baseline. Studies also need to ensure that true ‘progression’ is being measured, not influenced by detection bias.
3) Measurement of progression. Genetic studies typically have simple cross-sectional continuous or binary outcome measures. Progression studies will generally require more sophisticated models such as survival or trajectory analysis.
4) Availability of data. The final challenge is a practical one. Successful GWAS & MR studies have required huge consortia and large sample sizes. Therefore, it will be essential to integrate large studies with comparable progression measures.

We will discuss some of the potential solutions to these challenges in this area of considerable opportunity.

Keywords:
mendelian randomization, disease progression, methodology
Session 14: Application to social science and health behaviours
Chair: Dr Luisa Zuccolo, MRC Integrative Epidemiology Unit, UK
Causal association between Education and Coronary Heart Disease: a two-sample Mendelian randomization study

Presenter: Dr Taavi Tillmann
Department of Epidemiology & Public Health, University College London, UK. Email: t.tillmann@ucl.ac.uk

Co-authors: Taavi Tillmann [1]†, Julien Vaucher [2]†, Aysu Okbay [3], Hynek Pikhart [1], Anne Peasey [1], Ruzena Kubinova [4], Andrzej Pajak [5], Abdonas Tamosiunas [6], Sofia Malyutina [7], Krista Fischer [8], Giovanni Veronesi [9], Tom Palmer [10], Jack Bowden [11], George Davey Smith [11], Martin Bobak [1], Michael V Holmes [12], [13]
1 Department of Epidemiology & Public Health, University College London, UK. 2 Department of Internal Medicine, Lausanne University Hospital, Switzerland. 3 Department of Applied Economics, Erasmus University, Netherlands. 4 Department of Environmental Health Monitoring System, National Institute of Public Health, Prague, Czech Republic. 5 Department of Epidemiology and Population Studies, Institute of Public Health, Jagiellonian University, Krakow, Poland. 6 Institute of Cardiology, Lithuanian University of Health Sciences, Kaunas, Lithuania. 7 Institute of Internal and Preventive Medicine, Novosibirsk, Russia. 8 Estonian Genome Center, University of Tartu, Estonia. 9 Research Center in Epidemiology and Preventive Medicine, University of Insubria, Varese, Italy. 10 Department of Mathematics and Statistics, Lancaster University, Lancaster, UK. 11 MRC Integrative Epidemiology Unit at the University of Bristol, UK. 12 Clinical Trial Service Unit & Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, UK
13 MRC Population Health Research Unit at the University of Oxford, UK †=contributed equally

Abstract

OBJECTIVES: Education is observationally associated with a lower risk of coronary heart disease (CHD), however it is not known whether this association is causal. We asked: 1) Is there a causal association from Education to CHD?; 2) Are there causal associations from Education to the conventional cardiovascular risk factors?

METHODS: Using two-sample MR, we analysed data on 162 single nucleotide polymorphisms (SNPs) that have previously been associated with education. SNP-education associations were taken from the Social Science Genetic Association Consortium (n=319,600 after excluding sample overlap). SNP-CHD associations were taken from CARDIoGRAMplusC4D (63,746 CHD cases, 130,681 controls). Data on cardiovascular risk factors were taken from 6 other genetic consortia.

RESULTS: Genetic predisposition to 3.6 years of higher education was causally associated with lower odds of coronary heart disease (OR = 0.67, 95%CI: 0.59-0.77). This was unlikely to have been driven by pleiotropy (MR-Egger OR=0.54 [0.31-0.93]; median-MR OR=0.70 [0.58-0.85]), and was comparable to observational data (prevalence OR=0.73 [0.68 to 0.78]; incidence OR=0.80 [0.76 to 0.83]). In the reverse direction, genetic liability for CHD was not associated with adverse educational attainment (0.2 [-1.3 to 1.6] days of additional education, per 1-log unit increase in genetic risk of CHD). Higher education was causally associated with less smoking, reduced BMI and a favourable lipid profile.

CONCLUSIONS: Our findings suggest that education is causally associated with risk of CHD, offering support for policy interventions that increase education in order to improve population health. More research is needed on the underlying mechanisms. Mendelian randomization can be applied to social determinants of health.

Keywords: socioeconomic factors, educational status, cardiovascular disease, coronary artery disease, pleiotropy
Educational attainment and allostatic load in later life: Evidence using genetic markers

Presenter: Ms Xuejie Ding  
University of Oxford. Email: xuejie.ding@stx.ox.ac.uk

Co-authors: Xuejie Ding [1][2], Nicola Barban [1][3], Melinda Mills [1][3]  
1 Department of Sociology, University of Oxford, Oxford, OX1 3UQ  
2 St Cross College, University of Oxford, Oxford, OX1 3LZ  
3 Nuffield College, University of Oxford, Oxford OX1 1NF

Abstract  
Education is strongly correlated with health status in older adulthood. Yet whether the impact of expanding educational opportunities improves health remains unclear due to a lack of clarity over the causal relationship. This paper applies an instrumental variable (IV) approach to identify causality in relation to education on health as measured by allostatic load. Using the Health and Retirement Study (2008), we adopt two sets of IV instruments: (1) a polygenic score constructed from single-nucleotide polymorphisms (SNPs) associated with years of education; and, (2) quarter of birth. Using multiple instruments, we find a 15% reduction (seven times larger than the equivalent coefficient using OLS) in allostatic load per year of schooling. IV analyses were also used to identify potential pathways between education and health. Our IV analyses suggest that a higher level of education is causally related to better health in older adulthood.

Keywords:  
allostatic load, educational attainment, mendelian randomisation, polygenic score, quarter of birth
The causal effects of education on health: evidence from Mendelian randomization and the raising of the school leaving age

Presenter: Dr Neil Davies  
*University of Bristol. Email: neil.davies@bristol.ac.uk*

Co-authors: Neil M Davies [1,2], Matt Dickson[3], George Davey Smith[1,2], Gerard van den Berg[4], and Frank Windmeijer[1,4]

1 Medical Research Council Integrative Epidemiology Unit, University of Bristol, BS8 2BN, United Kingdom.  
2 School of Social and Community Medicine, University of Bristol, Barley House, Oakfield Grove, Bristol, BS8 2BN, United Kingdom.  
3 Department of Social and Policy Sciences, University of Bath, United Kingdom.  
4 Department of Economics, University of Bristol, Priory Road Complex, Bristol BS8 1TU, United Kingdom.

Abstract

Background
Instrumental variable analysis can potentially overcome a major limitation of observational studies – unmeasured and residual confounding. But it can only provide unbiased estimates of causal effects if the three core instrumental variable assumptions hold. In this study we compared two potential instruments for educational attainment (EA). First, a policy reform which raised the school leaving age to 16, and second an allele score constructed from genetic variants which associated with EA in a large GWAS.

Methods
We evaluated the IV assumptions using data from the participants of the UK Biobank (N 94,308). We assessed the strength of the instruments using a partial F-statistic. We used bias component plots to assess whether the proposed instruments were associated with 14 potential environmental confounders: location, maternal and paternal mortality, number of male and female siblings, whether the participant was breastfed, childhood height and weight, exposed to tobacco in utero, birthweight, and allele scores for a range of other phenotypes.

Results
Both the policy reform and the allele score were strongly associated with remaining in school (partial F-statistics were 295 and 1847 respectively). There was little evidence the policy reform was associated with any of the potential confounders except paternal longevity and childhood body size. Whereas the EA allele score was associated with 11 of the potential phenotypic confounders (measures of location, IMD of birth place, number of brother and sisters, breastfeeding, childhood height, maternal smoking in pregnancy, birthweight, and the BMI, birthweight and zinc allele scores).

Conclusion
The policy reform was largely independent of the potential confounders, and is a plausible instrument for education. The EA allele score associated with a range of potential confounders. Mendelian randomization analyses of social exposures, such as education, should be interpreted with caution.

Keywords: rosia, instrumental variable analysis, education, genomic confounding
Mendelian randomisation of alcohol intake and cardiovascular diseases in Chinese adults

Presenter: Dr Iona Millwood
University of Oxford. Email: iona.millwood@ndph.ox.ac.uk

Co-authors: Iona Y. Millwood [1], Robin G. Walters [1], Xue W. Mei [1], George Davey Smith [2], Rory Collins [1], Michael V. Holmes [1], Liming Li [3], Richard Peto [1], Zhengming Chen [1], on behalf of the China Kadoorie Biobank Collaborative Group

[1] Nuffield Department of Population Health, University of Oxford, UK
[2] MRC Integrative Epidemiology Unit, University of Bristol, UK
[3] Peking University Health Science Centre, Peking University, China

Abstract
Objectives: Moderate alcohol intake is associated with lower risks of cardiovascular diseases (CVD) in traditional epidemiology studies. We used functional variants in the alcohol-metabolising ADH and ALDH genes, which influence alcohol intake and are common in East Asians, to assess causality.

Methods: We genotyped ALDH2-rs671 (G>A) and ADH1B-rs1229984 (G>A) in 60,000 men and 90,000 women aged 30-79 years from the prospective China Kadoorie Biobank recruited from ten areas of China. Mendelian randomisation analyses associated alcohol intake predicted by genotype with CVD risk factors and incidence of stroke and ischaemic heart disease (IHD). Results were compared with traditional epidemiological analyses of self-reported alcohol intake and CVD outcomes.

Findings: rs671-A (allele frequency 0.20) and rs1229984-A (0.69) genotypes were associated with greatly reduced alcohol intake in men, and predicted a wide range of alcohol intake, from <1 to >25 drinks/week. In Mendelian randomisation analyses in men, genetically-predicted alcohol intake was associated with increased systolic blood pressure (SBP) and HDL-cholesterol, and with increased risk of stroke (8,000 events) in a dose-response manner consistent with the effects on SBP. There was little apparent effect on IHD (2,000 events). For both diseases, there was no causal evidence for the substantial protective effects of moderate alcohol intake which were seen in traditional epidemiological analyses of self-reported alcohol intake. Applying the same prediction to women, who drank little alcohol regardless of genotype, there was no association with CVD outcomes excluding the possibility of pleiotropic effects of the genetic instruments.

Conclusions: This large study in Chinese adults shows that alcohol intake at all levels increased risk of stroke, consistent with any beneficial effects of moderate drinking in traditional epidemiology studies being non-causal.

Keywords:
alcohol, cardiovascular disease, china
Explaining socioeconomic and health outcomes in non-experimental data: An instrumental variables approach using polygenic scores

Presenter: Prof. Dr. Philipp Koellinger
Center for Neurogenomics and Cognitive Research, Department of Complex Trait Genetics, VU University Amsterdam, 1081 HV Amsterdam, The Netherlands. Email: p.d.koellinger@vu.nl

Co-authors: Thomas A DiPrete [1], Casper AP Burik [2], Philipp D Koellinger[2]
1 Department of Sociology, Columbia University, New York, NY 10027
2 Center for Neurogenomics and Cognitive Research, Department of Complex Trait Genetics, VU University Amsterdam, 1081 HV Amsterdam, The Netherlands

Abstract
Estimating causal effects with non-experimental data is of central importance across multiple fields of scientific inquiry. Here, we propose genetic instrumental variables (GIV) regression for this purpose. GIV regression utilizes large-scale genome-wide association studies (GWAS) that now allow constructing predictive polygenic scores (PGS) for many human traits. Our approach is based on the idea that adding the true PGS for the outcome to a regression model would effectively eliminate bias arising from a genetic correlation between the outcome and an exposure of interest. We argue that without the true PGS in the structural model, using genes as instrumental variables (IVs) as proposed in Mendelian Randomization (MR) is problematic due to pleiotropic effects that invalidate IV regression. However, PGS capture only a small fraction of the heritability of most traits because GWAS are estimated in finite sample sizes that yield noisy estimates of the effects of each SNP. We argue and empirically demonstrate that it is possible to correct attenuation bias by splitting the GWAS sample to obtain several PGS (i.e. multiple indicators) in the prediction sample that can be used as instruments for each other in IV regression. Then we extend the approach to the problem of estimating causal effects and gene-environment interactions with non-experimental data. We find that our approach yields more accurate results than alternative approaches such as standard linear regression or MR. Our approach produces estimates of the chip heritability of educational attainment (EA) that are consistent with GREML-based heritability estimates, and, unlike the results using MR, GIV-based estimates find that the positive correlation between body height and EA is primarily due to genetic effects. In a negative control, we find no effect of educational attainment on body height.

Keywords:
genetic instrumental variable regression (giv), enhanced mendelian randomization, gwas, polygenic scores
Session 15: Methodology
Chair: Dr Frank Dudbridge, Department of Health Sciences, University of Leicester, UK
Semiparametric methods for estimation of a non-linear exposure-outcome relationship using instrumental variables in Mendelian randomization

Presenter: Mr James Staley
University of Bristol. Email: js16174@bristol.ac.uk

Co-authors: James R Staley [1,2], Stephen Burgess [2,3]
1 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN
2 Cardiovascular Epidemiology Unit, University of Cambridge, Cambridge, CB1 8RN
3 MRC Biostatistics Unit, Cambridge Institute of Public Health, Cambridge, CB2 0SR

Abstract
Background
Mendelian randomization, the use of genetic variants as instrumental variables (IV), can test for and estimate the causal effect of an exposure on an outcome. Most IV methods for estimating the association between an exposure and outcome implicitly assume that the relationship is linear. However, in practice this assumption may not hold. Indeed, often the primary question of interest is to assess the shape of this relationship.

Methods
We present two novel IV approaches for investigating the shape of the exposure-outcome relationship in individual-level concomitant data (i.e. in one sample): a fractional polynomial method and a piecewise linear method. These methods rely on dividing the population into strata using the exposure distribution, and estimating a causal effect, referred to as a localized average causal effect (LACE), in each stratum of the population. The fractional polynomial method performs meta-regression across these LACE estimates. The piecewise linear method estimates a continuous piecewise linear function, the gradient of which is the LACE estimate in each stratum. These approaches were tested using simulations and were used to investigate the effect of body mass index (BMI) on systolic blood pressure (SBP) and diastolic blood pressure (DBP) in UK-Biobank. These methods are available in the nlmr R package (https://github.com/jrs95/nlmr).

Results
Our simulations suggest that both modelling approaches yield reasonable model fits to a variety of underlying data generating models. In particular, the fractional polynomial method performed well when the underlying effect was curvilinear, while the piecewise linear approach performed well under a threshold relationship. Using these methods we identified strong non-linear causal effects of BMI on DBP and SBP in UK-Biobank.

Conclusion
These novel IV approaches can be used to investigate the shape of exposure-outcome relationships in the context of Mendelian randomization.

Keywords:
non-linear relationship, fractional polynomials, piecewise linear function, localized average causal effect, uk-biobank
A survey of causal inferences via Mendelian randomization across the landscape of available GWAS

Presenter: Ms. Shirbi Ish-Shalom

Biomedical Informatics, Stanford University School of Medicine. Email: shirbi@stanford.edu

Co-authors: Shirbi Ish-Shalom [1], Craig Hyde [2], Eric Fauman [3], Stephen Burgess [4][5], Daniel Ziemek [6]

1 Biomedical Informatics, Stanford University School of Medicine, Stanford, CA 94305, USA, 2 Clinical Research Statistics, WRD, Pfizer, Cambridge, MA, 3 Computational Target Validation, WRD, Pfizer, Cambridge, MA, 4 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK, 5 Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK, 6 Inflammation & Immunology, WRD, Pfizer, Germany

Abstract

Understanding causal genetic relationships between complex human traits and diseases can provide insights into how we identify and, potentially, modify risk of disease. By developing a comprehensive human trait network, we can leverage the causal and subsumption structure to understand the underlying local and global trait relationships. However, genetic intricacies (such as pleiotropy, linkage disequilibrium, and SNP-matching), varying sequencing methods, and limited access to individual-level genetic data create significant roadblocks to building such a network. Recently, regression methods in Mendelian randomization analyses that can identify violations in pleiotropy assumptions have been developed to use summary-level genetic data to obtain estimates of the causal effects between traits. Using over one thousand genome-wide association studies, we build a phenotypic trait network that analyzes over one million trait-trait causal effect relationships using the MR-Egger regression method. We see confirmation for many published trait relationships, but also discover some potentially novel genetic associations. This framework for understanding the genetic architecture of disease can be leveraged to generate hypotheses about increased disease recognition and lowered disease risk by identifying novel intervention strategies.

Keywords:
mr, metabolomics, proteomics, meta-analysis survey, gwas
Efficient And Accurate Causal Inference With Hidden Confounders From Genome-Transcriptome Variation Data

Presenter: Dr Tom Michoel
The University of Edinburgh. Email: tom.michoel@roslin.ed.ac.uk

Co-authors: Lingfei Wang [1], Tom Michoel [1]
The Roslin Institute, The University of Edinburgh, Easter Bush, Midlothian, EH25 9RG

Abstract
Mapping gene expression as a quantitative trait using whole genome-sequencing and transcriptome analysis allows to discover the functional consequences of genetic variation. We developed a novel method and ultra-fast software Findr for highly accurate causal inference between gene expression traits using cis-regulatory DNA variations as causal anchors, which improves current methods by taking into account hidden confounders and weak regulations. Findr outperformed existing methods on the DREAM5 Systems Genetics challenge and on the prediction of microRNA and transcription factor targets in human lymphoblastoid cells, while being nearly a million times faster. Findr is publicly available at https://github.com/lingfeiwang/findr.

Keywords:
causal inference, eqtl, genome, transcriptome
Power asymmetry should give us pause when interpreting results from summary statistic based Mendelian-randomisation.

Presenter: Dr. Michel G Nivard
Department of biological psychology, faculty of behavioral and movement sciences, VU University, Amsterdam. Email: m.g.nivard@gmail.com

Co-authors: Michel G Nivard, Conor, V Dolan, Eco de Geus & Dorret I Boomsma
Department of biological psychology, faculty of behavioral and movement sciences, VU University, Amsterdam

Abstract
The object is to determine, in the context of summary statistic based Mendelian randomization studies, the consequences of asymmetric power, where the power to detect the effect of A on B is greater than the power to detect the effect of B on A. Specifically we discuss 3 classes of traits frequently considered in MR studies: 1) Omics, for example gene expression or CpG methylation, characterized by the availability of a strong instrument and small sample sizes; 2) endophenotypes, for example metabolites, characterized by multiple strong instruments and moderate sample sizes; and 3) complex traits for which many weak instruments are available, and which are characterized by an infinitesimal genetic architecture and large sample sizes.

Our simulations will lay bare an issue which arise when testing the causal relations between gene expression, CpG methylation, metabolites and complex traits. We demonstrate through simulation that it is easier to detect causal effects when the cause has a limited number of genetic determinants, and the outcome is measured in a large sample. Due to the asymmetry in power, and the available summary data, one is far more likely to detect the effects of gene expression or CpG methylation on metabolites, and the effects of these on complex traits than reverse causal effects. Ignoring the power asymmetry can lead to an overly reductionist interpretation of the web of causal relations between transcription, the epigenome, metabolites and complex traits. Our analysis highlights the need to jointly consider summary statistics based Mendelian-randomisation and other (genetic) techniques which can estimate causal relations.

Keywords:
mendelian-randomisation, conceptual considerations, power,
Abstract

Objective: To test for violation of the InSIDE assumption when implementing MR-Egger analysis.

Methods: In an MR study performed using the ratio (or Wald) method with J uncorrelated genetic variants as instruments, a subgroup of P instruments that share a common pleiotropic pathway are identified by their association to a secondary phenotype known to directly affect the outcome. Performing MR-Egger analysis with J, P and J-P instruments should all give the same causal effect estimate if the necessary assumptions, including InSIDE, hold. Any significant discrepancy in the estimates generated would suggest violation of the requisite assumptions. The applicability of this approach is exampled in an MR study investigating the causal effect of age at menarche on forced vital capacity (FVC, a measure of lung function) in adolescence, where J is all 122 selected instruments for age at menarche, P is the 61 instruments most significantly associated with height (p<0.017), and J-P is the 61 instruments least associated with height (p>0.017). Instrument-phenotype associations are estimated using GWAS meta-analysis results.

Results: Performing MR-Egger regression analysis in the described example with J, P and J-P instruments gives differing causal effect estimates on FVC per year increase in age at menarche, of 74mL (95% confidence interval -130mL to 278mL), -23mL (95% CI -277mL to 230mL) and 365mL (95% CI -30mL to 759mL), respectively.

Conclusion: This approach can be used to test whether any of the instruments are violating the necessary assumption for MR-Egger. It is important to appreciate that selection of P based on the likely nature of pleiotropy, rather than randomly, may better determine the degree to which InSIDE is violated; this can be exploited to highlight differences in the effect estimates generated by MR-Egger. Furthermore, this method is only useful for detecting when the assumptions of MR-Egger are being violated, not evidence that they are being held.

Keywords:
mendelian randomization, pleiotropy, mr-egger, inside
Improving the accuracy of two-sample summary data Mendelian randomization: moving beyond the NOME assumption

Presenter: Jack Bowden
MRC Integrative Epidemiology Unit. Email: jack.bowden@bristol.ac.uk

Co-authors: Jack Bowden [1], Fabiola Del Greco M [2], Cosetta Minelli [3]
John Thompson [4] and George Davey Smith [1]
[1] MRC Integrative Epidemiology Unit, University of Bristol
[2] Center for Biomedicine, EURAC research, Bolzano, Italy

Abstract
Two-sample summary data Mendelian randomization (MR) incorporating multiple genetic variants is becoming an ever more popular technique for assessing causality in epidemiological research within the framework of a meta-analysis. If all genetic variants satisfy the instrumental variable (IV) assumptions, then their individual causal ratio estimates should be homogeneous. Observed heterogeneity, therefore, supports the notion that a portion of variants violate the IV assumptions due to pleiotropy. Heterogeneity is typically assessed, in the first instance, using Cochran’s Q statistic. MR-Egger regression can be usefully applied, in addition, when heterogeneity is detected and is thought to meaningfully bias the analysis. Heterogeneity around the MR-Egger fit can be assessed using Rucker’s Q’ statistic.

Both Cochran’s Q and Rucker’s Q’ require an approximation for the variance of each ratio estimate in order to be calculated. We show that the most popular ‘1st order’ approximation can lead to an inflation in the chances of detecting heterogeneity when in fact it is not present (a type I error). Conversely, an ostensibly more accurate ‘2nd order’ approximation can dramatically increase the chances of failing to detect heterogeneity, when it is truly present (a type II error). We derive a new ‘modified 2nd order’ approximation for the variance to mitigate both of these adverse affects.

Using Monte-carlo simulations, we show that the modified 2nd order approximation dramatically outperforms both its 1st and 2nd order counterparts, especially in the presence of weak instruments. We illustrate the utility of the new methods using data from a recent two sample summary data MR analysis to assess the causal role of systolic blood pressure on coronary heart disease risk.

In conclusion, modified 2nd order weighting should be used as standard within two sample summary data MR studies to quantify and correct for heterogeneity.

Keywords:
two sample summary data mr, ivw, mr-egger, cochran’s q, rucker’s q'
Session 16: Pleiotropy
Chair: Dr Cosetta Minelli, Population Health and Occupational Disease, Imperial College, London, UK
Pleiotropy-robust Mendelian Randomization

Presenter: Dr Hans van Kippersluis
Erasmus University Rotterdam. Email: hvankippersluis@ese.eur.nl

Co-authors: Hans van Kippersluis [1], Cornelius A. Rietveld [2]
1 Erasmus School of Economics, Erasmus University Rotterdam, 2 Erasmus School of Economics, Erasmus University Rotterdam

Abstract
Background: The potential of Mendelian Randomization studies is rapidly expanding due to (i) the growing power of GWAS meta-analyses to detect genetic variants associated with several exposures, and (ii) the increasing availability of these genetic variants in large-scale surveys. However, without a proper biological understanding of the pleiotropic working of genetic variants, a fundamental assumption of Mendelian Randomization (the exclusion restriction) can always be contested.

Methods: We build upon and synthesize recent advances in the literature on instrumental variables (IVs) estimation that test and relax the exclusion restriction. Our Pleiotropy-robust Mendelian Randomization (PRMR) method first estimates the degree of pleiotropy, and in turn corrects for it. If a subsample exists for which the genetic variants do not affect the exposure, the selection into this subsample is not a joint consequence of the IV and the outcome, and pleiotropic effects are homogenous, PRMR obtains unbiased estimates of causal effects.

Results: Simulations show that existing MR methods produce biased estimators for realistic forms of pleiotropy. Under the aforementioned assumptions, PRMR produces unbiased estimators. We illustrate the practical use of PRMR by estimating the causal effect of (i) tobacco exposure on Body Mass Index (BMI); (ii) prostate cancer on self-reported health, and (iii) educational attainment on BMI in the UK Biobank data.

Conclusions: PRMR allows for instrumental variables that violate the exclusion restriction due to pleiotropy, and corrects for pleiotropy in the estimation of the causal effect. If the degree of pleiotropy is unknown, PRMR can still be used as a sensitivity analysis.

Keywords:
mendelian randomization, pleiotropy, plausibly exogenous
Improving the analysis and interpretation of two-sample summary data Mendelian randomization via Radial plot regression

Presenter: Jack Bowden
MRC Integrative Epidemiology Unit. Email: jack.bowden@bristol.ac.uk

Co-authors: Jack Bowden [1], Fabiola Del Greco M [2], Cosetta Minelli [3], John R Thompson [4] and George Davey Smith [1]

1 MRC Integrative Epidemiology Unit, University of Bristol.
2 Center for Biomedicine, EURAC research, Bolzano.
3 Population Health and Occupational Disease, NHLI, Imperial College, London.
4 Department of Health Sciences, University of Leicester, Leicester.

Abstract
Summary data furnishing a Mendelian randomization study are often visualized with the aid of a scatter plot, in which single nucleotide polymorphism (SNP)-outcome associations are plotted against the SNP-exposure associations to provide an immediate picture of the causal effect estimate for each individual variant. It is also convenient to overlay the standard inverse variance weighted (IVW) estimate of causal effect as a fitted slope, to see whether an individual SNP provides evidence that supports, or conflicts with, the overall consensus. Unfortunately, the traditional scatter plot is not the most appropriate means to achieve this aim whenever SNP-outcome associations are estimated with varying degrees of precision and this is reflected in the analysis. We propose instead to use a small modification of the scatter plot - the Radial plot - for the presentation of data and results from an MR study, which enjoys many advantages over the original method. On a practical level it removes the need to recode the genetic data and enables a more straightforward detection of outliers and influential data points. Its use extends beyond the purely aesthetic, however, to suggest a more general regression modelling framework to operate within when conducting an MR study. This increases the flexibility to incorporate different assumptions about the variance of individual ratio estimates, both in standard and in pleiotropy robust regression models such as MR-Egger regression. It also facilitates simple comparisons of the relative goodness-of-fit between such models. We illustrate the methods using data from a two-sample Mendelian randomization study to probe the causal effect of systolic blood pressure on coronary heart disease risk.

Keywords:
mendelian randomization, two-sample summary data mr, radial plot, ivw estimate, mr-egger regression
Mendelian randomization with Egger pleiotropy correction and weakly informative Bayesian priors

Presenter: Amand Floraan Schmidt
UCL. Email: amand.schmidt@ucl.ac.uk

Co-authors: A F Schmidt [1] and F Dudbridge [2]
[1] Institute of Cardiovascular Science, Faculty of Population Health, University College London, London WC1E 6BT, United Kingdom.
[2] Department of Health Sciences, University of Leicester, Leicester LE1 7RH, United Kingdom.

Abstract
Background The MR-Egger (MRE) estimator has been proposed to correct for unbalanced pleiotropic effects of genetic instruments in an instrumental variable (IV) analysis. Power of this method is considerably lower than that of conventional estimators, limiting its applicability. Here we propose a novel Bayesian implementation of the MR-Egger (BMRE) and explore the utility of applying weakly informative priors on the intercept term (the pleiotropy estimate) to increase power of the IV (slope) estimate.
Method Simulation study to compare the performance of different IV estimators. Scenarios differed in the presence of a causal effect, the presence of pleiotropy, and violations of the “Instrument Strength Independent of Direct Effect” (InSIDE) assumption. Based on the empirical example of plasma urate and coronary heart disease (CHD) we present an approach to explore the presence and relevance of unbalanced pleiotropy.
Results A weakly informative prior on the intercept term increased power of the slope estimate while maintaining type 1 error rates close to the nominal value of 0.05. Under the InSIDE assumption, performance was not affected by the presence or absence of pleiotropy. Violation of the InSIDE assumption biased all estimators, and impacted the BMRE more than the MRE. In our empirical example urate (per SD) increased the risk of CHD: OR 1.14 (95%CI 1.03; 1.27).
Conclusion Depending on the strength of the prior, the BMRE has more power at the cost of an increased susceptibility to InSIDE assumption violations. As such the BMRE is a compromise between the MRE and conventional IV estimators.

Keywords:
epidemiology methods, bayesian analysis, mendelian randomization analysis, statistics
Demystifying causal effect heterogeneity of composite risk-factors in multi-instrument Mendelian Randomisation studies using a novel Bayesian feature selection algorithm

Presenter: Dr Christopher Foley
MRC Biostatistics Unit, University of Cambridge. Email: christopher.foley@mrc-bsu.cam.ac.uk

Co-authors: Christopher N Foley [1], Stephen Burgess [1]
1 MRC Biostatistics Unit, University of Cambridge

Abstract
Mendelian Randomisation (MR) uses single nucleotide variants as 'naturally-randomised' instrumental variables to assess a causal link between a risk-factor and an outcome (e.g. disease). MR has thrived over recent years with the expanding success of genome-wide association studies (GWAS). The wide-spread availability of summary association statistics from hundreds of GWAS, which have identified ~32,000 trait associated loci, have allowed for MR analyses which aggregate effect estimates across hundreds of instruments and multiple risk-factors. However, a key assumption of MR is the exclusion restriction: a genetic instrument only acts on the outcome through the risk-factor. Violations of this, when a locus is associated with multiple traits (pleiotropy), is common in human genetics; leading to bias of causal effect estimates and effect heterogeneity in aggregated analyses.

We focus on the assessment of composite risk-factors (CRFs), i.e. risk-factors that are composed of several lower-level risk-factors (LLRFs), e.g. BMI is a CRF computed from two LLRFs: body weight and height. If a LLRF is associated with the outcome directly as well as being mediated through the CRF the MR estimates suffer symptoms of pleiotropy. We take advantage of this detail, however, and by blending a Bayesian Model Averaging technique with a novel feature selection algorithm we: 1) identify the number of latent LLRFs 2) match each instrument to a LLRF group and either 3a) estimate a CRF causal effect or 3b) interpret results when all instruments are invalid. Broadly, every risk-factor is a CRF. Hence, our hypothesis-free method is widely applicable. We illustrate the approach using BMI and inflammation CRF examples.

Keywords:
causal effect heterogeneity, multiple instruments, pleiotropy, bayesian model average, feature selection
An examination of multivariable Mendelian Randomization: Translating and extending econometric theory from the single sample to the two sample summary data setting

Presenter: Eleanor Sanderson
MRC Integrative Epidemiology Unit, University of Bristol. Email: eleanor.sanderson@bristol.ac.uk

Co-authors: Eleanor Sanderson [1], George Davey Smith [1], Jack Bowden [1]
1 MRC Integrative Epidemiology Unit, University of Bristol

Abstract
Mendelian Randomisation (MR) is a powerful tool in epidemiology, which uses genetic variants as instrumental variables (IVs) to estimate the causal effect of an exposure on an outcome in the presence of unobserved confounding. This can be extended to Multivariable MR (MVMR) to estimate the effect of two or more exposures on an outcome. Estimation of IV models with multiple exposure variables and individual level data on a single sample is well-established and frequently used in Economics. We demonstrate its application to a Mendelian randomisation analysis and exploit recent developments in Econometric IV theory to show how violations of the IV assumptions due to weak instruments or pleiotropy can be assessed.

We then examine the extent to which this established econometric methodology can be applied to MVMR in the two-sample summary data setting. We find that that summary data alone does not contain the all of the information required to be able to robustly test for weak and pleiotropic instruments. Several strategies are proposed to mitigate the effect of this missing information, the properties of which are assessed through simulation studies.

Keywords:
multivariable mendelian randomisation
Session 17: Applications in cancer
Chair: Professor Richard Martin, MRC Integrative Epidemiology Unit, UK
Applications of MR in cancer

Presenter: Associate Professor Stuart MacGregor
QIMR Berghofer Medical Research Institute. Email: stuart.macgregor@qimrberghofer.edu.au

Co-authors: Stuart MacGregor [1], Jue Sheng Ong [1], Upekha Liyanage [1], Jiyuan An [1], Matthew Law [1], Puya Gharahkhani [1]
1 QIMR Berghofer Medical Research Institute, Herston, Brisbane, Queensland 4029, Australia.

Abstract
The identification of modifiable risk factors is crucial in public health efforts to control cancer. Observational studies have been widely used to identify putative cancer risk factors but issues such as confounding and reverse causality make interpretation difficult. Mendelian Randomization (MR) can ameliorate some of these issues and enable better assessment of the true relationship between risk factor and cancer risk. I will describe work harnessing large scale genetic data to make causal inferences about the role of a variety of putative cancer risk factors. Data will be drawn from a range of GWAS consortia and biobanks. I will focus on cancers of the skin, esophagus and ovary, as well as considering overall cancer risk. As proof of principle I will show concordance between MR and observational findings for the effect of increased height on cancer risk. Modifiable risk factors investigated include obesity, fatty acid levels, cholesterol levels, vitamin D levels, as well as coffee and alcohol consumption. I will discuss the findings from these MR studies, giving due consideration to the MR assumptions and statistical analysis issues.

Keywords:
causal inference, diet, obesity, cancer risk
Solving complex relationships between metabolic parameters and renal cell carcinoma risk: from Mendelian randomization to matrix analysis.

Presenter: Mr Robert Carreras-Torres
Section of Genetics, International Agency for Research on Cancer (IARC), Lyon, France. Email: carrerasr@fellows.iarc.fr

Co-authors: Robert Carreras-Torres [1], Gibran Hemani [2], Philip C Haycock [2], Caroline Relton [2], Richard Martin [2], George Davey Smith [2], Ghislaine Scelo [1], Mattias Johansson [1], Paul Brennan [1].
1 Section of Genetics, International Agency for Research on Cancer (IARC), Lyon, France.
2 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK.

Abstract
Obesity and metabolic parameters have been identified as risk factors for renal cell carcinoma (RCC) through Mendelian randomization (MR) analyses. Each standard deviation (SD) increase in body mass index (BMI), waist-to-hip ratio (WHR) and fasting insulin were predicted to increase RCC risk between 44-66%. However, the genetic instruments for these parameters were correlated, suggesting the possibility of mediated effects between the risk factors. We analyze the matrix of two-sample MR total causal effects (direct plus indirect effects) between these parameters and RCC risk (10,784 RCC cases and 20,406 controls). A method to orthogonalise the matrix of total effects to obtain direct effects is calculating the inverse matrix. This inversion method appeared to be reasonably accurate when analyzing the interaction of a few parameters. Standard errors of direct effects were obtained by bootstrapping 1000 times the normally distributed total effects. We analyzed two interactive models: one reflecting the relationship between BMI, fasting insulin and RCC risk; and the other between WHR, fasting insulin and RCC risk. Preliminary results indicated that obesity parameters and fasting insulin were independent risk factors for RCC risk. The first model suggested that RCC risk was increased by 45% (Odds Ratio (OR); 95% Confidence Interval (95%CI) = 1.45; 1.20-1.74) per SD increase in BMI and by 66% (OR = 1.66; 95%CI: 1.13-2.46) per SD increase in fasting insulin. Thus, the 19.1% of the total effect of BMI on RCC risk corresponded to an indirect effect mediated by fasting insulin. The second model indicated that each SD increase in WHR and fasting insulin were associated with 32% and 55% increases in RCC risk, respectively (OR; 95%CI = 1.32; 0.96-1.82 for WHR and OR; 95%CI = 1.55; 1.02-2.36 for fasting insulin), indicating some reciprocal mediation.
This study provides important evidence on the causal relevance of metabolic components in kidney cancer etiology.

Keywords:
renal cell carcinoma, medelian randomization. direct and indirect effects.
Session 18: Applications
Chair: Dr Beate St Pourcain, Max Planck Institute for Psycholinguistics, The Netherlands
The causal map of 150 complex traits and diseases: a first draft

Presenter: Dr Gibran Hemani
MRC Integrative Epidemiology Unit. Email: g.hemani@bristol.ac.uk

Co-authors: Gibran Hemani [1], Philip Haycock [1], Jie Zheng [1], Tom Gaunt [1], George Davey Smith [1]
1 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN

Abstract
It is now somewhat trivial to use mrbase.org to obtain pairwise 2-sample Mendelian randomisation (2SMR) estimates between hundreds of traits. This raises two fundamental questions: 1) How can we navigate through the myriad of available 2SMR methods to obtain the most reliable causal effect estimate for each relationship; and 2) What is the utility of such a matrix of effects?

Through simulations we explored how the reliability of a causal effect estimate between two traits can be improved if we have data on many other traits. We demonstrate that when GWAS statistical power is high, and traits are related in complex networks, it is very likely that some instruments will be invalid because they are primarily acting on confounders or are having a reverse causal effect. Using the Rucker framework to account for horizontal pleiotropy does not fully avoid this problem, nor do median or mode based estimators. Using an iterative approach to navigate through the network of traits to update pairwise estimates by omitting instruments that are likely to be unreliable does substantially reduce bias and false discovery rates, and can often improve power by reducing heterogeneity.

We used the approach to construct a first draft of the map of putative causal relationships between 150 complex traits and diseases. After Bonferroni correction we found many hundreds of associations. Using this matrix can substantially improve power in MR, because the number of putative exposure-outcome associations increases dramatically when allowing for causal chains to include multiple intermediate traits. Subsets of the matrix can be orthogonalised to decompose the total effect estimates into direct and indirect relationships for an arbitrary number of traits. However, the empirical results from this analysis make it clear that apparently robust MR estimates can be obtained for biologically impossible trait relationships, and the automation of causal inference must be approached with caution.

Keywords:
systems biology, mediation, statistical power
Association of a LDL-cholesterol genetic instrument with cardiovascular risk factors and events in the China Kadoorie Biobank

Presenter: Ruth Boxall  
*University of Oxford. Email: ruth.boxall@ndph.ox.ac.uk*

Co-authors: Ruth Boxall [1], Michael V Holmes [1,2,3], Kuang Lin [2], Derrick A. Bennett [2], Iona Y Millwood [1,2], Robin G Walters [2] and Zhengming Chen [2] on behalf of the China Kadoorie Biobank working group

1 Medical Research Council Population Health Research Unit at the University of Oxford  
2 Clinical Trial Service Unit & Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, BHF Centre of Research Excellence, Big Data Institute Building, Roosevelt Drive, University of Oxford, UK  
3 National Institute for Health Research Oxford Biomedical Research Centre, Oxford University Hospital, Old Road, Oxford OX3 7LE, UK

Abstract

Objectives: LDL-cholesterol (LDL-C) is established as causal for coronary heart disease, but its role in other cardiovascular diseases (CVD) is less clear. We investigated the associations of a LDL-C genetic risk score on risk factors and CVD in the China Kadoorie Biobank (CKB).

Methods: CKB is a prospective study of 0.5M adults aged 30-79 yrs recruited in 2004-8 from 10 areas of China, with linkage to registries and electronic health records. Imputed genotype data for 27,903 individuals (including population-based and case-control samples) were available for 79 of 84 independent LDL-C associated SNPs identified by the Global Lipids Genetic Consortium (GLGC). A weighted genetic risk score (GRS) was created from these using GLGC effect estimates (in SDs). Linear and logistic regression was used to associate this with lipid measures (N=17,750), continuous traits and CVD sub-types.

Results: The LDL-C GRS was strongly associated with LDL-C (0.62mmol/l higher LDL-C per unit, \(P=1.2\times10^{-243}\)), higher apolipoprotein B (\(P=3.0\times10^{-268}\)), triglycerides (\(P=2.3\times10^{-10}\)) and lipoprotein(a) (\(P=1.7\times10^{-19}\)), and more weakly, with lower HDL-cholesterol (\(P=0.007\)) and apolipoprotein A1 (\(P=0.005\)). The GRS associated with higher carotid intima medial thickness (0.21SD per unit GRS; 95%CI: 0.14-0.28) and presence of plaque (OR 1.49; 95%CI: 1.24-1.79) but not with blood pressure or adiposity measures. The GRS associated with a higher risk of myocardial infarction (MI, OR=1.26, 95%CI: 1.01-1.55) and major coronary events (MCE, OR=1.22, 95%CI: 1.04-1.44). There was a trend to increased risk of ischaemic stroke (IS, OR=1.12, 95%CI: 1.00-1.26), but no association with haemorrhagic stroke (HS, OR=0.96; 95%CI: 0.86-1.08).

Conclusions: An LDL-C GRS associated with subclinical atherosclerosis, MI and MCE, and a similar trend with IS. No association with HS was identified. Power for detecting stroke subtype associations will be improved by extending this work to 100,000 genotyped participants.

Keywords:  
LDL-cholesterol, mendelian randomization, biobank
Mendelian randomization study compared to randomized controlled trial: a systematic review and meta-analysis

Presenter: Miss Shan LUO
School of Public Health, The University of Hong Kong. Email: luoshan@hku.hk

Co-authors: Shan Luo [1], Jie V Zhao [1], C Mary Schooling [1,2]
1 School of Public Health, The University of Hong Kong
2 City University of New York Graduate School of Public Health and Health Policy

Abstract
Randomized controlled trials (RCTs) are regarded as the ‘gold-standard’ for establishing causal relations. Mendelian randomization (MR), using genetic variants as instrumental variables, is analogous to the randomization process in RCTs, making it less susceptible to confounding. MR is increasingly used to assess causal effects from observational studies, but has been heavily criticized on methodological grounds, so the credence that should be given to MR studies is currently uncertain.

A systematic review and meta-analysis was conducted to assess the validity of MR studies by comparing the concordance rate in terms of direction of effect between MR studies and RCTs. Two people searched 3 databases to March 2017 using “Mendelian randomization” or “Mendelian randomisation”. We then searched for RCTs or systematic reviews and meta-analysis of RCTs addressing the same research questions. As necessary MR studies or RCTs on the same topic were meta-analyzed together. Concordance between MR studies and RCTs were assessed by Cohen’s kappa.

Among 113 research topics, 75 had the same direction of effect for MR studies and RCTs, that is both MR and RCTs had positive, negative or null associations. The Cohen’s kappa was 0.319, with 95% confident interval (CI) (0.144, 0.475). In sensitivity analysis excluding nutrition trials, the Cohen’s kappa was 0.494 with 95%CI (0.297, 0.651).

Our findings provide evidence of concordance of direction of effects from MR studies and RCTs. The main reasons for differences include violations of the MR assumptions, RCTs having off-target effects and possibly different effects of endogenous and exogenous exposures. This validation of MR studies clarifies that MR studies as ‘natural’ RCT, with implication for evidence based practice and policy. Incorporating MR studies before and integrated into RCTs at the design stage is recommended to improve allocation of resources, etiologic understanding and causal inference.

Keywords:
mendelian randomization, randomized controlled trials, systematic review, meta-analysis, concordance analysis
Do blood lipid levels influence bone mineral density? Findings from a Mendelian randomization study.

Presenter: Dr. Jie Zheng
MRC Integrative Epidemiology Unit. Email: jie.zheng@bristol.ac.uk

Co-authors: Jie Zheng[1], Marie-Jo Brion[2], John Kemp[2], Nicole Warrington[2], Gibran Hemani[1], Tom G. Richardson[1], Zhen Qiao[2], Philip Haycock[1], Jon Tobias[1,3], David Evans[1,2]

1 Medical Research Council (MRC) Integrative Epidemiology Unit, School of Social & Community Medicine, University of Bristol, Bristol, United Kingdom; 2 University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Queensland, Australia; 3 Musculoskeletal Research Unit, University of Bristol, Bristol, UK

Abstract
Treatment with statins is associated with increased bone mineral density (BMD) and reduced fracture risk. Previous laboratory studies suggest that the direct effect of statins on osteoblasts and/or osteoclasts underlie this relationship. However, few studies have investigated an alternative possibility that part of the beneficial effects of statins on bone is mediated by lowering levels of low density lipoprotein cholesterol (LDL-C). In order to investigate the relationship between plasma lipids and BMD, we performed a two-sample Mendelian randomization (MR) study.

We utilized 184 SNPs robustly associated with plasma lipid levels and 239 SNPs associated with heel BMD from 142,487 UK Biobank individuals. We performed univariate MR analyses on LDL-C, HDL-cholesterol (HDL-C) and triglyceride levels (TG), as well as multivariable MR and a range of sensitivity analyses to account for horizontal pleiotropy among the different lipid fractions. To test whether the effect of statins on BMD was mediated by lowering lipid levels, MR was repeated with and without the HMGCR SNPs in order to mimic the effect of statin use. We also performed bi-directional MR to examine the possibility of reverse causality between BMD and blood lipids.

Univariate MR using SNPs associated with LDL-C, HDL-C or TG provided evidence for a causal effect of LDL-C on BMD (β=−0.077, P=2x10^-5). Multivariable analysis suggested that the effect of LDL-C on BMD was independent of HDL-C and triglycerides. MR analyses excluding the HMGCR SNPs were consistent with a causal effect of LDL-C on BMD (β=−0.102, P=3x10^-5) as were MR analyses for 5 HMGCR SNPs (beta=−0.11, P=0.00031). Bidirectional MR provided no evidence for a causal effect of BMD on blood lipids.

Our results suggest lower LDL-C is causally related to increased BMD. Further studies are justified to explore the mechanisms by which lower LDL-C improves BMD and to examine the potential role of modifying blood lipids in treating osteoporosis.

Keywords:
blood lipids, bone mineral density, statin, osteoporosis, mendelian randomization
Body mass index and inflammatory skin disease: is there a causal relationship?

Presenter: Dr Ashley Budu-Aggrey

Medical Research Council (MRC) Integrative Epidemiology Unit, University of Bristol, UK

Co-authors: Ashley Budu-Aggrey[1,2,[*], Ellen H Modalsli[3,4,[*]], Sarah Watkins[1,2,[*]], Gunnhild Aaberge Vie[3], Ben Michael Brumpton[1,5,6], Bjørn Olav Åsvold[5,7], Mari Løset[4,5], Maiken E Gabrielsen[5], Lars Fritsche[5,8], Kristian Hveem[5], Cristen Willer[9], Pål Richard Romundstad[3], Kjersti Danielsen[10,11], Anne-Sofie Furberg[10,12], Sara J Brown[12,13], Marit Saunes[3,14] and Lavinia Paternoster[1,2]

Abstract

Background: Eczema (atopic dermatitis) and psoriasis are common inflammatory skin diseases which may be associated with obesity. We investigated the causal relationship between body mass index (BMI) and these two conditions using Mendelian Randomization (MR).

Methods: Firstly, we reviewed the literature for epidemiological evidence of association between obesity and either eczema or psoriasis. We then tested for causality among 161,459 participants from the UK Biobank and the Nord-Trøndelag Health Study (HUNT). We derived a BMI genetic instrument from 97 independently associated single nucleotide polymorphisms (SNPs) that have been previously reported. These were used to create a weighted genetic risk score to perform one-sample MR with the two stage least squares method. We then performed two-sample MR (inverse-variance weighting) with published BMI and eczema association data. The one-sample and two-sample MR estimates were meta-analysed using a fixed-effect model.

Preliminary results: Evidence in the literature suggested an observational association of higher weight with eczema and psoriasis. One-sample MR gave causal odds ratio (OR) estimates of 1.37 (95% CI 0.92–2.04) for psoriasis and 1.28 (95% CI 0.98–1.69) for eczema per 1 standard deviation (SD) increase in BMI (1.06 and 1.04 per kg/m² respectively). Also for eczema, two-sample MR gave an OR of 1.07 (95% CI 0.92–1.24) per 1-SD increase in BMI, and 1.11 (95% CI 0.98–1.27) upon meta-analysis with the one-sample MR estimates.

Conclusions: Our initial findings provide weak evidence of a causal effect of BMI on eczema and psoriasis risk. Additional samples will increase the power of the study and improve the precision of the effect estimates. Various sensitivity analyses will also be conducted to investigate potential sources of bias, including pleiotropy. A causal effect of BMI on eczema and psoriasis will uncover the pathomechanisms involved, also informing the prevention and treatment of disease.

Keywords: mendelian randomization, causality, eczema, psoriasis, bmi
Uncoupling “healthy obesity” from “unhealthy obesity” genetic variants to provide a tool for Mendelian randomization studies

Presenter: Dr Hanieh Yaghootkar
Institute of Biomedical and Clinical Science, University of Exeter Medical School, Royal Devon and Exeter Hospital, Barrack Road, Exeter, EX2 5DW, UK. Email: h.yaghootkar@ex.ac.uk

Co-authors: Hanieh Yaghootkar, Yingjie Ji, Jess Tyrrell, Sam E. Jones, Robert Beaumont, Marcus A. Tuke, Kate S. Ruth, Rachel M. Freathy, Anna Murray, Mike N. Weedon, Andrew R. Wood, Timothy M. Frayling
Institute of Biomedical and Clinical Science, University of Exeter Medical School, Royal Devon and Exeter Hospital, Barrack Road, Exeter, EX2 5DW, UK

Abstract
It is not clear if higher adiposity itself, or the adverse metabolic effects of higher adiposity, increases susceptibility to non-metabolic diseases. We aimed to identify genetic variants associated with “healthy obesity” to test the effect of more favourable higher BMI on non-metabolic outcomes.

We performed a GWAS to identify genetic variants associated with body fat % in 220,000 individuals from UK Biobank and published studies. Next, we used a multivariate model that jointly tested the effects of these variants on biomarkers of "healthy obesity": higher HDL-C, adiponectin and SHBG; and lower Triglycerides, fasting insulin and ALT. We defined "healthy obesity" alleles as those which had a multivariate p < 5x10^-4 and were associated with the multivariate model stronger than body fat % alone. The latter criterion was set to uncouple "healthy obesity" from "unhealthy obesity" alleles. To validate the variants, we calculated the protective effect of "healthy obesity" genetic score against metabolic diseases. We performed Instrumental Variable analysis (IV) to estimate the causal effect of "healthy obesity" on non-metabolic outcomes available in UK Biobank.

We identified 11 "healthy obesity" genetic variants. The multivariate model was able to uncouple "healthy obesity" (e.g. PPARG allele) from "unhealthy obesity" alleles (e.g. FTO allele). In the genetic score analysis, each allele was associated with a 0.068 Kg/m2 higher BMI (p = 2x10^-62) but lower risk of type 2 diabetes (0.96 OR; 2x10^-21), lower heart disease (0.98 OR; 6x10^-9) and lower systolic blood pressure (-0.156 mmHg; 1x10^-7). The IV analysis did not support causal association with depression, bone mineral density, bowel disease, cancer (any type) and osteoporosis (all p > 0.3), but we could not exclude observational estimates of the associations.

The identification of more "healthy obesity" variants will provide more power to test the role of higher adiposity without its adverse metabolic effects.

Keywords: healthy obesity, multivariate test, uk biobank, body fat percentage
Session 19: Invalid instruments
Chair: Dr Nic Timpson, MRC Integrative Epidemiology Unit, UK
Mendelian randomisation across a range of exposure

Presenter: Prof Frank Dudbridge
Dept Health Sciences, University of Leicester. Email: frank.dudbridge@leicester.ac.uk

Co-authors: Frank Dudbridge [1]
Dept Health Sciences, University of Leicester

Abstract
I review some applications in which it is of interest to evaluate the gene-outcome association within subgroups of individuals having certain values of the exposure. In the first situation, subgroups with constant exposure are identified, in which there can be no gene-exposure association and therefore any gene-outcome association must arise from pleiotropy; lack of the latter association gives support for the exclusion restriction assumption. Examples include non-smokers, who by definition have no smoking consumption, and Asian females, who by culture tend to consume no alcohol. However one must be careful to ensure that such constant exposures are caused by, not a cause of, the subgroup definition. A second, related situation concerns inference of non-linear causal effects, in which it is useful to estimate causal effects within distinct ranges of exposure. Stratifying analyses directly on exposure opens a collider bias through unobserved confounders. This may be avoided by stratifying instead on potential values of the exposure for a fixed genetic value. While this approach can reveal evidence of non-linear effects, its current limitations include the need for individual-level data, assumption of homogeneous genetic effects, and problems of interpretation for the resulting estimates of localised causal effects.

Keywords:
collider bias, non-linear effect, stratified analysis
Linear Slichter Regression: using gene-environment interactions to correct for pleiotropic bias in Mendelian randomization analyses

Presenter: Mr Wes Spiller
MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK. Email: wes.spiller@bristol.ac.uk

Co-authors: Wes Spiller [1], Jack Bowden [1], David Slichter [2], George Davey Smith [1]
[1] MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
[2] Economics Department, Binghamton University, Binghamton, New York, USA.

Abstract
Mendelian randomization is a popular approach to examining causal relationships in epidemiology, however, it remains controversial as causal effect estimates may exhibit pleiotropic bias. Methods such as MR-Egger regression have proven effective in a two-sample summary MR setting, but require the use of many genetic instruments. Such methods are often inappropriate for individual level data, where it is standard practice to combine variants into allelic scores to overcome weak instrument bias. Recent work (Slichter 2014, Cho et al, 2015) has highlighted the potential use of gene-environment interactions in detecting pleiotropic bias, by identifying target population subgroups where treatment assignment is independent of instrument status. For such groups one would expect the instrument and outcome to be independent, whilst an observed association serves as evidence of pleiotropy. We present linear Slichter regression (LSR) as a formal statistical method to identify and correct for pleiotropic bias using gene-environment interactions. This is achieved within a linear regression framework, regressing the instrument-outcome associations upon the instrument-treatment associations for each level of the interaction covariate. This yields a corrected causal effect estimate, and an estimate of average pleiotropic effect. The technique can be applied with a single instrument. Moreover, the instrument and treatment need not be strictly independent for any covariate group, as long as there is some variation in the strength of dependency between groups. We illustrate the effectiveness of LSR using simulations and data from UK Biobank to assess the role of alcohol consumption upon systolic blood pressure.

Keywords:
mendelian randomization, invalid instruments, pleiotropy, slichter regression, gene-environment interaction.
Session 20: **Applications**
Chair: *Dr Sarah Lewis, MRC Integrative Epidemiology Unit, UK*
Causal epigenome-wide association study identifies CpG sites with evidence of a mediatory role in cardiovascular disease risk

Presenter: Tom Richardson
MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN, UK. Email: Tom.G.Richardson@bristol.ac.uk

Co-authors: Tom G Richardson [1], Jie Zheng [1], George Davey Smith [1], Nicholas J Timpson [1], Tom R Gaunt [1], Caroline L Relton [1], Gibran Hemani [1]

1 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN, UK

Abstract
The extent to which genetic influences on complex traits and disease are mediated by changes in DNA methylation levels has not been systematically explored. We developed an analytical framework that integrates genetic fine mapping and Mendelian randomization with epigenome-wide association studies to evaluate the causal relationships between methylation levels and 14 cardiovascular disease traits.

We identified 10 genetic loci known to influence proximal DNA methylation which were also associated with cardiovascular traits (P < 3.83 x 10^{-08}). Bivariate fine mapping suggested that the individual variants responsible for the observed effects on cardiovascular traits at the ABO, ADCY3, ADIPOQ, APOA1 and IL6R loci were likely mediated through changes in DNA methylation. Causal effect estimates on cardiovascular traits ranged between 0.109-0.992 per standard deviation change in DNA methylation and were replicated using results from large-scale consortia.

Functional informatics suggests that the causal variants and CpG sites identified in this study were enriched for histone mark peaks in adipose tissue and gene promoter regions. Integrating our results with expression quantitative trait loci data we provide evidence that variation at these regulatory regions is likely to also influence gene expression at these loci. Lastly, we demonstrate the further potential of our framework using two-sample Mendelian randomization to evaluate the causal relationship between DNA methylation and complex traits using summary statistics.

Keywords:
epigenetics, dna methylation, mediation, cardiovascular disease, causal pathway
Can MR using structural equation modelling inform us about the fetal origins hypothesis?

Presenter: Dr Nicole Warrington
The University of Queensland Diamantina Institute, The University of Queensland, Translational Research Institute, Brisbane, 4102, Australia. Email: n.warrington@uq.edu.au

Co-authors: Nicole M Warrington [1] and David M Evans [1,2,3]
[1] The University of Queensland Diamantina Institute, The University of Queensland, Translational Research Institute, Brisbane, 4102, Australia
[2] MRC Integrative Epidemiology Unit, University of Bristol, Bristol BS8 2BN, UK
[3] School of Social and Community Medicine, University of Bristol, Bristol BS8 2BN, UK

Abstract
Objectives: Low birth weight is robustly associated with increased risk of a range of cardio-metabolic diseases in later life, including type 2 diabetes, hypertension and dyslipidaemia. Recently, there have been investigations into whether this relationship is causal using Mendelian randomization. However, these studies have used only the original 7 known birth weight associated variants and failed to account for pleiotropy or potential effects contributed by the mother’s genotype on the intrauterine environment. The aim of our study was to assess whether the relationship between birth weight and systolic blood pressure was causal while accounting for these additional complexities.

Methods: We use a structural equation modelling approach in the UK Biobank with the 60 genome-wide significant variants for birth weight. Our method enables us to model both grand-maternal and offspring genotypes (which are absent in the UK Biobank) as latent factors, to estimate maternal and fetal effects of each genetic variant on birth weight, to estimate the direct effect of the variants on systolic blood pressure (i.e. pleiotropic effects) and the direct effects of grand-maternal variants on systolic blood pressure (i.e. the intrauterine environment) in the same statistical model.

Results: The results from this method suggest that there is no causal relationship between birth weight and systolic blood pressure, even though standard Mendelian Randomization techniques suggest a negative relationship. We will discuss the potential reasons for this discrepancy.

Conclusions: Although there are increasing numbers of publically available summary statistics from genome-wide association studies and well developed methods for applying these for Mendelian randomization analyses, the fetal origins hypothesis is one area that needs more complex methodologies to appropriately assess the causal relationship between birth weight and later life disease.

Keywords: structural equation modelling, birth weight, fetal origins, systolic blood pressure, intrauterine effect
Posters session
Even numbered posters will be formally presented on Tuesday 11\textsuperscript{th} July and odd numbered posters will be formally presented on Wednesday 12\textsuperscript{th} July.
**P1 A Note on the Use of Egger Regression in Mendelian Randomization Studies**

**Presenter:** Eric A.W. Slob M.Sc.  
*Erasmus University Rotterdam. Email: e.a.w.slob@ese.eur.nl*

**Co-authors:** Eric A.W. Slob [1,2], Patrick J.F. Groenen [1,3], A. Roy Thurik [1,2,4], Cornelius A. Rietveld [1,2]  
1 *Erasmus University Rotterdam Institute for Behavior and Biology, Erasmus School of Economics, Rotterdam, the Netherlands.*  
2 *Department of Applied Economics, Erasmus School of Economics, Rotterdam, the Netherlands.*  
3 *Econometric Institute, Erasmus School of Economics, Rotterdam, the Netherlands.*  
4 *Montpellier Business School, Montpellier, France.*

**Abstract**  
For a genetic variant to be a valid instrument in Mendelian randomization (MR) studies, three assumptions need to hold: (i) The genetic variant is associated with the exposure of interest (relevance assumption); (ii) The genetic variants is independent of all confounders (independence assumption); (iii) The genetic variants only effects the outcome through the exposure of interest (exclusion restriction). Without specific knowledge about the biological mechanisms affected by genetic variants, it is virtually impossible to prove that the exclusion restriction holds for a specific genetic variant. For example, genetic variants may have pleiotropic effects on both the exposure and the outcome through different biological pathways. Bowden and colleagues recently proposed to use Egger regression to correct for pleiotropic effects of genetics variants. Using simulations they show that MR-Egger provides unbiased estimates of causal effects if pleiotropy is balanced. Also in case of directional pleiotropy MR-Egger performs well, but only as long as the instrument-exposure and instrument-outcome associations are independent. This so-called “InSIDE” assumption is a relaxation of the exclusion restriction. Nevertheless, MR-Egger is currently often used as a robustness check on results obtained with regular Mendelian randomization analysis without proper discussion whether the InSIDE assumption holds. Here, we derive the bias of the MR-Egger estimator and show that if the InSIDE assumption does not hold the bias can be larger than the bias of the regular Inverse-Variance Weighting (IVW) estimator. We also show that in the illustrative analysis by Bowden and colleagues the InSIDE assumption does not hold, and that it is impossible in this example to evaluate whether the MR-Egger is less biased than the IVW estimator. Hence, we conclude that the use of MR-Egger as robustness check of IVW estimates is easily prone to unwarranted conclusions about the causal effect estimate.

**Keywords:**  
mendelian randomization, egger regression, exclusion restriction
P2 mrrobust: A Stata package implementing MR-Egger regression type analyses

Presenter: Dr Tom Palmer

Dept. of Maths and Statistics, Lancaster University. Email: tom.palmer@lancaster.ac.uk

Co-authors: Tom M Palmer [1], Wesley Spiller [2], Neil M Davies [2]

1 Department of Maths and Statistics, Lancaster University, Lancaster.
2 MRC Integrative Epidemiology Unit, University of Bristol, Bristol.

Abstract

MR-Egger regression analyses are becoming increasingly common in Mendelian randomization studies (MR) (Bowden et al. 2015). MR-Egger analyses use summary level data, as reported by genome-wide association studies. Such data is conveniently available from the MR-base platform (Hemani et al. 2016).

MR-Egger and related methods treat a multiple instrument MR analysis as a meta-analysis across the multiple genotypes. In the MR-Egger approach, bias from the pleiotropic effects of the multiple genotypes is treated as small study reporting bias in meta-analysis. They represent an important quality control check for any MR analysis incorporating multiple genotypes.

We implemented several of these methods (inverse-variance weighted [IVW], MR-Egger and weighted median approaches, as well as a relevant plot) in a package for Stata called mrrobust (pleiotropy robust methods for MR). There are also implementations of these methods in R (Yavorska and Burgess 2016).

mrrobust is freely available from https://github.com/remlapmot/mrrobust, which includes instructions on how to install the package from within Stata. We plan to add features overtime.

References


Yavorska O, Burgess S. MendelianRandomization: Mendelian Randomization Package. 2016, version 0.2.0.
https://CRAN.R-project.org/package=MendelianRandomization

Keywords:
mr-egger, mrrobust, multiple genotypes, pleiotropy, stata,
P3 Mental and Physical Wellbeing: Testing Causal Pathways with Mendelian Randomisation

Presenter: Robyn Wootton
University of Bristol, IEU and School of Experimental Psychology. Email: robyn.wootton@bristol.ac.uk

Co-authors: Robyn E Wootton12, Amy Taylor12, Oliver S P Davis2, George Davey-Smith2, Marcus Munafo12, Nicholas J Timpson2, Claire M A Haworth12
1 School of Experimental Psychology, University of Bristol
2 MRC Integrative Epidemiology Unit, University of Bristol, Bristol

Abstract
Observational evidence suggests a bi-directional association between mental and physical wellbeing. Research suggests high mental wellbeing predicts better cardiovascular health and a healthier BMI as well as better health predicting greater wellbeing. This study used Mendelian Randomisation (MR) to investigate whether or not these associations are causal. We hypothesised a bi-directional, causal effect of BMI and cardiovascular health on subjective wellbeing. Two-sample MR was conducted using summary GWAS statistics on MR BASE. No causal effect of wellbeing was found on the outcomes. However, there was a significant effect of BMI on subjective wellbeing. This effect did not show evidence of directional pleiotropy. We are currently conducting further sensitivity analyses including a one-sample MR in UK Biobank to get around problems of sample overlap. We are also planning to conduct MR analyses which allow for non-linear effects. This is because previous research between BMI and depression has suggested that the association is non-linear such that extremely high and extremely low BMI both predict higher instance of depression. Our results to date suggest a causal pathway from higher BMI to lower wellbeing, which has significant implications for mental health given the ever-rising levels of obesity. We speculate that this causal pathway involves social mediated traits, such as self-esteem and bullying.

Keywords:
subjective wellbeing, cardiovascular health, body mass index
Abstract
This poster will illustrate how the conceptual and analytic advances offered by Mendelian Randomization, and the opportunities afforded by the rapid development of digital and biological resources, could contribute improved evidence and methods to healthcare policy, health technology appraisal and health economics more generally.

The evaluation of methods to optimise management of chronic health conditions requires accurate estimates of the effect of these conditions on healthcare cost and quality of life. Current methods to produce these estimates tend to rely on observational methods that are prone to bias. Randomized controlled trials provide higher quality evidence, but are limited in their ability to provide information beyond the end of trial follow-up, and are often challenged by issues of feasibility, cost, and ethics. New ways of generating evidence of this nature are now much needed.

Mendelian Randomization allows inferences to be made about the causal effects of health conditions on outcomes by treating common genetic variants as instrumental variables. Already well established in genetic epidemiology, the possible uses for Mendelian Randomization in healthcare policy and technology appraisal are numerous and growing but remain largely unexploited. Datasets with linked genotypic, cost and quality of life information are now becoming available, and will facilitate the analysis of cost and quality of life data as outcomes within this methodological framework.

Potential applications for novel causal evidence from Mendelian Randomization study designs will be considered in relation to policy making, cost-effectiveness appraisals of healthcare interventions and technologies, the evaluation of screening and preventative programmes, and the setting of research priorities. These applications, and their possible pitfalls, will be described by reference to a number of prevalent chronic conditions, including obesity and coronary artery disease.

Keywords:
economics, policy, costs, quality of life,
P5 The identification of bi-directional association using Mendelian randomization may offer evidence of shared genetic aetiology rather than bi-directional causal effect.

Presenter: Dr Dipender Gill  
*Imperial College London. Email: dipender.gill@imperial.ac.uk*

Co-authors: Dipender Gill [1]
[1] Department of Clinical Pharmacology and Therapeutics, Imperial College London, London, W2 1NY

Abstract
Objective: Bi-directional Mendelian randomization (MR) has been offered as a method for investigating bi-directional causal effect in the presence of an observed association between two phenotypes. The objective of this study is to investigate whether such an approach should also be considered as an investigation of shared genetic aetiology.

Methods: Two bi-directional MR studies are considered; these explore the association between cannabis use and schizophrenia risk, and obesity and vitamin D status, respectively. The results of these studies are reviewed, paying particular attention to the plausibility of shared genetic aetiology as an explanation for the identified associations.

Results: The bi-directional MR study of cannabis use and schizophrenia risk showed that the instruments for cannabis use are associated with schizophrenia risk and that the instruments for schizophrenia risk are associated with cannabis use. While the published study interprets these results to suggest bi-directional causal effect, these findings may also be explained by a shared genetic aetiology for cannabis use and schizophrenia risk. The bi-directional MR study of obesity and vitamin D status only identified an association between the instruments for obesity and vitamin D status. The lack of association between the instruments for vitamin D status and obesity is not suggestive of shared genetic aetiology as an explanation for the findings of this study.

Conclusion: The findings of bi-directional MR should be interpreted in light of the possibility that shared genetic aetiology may explain any identified bi-directional association. Shared genetic aetiology in itself does not infer causal effect.

Keywords: mendelian randomization, bi-directional, shared genetic aetiology, causal effect
Abstract
Characterizing the causality between exposure and outcome is an important task in various economical and biological studies. Mendelian Randomization, specially, uses genetic variants as instruments to measure causal effects in epidemiological research. However, conventional instrumental variable methods rely on the assumptions of valid instruments, which may not be true in real concerns.

In this paper, we adopt the Bayesian framework and use hierarchical empirical Bayes models to incorporate the invalid effects of instruments. Our model also allows a fraction of the instruments to be invalid by specifying a Gaussian mixture prior. The theoretical performance and algorithm implementations will be derived and explained. The proposed method is demonstrated in various simulation settings and on real datasets with great accuracy.

Keywords:
mendelian randomization, instrumental variable analysis, hierarchical model, empirical bayes
Assessing the causal role of adiposity on disordered eating in childhood, adolescence and adulthood: a Mendelian randomization analysis

Presenter: Miss Zoe Reed
MRC Integrative Epidemiology Unit, School of Social and Community Medicine, Faculty of Health Sciences, University of Bristol, Bristol, United Kingdom. Email: zoe.reed@bristol.ac.uk

Co-authors: Zoe E. Reed[1], Nadia Micali[2,3], Cynthia M. Bulik[4,5,6], George Davey Smith[1], Kaitlin H. Wade[1]

1 MRC Integrative Epidemiology Unit, School of Social and Community Medicine, Faculty of Health Sciences, University of Bristol, Bristol, United Kingdom
2 University College London (UCL) Institute for Child Health, 30 Guilford Street, London, WC1N 1EH, United Kingdom
3 Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, US
4 Department of Psychiatry, University of North Carolina, Chapel Hill, NC, USA, 27599-7160
5 Department of Nutrition, University of North Carolina, Chapel Hill, NC, USA, 27599-7160
6 Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Abstract
Observational studies suggest that higher body mass index (BMI) is associated with an increased risk of disordered eating patterns. However, the direction of causality remains uncertain. The aim of this study was to use Mendelian randomization (MR) to assess the causal direction between BMI and disordered eating patterns in childhood and adolescence in the Avon Longitudinal Study of Parents and Children (ALSPAC, N=4473). The disordered eating patterns used can be interpreted on the standard deviation (SD) scale, with a higher score indicating higher levels of engaging in that disordered eating pattern. In addition, we used a two-sample MR approach and publically available data to assess the causal direction of association between BMI and eating disorders (EDs), specifically bulimia nervosa (BN) and anorexia nervosa (AN), in adulthood. Results indicated that higher BMI at age 7 causes higher levels of binge eating/overeating in males (difference in binge eating/overeating score per unit increase (kg/m2) in BMI: 0.30SD; 95% CI: 0.20, 0.40; P=9.17x10^-09) and females (difference: 0.17SD; 95% CI: 0.09, 0.25; P=4.98x10^-05) and both weight/shape concern and weight control behaviours in males (difference: 0.39SD; 95% CI: 0.29, 0.50; P=1.40x10^-12) and females (difference: 0.32SD; 95% CI: 0.24, 0.41; P=1.07x10^-12), as well as food restriction in males at age 13 (difference: 0.09SD; 95% CI: 0.0004, 0.19; P=0.05). Furthermore, our results suggested that a higher level of binge eating/overeating in males at age 13 causes higher BMI at age 17 (difference in mean BMI per SD increase in the binge eating/overeating score: 4.11kg/m2; 95% CI: 0.25, 7.96; P=0.04). We found no causal effect of BMI on BN or AN (or the reverse) in adulthood. Future research should investigate ways to break the causal chain between BMI and disordered eating in childhood and adolescence. Such work may help to inform and prevent disordered eating problems in later adolescence.

Keywords:
body mass index, disordered eating, alspac, early life
P8 Appraising the Causal Relevance of DNA Methylation on Lung Cancer Risk

Presenter: Mr Thomas Battram
MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK. Email: tb13101@bristol.ac.uk

Co-authors: Thomas Battram [1], Rebecca Richmond [1], Laura Baglietto [2], Philip Haycock [1], Gibran Hemani [1], Florence Guida [3], Robert Carreras-Torres [4], Andrew Teschendorff [5], George Davey Smith [1], Paolo Vineis [3], Gianluca Severi [2], Mattias Johansson [4], Caroline Relton [1]

1 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
2 Centre for Epidemiology and Biostatistics, University of Melbourne, Melbourne, Australia
3 MRC-PHE Centre for Environment and Health, Imperial College London, London, UK
4 International Agency for Research on Cancer, Lyon, France
5 UCL Cancer Institute, University College London, London, UK.

Abstract

Lung cancer causes the most cancer-related deaths. DNA methylation changes have been implicated during carcinogenesis and clonal expansion and, in peripheral blood, have been identified in relation to lung cancer risk. These epigenetic marks are sensitive to reverse causation, being affected by cancer processes, and associations are prone to confounding. Thus, the observational analysis conducted so far gives little indication of the causal nature of these methylation changes.

We performed a fixed effects meta-analysis of epigenome wide association studies of lung cancer in four prospective cohorts (918 cases and controls), which identified differential methylation at 16 CpG sites (FDR<0.05).

Two-sample Mendelian randomisation (MR) was used to analyse the causal effect of this differential methylation. 14 CpG sites were available for instrumentation using methylation quantitative trait loci (meQTL) (n=837).

Genetically increased methylation at multiple CpG sites correlated with lower log(ORs) in different lung cancer subtypes (total N=75882), including associations at cg03636183 in F2RL3 with overall lung cancer (-0.07 [95% CI, -0.14 to -0.01]), and cg23771366 in PRSS23 with squamous cell lung cancer (-0.11 [95% CI, -0.22 to -0.01]). However, none of the associations was robust to multiple testing correction.

Additionally, methylation levels at these sites within lung cancer tissue and adjacent healthy tissue was examined. This revealed a decrease in methylation at cg23771366 in PRSS23 to be associated with lung adenocarcinoma (P=4x10^-9) and squamous cell lung cancer (P=0.033), but no concordant changes were observed for the other sites where there was suggestive evidence for causal effects.

This study provides a framework for evaluating causal relationships between DNA methylation variation and cancer risk using MR. The data presented provide limited evidence that DNA methylation detected in peripheral blood plays a major causal role in lung cancer development.

Keywords:
lung cancer, methylation, ewas, mr
The causal relationship between Body Mass Index and cardiac structure and function in young adults using Mendelian Randomisation

Presenter: Hannah Taylor
UCL. Email: hannah.taylor.15@ucl.ac.uk

Co-authors: Hannah C M Taylor [1], Dr Kaitlin Wade [2], Dr Chloe M Park [1], Dr Nicholas J Timpson [2], Dr Abigail Fraser [2], Dr Laura Howe [2], Prof. Debbie-Lawlor [2], Prof. George Davey Smith [2], Prof. Nishi Chaturvedi [1], Prof. Alun D Hughes [1]

1 Cardiometabolic Phenotyping Group, Institute of Cardiovascular Science, Faculty of Pop Health Sciences, University College London, London, W1T 7HA.
2 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN.

Abstract

a) Introduction: Increased BMI is associated with adverse cardiac effects in observational studies; however, bias and reverse causation prevent causal inference. We determined causal relationships between higher BMI and cardiac measures in young adults using Mendelian Randomisation.

b) Methods: We estimated the causal effect of higher body mass index (BMI) on measures of cardiac structure and function using Mendelian randomisation (MR) from 1571 young adults in the Avon Longitudinal Study of Parents and Children (age 17.8y; 51% male). The causal effect of BMI on detailed cardiovascular outcomes was estimated using a Genetic Risk Score (comprising 97 independent genetic variants) as an instrumental variable. Results are presented as mean differences in each cardiovascular outcome (95% confidence interval) per unit increase in BMI (kg/m2).

c) Summary of Results: Higher BMI caused higher systolic and diastolic blood pressure (BP) (0.8mmHg (0.3, 1.8) and 0.3mmHg (0.0002, 0.6), respectively), increased left atrial size (0.08cm (0.05, 0.11)), increased end diastolic volume (2.0ml (0.6, 3.5)), left ventricular mass indexed to height2.7 (1.07g/m2.7 (0.62, 1.52)) and cardiac output (79.0ml/min (18.2, 139.9)) (p values all <0.05). The elevated cardiac output was likely attributable to a higher stroke volume (1.0ml (0.1, 1.9)) as results showed no evidence of a causal effect of BMI on heart rate (-0.1bpm (-0.5, 0.4)). Other cardiac measures, including total arterial compliance, systemic vascular resistance and measures of systolic and diastolic function, were only marginally or not affected.

d) Conclusions: In young adults increased BMI causes elevated left ventricular mass, atrial size and BP. Higher BP is wholly attributable to elevated stroke volume at this age. Other associations observed in observational analyses may represent confounding or reverse causation.

Keywords:
cardiovascular health, young adulthood, mendelian randomisation
P10 Causal association of body mass index with estimated glomerular filtration rate: One- and two-sample Mendelian randomization studies

Presenter: Dr Minsu Park

Seoul National University College of Medicine. Email: minsu.park51@gmail.com

Co-authors: Minsu Park [1], Kwangsoo Kim [2], Hyung-Jin Yoon [1]

[1] Department of Biomedical Engineering, Seoul National University College of Medicine, Seoul, Republic of Korea
[2] Division of Clinical Bioinformatics, Biomedical Research Institute, Seoul National University Hospital, Seoul, Republic of Korea

Abstract

Estimated glomerular filtration rate (eGFR), which is calculated using the Chronic Kidney Disease Epidemiology Collaboration equation based on serum creatinine level, age, ethnicity, and gender, has been used to measure the level of kidney function and determine the stage of kidney disease. Body mass index (BMI) is widely used measure for estimating body fat mass based on height and weight, and elevated BMI is associated with decreased eGFR levels, but whether this association is causal remains uncertain. To investigate causal effects on one-sample setting, we performed Mendelian randomization (MR) analyses in the Ansan-Ansung cohort of 7,829 participants at baseline obtained from Korea Centers for Disease Control and Prevention. In this study, 75 out of 2,147 single nucleotide polymorphisms (SNPs), extracted from genome-wide association studies (GWAS; Plink1.9) with 64K SNPs, were satisfied three assumptions to be instrumental variables. Also, to perform two-sample MR study, summarized data on the association of these variants on traits of interest were extracted from different public databases. We applied multifarious estimation approaches of causal effects not only some invalid some valid instrumental variables estimator (sisVIVE) and two-stage least squares (TSLS) with an unweighted and a weighted allele score on the one-sample MR study but also inverse variance weighted (IVW), MR Egger regression, and weighted median estimator on the two-sample setting. Using multivariate regression analysis adjusted 13 founders, BMI was associated with a -0.377 mL/min/1.73m² lower eGFR level. However, employing aforementioned methods for estimating causal effects on one- and two-sample settings, our results do not support the causal effect of BMI on eGFR levels and it may be caused by the existence of confounders or these combinations on the path of BMI to eGFR.

Keywords:
body mass index, estimated glomerular filtration rate, one- and two-sample mendelian randomization, causal inference, instrumental variables
P11 A Mendelian randomization study of vitamin B12 and its relationship with body mass index, cardiometabolic risk factors and disease

Presenter: Gunn-Helen Moen
Department of Endocrinology, Morbid Obesity and Preventive Medicine, Oslo University Hospital, Oslo, Norway & Faculty of Medicine, University of Oslo, Institute of Clinical Medicine, Oslo, Norway.
Email: g.h.o.moen@studmed.uio.no

Co-authors: Gunn-Helen Moen [1, 2], Elisabeth Qvigstad [1], Kåre I Birkeland [1,2,3], Christine Sommer[1]
1 Department of Endocrinology, Morbid Obesity and Preventive Medicine, Oslo University Hospital, Oslo, Norway,
2 Faculty of Medicine, University of Oslo, Institute of Clinical Medicine, Oslo, Norway,
3 Department of Transplantation Medicine, Oslo University Hospital, Oslo, Norway

Abstract
Low serum levels of vitamin B12 (vB12) have been associated with increased body mass index (BMI) and adverse metabolic outcomes in several observational studies. However, it is unclear if the associations reflect a causal effect of vB12 on adiposity or latent confounding.

In this study we investigated a possible causal relationship between vB12 and BMI, leptin (reflecting fat mass), height and hip circumference (reflecting body size) and several adverse cardiometabolic outcomes such as type 2 diabetes and cardiovascular disease. By using publicly available summary results data we have large sample sizes and sufficient power to explore these relationships.

We selected 11 variants robustly associated with vB12 (rs2336573 (CD320), rs1131603 (TCN2), rs3742801 (ABCD4), rs2270655 (MMAA), rs12272669 (MMACHC), rs7788053 (FUT6), rs602662 (FUT2), rs1801222 (CUBN), rs41281112 (CLYBL), and rs1141321 (MUT)) from a previous genome-wide association study of 45576 individuals and performed two sample MR analyses of vB12 with BMI and a variety of other cardiometabolic risk factors and diseases. To investigate the robustness of our results we performed sensitivity analyses using MR Egger regression, weighted median and mode estimators, and analyses excluding the variant at FUT2 due to potential pleiotropy. We will discuss these results and our thoughts on the potential of B12 interventions to decrease the risk of cardiometabolic disease.

Keywords:
mendelian randomization, vitamin b12, bmi, cardiometabolic risk
P12 The effect of TSH and fT4 on bone mineral density: A two-sample Mendelian randomization study

Presenter: Dr Nicolen van Vliet
Department of Internal Medicine, section of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands. Email: n.a.van_vliet@lumc.nl

Co-authors: Nicolen A van Vliet [1], Raymond Noordam [1], Diana van Heemst [1]
1 Department of Internal Medicine, section of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands

Abstract
Thyroid hormones are crucial for bone homeostasis. In adults with hyperthyroidism lower bone mineral density (BMD) and higher fracture risk was found. Interestingly, a similar pattern of bone fragility was found in patients with subclinical hyperthyroidism and in the femoral neck of post-menopausal women with thyroid status at the upper end of the normal range. However, the respective roles of thyrotropin (TSH) and free thyroxine (fT4) cannot be resolved from these observational data due to the reciprocal relationship between thyroid hormones and TSH. To investigate whether circulating levels of TSH or fT4 in the normal range are causally related to BMD, we conducted a two-sample Mendelian randomization study. In this study we used 20 genetic variants for circulating levels of TSH and 3 genetic variants for circulating levels of fT4 previously identified in a genome-wide association study. For the analyses, we used summary level data of a genome-wide association study on BMD of the femoral neck (N=32735), the lumbar spine (N=28498) and the forearm (N=8143) in cohorts of European descent from the GEFOS Consortium. Associations between the combined genetic instruments and the BMD traits were modeled using Inverse Variance Weighted, MR Egger and Weighted Median analyses. We found no evidence for an association between genetically determined one SD increase in circulating TSH levels and femoral neck BMD (IVW +0.003 SD (95%C.I. -0.048;0.053)), lumbar spine BMD (IVW +0.010 SD (95%C.I. -0.049;0.069)), or forearm BMD (IVW -0.020 SD (95%C.I. -0.143;0.103)) nor between genetically determined one SD increase in circulating fT4 levels and femoral neck BMD (IVW -0.038 SD (95%C.I. -1.318;1.243)), lumbar spine BMD (IVW +0.024 SD (95%C.I. -2.101;2.150)), or forearm BMD (IVW +0.017 SD (95%C.I. -0.490;0.565)). The results did not materially change in the other analyses. Therefore, we conclude that neither circulating levels of TSH nor of fT4 are causally implicated in BMD.

Keywords: bone mineral density, two-sample mendelian randomization, tsh, ft4
P13 The effect of lung function and coronary artery disease risk: A two-sample Mendelian randomization study

Presenter: Dr Shiu Lun Au Yeung
School of Public Health, Li Ka Shing Faculty of Medicine, the University of Hong Kong, HKSAR, China. Email: ryanaysl@connect.hku.hk

Co-authors: Shiu Lun Au Yeung [1], Maria C Borges [2], Debbie A Lawlor [2,3]
1 School of Public Health, Li Ka Shing Faculty of Medicine, the University of Hong Kong, HKSAR, China
2 School of Social and Community Medicine, University of Bristol, Bristol, BS8 2BN
3 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN

Abstract
Background: Lung function, assessed by forced expiratory volume in 1 second (FEV1) and forced vital capacity (FVC), is inversely associated with coronary artery disease (CAD), but these associations could be due to confounding by height and smoking, or reverse causality. We conducted a two-sample Mendelian randomization study, using publicly available data from relevant genome wide association studies (GWAS), to test evidence for a causal effect of FEV1 or FVC on CAD.

Methods: We used the most recent GWAS on lung function to extract genetic instruments related to FEV1 and FVC (n=94,162). Data on the association between genetic instruments and CAD were obtained from CARDIoGRAMplusC4D 1000 Genomes based GWAS (60,801 CAD cases and 123,504 controls). We used inverse variance weighting (IVW) with fixed effects to obtain the effects between FEV1, FVC and CAD. Sensitivity analyses included the weighted median method and MR-Egger method.

Results: Using IVW, each standard deviation (SD) greater FEV1 led to a relative reduction in CAD of 32% (Odds ratio (OR): 0.68 per SD; 95% CI: 0.57 to 0.81). FVC did not appear to affect CAD (OR: 0.91 per SD; 95% CI: 0.71 to 1.16). Estimates were similar when we used weighted median method or MR-Egger method.

Conclusion: Our study supports the causal role of FEV1, but not FVC, on CAD.

Keywords:
lung function, coronary artery disease, mendelian randomization
P14 Evaluating the causal association of vitamin D with pregnancy-induced hypertension and pre-eclampsia using a Mendelian Randomization approach

Presenter: Dr Maria C. Magnus
MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, United Kingdom; School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom; Division of Mental and Physical Health, Bristol, United Kingdom. Email: maria.christine.magnus@bristol.ac.uk

Co-authors: Maria Magnus [1,2,3], Kozeta Miliku [4], Christian Page [3,5], German Tapia [3], Debbie A Lawlor [1,2], Janine F Felix [4], Per Magnus [6], Wenche Nystad [3], Lars Christian Stene [3], Vincent WV Jaddoe [4], Carolina Bonilla [1,2], Abigail Fraser [1,2]

1 MRC Integrative Epidemiology Unit at the University of Bristol, Bristol
2 School of Social and Community Medicine, University of Bristol, Bristol
3 Division for Mental and Physical Health, Norwegian Institute of Public Health, Oslo, Norway
4 Erasmus University Medical Centre, Rotterdam, the Netherlands
5 Oslo Centre for Biostatistics and Epidemiology, Oslo University Hospital, Oslo, Norway
6 Domain for Health Data and Digitalisation, Norwegian Institute of Public Health, Oslo, Norway

Abstract
Background: Higher vitamin D levels might reduce the risk of pre-eclampsia (PE). Unmeasured confounding is inherent in observational studies, and trials to date have been small and heterogeneous in terms of dose and duration of vitamin D supplementation.

Methods: We conducted a Mendelian randomization study to estimate the causal effect of 25-hydroxyvitamin D (25(OH)D) on any pregnancy-induced hypertension (PIH) and PE. A total of 8,623 women participating in ALSPAC, Generation R and MoBa who had a singleton pregnancy and were normotensive before pregnancy were included. Four single nucleotide polymorphisms (SNPs) in genes associated with vitamin D synthesis (rs10741657 and rs12785878) and metabolism (rs6013897 and rs2282679) were used as instruments, both separately and combined into synthesis, metabolism, and total genetic risk scores. The observational analysis used multinomial logistic regression, the Mendelian randomization analysis used instrumental variable probit regression, and a random effects meta-analysis was used to combine results from all cohorts.

Results: There was a threshold association between 25(OH)D and PE (177 cases) in the observational analysis, with an adjusted RR of 2.02 (1.07, 3.81) when comparing the women with a 25(OH)D level <25 nmol/L to those ≥75 nmol/L. There was no association with PIH (879 cases). rs2282679, in addition to the metabolism and total genetic risk scores, showed modest evidence of a positive association with PE per risk allele (i.e. allele associated with lower 25(OH)D levels). The Mendelian randomization analysis using the total genetic risk score as the instrument indicated that the estimates of the causal associations of 25(OH)D with PIH or PE were OR 1.08 (95% CI: 0.98, 1.18) and 0.90 (95% CI: 0.79, 1.04) per 10% increase in 25(OH)D, respectively.

Conclusion: We did not find strong evidence to support a causal role for 25(OH)D on the risk of PIH or PE.

Keywords: mendelian randomization, pre-eclampsia, pregnancy-induced hypertension, vitamin d
P15 Using Mendelian randomization to look for a causal relationship between iron status and severe infection in African children

Presenter: John M Muriuki
KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya. Email: jmuriuki@kemri-wellcome.org

1 KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya
2 Centre for Clinical Vaccinology and Tropical Medicine and the Jenner Institute Laboratories, University of Oxford, Oxford, UK
3 Department of Clinical Biochemistry, Oxford University Hospitals, Oxford, UK
4 Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK
5 Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK
6 MRC/Respiratory and Meningeal Pathogens Research Unit, University of the Witwatersrand, Johannesburg, South Africa
7 MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol, UK
8 MRC/UVRI Uganda Research Unit on AIDS, Entebbe, Uganda
9 London School of Hygiene & Tropical Medicine, London, UK
10 Department of Medicine, Imperial College, London, UK
11 Department of Paediatrics, University of Oxford, Oxford, UK

Abstract
Background: Iron deficiency is common in African children and is associated with impaired child development. Similarly, severe malaria and bacteraemia remain important public health problems in Africa. However, there is concern that improving iron status might increase the risk of infection in children since pathogens require iron for growth. In this study, we will apply Mendelian randomization (MR) to investigate whether iron status is causally associated with the risks of severe malaria and bacteraemia.

Methods: The study involves two steps. Firstly, in order to identify genetic variants for use as instrumental variables (IVs), we will conduct the first genome wide association study (GWAS) of iron status in African populations. We have measured a range of iron markers, including hepcidin, in 3867 children, 6 months to 5 years of age, enrolled in community cohorts in Kenya, Uganda, Burkina Faso and South Africa. Secondly, genetic variants, validated as IVs, will be genotyped in stored samples from large case-control studies of severe malaria (n=12,500 cases and similar number of controls from the Kilifi Biobank and the MalariaGEN Consortium) and of blood-culture confirmed bacteraemia (n=5000 cases and similar number of controls from the Kilifi Biobank) and MR analyses will be applied.

Results: The study is in progress. Preliminary data from our pilot study indicate that there are common genetic variants unique to African populations that alter iron status. Measurements of iron status have been completed and the GWAS is in progress. We anticipate finding novel genetic variants that reliably alter iron status in African populations that we can take forward for MR analyses to investigate the causal role of iron status in severe malaria and bacteraemia.

Conclusion: This study will apply the MR approach to a critically important public health problem for African children, the relationship between iron and infection.

Keywords: iron, malaria, bacteraemia, mr, children
P16 Association between Vitamin D and Colorectal cancer: A 2-Sample Mendelian Randomization Approach

Presenter: Miss Kawthar Al-Dabhani
Department of Biostatistics and Epidemiology, School of Public Health, Imperial College London, London, UK. Email: ka1813@ic.ac.uk

Co-authors: Konstantinos K Tsilidis[1,2], Marc Gunter[4], Ioanna Tzoulaki[1,2,3]
1 Department of Biostatistics and Epidemiology, School of Public Health, Imperial College London, London, UK; 2 Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece; 3 MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, UK; 4 Nutritional Epidemiology Group, International Agency for Research on Cancer

Abstract
Colorectal cancer (CRC) is the third most common cancer in the world with approximately 70-80% of the variation in risk potentially explained by environmental factors. Observational studies have reported that higher levels of serum vitamin D are associated with reduced risk of CRC. Mendelian randomization analysis can be used to provide more robust evidence for causality.

Summary-level data was obtained from a previously published study to ascertain the association estimates of the genetic variants (rs12785878, rs1993116, rs2282679, and rs17217119) associated with circulating concentrations of 25-hydroxyvitamin D. Individual-level data was obtained from 126,947 UK Biobank participants with available genetic data, in order to estimate the association between the aforementioned genetic variants with CRC risk. Two different approaches of Mendelian randomization for summary data were used in this study were: the inverse-variance weighted approach and the likelihood-based approach. These methods were applied to estimate causal inference of the association between the allele risk score, the metabolism allele score (DHCR7 and CYP2R1)), and the synthesis allele (CYP24A1 and GC)) with CRC risk.

Among the 126,947 UK Biobank participants, there were 712 prevalent and incident cases of CRC. No association was found for any of the four genetic variants of 25-hydroxyvitamin D and genetic risk scores with CRC risk in the UK Biobank population. There was no evidence of a causal association between the genetic variants of 25-hydroxyvitamin D and the risk of CRC (OR: 1.01, 95%CI: 0.97-1.04).

Our results do not provide evidence for a causal association between 25-hydroxyvitamin D concentration and the risk of CRC. This may have been due to the low power in this study as a result of the small number of CRC cases in the UK Biobank and the limitations of the Mendelian randomisation assumptions.

Keywords:
vitamin d, colorectal cancer, mendelian randomisation
P17 Altered amino acid metabolism for arginase activity in metabolic syndrome: the Korean Genome and Epidemiology Study (KoGES)

Presenter: BS Garam Jo
Department of Public Health Sciences, BK21PLUS Program in Embodiment: Health-Society Interaction, Graduate School, Korea University, Seoul 02841, Republic of Korea. Email: grcho94@korea.ac.kr

Co-authors: Garam Jo, So-Young Kwak, Yoonsu Cho, Min-Jeong Shin

Abstract
BACKGROUND: The upregulated arginase activity has been implicated in various metabolic disorders such as obesity, type 2 diabetes and endothelial dysfunction, mainly through the competition with endothelial NO synthase (eNOS) for L-arginine. The aim of the current study was to test if there are alterations in the ratios of amino acids reflecting arginase activity (i.e. ornithine/citrulline, proline/citrulline, and ornithine/arginine) in metabolic syndrome (MetS) in Korean adults, and that this change is associated with the components of MetS. METHODS: A total of 1,998 subjects without diabetes and cardiovascular disease at the Ansan-Ansung cohort (2005 to 2006) from the Korean Genome and Epidemiology Study (KoGES) were investigated. MetS was defined based on the National Cholesterol Education Program Adult Treatment Panel III. We used logistic regression models for statistical analysis. Multivariate logistic regression and correlation analysis was performed to indentify the association between metabolites and MetS components. RESULTS: The subjects with MetS had altered amino acid metabolite profiles, with higher ratios of ornithine/citrulline [OR 1.71, CI 1.15, 2.54, p=0.008], than those without MetS after adjusted for age, sex, body mass index, and smoking/drinking status. In addition, ornithine/arginine ratio reflecting arginase activity was significantly higher in MetS than in non-MetS after the same adjustment [OR 2.57, CI 1.69, 3.90, p<0.001], whereas no significant association in proline/citrulline was observed. These ratios were significantly correlated with the components of MetS, specifically positively with waist circumference (p<0.001), blood triglyceride (p<0.001), fasting glucose (p<0.001) and negatively with HDL-cholesterol (p=0.012). CONCLUSION: The altered ratios of the amino acids metabolites reflected the increased arginase activity in MetS, indicating a shift in the balance of arginase/nitric oxide synthase under metabolic abnormalities.

Keywords: metabolic syndrome, arginase activity, amino acids metabolism
P18 Phenotypic manifestation of genetic risk of multiple sclerosis in the general population

Presenter: Dr Ruth Mitchell
MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN. Email: r.mitchell@bristol.ac.uk

Co-authors: George Davey-Smith [1], Lavinia Paternoster [2]
[1] MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN
[2] MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN

Abstract
Multiple sclerosis (MS) an inflammatory demyelinating disorder of the central nervous system with diverse manifestations and comorbid conditions. Genome-wide association studies have identified over 110 non-MHC genetic variants that contribute to MS susceptibility. However, limited evidence explains how genetic risk for MS is manifest in the general population. To investigate this, we generated a weighted genetic risk score (GRS) based on the most up-to-date GWAS from the International Multiple Sclerosis Genetics Consortium. Regression analyses were used to assess whether the MS GRS associated with a range of phenotypes previously been linked to MS. In the Avon Longitudinal Study of Parents and Children (ALSPAC), in the children, the MS GRS is suggestively associated with a trait score for autism (Odds Ratio (OR) per SD increase in GRS 1.36; 95% CI, 1.11-1.62), severe allergy (OR per SD increase in GR, 0.81, 95% CI, 0.85-0.97), and weight at age 7 (OR per SD increase in GRS 0.13; 95% CI 0.02-0.25). In mothers, the MS GRS was suggestively associated with chronic fatigue syndrome (OR per SD increase in GRS 0.79; 95% CI 0.63-0.99, p=0.04); this remained associated in UK Biobank (OR per SD increase in GRS 1.12; 95% CI 1.03-1.22, p=0.01). The analysis performed in UK Biobank also showed that hypothyroidism (OR per SD increase in GRS 1.08; 95% CI 1.05-1.11) and eosinophil percentage (OR per SD increase in GRS 0.03; 95% CI, 0.02-0.041) are positively associated and monocyte percentage (OR per SD increase in GRS -0.034; 95% CI -0.049- -0.018) is negatively associated with the MS GRS. These findings suggest polygenic overlap between genetic variants associated with MS and these phenotypes that are a manifestation of this genetic risk. Increased understanding of the genetic overlap between different conditions helps to understand the underlying mechanisms of these diseases, and may also provide useful information into repositioning and off-target effects of current therapies.

Keywords:
genetic risk score, multiple sclerosis
P19 Metabolic signatures associated with the risk of type 2 diabetes in Korean adults: the Korean Genome and Epidemiology Study (KoGES)

Presenter: BS So-Young Kwak
Department of Public Health Sciences, BK21PLUS Program in Embodiment: Health-Society Interaction, Graduate School, Korea University, Seoul 02841, Republic of Korea. Email: joilfille@korea.ac.kr

Co-authors: So-Young Kwak, Garam Jo, Min-Jeong Shin
Department of Public Health Sciences, BK21PLUS Program in Embodiment: Health-Society Interaction, Graduate School, Korea University, Seoul 02841, Republic of Korea

Abstract
BACKGROUND: The identification of metabolic alterations in type 2 diabetes (T2D) is useful in elucidating pathophysiology and classifying high-risk individuals. In the present study, we prospectively examined the associations between serum metabolites and the risk of T2D in a Korean community-based cohort (Ansan-Ansung cohort). METHODS: Data were derived from 1,983 participants (43–74 years) without diabetes, cardiovascular disease and cancer at baseline whose metabolite profile data were available. Targeted metabolomics (186 metabolites) of fasting plasma samples including acylcarnitines, amino acids, amines, and phospholipids was performed by liquid chromatography/tandem mass spectrometry (LC-MS/MS) and flow injection analysis-tandem mass spectrometry (FIA-MS/MS). We estimated the multivariable-adjusted hazard ratios (HRs) and 95% confidence intervals (CIs) using Cox regression analysis. RESULTS: During the follow-up period of 8 years, we identified 273 cases of incident T2D. Of all the metabolites measured, 36 metabolites were significantly associated with T2D risk after correction for multiple testing (FDR p<0.05). Specifically, serum alanine, isoleucine, proline, valine and phosphatidylcholine diacyl (C32:0, C32:1, C34:1, C36:1, C36:5, C40:5) were significantly positively associated with T2D risk after adjusting for age, sex, energy intake, body mass index, physical activity, smoking, drinking, income, education and hypertension. In contrast, lyso-phosphatidylcholine acyl C18:2, and other glycerophospholipids were significantly inversely related to T2D risk after the same adjustment. They were further correlated with T2D relevant risk factors such as index for insulin resistance and insulin sensitivity. CONCLUSION: Our results indicate that metabolic alterations, including amino acids, and phospholipids are associated with the risk of T2D in Korean adults.

Keywords:
type 2 diabetes, metabolomics, alanine, phosphatidylcholine diacyl
P20 Bayesian Model Averaging to Derive Multi-SNP Mendelian Randomization Instruments from Meta-GWAS Summary Statistics

Presenter: Dr Apostolos Gkatzionis  
MRC Biostatistics Unit, University of Cambridge, Cambridge, CB2 0SR. Email: apostolos.gkatzionis@mrc-bsu.cam.ac.uk

[1] MRC Biostatistics Unit, University of Cambridge, Cambridge, CB2 0SR, [2] MRC Biostatistics Unit and Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, CB2 0SR, [3] MRC Biostatistics Unit, University of Cambridge, Cambridge, CB2 0SR.

Abstract  
Recently, there has been a growing trend of Mendelian Randomisation analyses based on summary data from large consortia meta-GWAS (Genome Wide Association Studies). These studies seek to leverage the power harnessed by meta-analysis of large numbers of individuals but are often complicated by the availability of summarised data only; typically one-at-a-time tests of each variant and trait. Despite the power of meta-GWAS to identify large numbers of variants robustly associated with a trait of interest, these Mendelian Randomization analyses are often based on a small number of variants, pruned for independence. The JAM algorithm (Joint Analysis of Marginal summary statistics, Newcombe, Conti and Richardson, 2016) is a recently proposed algorithm for identifying genetic variants associated with a specific trait in summary data from genetic association studies. The algorithm facilitates Bayesian model selection and model averaging via a Reversible Jump MCMC procedure. In this poster, we will present some early research in the direction of implementing the JAM algorithm in the context of Mendelian randomization. The purpose of model selection in this context is to identify those genetic variants that are more suitable for inclusion in a Mendelian randomization study with multiple instruments. Advantages of the JAM algorithm are that its implementation accounts for genetic correlations despite only requiring access to the marginal one-at-a-time summary statistics and scalability to build multi-SNP instruments averaged over large numbers of variants. Simulations will be presented to illustrate the use of the algorithm and some possible extensions will also be discussed.

Keywords: mendelian randomization, jam algorithm, model selection, summary statistics, multiple instruments.
P21 Associations of genetic determinants of serum vitamin B12 and folate concentrations with hay fever and asthma: a Mendelian randomization meta-analysis

Presenter: Dr. Tea Skaaby

Research Centre for Prevention and Health, Capital region of Denmark, Glostrup, Denmark. Email: tea.skaaby.01@regionh.dk

Co-authors: Tea Skaaby [1], Amy E. Taylor [2,3], Rikke K. Jacobsen [1], Line T. Møllehave [1], Nele Friedrich [1,4], Betina H. Thuesen [1], Daniel Mønsted Shabanzadeh [1,5], Lavinia Paternoster [2], Uwe Völker [6], Matthias Nauck [4], Henry Völzke [7], Marcus Munafò [2,3], Torben Hansen [8], Oluf Pedersen [8], Torben Jørgensen [1,9,10], Niels Grarup [8], Allan Linneberg [1,11,12]

1 Research Centre for Prevention and Health, Centre for Health, Capital Region of Denmark, Copenhagen, Denmark. 2 MRC Integrative Epidemiology Unit (IEU) at the University of Bristol, Bristol, UK. 3 UK Centre for Tobacco and Alcohol Studies, School of Experimental Psychology, University of Bristol, Bristol, UK. 4 Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Germany. 5 Bispebjerg University Hospital, Digestive Disease Center, Bispebjerg Bakke 23, DK-2400 Copenhagen. 6 Interfaculty Institute for Genetics and Functional Genomics, University Medicine and Ernst-Moritz-Arndt University Greifswald, Germany. 7 Institute for Community Medicine, University Medicine Greifswald, Germany. 8 The Novo Nordisk Foundation Center for Basic Metabolic Research, Section on Metabolic Genetics, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. 9 Department of Public Health, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark. 10 Faculty of Medicine, Aalborg University, Denmark. 11 Department of Clinical Experimental Research, Rigshospitalet, Glostrup, Denmark. 12 Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.

Abstract

Background: Studies of the effect of vitamin B12 and folate on risk of hay fever and asthma have shown inconsistent results that may be biased by confounding and reverse causation. We used a Mendelian randomization approach to examine a potential causal effect of vitamin B12 and folate on hay fever, asthma, and selected biomarkers of allergy by using eleven vitamin B12-associated single nucleotide polymorphisms (SNPs) and two folate-associated SNPs as un-confounded markers.

Methods: We included 162,736 participants from nine population-based studies including the UK Biobank. Results were combined in instrumental variable and meta-analyses and effects expressed as odds ratios (ORs) or estimates with 95% confidence interval (CI).

Findings: Instrumental variable analyses showed little evidence for associations between serum B12 and hay fever: OR=1.02 (95% CI: 0.98, 1.05), asthma: OR=0.99 (95% CI: 0.95, 1.04), allergic sensitization: OR=1.02 (95% CI: 0.74, 1.40), and log-transformed serum IgE: β=0.100 (95% CI: 0.096, 0.296) per 100 pg/ml B12. Similarly there was little evidence for association between serum folate and hay fever: OR=0.74 (95% CI: 0.45, 1.21), asthma: OR=0.80 (95% CI: 0.43, 1.49), allergic sensitization: OR=1.92 (95% CI: 0.11, 33.45), but some evidence for association with log-transformed serum IgE: β=2.00 (95% CI: 0.43, 3.58) per 10 ng/ml serum folate.

Conclusions: Except for some evidence of a positive association between serum folate and serum total IgE, our results did not support the hypothesis that serum levels of vitamin B12 and folate are causally related to hay fever, asthma, or biomarkers of allergy.

Keywords: serum specific ige, hay fever, asthma, allergic sensitization
P22 Vitamin D and oral and pharyngeal cancer risk: A Mendelian randomisation study.

Presenter: Mr Tom Dudding
MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN. Email: Tom.Dudding@bristol.ac.uk

Co-authors: Steve Thomas [1], Paul Brennan [2], Mattias Johannson [2], Richard Martin [3], Nic Timpson [3]
1 School of Oral and Dental Science, University of Bristol, Bristol, BS1 2LY
2 Genetic Epidemiology Group, International Agency Research on Cancer, 69372 Lyon
3 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN

Abstract
Circulating 25-hydroxyvitamin D may reduce the risk of oral and pharyngeal cancers. The most recent well-powered prospective study reported a strong protective effect in a European population. The evidence is not consistent however, and this may reflect the influence of confounding factors. The availability of genetic data in a large oral and pharyngeal cancer case-control collaboration (N=11,438) permits the use of genetic variants to (1) proxy 25-hydroxyvitamin D (where serum measures are not available) and (2) conduct Mendelian randomization to assess for a causal association.

5,454 Oral and pharyngeal cancer cases and 5,984 matched controls with genome wide genetic data were included in the study. 8 genetic variants identified in a recent genome-wide association study were used as instrumental variables for circulating 25-hydroxyvitamin D. A causal odds ratio for a 1 standard deviation increase in log 25-hydroxyvitamin D was estimated using a 2-stage estimator in individual level participant data. The first stage effect sizes were taken from the recent genome wide association study and odds ratio standard errors were adjusted using a bootstrap method.

In a meta-analysis of the 3 geographic regions 25-hydroxyvitamin D was not causally associated with a reduced risk of oral and pharyngeal cancers combined (odds ratio (OR) = 1.02, 95% CI: 0.84-1.24) or separately (oral: OR = 0.91, 95% CI: 0.72-1.15; pharynx: OR = 1.11, 0.86 – 1.43). Estimates by region are consistent for oral cancer but a risk increasing association was identified in the South American population (OR = 3.77, 95% CI 1.51 – 9.42). Causal estimates suggest any effect of vitamin D on the risk of oral and pharyngeal cancer is small smaller than the observational association. It is likely that this association is driven by a confounding effect of smoking. Further work is required to investigate the association between 25-hydroxyvitamin D and pharyngeal cancer in the South American population.

Keywords:
mendelian randomisation, cancer, vitamin d,
P23 Using activity-monitor derived measures of sleep patterns and bidirectional Mendelian randomisation to dissect the causal relationship between sleep and obesity related traits

Presenter: Dr Samuel E Jones
University of Exeter Medical School. Email: samuel.e.jones@exeter.ac.uk

Co-authors: Samuel E Jones [1], Robin Beaumont [1], Hanieh Yaghootkar [1], Marcus A Tuke [1], Katherine S Ruth [1], Rachel M Freathy [1], Anna Murray [1], Jessica Tyrrell [1], Vincent Van Hees [2], Melvyn Hillsdon [3], Timothy M Frayling [1], Andrew R Wood [1], Michael N Weedon [1]

1 Genetics of Complex Traits, University of Exeter Medical School, University of Exeter, UK
2 Netherlands eScience Center, Amsterdam, Netherlands
3 Department of Sport and Health Sciences, University of Exeter, UK

Abstract

Aims: There are strong associations between disrupted sleep patterns, and obesity and type 2 diabetes. The causal nature of each association is unclear. Recent studies have identified genetic variants associated with sleep pattern phenotypes, including chronotype and sleep duration. These studies were based on self-report data which can be subject to reporting biases. We used actigraphy data on 103,000 UK Biobank individuals to derive objective measures of sleep patterns and test causal links with obesity and type 2 diabetes.

Methods: We derived a range of sleep pattern measures from the activity monitor data using a validated sleep algorithm, including sleep duration, sleep mid-point and sleep efficiency, as well as L5 (least-active 5 hours) and M5 (most-active) timing. We performed GWAS of these measures in ~26,000 white-European UK Biobank participants with both genetic and valid actigraphy data. We then used known BMI and type 2 diabetes variants as instruments to test their causal effect on the derived sleep characteristics.

Results: There were strong observational associations between sleep patterns and BMI and type 2 diabetes (T2D). For example, sleep efficiency was strongly associated with BMI (P=2x10^-225) and T2D (P=2.9x10^-10) after adjusting for BMI. BMI and T2D genetic risk scores consisting of 69 and 55 published loci, respectively, were not associated with any of the derived sleep patterns, except for BMI and M5 timing (-12mins per S.D. increase in BMI; P=0.018).

We found no evidence that BMI or T2D are causal to any of the sleep and activity characteristics we derived from the actigraphy data (all MR-Egger P>0.05).

Conclusions: We performed the first large-scale analysis of objective measures of sleep patterns and used them in MR analyses. Our study shows that despite strong observational associations, causality cannot be determined. The full UK Biobank release will provide a better opportunity to assess causality.

Keywords:
uk biobank sleep actigraphy obesity
P24 Measures of adiposity and blood-based metabolomics: a comparison of findings from observational and Mendelian randomization analyses

Presenter: Ms Zammy Fairhurst-Hunter
University of Oxford. Email: zammy.fairhurst-hunter@keble.ox.ac.uk

Co-authors: Zammy Fairhurst-Hunter [1,3], Cecilia Lindgren [2], Iona Millwood [3], Robert Clarke [3], Matthew Arnold [3], Christiana Kartsonaki [3], Robin Walters [3], Mark McCarthy [1], Zhengming Chen [3], Michael V Holmes [3]
1 Wellcome Trust Centre for Human Genetics, University of Oxford
2 The Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford
3 Clinical Trial Service Unit & Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford

Abstract
Aim: To clarify the observational and causal associations of measures of adiposity with the metabolome.

Methods: 225 blood-based metabolites were measured in 4,577 participants of the China Kadoorie Biobank (CKB) using NMR-spectroscopy. Externally-weighted genetic risk scores (GRS) based on published GWAS-identified SNPs were derived for adiposity traits: body mass index (BMI), waist and hip circumference (WC, HC), waist-hip ratio (WHR), WC adjusted for BMI (WCadjBMI), HC adjusted for BMI (HCadjBMI) and WHR adjusted for BMI (WHRadjBMI). The causal effects of adiposity traits on the metabolome were estimated using two-stage least squares regressions. Observational effect estimates were estimated using linear regression.

Results: The proportions of variance of each adiposity trait explained by the GRS were: 1.7% for BMI (97 SNPs); 1.0% WC (50 SNPs); 0.7% HC (61 SNPs); 1.7% WHR (36 SNPs); 0.1% WCadjBMI (78 SNPs); 1.0% HCadjBMI (95 SNPs); 5.0% WHRadjBMI (51 SNPs).

Observationally, almost all adiposity traits were adversely associated with cardiometabolic risk factors such as; lipoprotein particle concentration, amino acids, fatty acids, inflammatory markers and markers of glucose metabolism (number of metabolites significantly associated at p<0.0007: BMI=197, WC=198, HC=191, WHR=195, WCadjBMI=159, HCadjBMI=6, WHRadjBMI=163).

Observational metabolomic profiles associated with adiposity were mirrored in the causal effect estimates for BMI, WC and HC (observational vs MR correlation: BMI=0.94, WC=0.95, HC=0.75). However, for WHR and the BMI adjusted traits the concordance was lower (correlation: WHR=0.31, WCadjBMI=0.49, HCadjBMI=0.17, WHRadjBMI=0.15).

Conclusions: Adiposity traits have wide-ranging causal relationships with the metabolome, often consistent with observational associations. Future GWAS in CKB will improve genetic variant selection for MR analyses, facilitating the dissection of the relationship of different adiposity traits with the metabolome.

Keywords: adiposity, metabolomics
P25 Assessing causality in the association between maternal pre-pregnancy obesity and child neurodevelopment: observational and mendelian randomization analyses

Presenter: Mariona Bustamante
ISGlobal. Email: mariona.bustamante@isglobal.org

Co-authors: Maribel Casas [1], Sílvia Alemany [1], David Martínez [1], Mariona Bustamante [1], Jordi Sunyer [1], Debbie A Lawlor [2] and Martine Vrijheid [1], on behalf of the CHICOS consortium
1 ISGlobal, Centre for Research in Environmental Epidemiology (CREAL); Universitat Pompeu Fabra (UPF); CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain
2 Medical Research Council Integrative Epidemiology Unit at the University of Bristol, Bristol, United Kingdom & School of Social and Community Medicine, University of Bristol, Bristol, UK

Abstract
Background: Observational studies have shown that high maternal pre-pregnancy body mass index (BMI) is associated with delayed infant neuropsychological development. It is unclear whether these associations are due to causal intrauterine effects or residual confounding by socioeconomic and related factors.
Objective: We aimed to determine whether maternal pre-pregnancy BMI causally influences child neurodevelopment by performing: (i) multivariable (MV) regression analysis in 24 birth cohorts (>100,000 mother/infant pairs), (ii) a negative control study comparing paternal to maternal MV associations, and (iii) Mendelian randomization (MR) analysis in 10 of these cohorts (>10,000 mother/infant pairs).
Methods: For the MR analyses we will use a maternal BMI weighted allele score using 32 SNPs and their beta coefficients from independent (of our cohorts) GWAS studies, as our instrumental variable. We will assess effects with infant cognition, psychomotor, and behaviour outcomes, including attention-deficit hyperactivity disorder and autism symptoms from infancy to 11 years of age. We will undertake a range of sensitivity analyses to test for violation of assumptions via fetal/paternal genotype and other sources of horizontal pleiotropy.
Results: Data from the majority of cohorts have now been collated and harmonized. By the time of the conference we will have MR and negative control results to present. Currently our preliminary observational analysis, using data from the INMA (Environment and Childhood) Spanish cohort (n=1402), show that each kg/m2 increase of maternal pre-pregnancy BMI is associated with a reduction of 0.23 infant cognitive scores (95% CI: -0.42, -0.04) at the age of 1.5 years, with adjustment for multiple measures of socioeconomic position.
Conclusion: Our study, which will triangulate MR with a MV analyses and a negative control study, will help disentangle whether there is an intrauterine effect of maternal pre-pregnancy obesity on offspring neurodevelopment.

Keywords: mother-child, obesity, pregnancy, neurodevelopment, mendelian randomization
P26 Genomic Resources for Mendelian Randomisation in the China Kadoorie Biobank

Presenter: Dr Robin G Walters

University of Oxford. Email: robin.walters@ndph.ox.ac.uk

Co-authors: Robin G. Walters [1], Kuang Lin [1], Iona Y. Millwood [1,2], Michael V. Holmes [1,3], Zhengming Chen [1], for the China Kadoorie Biobank Collaborative Group

1 Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF
2 MRC Population Health Research Unit, University of Oxford, Oxford OX3 7LF
3 NIHR Oxford Biomedical Research Centre, Churchill Hospital, Oxford OX3 7LE

Abstract

Objectives: China Kadoorie Biobank (CKB) is a prospective study of 512,891 adults from 10 areas of China, aged 30-79 years at recruitment in 2004-2008, with follow-up provided by linkages to registries and electronic health records. Dense genotyping of the cohort will provide opportunities for assessing causality of risk factors through Mendelian randomisation, and for investigating the impact of inactivation or inhibition of specific gene products.

Methods: A custom-designed Affymetrix Axiom® array was used to genotype 700,694 variants in 32,865 CKB participants. An updated version of the array (improved and revised on the basis of initial data) with 803,347 variants was used for an additional 69,935 participants. After SNP and sample QC, imputation into the full 1000 Genomes Phase 3 reference panel was performed using SHAPEIT3 and IMPUTE4.

Results: Samples from a total of 100,908 participants passed QC. For the initial array design, 589,033 variants passed QC, enabling imputation with mean certainty >0.9 of 9,632,982 variants with MAF >0.005, including 86.7% of variants with MAF >0.05 in the reference dataset. In addition, both array designs include >80,000 SNPs putatively identified as loss-of-function or missense variants, many not present at appreciable frequency in non-East Asian populations. >90% of LoF/missense variants pass QC, 87% having >20 carriers in this dataset. In addition, many of the remainder only failed QC in individual genotyping batches and can still be used for analyses in a subset of samples.

Conclusions: These genomics resources in CKB are enabling Mendelian randomisation studies, including of alcohol, blood pressure, adiposity, diabetes, and blood lipids; and phenome-wide analyses of the impact of modulating levels and/or activity of major potential or actual drug targets, including the lipoprotein-related PLA2G7, CETP, PCSK9, HMGCR, and NPC1L1 proteins. Such studies will be enhanced by the planned genotyping of the rest of the CKB cohort.

Keywords:
genomics, biobank, mr
P27 Understanding the causal effects of iron metabolism on chronic disease outcomes using Mendelian Randomization

Presenter: Dr Anna Ramond
University of Cambridge, Department of Public Health and Primary Care. Email: ar691@medschl.cam.ac.uk

Co-authors: Anna Ramond [1], Daniel Freitag [2], James Staley [3], Adam Butterworth [1], Stephen Burgess [1,4], Emanuele Di Angelantonio [1]
[1] Department of Public Health and Primary Care, University of Cambridge, United Kingdom
[2] Cardiovascular Data Sciences, Bayer Pharmaceuticals, Cologne Area, Germany
[3] MRC Integrative Epidemiology Unit, University of Bristol, United Kingdom
[4] MRC Biostatistics Unit, University of Cambridge, United Kingdom

Abstract
Iron metabolism has been implicated in the pathogenesis of various common chronic diseases, including cardiovascular disease, various forms of cancer and neurodegenerative diseases. While the observational evidence is indicative of an effect of iron metabolism in the pathology of several of these diseases, the potential causal effect remains unclear.

We have used an instrumental variable analysis approach, also known as mendelian randomization, to investigate the causal effects of four iron biomarkers: serum iron, serum ferritin, transferrin and transferrin saturation, on common chronic diseases, in particular, coronary heart disease (CHD), ischaemic stroke, colorectal cancer, lung cancer, Alzheimer’s disease (AD) and Parkinson’s disease (PD). We selected 7 single nucleotide polymorphisms (SNPs) associated with these four iron biomarkers at genome wide significance level ($p<5.10^{-8}$) for use as instrumental variables. We used summary statistics from large GWAS of CHD, ischaemic stroke, colorectal cancer, lung cancer, PD and AD, to investigate the effects of the four iron biomarkers on these disease outcomes using mendelian randomization.

Results of the mendelian randomization analysis showed evidence of an increased risk of colorectal cancer with increased levels of ferritin and transferrin saturation, whereas there was suggestion of a decreased risk of Parkinson’s disease and lung cancer with increasing levels of serum iron and transferrin saturation. There was no evidence of an association of CHD, ischaemic stroke or AD with any of the iron biomarkers.

We found evidence suggesting that iron biomarkers are causally associated with colorectal cancer, and potentially with lung cancer and PD. Future work will focus on multivariate mendelian randomization approaches to determine whether these iron biomarkers are independently associated with disease risk in order to improve our understanding of the underlying etiological mechanisms.

Keywords:
iron metabolism, mendelian randomization, genetic epidemiology
**P28 Sensitivity analysis in Mendelian randomization using negative controls**

**Presenter:** Dr Dipender Gill  
*Imperial College London. Email: dipender.gill@imperial.ac.uk*

**Co-authors:** Dipender Gill  
*Imperial College London*

**Abstract**

**Objectives:** In Mendelian randomization studies, a negative control may be used to investigate the specificity of any identified causal effect to the intermediate phenotype or exposure of interest. For example, such a negative control may involve using SNP-outcome estimates from a population that does not experience the intermediate phenotype or exposure.

**Methods:** In an MR study investigating the causal effect of age at menarche on time spent in education, we perform such a negative control analysis to investigate whether the identified effect of age at menarche on time spent in education in women is indeed driven by mediation through age at menarche and not via some alternative pathway. As a negative control, we estimate this association in men using SNP-time spent in education estimates for the 122 age at menarche instruments. Since men do not undergo menarche, the intermediate phenotype under investigation, lack of any association in men would provide further evidence that an association in women is due to a causal effect of age at menarche on time spent in education.

**Results:** Inverse-variance MR, MR-Egger and weighted median estimator analyses all suggest a causal effect of age at menarche on time spent in education in women. The negative control sensitivity analysis using SNP-time in education estimates from men confirms that this effect is specific to women, who undergo menarche, and not generalizable to men, despite the age at menarche SNPs having significant overlap with puberty in both males and females.

**Conclusions:** The use of a negative control in MR analysis can be an effective way of confirming the specificity of any identified causal effect to the intermediate phenotype or exposure of interest.

**Keywords:** mendelian randomization, negative control
Abstract
Metabolism refers broadly to the set of life-sustaining chemical transformations within cells and involves the conversion of food into energy and the building blocks of proteins, lipids, nucleic acids, and carbohydrates. Cancer cells co-opt metabolism to such an extent that the reprogramming of cellular metabolism has been deemed a hallmark of cancer. However, human studies of the metabolic basis of prostate cancer (PCa) have yet to determine which metabolic changes associated with PCa are a cause or a consequence of tumour development and progression. To address this, in the largest study to date of metabolites in PCa, we analysed nuclear-magnetic resonance (NMR)-measured metabolites among 2335 PCa cases (blood collected at diagnosis) and 2705 controls in ProtecT and compared the observational findings with findings from two-sample Mendelian randomization (MR), where genetic instruments for the metabolites were obtained from a large GWAS of metabolites (~25K subjects) and summary data of the effect of these instruments on PCa was obtained from the PRACTICAL consortium (~45K cases and ~28K controls). In models using robust standard errors, regressing each metabolite logistically on PCa status and adjusting for age and family history of PCa, 37 of 228 metabolites (representing a handful of amino acids and various lipid species) were below our multiple-testing corrected threshold ($p<0.0014$). None of these top metabolic findings were confirmed with MR, suggesting they are not causal, and instead point to their potential as screening biomarkers. Examination of the top MR findings revealed a suggestively causal role in PCa early detection/initiation for creatinine (Maximum likelihood method odds ratio=1.40; 95% CI: 1.20-1.66), a metabolite that was not among our top findings observationally, but confirming an association observed in a previous prospective study of serum creatinine and PCa.

Keywords:
metabolomics, prostate cancer, biomarkers
Abstract

Background: Previous studies have suggested that taller stature may causally improve socioeconomic status through discrimination against shorter people, but the nature of the relationship remains unknown. We used Mendelian randomization analysis to assess the causal effect of taller stature on socioeconomic status in middle-aged Korean population.

Methods: Data including height and socioeconomic status (i.e. income and education) were collected from 156,701 Koreans, aged 20 years or older, who took part in the Korean Cancer Prevention Study-II (KCPS-II) Biobank cohort. Weighted genetic risk score was calculated using 4 SNPs (rs6902771, rs724016, rs7567851, and rs806794) among 3,280 men and 1,630 women, participants randomly selected from KCPS-II Biobank subcohort.

Results: Both SNPs and weighted genetic risk score (F-statistic=14.8) were strongly associated with height. Simultaneously, height was associated with both income and education levels in an ordinary least-squares analysis, with or without adjusting for smoking status. However, in 2-stage least-squares Mendelian randomization analysis, no causal relationship between height and socioeconomic status was found.

Conclusions: There is no evidence that height level is causally associated with socioeconomic status in Koreans. Therefore, shorter height is not a disadvantage to socioeconomic status.

*This study was funded by a grant of the Korean Health Technology R&D Project, Ministry of Health & Welfare, Republic of Korea (HI14C2686) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2015R1D1A1A01059651).

Keywords:
height, smoking, socioeconomic status, causality
P31 Two-sample epigenetic Mendelian randomization of asthma and DNA methylation in childhood and adolescence.

Presenter: Ryan Arathimos
School of Social and Community Medicine, University of Bristol, Bristol, UK. Email: ryan.arathimos@bristol.ac.uk

Co-authors: Ryan Arathimos[1,2], Matthew Suderman[1,2], Gemma C Sharp[1,2], Kimberley Burrows[1,2], Raquel Granell[1], Kate Tilling[1,2], Susan Ring[1,2], Tom R Gaunt[1,2], John Henderson[1], Rebecca Richmond[1,2] & Caroline L Relton[1,2]
1 School of Social and Community Medicine, University of Bristol, Bristol, UK
2 Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK

Abstract
Asthma heritability has only been partially explained by genetic variants and is known to be sensitive to environmental factors, implicating epigenetic modifications such as DNA methylation in its pathogenesis. However, DNA methylation could be both a cause and an effect of asthma. We used the genetic architecture of DNA methylation to leverage bi-directional Mendelian randomization (MR) to establish causality and to infer the direction of association between DNA methylation and asthma.

Using data from the Avon Longitudinal Study of Parents and Children (ALSPAC), we assessed associations of asthma with DNA methylation at 7.5 years and 16.5 years, at over 450,000 CpGs in peripheral blood of 1000 participants. We tested the associated CpGs (FDR-adjusted P-value <0.05) using bi-directional MR. We identified instrumental variables (IVs) for asthma from the GWAS catalog and cis-SNPs for CpGs from ARIES-ALSPAC.

We identified 302 CpGs associated with current asthma at 7.5 years and 2 CpGs associated with current asthma at 16.5 years. Associations appeared to be driven by higher eosinophil counts in asthmatics. All associations attenuated when adjusted for eosinophil and neutrophil cell-count estimates. Two-sample bi-directional MR indicated a causal effect of asthma on DNA methylation at several CpGs at 7.5 years and one CpG which caused asthma. However, these associations did not survive our multiple testing threshold (adjusted for correlation between comethylated CpGs). At 16.5 years, there was no evidence of a causal effect of asthma on the 2 CpGs associated with asthma. Neither of the 2 CpGs had available cis-SNPs that could be used as IVs. We were therefore unable to test the causal association in the reverse direction.

The majority of observed associations between asthma and DNA methylation are driven by higher eosinophil cell-counts in asthma cases, acting as an intermediate phenotype. There was no strong evidence of a causal effect in either direction.

Keywords:
alspac, asthma, wheeze, dna methylation, mendelian randomization
P32 The causal effect of BMI on the metabolome profile in childhood in the Avon Longitudinal Study of Parents and Children (ALSPAC)

Presenter: Miss Mary Ward
MRC Integrative Epidemiology Unit, University of Bristol. Email: Mary.Ward@bristol.ac.uk

Abstract
Introduction: Numerous studies have observed strong associations between adiposity and the metabolome, however few of these studies investigated the direction of the causal pathway in these relationships. We used Mendelian randomization (MR) to investigate whether BMI has a causal effect on the metabolome in childhood.

Methods: We performed cross-sectional linear regression analysis to investigate the relationship between BMI and 160 metabolite measures in children age 7 years in ALSPAC (n=5400 children). Having observed associations between BMI and several of the metabolite measures, we used a BMI allele score to perform MR analysis to assess whether our observed associations between BMI and metabolites represent a causal effect of BMI on the metabolome.

Results: We observed cross-sectional associations in the 7-year-olds between BMI and 108 metabolite measures (p < 0.001). We observed associations for almost all of the VLDL concentration measures, the majority of the HDL concentration measures, and several other measures including diacylglycerol, cholines, apolipoproteins, fatty acids and amino acids. The causal effect estimates from MR analyses were mostly directionally consistent with the observational estimates, however the confidence intervals of the causal effect estimates were wide and most spanned zero.

Conclusions: We observed strong evidence of associations between BMI and several metabolite measures in 7-year-old children. The results of our MR analyses suggest that BMI may have a causal effect on some components of the metabolome, however the evidence is not conclusive.

MW is supported by Wellcome Trust 099873/Z/12/Z.

Keywords:
bmi, metabolome, childhood
P33 An online tool to aid the design of recall by genotype (RbG) studies

Presenter: Dr Laura Corbin
MRC Integrative Epidemiology Unit, University of Bristol. Email: laura.corbin@bristol.ac.uk

Co-authors: Laura J. Corbin [1], Vanessa Y. Tan [1], David A. Hughes [1], Kaitlin H. Wade [1], Katherine E. Tansey [2], Frank Dudbridge [3], Nicholas J. Timpson [1]
1 Medical Research Council (MRC) Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN, UK
2 Core Bioinformatics and Statistics Team, College of Biomedical & Life Sciences, Cardiff University, Cardiff, CF10 3XQ, UK
3 Department of Health Sciences, University of Leicester, Leicester, LE1 7RH, UK

Abstract
Genome-wide association studies have been useful in identifying common genetic variants related to a variety of complex traits and diseases; however, they are often limited in their ability to inform us on underlying biology. Whilst bioinformatics analyses, cell studies, animal models and applied genetic epidemiology have provided some understanding of genetic associations or causal pathways, there is a need for new genetic studies that elucidate causal relationships and mechanisms in a cost-effective and statistically efficient fashion. Recall-by-genotype (RbG) study design, based on either single (RbGsv) or multiple variants (RbGmv), enables genotype-directed deep-phenotyping and improvement in drawing causal inferences. Dependent on the nature of the genetic variation in question, the sample type, participant recruitment opportunities and the outcomes of interest, there will be optimal conditions for either RbGsv or RbGmv study designs. We have developed an online tool to aid the design of new RbG studies. This tool allows users to consider the efficacy of the RbG approach as compared to a more traditional study design through power estimation. We have developed a user-friendly interface that allows researchers to enter parameters specific to their study in order to generate meaningful predictions of power under a range of scenarios. For RbGsv scenarios, users may choose to recruit either homozygotes only or major homozygotes and heterozygotes. The latter of which may be beneficial in cases where the minor allele is rare. For RbGmv scenarios, users can choose between a simulation approach that takes into account properties of the variants used and an analytical approach. In summary, the freely available online app (available at: http://bit.ly/rbgplanner) is a tool that will help researchers plan future RbG studies.

Keywords:
recall by genotype, genetic risk score, polygenic risk score
P34 The association between circulating Vitamin D and IGFBP-3: observational and causal estimates from independent sources.

Presenter: Dr Vanessa Y. Tan
MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, United Kingdom. Email: Vanessa.tan@bristol.ac.uk
Co-authors: Vanessa Y. Tan[1,3]*, Kalina M. Biernacka[2,6]*, Tom Dudding[1,3], Carolina Bonilla[1,3], Rebecca Gilbert[3], Robert Kaplan[4], Qi Qibin[4], Alexander Teumer[5], Richard M. Martin[1,3,6], Claire M. Perks[2], Nicholas J. Timpson[1,3]*, Jeff M.P. Holly[2]*.
1MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, United Kingdom 2IGFs & Metabolic Endocrinology Group, School of Clinical Sciences, Learning & Research Building, Southmead Hospital, Bristol, United Kingdom 3School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom 4Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, 10461, USA 5Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, 17475, Germany 6National Institute for Health Research (NIHR) Bristol Biomedical Research Centre, University Hospitals Bristol NHS Foundation Trust and the University of Bristol.

Abstract
a) Circulating insulin-like growth factor binding protein 3 (IGFBP-3) is a risk factor for cancer. There is strong evidence from preclinical studies that active vitamin D (1,25(OH)2D) regulates IGFBP-3 expression. However, the relationship between the most commonly used marker of active vitamin D (the inactive precursor (25(OH)D)) and IGFBP-3 has not been sufficiently assessed in population-based studies. We aimed to investigate the relationship between 25(OH)D and IGFBP-3 both observationally and causally using data from the Prostate Testing for Cancer and Treatment (ProtecT) study and independently using the genome-wide association study (GWAS) results for 25(OH)D.
b) Individual-level data from the ProtecT study were used for an observational analysis and an MR analysis to assess the relationship between 25(OH)D and IGFBP-3. To validate work in the ProtecT study and increase analytical power, summary-level data from the IGF working group of the CHARGE consortium (n=18,995) was used in a two-sample MR framework to assess the causal relationship between 25(OH)D and IGFBP-3.
c) In cases and controls combined (1,366 cases and 1,071 controls), a standard deviation (SD) increase (7.8ng/ml) in 25(OH)D was associated with a 0.1 SD (95% CI: 0.05, 0.14; p<0.001) increase in IGFBP-3 levels. MR analyses in ProtecT study were not suggestive of a causal effect. Two-sample MR analysis using data from the CHARGE consortium found little evidence for a causal effect of circulating 25(OH)D on IGFBP-3 (per SD increase in 25(OH)D the change in IGFBP-3 was 0.02 SD; 95% CI: -0.06, 0.01; p=0.61).
d) There is evidence that that circulating 25(OH)D is associated with circulating IGFBP-3, but MR analyses indicated these findings were unlikely to be causal. Findings here are limited by the nature of instrumentation of 25(OH)D and the utility of circulating measures, but are important as they suggest that circulating 25(OH)D levels is unlikely to be causally related to circulating IGFBP-3.

Keywords: vitamin d, 25-hydroxyvitamin d, insulin-like growth factor binding protein, prostate cancer, mendelian randomization
P35 Examining the genetic influences on educational attainment and the validity of value-added scores as measures of progress in educational research

Presenter: Mr Tim Morris
School of Geographical Sciences, University of Bristol, University Road, Bristol, BS8 1SS. Email: tim.morris@bristol.ac.uk

Co-authors: Neil M Davies [1], Danny Dorling [2], George Davey Smith [1]
1 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN
2 School of Geography and the Environment, University of Oxford, Oxford, OX1 3QY

Abstract
Study objectives: In this study, we estimate (i) the heritability of educational attainment at three time points throughout the compulsory educational lifecourse; (ii) the heritability of value-added measures of educational progress built from test data to determine their ability to control for between individual genomic differences; and (iii) the extent to which value-added measures built from teacher rated ability may be biased due to measurement error.

Methods: We utilise a genome wide approach using generalized restricted maximum likelihood to determine the total phenotypic variance that can be explained by common genetic variation across the genome within a sample of unrelated individuals from a UK birth cohort, the Avon Longitudinal Study of Parents and Children.

Results: Our findings suggest that genetic heritability of educational attainment measured using point score test data increases with age from 47% at age 11 to 58% at age 14 and 61% at age 16. Additionally, we find that in contrast to previous findings based upon data using teacher rated ability as a measure of attainment, value-added measures based upon point score test data successfully control for time-invariant between individual differences and cannot be explained by common genetic variation. Our results also suggest moderate heritability (36%) of value-added measures built from teacher rated ability, suggesting that previous findings may have been biased by measurement error present within teacher rated measures of ability.

Conclusions: Our findings suggest that while the heritability of attainment increases throughout the education lifecourse, value-added measures based upon point score test data successfully control for genomic differences between individuals and offer measures of student attainment that reflects purely environmental influences.

Keywords:
education, genetics, value added scores, gcta
P36 Does cocoa protect against ischemic heart disease?

Presenter: C Mary Schooling  
CUNY School of Public Health. Email: mary.schooling@sph.cuny.edu

Co-authors: C Mary Schooling [1], Ghada Soliman [1]  
1 CUNY School of Public Health, New York, NY10027, USA

Abstract  
Despite great progress in reducing cardiovascular disease mortality rates in Western settings, cardiovascular disease remains poorly understood with repeated drug failures, and many genetic targets unexploited. The adenosine A2a receptor (ADORA2A) gene is strongly associated with ischemic heart disease (IHD) and is the target of many existing drugs and nutraceuticals, including caffeine and a key constituent of cocoa, i.e., theobromine, which was formerly used to treat hypertension and angina. In meta-analysis of randomized controlled trials, cocoa reduces blood pressure, but effects on IHD and myocardial infarction (MI) have not been similarly assessed. We assessed if theobromine was causally associated with IHD or MI using Mendelian randomization. To exclude an effect via lipids, associations with low-density lipoprotein, high-density lipoprotein cholesterol and triglycerides were also assessed. Single nucleotide polymorphisms (SNP) strongly and independently predicting blood theobromine were taken from a genome wide association study of 7,824 adults of European descent, and applied to a large, extensively genotyped IHD and MI case (n= 60,801 for IHD, 43,676 for MI) control (n= 128,199) study largely in Europeans, i.e., CARDioGRAMplusC4D 1000 Genomes, and to a large (n=188,577) cross-sectional study of lipids, i.e., Global Lipids Genetics Consortium. SNP-specific Wald estimates were combined using inverse variance weighting. Based on 4 SNPs, theobromine was not clearly associated with IHD (odds ratio (OR) 0.75, 95% confidence interval (CI) 0.5 to 1.01, but was inversely associated with MI (OR 0.71, 95% CI 0.51 to 0.99). Theobromine was not associated with lipids. One of the SNPs predicting theobromine also predicted caffeine which could indicate a pleiotropic effect on IHD and MI, however previous Mendelian randomization studies have found no effect of coffee on IHD. As such, these results add weight to the argument that cocoa might have cardio-protective effects.

Keywords:  
theobromine, cocoa, ihd, mi
P37 Pubertal Timing and Antisocial Behaviour in Girls: Distinguishing pleiotropy, phenotypical correlation and shared genetic variance of complex traits

Presenter: Dr Martin Steppan
University of St Andrews (PhD Student); Universitäre Psychiatrische Kliniken (Basel, Switzerland).
Email: ms307@st-andrews.ac.uk

Co-authors: Martin Steppan [1,2]
1 University of St Andrews; 2 Universitäre Psychiatrische Kliniken Basel, Switzerland

Abstract
Background: Mendelian Randomization (MR) has received attention as a powerful technique to make causal inferences in epidemiology. Ziegler et al. (2015) argue in a recent paper that not MR, but the use of path models is well suited to draw causal inferences. Earlier research has shown that there is an association between antisocial behaviour and early pubertal timing, which both represent complex traits.

Methods / Data: The Avon Longitudinal Study of Parents and Children (ALSPAC) provides data on 15,247 pregnancies, including genomic data on 8,365 children. Data on 71 SNP’s were used which had been predictive of pubertal timing in earlier studies. Path models were applied using AMOS and R. LD Score regression according to Bulik-Sullivan et al. (2015) was used to cross-validate findings on shared genetic variance of antisocial behaviour, psychiatric disorders, BMI and pubertal timing.

Results: Preliminary results show that a genetic risk score for early puberty was significantly associated with higher BMI, but not with a general factor of psychiatric disorders. Several small correlations (r<.10) of this risk score with psychiatric disorders emerged (e.g. Conduct Disorder at age 7), but these correlations do not persist after controlling for multiple testing. LD Score Regression corroborates these results, i.e. there is a significant genetic correlation of pubertal timing with BMI, but not with common psychiatric disorders or antisocial behaviour.

Conclusions: The results are in line with earlier findings that pubertal timing has shared genetic roots with BMI in girls. The results are at odds with some studies that highlight an association of pubertal timing and psychopathology in girls. The non-significance of MR approaches or LD score regression suggest that these findings may be transient (only occur during puberty) or are caused by confounders (e.g. BMI, smoking).

Keywords: pubertal timing, antisocial behaviour, mendelian randomization, conduct disorder, general psychopathology
Abstract
Introduction: Higher dietary phylloquinone intake is associated with lower type 2 Diabetes (T2D) incidence. However, observational studies have to be interpreted carefully since there is a chance of reverse causation and confounding. The aim of this study is to investigate the causal effect of plasma phylloquinone levels on T2D incidence via a Mendelian Randomization approach.

Methods: We used data from the European Prospective Investigation into Cancer and Nutrition (EPIC)-InterAct case-cohort study comprising 10,071 diabetes cases and 13,309 subcohort members from eight European countries. A weighted genetic risk score (wGRS) was made of four SNPs (rs2192574, rs6862071, rs4645543 and rs2108622) known to be related to plasma phylloquinone levels in a GWAS study. We assessed the association between the wGRS and T2D incidence using a Prentice-weighted Cox regression analysis. Estimates from this Cox regression analysis, and estimates for the relation of the SNPs with plasma phylloquinone were used for inverse-variance weighted (IVW) analysis to obtain a hazard ratio (HR) for the unconfounded relation between plasma phylloquinone levels and T2D incidence. All analyses were adjusted for sex, center, principal components of ancestry, genetic platform, triglycerides and hours fasting.

Results: The median follow-up time was 10.9 years. The wGRS was unrelated to potential confounders of the observational relation between phylloquinone and T2D. A higher wGRS reflects higher plasma phylloquinone levels according to a previous GWAS (β 0.44 (0.06; 0.82). A higher wGRS related to a lower T2D incidence: HR 0.87 (0.78; 0.97) per point higher of the wGRS. The hazard ratio for the unconfounded relation between plasma phylloquinone levels and T2D incidence, resulting from the IVW analysis, was 0.87 (0.78; 0.97).

Conclusion: Our study suggests the association between plasma phylloquinone levels and T2D incidence may be causal.

Keywords:
phylloquinone, type 2 diabetes incidence
P39 Unravelling the obesity paradox using mendelian randomisation in UK Biobank

Presenter: David A Jenkins

Farr Institute for Health Informatics Research Division of Informatics, Imaging and Data Science, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK. Email: david.jenkins-5@manchester.ac.uk

Co-authors: David A Jenkins [1], Jack Bowden [2], Heather A Robinson [1], Ruth Loos [3], Martin Rutter [4][5], Mathew Sperrin [1]

1 Farr Institute for Health Informatics Research Division of Informatics, Imaging and Data Science, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK
2 Medical Research Council Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol, UK
3 Charles Bronfman Institute for Personalized Medicine, Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA
4 Division of Diabetes, University of Manchester, Manchester, UK
5 Manchester Diabetes Centre, Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK

Abstract

The ‘obesity paradox’ refers to the commonly observed epidemiological finding that, in some disease populations, being overweight or obese is associated with better survival than being a ‘healthy weight’. The UK biobank is a large cohort study with over 500,000 subjects that combines healthcare records, patient characteristics and even individual genetic information. This data, approached with mendelian randomisation (MR) methods, could help us in understanding the obesity paradox.

There are many methodological challenges in applying MR, particularly to ‘time to event’ data. These are very often overlooked but can have huge impact to inferences. Creating an appropriate instrumental variable using the properties of mendelian randomisation can also be challenging and violation of assumptions may occur. One example of such an instrument is a genetic risk score developed using external information, or by splitting a single set of data into a two sample data set. Depending on the data available, a risk score can either be developed using an ‘allele score’ method that can be applied by a one or two sample approach, or by a ‘weighted variant’ method. Once an instrument has then been developed, more challenges arise in assessing its validity. Methods using meta-analytic techniques, known as, ‘MR-Egger’ and the inverse-variance-weighted average do this, but the challenge is how to interpret and change the instrument if needed.

We plan to use UK biobank data by applying MR within subpopulations of patients with diabetes and cardiovascular disease, to better understand the obesity paradox. To do this, we will use multiple methods to create a genetic risk score for obesity that can be used in our analysis. We will then assess each of the genetic instruments validity, and carefully consider the assumptions and interpretability of each approach.

Keywords:

obesity paradox, diabetes, genetic risk score, instrumental variable