WHAT EXCITES THE FACULTY

LOICYENGO



What 'Love' got to do with it?

International Statistical Genetics Workshop

Boulder, March 8th 2019

Loïc Yengo







Queensland Brain Institute

Couples often assort on...



Education (r ~ 0.5) Height (r ~ 0.2 - 0.3)

Resemblance of spouses = Assortative Mating (AM)

But also on...

- Body mass index
- Personality related traits
- Political preferences
- Diseases susceptibility
- Etc.

Evolutionary consequences of AM

- Fisher (1918), Wright (1921), Crow and Kimura (1970), Gimelfarb (1981, 1984).
- AM induces a **positive correlation between trait increasing alleles**.
- <u>Consequence</u>: increase of genetic variance (heritability).
- $var(a_1X_1 + a_2X_2) = a_1^2var(X_1) + a_2^2var(X_2) + 2a_1a_2cov(X_1,X_2)$

Why do we care about AM?

 Increase in genetic variance may lead to more extreme phenotypes in the population.

• AM create structures in the population that may engender health and socioeconomic inequalities.

How do we quantify it?

Chromosome	SNP	Tested Allele (A1)	Beta
1	rs1778789	А	0.01
2	rs67388911	G	-0.1
3	rs8800309	С	0.11
4	rs17777893	Т	0.71

Summary Statistics from a reference GWAS

Individual level genotypes from a study **independent** of the reference GWAS

	ID_1	ID_2	•••	ID_N
rs1778789	AA	AT		TT
rs67388911	GA	AA		GA
rs8800309	CC	CC		CC
rs17777893	GT	ТТ		GG

How do we quantify it?

Chromosome	SNP	Tested Allele (A1)	Beta
1	rs1778789	А	0.01
2	rs67388911	G	-0.1
3	rs8800309	С	0.11
4	rs17777893	Т	0.71

Summary Statistics from a reference GWAS

Individual level genotypes from a study independent of the reference GWAS

	ID_1	ID_2		ID_N
rs1778789	AA	AT	0 0 0	TT
rs67388911	GA	AA		GA
rs8800309	СС	СС		СС
rs17777893	GT	TT		GG

	Polygenic Score on odd chromosomes (So)	Polygenic Score on even chromosomes (Se)
ID_1	Beta(rs1778789) #A + Beta(rs8800309) #C	Beta(rs67388911) #G + Beta(rs17777893) #T
ID_2	Beta(rs1778789) #A + Beta(rs8800309) #C	Beta(rs67388911) #G + Beta(rs17777893) #T
•••		
ID_N	Beta(rs1778789) #A + Beta(rs8800309) #C	Beta(rs67388911) #G + Beta(rs17777893) #T

 θ = S(o)~S(e) (regression) = cov(Se,So)/var(Se)

Significant GPD for height and Educational attainment



No signal for BMI, convergence?

What to know more?

human behaviour

LETTERS https://doi.org/10.1038/s41562-018-0476-3

Imprint of assortative mating on the human genome

Loic Yengo ¹*, Matthew R. Robinson^{1,2}, Matthew C. Keller ³, Kathryn E. Kemper¹, Yuanhao Yang¹, Maciej Trzaskowski¹, Jacob Gratten^{1,4}, Patrick Turley^{5,6}, David Cesarini^{7,8,9}, Daniel J. Benjamin ^{7,10,11}, Naomi R. Wray^{1,12}, Michael E. Goddard^{13,14}, Jian Yang ^{1,12} and Peter M. Visscher ^{1,12*}

Australian Government

Australian Research Council





Australian Government

* National Health and Medical Research Council





ACKNOWLEDGEMENTS







MEIKE BARTELS





The Exposome



Network Integration



Network Integration



Network Integration



NICK MARTIN

Human Reproduction, Vol.25, No.6 pp. 1569-1580, 2010

Advanced Access publication on April 8, 2010 doi:10.1093/humrep/deq084

human reproduction **ORIGINAL ARTICLE Reproductive genetics**

A genome wide linkage scan for dizygotic twinning in 525 families of mothers of dizygotic twins

Jodie N. Painter^{1,*}, Gonneke Willemsen², Dale Nyholt¹, Chantal Hoekstra², David L. Duffy¹, Anjali K. Henders¹, Leanne Wallace¹, Sue Healey¹, Lisa A. Cannon-Albright³, Mark Skolnick³, Nicholas G. Martin¹, Dorret I. Boomsma^{2,†}, and Grant W. Montgomery^{1,†}



"Affected" sister-pair linkage study



QIMR 3,251 Mothers of DZ Twins 27,534 controls

MODZT association results - HRCr1.1 (overall) Scott Gordon GCTA MLMA LOCO analysis (Manhattan plot) filtered for MAF >= 0.01, Rsq >= 0.9 **FSHB** 15 -log10(P) 10 SMAD3 ZFMP1 ADRB2 ŝ 0 ю 20 21 22 22 ŝ -4 ß თ 0 15 16 1 7 18 3 с 4 -

Chromosome



Hamdi Mbarek

15 genes for DZT: 13 new; 13 have obvious function

					N	10DZT	MODZT+UKB (fixed effect)	MODZT+UKB (N-Weighted)
Ν	Genes	Chr	Start	Stop	Gene based P	GWAS top SNP P	Gene based P	Gene based P
1	FSHB	11	30252563	30256808	1.76E-10	7.31E-22	2.56E-11	1.31E-09
2	SMAD3	15	67356101	67487533	1.69E-07	1.91E-11	1.97E-08	3.69E-07
3		5			0.22	2.08E-08	0.44	0.42
4		8			3.61E-07	6.5E-09	2.34E-08	2.2E-07
5		16			2.05E-09	5.02E-10	6.16E-10	5.64E-07
6		5			5.78E-07		4.67E-07	2.35E-04
7		12			2.92E-08		2.27E-09	7.59E-08
8		17			6.10e-07		4.75E-07	2.84E-04
9		1			2.31E-05		2.59E-06	1.35E-04
10		2			8.56E-05		1.04E-06	1.36E-06
11		2			1.4E-04		1.48E-06	1.28E-06
12		8			2.03E-05		1.22E-06	6.75E-06
13		15			7.31E-06		6.35E-07	2.12E-05
14		2			0.024		1.45E-04	1.34E-06
15		4			0.022		1.73E04	2.72E-06

How well do these (PRS) predict

- DZT in an out population (e.g. Iceland)
- Large differences in DZT rate between Africans, Europeans, Asians
- Large differences between early age (18) and late (37) MODZT
- Other female reproductive traits (menarche, menopause, age 1st kid, #kids)
- Female infertility
- Reproductive cancers (breast, ovarian, testicular, prostate)
- Evidence for selection at these loci (eugenic?, dysgenic?)

HILARY MARTIN

Convergence of Mendelian and complex disease genetics _{Hilary Martin}

Wellcome Sanger Institute (near Cambridge, UK)

Mendelian versus complex disease genetics: a false dichotomy?

	Rare variants	Common variants
Mendelian diseases	Major role	?
Complex diseases	Some well-known examples (e.g. MODY), but don't account for much heritability	Major role

Common variants contribute to rare neurodevelopmental disorders previously assumed to be Mendelian

- ~7000 patients with rare, severe neurodevelopmental disorders in the Deciphering Developmental Disorders study versus ~9000 controls
- h²_{SNP} ~8% (for MAF>5%)
- replicated with polygenic scores in 1270 cases from Australia versus 1700 controls



Polygenic transmission-disequilibrium test in autism

- Polygenic scores for autism, schizophrenia and educational attainment are significantly overtransmitted to autism patients, regardless of their IQ
- Still see over-transmission of autism and schizophrenia PRS to probands with a large-effect *de novo* mutation



Common variants modifying penetrance of rare variants affecting cognition?

- rare loss-of-function variants in highly constrained genes increase risk of severe ID/DD and impact cognition in the general population
- are people in healthy population cohorts protected against such variants by polygenic background?
- preliminary work in INTERVAL, a cohort of healthy British blood donors



ongoing collaboration with Andrea Ganna, Sali Farhan at Broad to expand sample size!

Mari Niemi, Eugene Gardner

Common variants modifying penetrance of rare variants in breast cancer

Predicted cancer risks by percentile of the polygenic risk scores



Evaluation of Polygenic Risk Scores for Breast and Ovarian Cancer Risk Prediction in BRCA1 and BRCA2 Mutation Carriers 👌

Karoline B Kuchenbaecker, Lesley McGuffog, Daniel Barrowdale, Andrew Lee, Penny Soucy, Joe Dennis, Susan M Domchek, Mark Robson, Amanda B Spurdle, Susan J Ramus, ... Show more

JNCI: Journal of the National Cancer Institute, Volume 109, Issue 7, 1 July 2017, djw302, https://doi.org/10.1093/jnci/djw302

Open questions

- How much do low-frequency variants (e.g. MAF 0.1-5%) contribute to heritability of rare neurodevelopmental disorders?
- To what extent do common variants contribute to other rare disorders? e.g. heart defects
- Do polygenic scores and rare variants act additively on traits, or is there an interaction?
- Can we increase power to detect rare variants by conditioning on polygenic risk scores?
- To what extent can we improve polygenic prediction by incorporating rare variants? (as individual variants, or in aggregated groups)

Acknowledgements



Mari Niemi



Eugene Gardner



Jeff Barrett



Matt Hurles

For cognition and schizophrenia in particular, rare and common variant signals are enriched in highly constrained genes

High constrained genes: significantly depleted of loss-of-function variants in healthy people nature neuroscienc Brief Communication Published: 03 October 201



Article | Published: 03 October 2016

Increased burden of ultra-rare protein-altering variants among 4,877 individuals with schizophrenia

ovese 🖾, Menachem Fromer, Eli A Stahl, Douglas M Ruderfer. Kimberly Chambert, Mikael Landén, Jennifer L Moran, Shaun M Purcell, Pamela Sklar Patrick F Sullivan, Christina M Hultman & Steven A McCarroll

Bipolar

Social Interaction Other

Schizoprenia

Common variants contribute to rare neurodevelopmental disorders previously assumed to be Mendelian

- 429 Finnish intellectual disability (ID) cases versus 2195 controls
- heritability explained by polygenic risk scores for EA is the highest in mild ID (2.2%) but lower for more severe ID (0.6%)



MATT KELLER

Behavioral Genetics in the molecular age - answers to age old questions

> Matthew Keller CU Boulder
BG is interested in quantifying factors that cause individual differences

- Traditionally, we've done this using close relatives & twins
- Has been enormously successful and we've been able to make broad-brushed conclusions about genes & env.
- But these studies require strong assumptions about causes of relative similarity. For us who care about, e.g., VA:VNA, VF (familial variance from parental effects) VC (shared environmental effects), or AM (assortative mating), twin and family models can be biased and coarse
- We want ways to investigate old questions in new (and potentially less biased) ways
- Following are new ways to do this using molecular data

1. SNP data to estimate ~full VA'

- GREML is done on unrelated (distantly related) individuals we don't require assumptions about causes of similarity between close relatives.
- As we move to very large sequence reference panels (e.g., TOPMed; n ~ 100k) or to large sequence data itself, VA'_GREML should begin to approach VA. If VA'_GREML is much different than VA'_twin, something interesting is going on!
 - Upward biases in VA'_twin? De novo variants? Phenotypic heterogeneity?
- This is exciting: new and independent way of estimating VA. Via triangulation, we'll know ~ true VA soon.

2. SNP data to estimate allelic spectra of V(G) and COV(G)

- What is the mix of rare vs. common variants affecting trait variation? Important for study design and evolutionary interpretations.
- GREML models (previous slide) on imputed or sequence data, when done properly (multiple GRMs stratified by MAF and individual LD) can give us estimates of the allelic spectra
- These same methods can help us understand if rg is the same across the allelic spectra

3. SNP data to understand causes of spousal similarity

- Recent work* uses GREML and/or PRS to directly estimate the additive genetic correlation between mates
 - Spousal height correlation completely consistent with primary AM, but that of BMI is about ½ primary AM and ½ other factors (convergence? social homogamy?)
 - This is exciting: assumptions about the cause of spousal similarity can now be empirically tested!

4. SNP data to investigate familial transmission

- Passive G-E correlations arise when genetic effects are not random with respect to env. effects
 - Most commonly: env. of offspring a function of parental phenotypes
- Comparing between- and within-family GWAS estimates can provide estimates of the importance of G-E covariance
- Similarly, looking at relationship between PRS of transmitted genome vs. PRS of untransmitted genome as a way to test parenting effects on offspring.

AYSU OKBAY

Integrating polygenic scores into social sciences

Aysu Okbay





EA1 (N=126,559), 3 hits

EA2 (N=293,723), 74 hits

EA3 (N=1,131,881), 1,271

Educational attainment polygenic scores



Sample Size

Nearly as good as conventional controls in social science



Polygenic scores in social sciences



Gene × environment interactions



Example: Policy evaluation

<u>The Swedish educational</u> <u>reform:</u>

- Increased compulsory schooling from 7 to 9 years
- Delayed ability tracking
- Was rolled out gradually generating quasiexperimental variation

Questions:

- Did the reform affect individuals in different EA PGS bins heterogeneously?
- Did the reform achieve what it meant to achieve? More equality?

Preliminary result:

 higher ability females were more likely to obtain a high school degree, which is beyond the new minimum established by the reform.

ABDEL ABDELLAOUI

ABDEL ABDELLAOUI



a.abdellaoui@amc.nl

BEHAVIOR GENETICS Migration Mate Choice



a.abdellaoui@amc.nl



Association Between Autozygosity and Major Depression: Stratification Due to Religious Assortment

Abdel Abdellaoui • Jouke-Jan Hottenga • Xiangjun Xiao • Paul Scheet • Erik A. Ehli • Gareth E. Davies • James J. Hudziak • Dirk J. A. Smit • Meike Bartels • Gonneke Willemsen • Andrew Brooks • Patrick F. Sullivan • Johannes H. Smit • Eco J. de Geus • Brenda W. J. H. Penninx • Dorret I. Boomsma



- Educational attainment was significantly associated with *F*_{rob} (inbreeding)
- But parental education was *much* more significantly associated with *F*_{roh}
- Why?
 - Higher educated parents **migrated** significantly more often and greater distances
 - There is strong **assortative mating** for educational attainment

RESEARCH ARTICLE

Educational Attainment Influences Levels of Homozygosity through Migration and Assortative Mating

Abdel Abdellaou^{1,2}*, Jouke-Jan Hottenga¹, Gonneke Willemsen^{1,3}, Meike Bartels^{1,2,3}, Toos van Beijsterveldt¹, Erik A. Ehli¹, Gareth E. Davies⁴, Andrew Brooks⁵, Patrick F. Sullivan⁶, Brenda W. J. H. Penninx^{2,1,3}, Eo J. de Geus^{1,2,3}, Dorret I. Boomsma^{1,2,3} Correlation between offspring ancestry and geography significantly decreased as parental education increased



RESEARCH ARTICLE

Educational Attainment Influences Levels of Homozygosity through Migration and Assortative Mating

Abdel Abdellaoui^{1,2}*, Jouke-Jan Hottenga¹, Gonneke Willemsen^{1,3}, Meike Bartels^{1,2,3}, Toos van Beijsterveld¹, Erik A. Ehli¹, Gareth E. Davles⁴, Andrew Brooks⁸, Patrick F. Sullivan⁶, Brenda W. J. H. Pennix^{2,3,7}, Eco J. de Geus^{1,2,3}, Dret1. Boomsma^{1,2,3} Assortative mating for educational attainment is measurable at the genetic level (with polygenic scores)





Assortative mating on educational attainment leads to genetic spousal resemblance for polygenic scores

David Hugh-Jones^{a,*,1}, Karin J.H. Verweij^{b,1}, Beate St. Pourcain^c, Abdel Abdellaoui^{b,*}

^a Department of Economics, University of East Anglia, Research Park, Courtyard B, Norwich NR4 7TJ Norwich, England, United Kingdom ^b Department of Biological Psychology, VU University, van der Boechorstraat 1, 1081 BT Amsterdam, The Netherlands ^c Max Planck Institute for Psycholinguistics, Wunddam 1, 625 ZD Nijmegen, The Netherlands

Genetic evidence of assortative mating in humans

Matthew R. Robinson^{1*}, Aaron Kleinman², Mariaelisa Graff³, Anna A. E. Vinkhuyzen¹, David Couper⁴, Michael B. Miller⁵, Wouter J. Peyrot⁶, Abdel Abdellaoui⁷, Brendan P. Zletsch⁸, Ilja M. Nolte⁹, Jana V. van Vliet-Ostaptchouk³¹⁰, Harold Snieder⁹, The LifeLines Cohort Study¹, Genetic Investigation of Anthropometric Traits (GIANT) consortium¹, Sarah E. Medland¹¹, Nicholas G. Martin¹¹, Patrik K. E. Magnusson¹², William G. Iacono⁵, Matt McGue⁶, Kari E. North¹³¹³, Jian Yang¹¹⁴ and Peter M. Visscher¹³⁴⁴

Polygenic scores, before and after regressing out 100 PCs



GENETIC CONSEQUENCES OF SOCIAL STRATIFICATION IN GREAT BRITAIN

Spatial autocorrelation (i.e., geographic clustering) of 30 polygenic scores



- Uncorrected for PCs
- Corrected for 100 PCs

GENETIC CONSEQUENCES OF SOCIAL STRATIFICATION IN GREAT BRITAIN



GENETIC CONSEQUENCES OF SOCIAL STRATIFICATION IN GREAT BRITAIN



GENETIC CONSEQUENCES OF SOCIAL STRATIFICATION IN GREAT BRITAIN





a.abdellaoui@amc.nl

KATRINA GRASBY

DAVID EVANS

Using Genetics to Investigate the Developmental Origins of Health and Disease

David Evans University of Queensland University of Bristol















Prevalence of Impaired Glucose Tolerance or T2D in 64 year old men in the UK



Hales et al. (1991)

GWAS Analysis of Birthweight in UKBB and EGG



- Birthweight GWAS reflects a mixture of maternal and fetal genetic effects
- Unrelated individuals*

Disentangling Mother and Child Effects on Birth Weight



Ē 🔁 🎔	(*) Twitter	■ Biological and clinical insigt S Using structural equation n S <= <hr/>			-	٥	×
$\leftarrow \ \ \rightarrow$	0 6	evansgroup.di.uq.edu.au/MGPC/				È	
Find on page	cranio	No results $\langle \rangle$ Options \checkmark					×

Maternal and Offspring Genetic Effects Power Calculator

	Power to detect maternal genetic effect (1 df test):					
Number of genotyped individuals with their own and their offspring's phenotype (0-10000000):	0.7832705					
10000	Power to detect offspring genetic effect (1 df test):					
Number of genotyped individuals with their own	0.7832705					
phenotype only (0-10000000):	Power to detect either maternal or offspring genetic effect (2 df test):					
0	0.9992226					
Number of genotyped individuals with their offspring's phenotype only (0-10000000):	Alpha level is the type 1 error rate for the genetic association study.					
Number of genetymed methor offenning poirs with	Maternal genetic effect is the proportion of variance in the trait explained by maternal genetic effects at the locus. Offspring genetic effect is the proportion of variance in the trait explained by offspring genetic effects at the locus.					
offspring phenotype (0-10000000):	If you have an effect size estimate and allele frequency for the genetic variant of interest, you can calculate the proportion of					
0	(2*P*(1-P)*b^2)/VAR(trait), where P=allele frequency, b=effect size estimate.					
Number of genotyped mother-offspring pairs with both maternal and offspring phenotype (0-10000000):	Residual correlation between maternal and offspring phenotype refers to the residual correlation between the phenotypes after the effect of the locus has been removed. This is likely to be close to the phenotypic correlation between maternal and offspring phenotypes as the majority of loci will have a small effect.					
0	ND. Malamal and offension genetic offense can call be estimated uniquely them there are providenced individuals the second					
O Type here to search	🗄 💼 🚖 🦳 🧱 🎵 🔯 🥼 😰 🗷 🎕 🚳 🐼 🖓 🕲 🖉 👘 🖉 👘 👘 🖓 🖽 👘 👘 👘					

Behavior Genetics https://doi.org/10.1007/s10519-018-9944-9

ORIGINAL RESEARCH



Calculating Power to Detect Maternal and Offspring Genetic Effects in Genetic Association Studies

Gunn-Helen Moen^{1,2} · Gibran Hemani^{3,4} · Nicole M. Warrington⁵ · David M. Evans^{3,4,5}

Received: 21 August 2018 / Accepted: 11 December 2018 © Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Offspring outcomes are a function of maternal genetics operating on the intrauterine and postmatal environment, offspring genetics and environmental factors. Partitioning genetic effects into maternal and offspring components requires genotyped mother–offspring pairs or genotyped individuals with phenotypic information on themselves and their offspring. We performed asymptotic power calculations and data simulations to estimate power to detect maternal and offspring genetic effects under a range of different study designs and models. We also developed the "Maternal and offspring genetic effects Power Calculator" (M-GPC), an online utility which allows users to estimate the power to detect maternal and offspring genetic effects in their own studies. We find that approximately 50,000 genotyped mother–offspring pairs will be required to detect realistically sized maternal or offspring genetic effects (>0.1% variance explained) with approximately bower (power >90%, $\alpha = 5 \times 10^{-8}$, two degree of freedom test), whereas greater than 10,000 pairs will be required to determine whether known genetic loci have maternal and/or offspring genetic effects (power >78%, $\alpha = 0.05$). The structural equation modeling framework espoused in this manuscript provides a natural method of combining different data structures including those present in large scale biobanks in order to maximize power to detect maternal and offspring genetic effects were the sample sizes required to detect maternal and within the range of current research consortia.

Introduction

Keywords Genetic association · GWAS · Maternal effects · Offspring effects · Fetal effects · Power

Edited by Tinca Polderman.

Nicole M Warrington and David M Evans joint senior authors.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10519-018-9944-9) contains supplementary material, which is available to authorized users.

Gunn-Helen Moen g.h.o.moen@studmed.uio.no

- ¹ Department of Endocrinology, Morbid Obesity and Preventive Medicine, Oslo University Hospital, Postbox 4959, Nydalen, 0424 Oslo, Norway
- ² Faculty of Medicine, Institute of Clinical Medicine, University of Oslo, Oslo, Norway
- ³ Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK
- ⁴ Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
- ³ Translational Research Institute, University of Queensland Diamantina Institute, Brisbane, Australia

Offspring phenotypes, including perinatal outcomes such as birthweight, length at birth and other anthropometric measurements, are thought to be a product of maternal genetics, offspring genetics and environmental factors. In this manuscript, we define maternal genetic effects as the causal influence of the maternal genotype on the offspring phenotype (Wolf and Wade 2009). Maternal genetic effects arise when the mother makes a contribution to the phenotype of her progeny over and above that which results from the genes she contributes to the zygote (Mather and Jinks 1982). Thus, our definition focuses on the effect of the maternal genome and is distinct from mitochondrial inheritance and genetic effects due to imprinting. In contrast, we define offspring genetic effects as those genetic effects on the offspring's phenotype that are directly mediated by the offspring's genome, which is comprised of 50% of alleles inherited from their mother and 50% from their father.

Published online: 02 January 2019

🕗 Springer

Maternal vs Fetal Effects for Birth Weight SNPs



Child Effect

ISEA

International Journal of Epidemiology, 2018, 1229–1241 doi: 10.1093/ije/dyy015 Advance Access Publication Date: 13 February 2018 Original article



Methods

Using structural equation modelling to jointly estimate maternal and fetal effects on birthweight in the UK Biobank

Nicole M Warrington,¹ Rachel M Freathy,^{2,3} Michael C Neale⁴ and David M Evans^{1,3,5,*}

¹University of Queensland Diamantina Institute, The University of Queensland, Translational Research Institute, Brisbane, QLD, Australia, ²Institute of Biomedical and Clinical Science, University of Exeter Medical School, University of Exeter, Royal Devon and Exeter Hospital, Exeter, UK, ³Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK, ⁴Virginia Institute for Psychiatric and Behavioral Genetics, Departments of Psychiatry and Human & Molecular Genetics, Virginia Commonwealth University, Richmond, VA, USA, ⁵School of Social and Community Medicine, University of Bristol, Bristol, UK

*Corresponding author. MRC Integrative Epidemiology Unit, University of Bristol, Oakfield House, Oakfield Grove, Clifton, BS82BN, UK. E-mail: dave.evans@bristol.ac.uk

Editorial decision 16 January 2018; Accepted 25 January 2018

Abstract

Background: To date, 60 genetic variants have been robustly associated with birthweight. It is unclear whether these associations represent the effect of an individual's own genotype on their birthweight, their mother's genotype, or both, Methods: We demonstrate how structural equation modelling (SEM) can be used to estimate both maternal and fetal effects when phenotype information is present for individuals in two generations and genotype information is available on the older individual. We conduct an extensive simulation study to assess the bias, power and type 1 error rates of the SEM and also apply the SEM to birthweight data in the UK Biobank study. Results: Unlike simple regression models, our approach is unbiased when there is both a maternal and a fetal effect. The method can be used when either the individual's own phenotype or the phenotype of their offspring is not available, and allows the inclusion of summary statistics from additional cohorts where raw data cannot be shared. We show that the type 1 error rate of the method is appropriate, and that there is substantial statistical power to detect a genetic variant that has a moderate effect on the phenotype and reasonable power to detect whether it is a fetal and/or a maternal effect. We also identify a subset of birthweight-associated single nucleotide polymorphisms (SNPs) that have opposing maternal and fetal effects in the UK Biobank.

Conclusions: Our results show that SEM can be used to estimate parameters that would be difficult to quantify using simple statistical methods alone.

©The Author(s) 2018. Published by Oxford. University Prace on behalf of the International Epidemiological Association. 1229 This is an Open Access anticle distributed under the terms of the Creative Commons Arthroticn Licence (#pp)/creative.commons.org/licence.wb/y4(4)), which permits urreaticited music, ultimative, and emission of the Creative Commons Arthroticn Licence (#pp)/creative.commons.org/licence.wb/y4(4)), which permits urreaticited music, ultimative, and anticipative and empoyed read.





Estimating Causal Effects of Maternal SBP on Offspring Birthweight

Maternal Exposure	Outcome	Number of	Bota	SE	D-value
Systolic blood	Maternal effect on BW	JINI 3	Deta	JL	I -value
pressure	(unconditional)	68	-0.0158	0.0022	6.13E-10
	Fetal effect on BW				
(mmHg)	(unconditional)	68	-0.0082	0.0020	1.43E-04
	Maternal effect on BW				
	(adjusted)	68	-0.0148	0.0021	1.98E-09
	Fetal effect on BW				
	(adjusted)	68	-0.0008	0.0019	6.70E-01

(ii) SNP_m ── → BW

BW

SNP_f -

- No confounding from offspring genotype
- Very large sample sizes

Using Maternal Effects to Investigate the Effect of Birth Weight on Later Life Outcomes









Relationship between maternal and child SNPs, offspring birthweight (BW) and systolic blood pressure (SBP). Each model provides an explanation for the negative <u>genetic correlation</u> between birthweight and SBP observed in Horikoshi et al (2016). SNPs may exert maternal and/or fetal genetic effects on birthweight. A red cross indicates a blocked path due to conditioning on offspring/maternal genotype. The dashed black path represents the possibility of genetic pleiotropy mediated through the fetal genome. The dashed red path represents possible postnatal effects of maternal SNPs. **G**FA

International Journal of Epidemiology, 2019, 1-15 doi: 10.1093/ije/dyz019 Original article



ETTER

doi:10.1038/nature1980

Genome-wide associations for birth weight and correlations with adult disease

Momoko Horikoshi^{1,2}, Robin N. Beaumont³, Felix R. Day⁴, Nicole M. Warrington^{5,6}, Marjolein N. Kooijman^{7,8,9} Momoko Horikoshi^{12,a}, Robin N. Beaumoni⁴, Felix R. Day⁴, Nicole M. Warrington^{5,as}, Marjolein N. Kozijman^{12,as}, Juan Fernandez, Tajes^{1,a}, Bjark enestral¹⁰, Natile N. and Zydan^{12,4}, Kyle I. Guiton^{11,4}). Nels Garap¹², Jonathan B. Bradfield¹¹, David P. Strachan¹⁴, Ruifang Li-Gas², Tarurweer S. Althuvalia^{12,4,5}, EdiX Kreiner¹⁸, Rico Rueed^{12,4,5}, Loo - Pekka Lytykiiane^{14,5}, Diatalia Vilor-Tejedor^{25,2,5,4}, Ying Wu², Elisbeth Thiering^{28,7}, Caral A. Wang², Christian T. Have^{12,4}, Jouke- Jan Hottenga²⁸, Natalia Vilor-Tejedor^{25,2,5,4}, Peter K. Joshi¹⁵, Elisen Tal Hali Boh¹⁵, Joann Natila^{12,5,5}, Riak Mitema²⁶, Anubha Mahajan, Elisbeth M. Nun Leewu^{4,6}, Rizmo Joro²⁵, Visilik Lagui-^{14,5,6}, Micheel Nodersch^{47,6}, Ousie A. Divef⁴⁴, Krim T. Zonderva^{1,5,6}, Mariona Bustamante^{25,2,6,4,4}, Med^{4,4,4,4,4}, Claudi H. 1, Tam^{16,7}, Wing Huing Tam²⁶, Josep M. Mercade^{47,4}, Annada I. Bennett², CHARGE Consortium Hematology Working Group[†], Santhi K. Ganesh⁵¹, Frank J. A. van Rool⁸, Samuel E. Jones³, Po-Ru Loh^{52,53}, Katherine S. Ruth³, Marcus A. Tuke³, Jessica Tyrrell^{3,54}, Andrew R. Wood³, Hanieh Yaghootkar³, Denise M. Scholtens⁴⁰, Atthentie S. kuthr, Marcius A. tuker, Jessies J Yrner¹⁴⁴, J Andrew K. Woodr, "Instinen Lagnootkar", Unitiee M. Schonersen, Luvinia Paternoster²⁵⁵, Inge Prokopenko²⁶, "Peter Koraca", Mustala Atalap²⁷, Star A. Willems⁴, Kallope Fanoutsopoulou²⁶ Xu Wang⁴, Lisbeth Carstensen²⁶, Frank Geller¹⁶, Katharina E. Schrauf⁴⁷, Mustala Matarina E. M. van Beijenerveldr²⁶, Gonneke Willemsen²⁷, Buril V. Rappel¹⁷, Gints E. Koryou^{25,40}, Carolia M. Star, T. Tiesler^{27,40}, Marie Stand¹⁷, Zoltan Kutalik^{20,40}, Stala Ionaka-Guarché¹, David M. Hougaard^{15,44}, Friman Sancher^{45,40}, David Greente^{45,46}, Johanne Wagge¹⁸, Mads V. Holegaard^{16,44}, Huoglinn G. de Haun¹, "First R. Rosendala¹⁰, Carolina Medina-Goner^{25,46}, Johanne M. King^{25,46}, Gibran Hemani^{55,56}, George McMahon⁵⁶, Neil R. Robertson^{1,2}, Christopher J. Groves², Claudia Langenberg⁴, Jian'an Luan⁴ Robert A. Scott⁴, Jing Hua Zhao⁴, Frank D. Mentch¹³, Scott M. MacKenzie⁴¹, Rebecca M. Revnolds⁶⁸, Early Growth Genetics Kobert A. Scott", ing ituz Zhao", irank D. Mentich", Scott M. MacKenze", Rebecca M. Reynolds", jany (orowth / enetecs (EGC) Consortium", William L. Lowe M⁴, Anke Fong^{43,3}, Thorkild I. A. Serensen^{45,56,425}, Harri Ninikoski^{56,73}, Cornelia M. van Duijn⁸, Wieland Kiess²³ Antie Korner^{43,3}, Thorkild I. A. Serensen^{45,56,425}, Harri Ninikoski^{56,73}, Katja Pahkala^{56,85}, Oli I. Raitakari^{36,79}, Eleftheria Zeggin⁵⁰, George V. Dedoussia⁵⁵, Yik-Ving Teo ^{31,86,8}, Seang-Mei Saw^{31,52} Mats Meblye^{30,83,84}, Harry Campbell¹⁹, James F. Wilson^{22,85}, Martino Vrijbeld^{22,33,1}, Eco J. C. N. de Geuz^{8,86}, Dorret I. Boomsma²⁸, Haja N. Kadarmideen⁸⁷, Jens-Christian Holm^{12,61}, Torben Hansen¹², Sylvain Sebert^{57,88,89} Lorrent Loomstand, Tapa A. Raulamiscon, J. Surv. Linestant non-, 10580 named, Styvina Selection, Statistical Andre, Statistica Andre, Statistical Andre, Charden Statistica, Statistical Andre, Statistical Andre, Statistical Andre, Andre, Andre, Statistical Andre, Charden Statistica, Charden Statistica, Andre, Statistical Andre, Statistical Andre, Andr Debbie A. Lawler^{85,26}, Timoto M. Frayling³, George Davey Smith^{55,56}, Andrew P. Morris^{1,21,4138}, Ken K. Ong^{4,19}8, Janine F. Feliz^{7,248}, Nicholas J. Timpson^{85,26}8, John R. B. Perry⁶8, David M. Evans^{3,55,68}, Mark I. McCarthy^{15,119} & Rachel M. Freutry^{15,21}

and maternal factors and in observational studies is reproducibly are in part the result of shared genetic effects and identify some of the associated with future risk of adult metabolic diseases including pathways through which these causal genetic effects are mediated. type 2 diabetes (T2D) and cardiovascular disease¹. These life-We combined GWAS data for BW from 153,781 individuals repcourse associations have often been attributed to the impact of an resenting multiple ancestries from 37 studies across three comp adverse early life environment. Here, we performed a multi-ancestry nents (Extended Data Fig. 1 and Supplementary Table 1); (i) 75,891 ne-wide association study (GWAS) meta-analysis of BW in individuals of European ancestry from 30 studies; (ii) 67,786 individuals genome-wide association study (GWAS) meta-anarysis of Diff in 153,781 individuals, identifying 60 loci where fetal genotype was of European ancestry from the UK Biobank; and (iii) 10,104 individuals associated with BW ($P < 5 \times 10^{-8}$). Overall, approximately 15% of diverse ancestries (African American, Chinese, Filipino, Surinamese of variance in BW was captured by assays of fetal genetic variation. Turkish and Moroccan) from six studies. Within each study, BW was Using genetic association alone, we found strong inverse genetic Z-score transformed separately in males and females after excluding correlations between BW and systolic blood pressure ($R_g = -0.22$, $P = 5.5 \times 10^{-13}$), T2D ($R_g = -0.27$, $P = 1.1 \times 10^{-6}$) and coronary where available. Genotypes were imputed using reference panels from artery disease ($R_g = -0.30^6$, $P = 6.5 \times 10^{-9}$). In addition, using large - the 1000 Genomes (1000G) Project² or combined 1000G and UK10K cohort datasets, we demonstrated that genetic factors were the projects3 (Supplementary Table 2). We performed quality control major contributor to the negative covariance between BW and future assessments to confirm that the distribution of BW was consistent cardiometabolic risk. Pathway analyses indicated that the protein across studies, irrespective of the data collection protocol, and products of genes within BW-associated regions were enriched for confirmed that self-reported BW in the UK Biobank showed genetic diverse processes including insulin signalling, glucose homeostasis, glycogen biosynthesis and chromatin remodelling. There was also BW in other studies⁴ (Methods). enrichment of associations with BW in known imprinted regions We identified 60 loci (of which 59 were autosomal) associated with $(P=1.9\times10^{-4})$. We demonstrate that life-course associations BW at genome-wide significance $(P < 5 \times 10^{-8})$ in either the European

A list of affiliations appears in the online version of this pape

248 | NATURE | VOL 538 | 13 OCTOBER 2016

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved

bioRxiv preprint first posted online Oct. 17, 2018; doi: http://dx.doi.org/10.1101/442756. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-8Y 4 0 international locense.

Title:

Maternal and fetal genetic effects on birth weight and their relevance to cardio-metabolic risk factors

Authors and affiliations:

Nicole M Warrington^{1,a}, Robin N Beaumont^{2,a}, Momoko Horikoshi^{3,4,5,a}, Felix R Day^{6,a}, Øyvind Helgeland^{7,8,9,a}, Charles Laurin^{10,11}, Jonas Bacelis¹², Shouneng Peng^{13,14}, Ke Hao^{13,14}, Biarke 10 Feenstra¹⁵, Andrew R Wood², Anubha Mahajan^{3,4}, Jessica Tyrrell^{2,16}, Neil R Robertson^{3,4}, N William Rayner^{3,4,17}, Zhen Qiao¹, Gunn-Helen Øiseth Moen^{18,19}, Marc Vaudel⁸, Carmen J 11 Marsit²⁰, Jia Chen²¹, Michael Nodzenski²², Theresia M Schnurr²³, Mohammad H Zafarmand^{24,25}, Jonathan P Bradfield^{26,27}, Niels Grarup²³, Marjolein N Kooijman^{28,29,30}, 13 14 Ruifang Li-Gao³¹, Frank Geller¹⁵, Tarunveer S Ahluwalia^{23,32,33}, Lavinia Paternoster¹⁰, Rico Rueedi^{34,35}, Ville Huikari³⁶, Jouke-Jan Hottenga^{37,38,39}, Leo-Pekka Lyytikäinen⁴⁰, Alana Cavadino^{41,42}, Sarah Metrustry⁴³, Diana L Cousminer^{44,45}, Ying Wu⁴⁶, Elisabeth Thiering^{47,48}, 15 16 Carol A Wang⁴⁹, Christian T Have²³, Natalia Vilor-Tejedor^{50,51}, Peter K Joshi⁵², Jodie N 17 Painter53, Ioanna Ntalla54, Ronny Myhre55, Niina Pitkänen56, Elisabeth M van Leeuwen25 18 Raimo Joro⁵⁷, Vasiliki Lagou^{3,58,59}, Rebecca C Richmond^{10,11}, Ana Espinosa^{60,61,62,63}, Sheila J Barton⁶⁴, Hazel M Inskip^{64,65}, John W Holloway⁶⁶, Loreto Santa-Marina^{60,67,68}, Xavier 19 20 Estivill^{69,70}, Wei Ang⁷¹, Julie A Marsh⁷¹, Christoph Reichetzeder^{72,73}, Letizia Marullo⁷⁴ 21 Berthold Hocher^{72,75}, Kathryn L Lunetta^{76,77}, Joanne M Murabito^{77,78}, Caroline L Relton^{10,11,79} 22 Manolis Kogevinas^{66,61,26,38}, ded Chatz¹⁰, Catherine Allard⁴¹, Luigi Bouchard^{13,82,83}, Marie-France Hiver^{84,85,86}, Ge Zhang^{37,88,89}, Louis J Muglia^{37,88,89}, Jani Heikkinen⁹⁰, Early Growth Genetics (EGG) Consortium⁵, Camilla S Morgen⁹¹, Antoine HC van Kampen^{92,39}, Barbera DC 23 24 25 van Schaik⁹², Frank D Mentch²⁶, Claudia Langenberg⁶, Jian'an Luan⁶, Robert A Scott⁶, Jing Hua Zhao⁶, Gibran Hemani^{10,11}, Susan M Ring^{10,11}, Amanda J Bennett⁴, Kyle J Gaulton^{3,94}, Juan Fernandez-Tajes³, Natalie R van Zuydam^{3,4}, Carolina Medina-Gomez^{18,29,95}, Hugoline G 27 de Haan³¹, Frits R Rosendaal³¹, Zoltán Kutalik^{34,96}, Pedro Marques-Vidal⁹⁷, Shikta Das⁴². 29 Gonneke Willemsen^{37,38,39}, Hamdi Mbarek^{37,38,39,98}, Martina Müller-Nurasyid^{99,100,101}. Marie Standl⁴⁷, Emil V R Appel²³, Cilius E Fonvig^{23,102,103}, Caecilie Trier^{23,102}, Catharina EM van 31 Beijsterveldt³⁸, Mario Murcia^{60,104}, Mariona Bustamante^{60,61,62}, Sílvia Bonas-Guarch¹⁰⁵, David 32 M Hougaard¹⁰⁶, Josep M Mercader^{105,107,108}, Allan Linneberg^{109,110}, Katharina E Schraut⁵² 33 Penelope A Lind⁵³, Sarah E Medland⁵³, Beverley M Shields¹¹¹, Bridget A Knight¹¹¹, Jin-Fang Chai¹¹², Kalliope Panoutsopoulou¹⁷, Meike Bartels^{37,38,39}, Friman Sánchez^{105,113}, Jacob Stokholm³², David Torrents^{105,114}, Rebecca K Vinding³², Sara M Willems²⁹, Mustafa Atalay⁵⁷ 35 Bo L Chawes³², Peter Kovacs¹¹⁵, Inga Prokopenko^{3,116}, Marcus A Tuke², Hanieh 27 38 Yaghootkar¹¹⁷, Katherine S Ruth², Samuel E Jones², Po-Ru Loh^{118,119}, Anna Murray², Michael N Weedon², Anke Tönjes¹²⁰, Michael Stumvoll^{115,120}, Kim F Michaelsen¹²¹, Aino-Maija 39 Eloranta⁵⁷, Timo A Lakka^{57,122,123}, Cornelia M van Duijn²⁹, Wieland Kiess¹²⁴, Antje 40 Körner^{115,124}. Harri Niinikoski^{125,126}, Katja Pahkala^{56,127}, Olli T Raitakari^{56,128}, Bo Jacobsson^{7,12} 41 Eleftheria Zeggini¹⁷, George V Dedoussis¹²⁹, Yik-Ying Teo^{112,130,131}, Seang-Mei Saw^{112,132} 42 Grant W Montgomery⁵³, Harry Campbell⁵², James F Wilson^{52,133}, Tanja GM Vrijkotte²⁴, 43 44 Martine Vrijheid^{60,61,62}, Eco JCN de Geus^{37,38,39}, M Geoffrey Hayes¹³⁴, Haja N Kadarmideen¹³ Jens-Christian Holm^{23,102}, Lawrence J Beilin¹³⁶, Craig E Pennell⁴⁹, Joachim Heinrich^{47,137}, Linda 45 S Adair¹³⁸, Judith B Borja^{139,140}, Karen L Mohlke⁴⁶, Johan G Eriksson^{141,142,143}, Elisabeth E 47

Widén⁹⁰, Andrew T Hattersley^{2,111}, Tim D Spector⁴³, Mika Kähönen¹⁴⁴, Jorma S Viikari^{145,146},

Original article

Elucidating the role of maternal environmental exposures on offspring health and disease using two-sample Mendelian randomization

David M Evans (), 1,2,3* Gunn-Helen Moen, 4,5 Liang-Dar Hwang, 1 Debbie A Lawlor^{2,3,6} and Nicole M Warrington¹

¹University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Queensland, Australia, ²Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK, ³Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK, ⁴Department of Endocrinology, Morbid Obesity and Preventive Medicine, Oslo University Hospital, Oslo, Norway, ⁵Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway and ⁶Bristol NIHR Biomedical Research Centre, Bristol, UK

*Corresponding author, University of Queensland Diamantina Institute, University of Queensland, Queensland, QLD 4102, Australia. E-mail: d.evans1@uq.edu.au.

Editorial decision 23 January 2019: Accepted 8 February 2019

Abstract

Background: There is considerable interest in estimating the causal effect of a range of maternal environmental exposures on offspring health-related outcomes. Previous attempts to do this using Mendelian randomization methodologies have been hampered by the paucity of epidemiological cohorts with large numbers of genotyped motheroffspring pairs.

Methods: We describe a new statistical model that we have created which can be used to estimate the effect of maternal genotypes on offspring outcomes conditional on offspring genotype, using both individual-level and summary-results data, even when the extent of sample overlap is unknown.

Results: We describe how the estimates obtained from our method can subsequently be used in large-scale two-sample Mendelian randomization studies to investigate the causal effect of maternal environmental exposures on offspring outcomes. This includes studies that aim to assess the causal effect of in utero exposures related to fetal growth restriction on future risk of disease in offspring. We illustrate our framework using examples related to offspring birthweight and cardiometabolic disease, although the general principles we espouse are relevant for many other offspring phenotypes.

Conclusions: We advocate for the establishment of large-scale international genetics consortia that are focused on the identification of maternal genetic effects and committed to the public sharing of genome-wide summary-results data from such efforts. This information will facilitate the application of powerful two-sample Mendelian randomization studies of maternal exposures and offspring outcomes

Birth weight (BW) has been shown to be influenced by both fetal between early growth phenotypes and adult cardiometabolic disease




- Two year postdoctoral position
- One year postdoctoral position
- PhD position



34th International Statistical Genetics Workshop – Boulder CO 2020

Monday	Tuesday	Wednesday	Thursday	Friday
Intro 2 Unix & R	plink & GRMs	GWAS & genetic data formats	LD score regression & genetic correlation	Simulation
Causes of Variation	ACE: from twins to GRMs	Multivariate Models	Developmental Models	Genomic SEM
Biometrical Genetics	Binary & Ordinal Data & Measurement	Genetic Factor Models	Direction of Causation DOC	Model Assumptions & Extended Pedigrees
OpenMx: Regression & tools 4 SEM	Multilevel Models	GREML in OpenMx	PRS & Mendelian Randomization MR	Statistical Power
ACE Model (full script + umx)	Heterogeneity: Sex & GxE interaction	GW-SEM	Combined DOC-MR	Lightning Rounds

Thank you for participating! See you next year?!