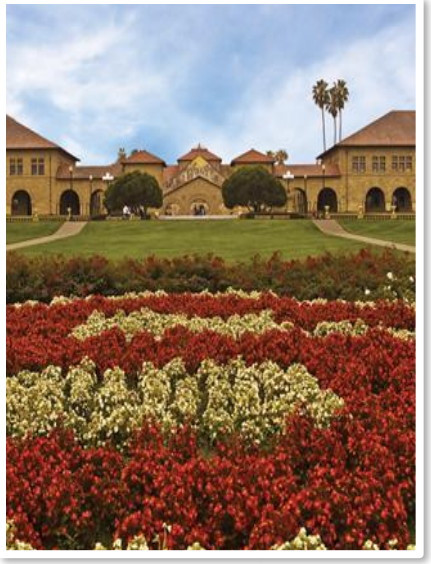
# Using genetic predictors

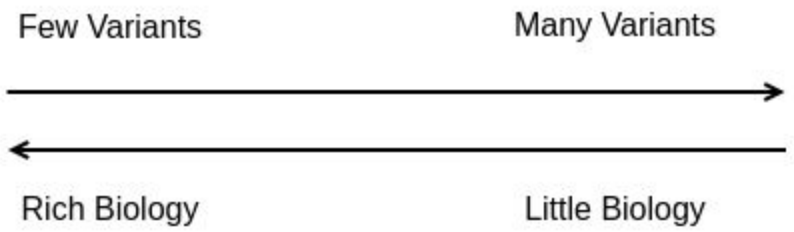
Oxford NCRM Summer School
Introduction to Using Molecular Genetic Data in the Social Sciences

BEN DOMINGUE
bdomingue@stanford.edu



# Genetic Predictors

(Predictive) Power/Mechanisms trade-off




- Always important to think about the trade-off when incorporating a genetic predictor.

# What can we do with limited sets of variants

Similar, perhaps, to earlier candidate gene era of research.

- Genes as Instruments
- Genes as modifiers

# What can we do with limited sets of variants

Similar, perhaps, to earlier candidate gene era of research.

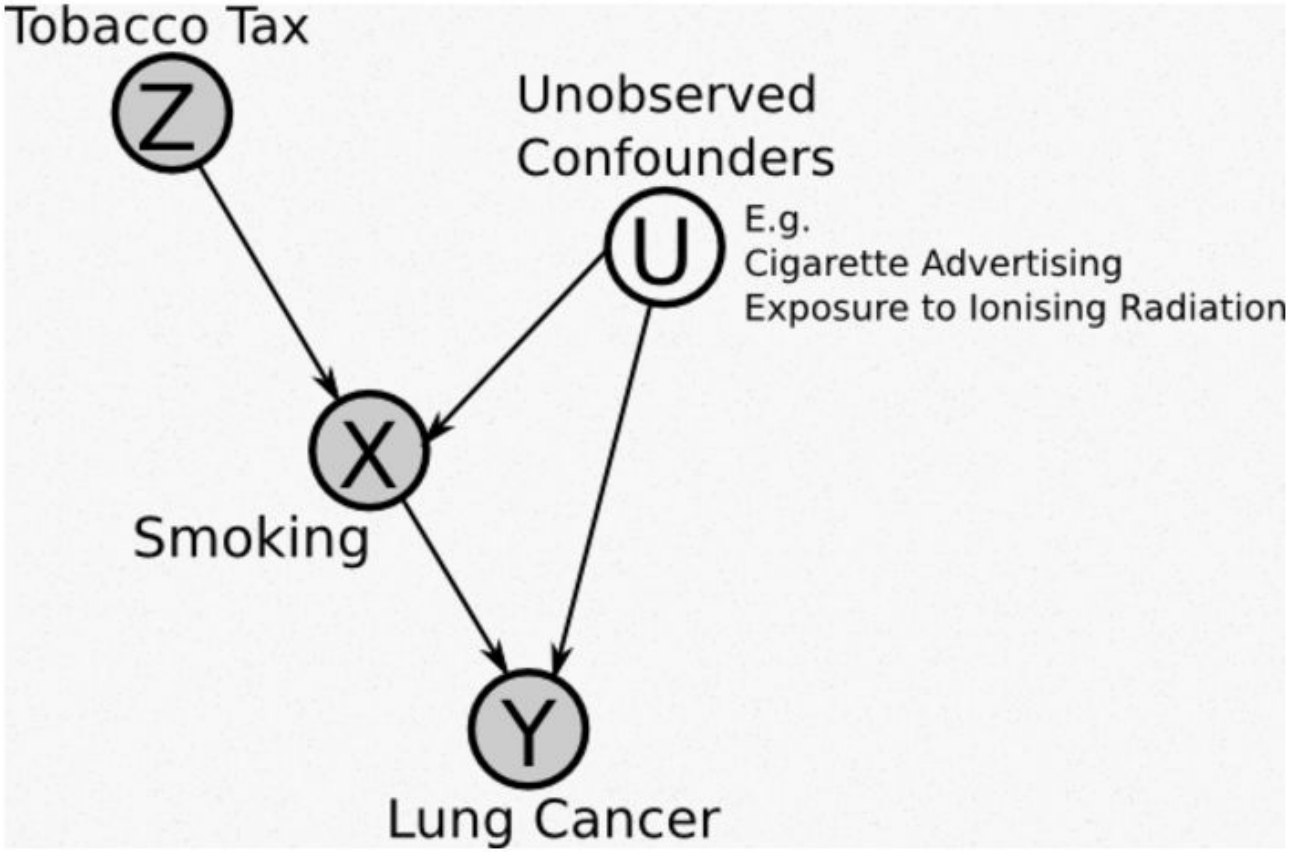
- Genes as Instruments
- Genes as modifiers

Crucial Question: What evidence do you have that the variants of interest are robustly associated with outcome?

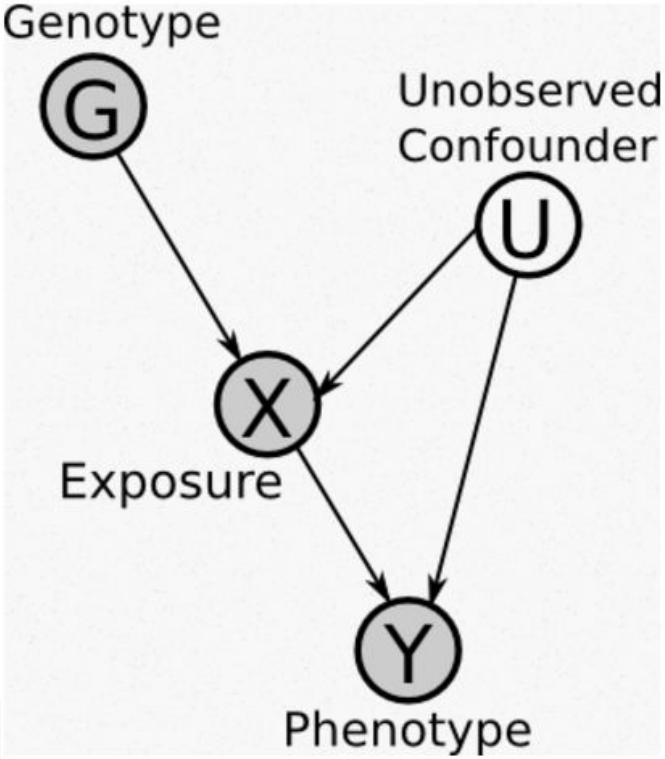
- How do they look in recent GWAS?

# Mendelian Randomization

- Basic idea: genes as instrumental variables




http://jamesmcm.github.io/blog/2014/08/17/mendelian/




## No Pleiotropy

Pleiotropy is the phenomena whereby one gene can affect many (even seemingly unrelated) phenotypes. Mendelian Randomisation makes the assumption of **no pleiotropy**.

In this case, this means that we assume the genotype is only influencing the phenotype via the considered exposure. I.e. ApoE2 **only** affects serum cholesterol levels, and cannot affect cancer risk by other, unobserved means.

This is a big assumption, and **prior knowledge** is necessary. If possible, using multiple, independent SNPs which act through the same path, can help to alleviate this issue, because, if they are all consistent then it is unlikely that they would all have other pathways causing the same change in phenotype. But note that they must be independent, and so cannot be in Linkage Disequilibrium.

# Alcohol Intake and Blood Pressure: A Systematic Review Implementing a Mendelian Randomization Approach

Lina Chen, George Davey Smith, Roger M Harbord, Sarah J Lewis

Published: March 4, 2008 • http://dx.doi.org/10.1371/journal.pmed.0050052

## Background

Alcohol has been reported to be a common and modifiable risk factor for hypertension. However, observational studies are subject to confounding by other behavioural and sociodemographic factors, while clinical trials are difficult to implement and have limited follow-up time. Mendelian randomization can provide robust evidence on the nature of this associatio by use of a common polymorphism in *aldehyde dehydrogenase 2* (*ALDH2*) as a surrogate for measuring alcohol consumption. *ALDH2* encodes a major enzyme involved in alcohol metabolism. Individuals homozygous for the null variant (**2*2*) experience adverse symptoms when drinking alcohol and consequently drink considerably less alcohol than wild-type homozygotes (**1*1*) or heterozygotes. We hypothesise that this polymorphism may influence the risk of hypertension by affecting alcohol drinking behaviour.

| Data Type | Reference | No. of Participants | Alcohol Exposure Groups, Grams of Ethanol per Day | Sex | Alcohol Status by Genotype[a] | | |
|---|---|---|---|---|---|---|---|
| | | | | | *1*1 | *1*2 | *2*2 |
| Categorical data | Saito et al., 2003 [28] | 335 | <18.6 | Male | 39 | 87 | 21 |
| | | | 18.6–37 | | 54 | 33 | 0 |
| | | | ≥37.1 | | 81 | 15 | 0 |
| | | | Ex-drinker | | 3 | 1 | 0 |
| | Yamada et al., 2002 [29] | 828 | Nondrinker | Male | 21 | 56 | 28 |
| | | | <2.26 | | 75 | 86 | 16 |
| | | | 2.27–47.4 | | 114 | 75 | 2 |
| | | | 47.5–71.1 | | 210 | 64 | 2 |
| | | | >71.1 | | 63 | 16 | 0 |
| Continuous data | Amamoto et al., 2002 [18] | 2,035 | N/D | Male | 26.3 | 15.5 | 0.66 |
| | | | N/D | Female | 1.4 | 0.5 | 0.03 |
| | Hashimoto et al., 2002 [13] | 133 | N/D | Male | 54.9 ± 13.1 | 51 ± 9.8 | — |
| | Mackenzie et al., 2005 [12] | 28 | N/D | Mixed | 23.7 ± 39.5 | 31.6 ± 39.5 | — |
| | Okayama et al., 1994 [14] | 159 | N/D | Male | 26.8 ± 19.1 | 17.6 ± 17.9 | — |
| | Takagi et al., 2001 [19] | 4,057 | N/D | Male | 23.7 ± 16.7 | 11.5 ± 15.8 | 1.1 ± 16.9 |
| | | | N/D | Female | 3.8 ± 6.0 | 1.2 ± 6.8 | 0.08 ± 6.5 |
| | Tsuritani et al., 1995 [20] | 403 | N/D | Male | 42.7 ± 28.8 | 22.9 ± 25.1 | 2.4 ± 3.6 |

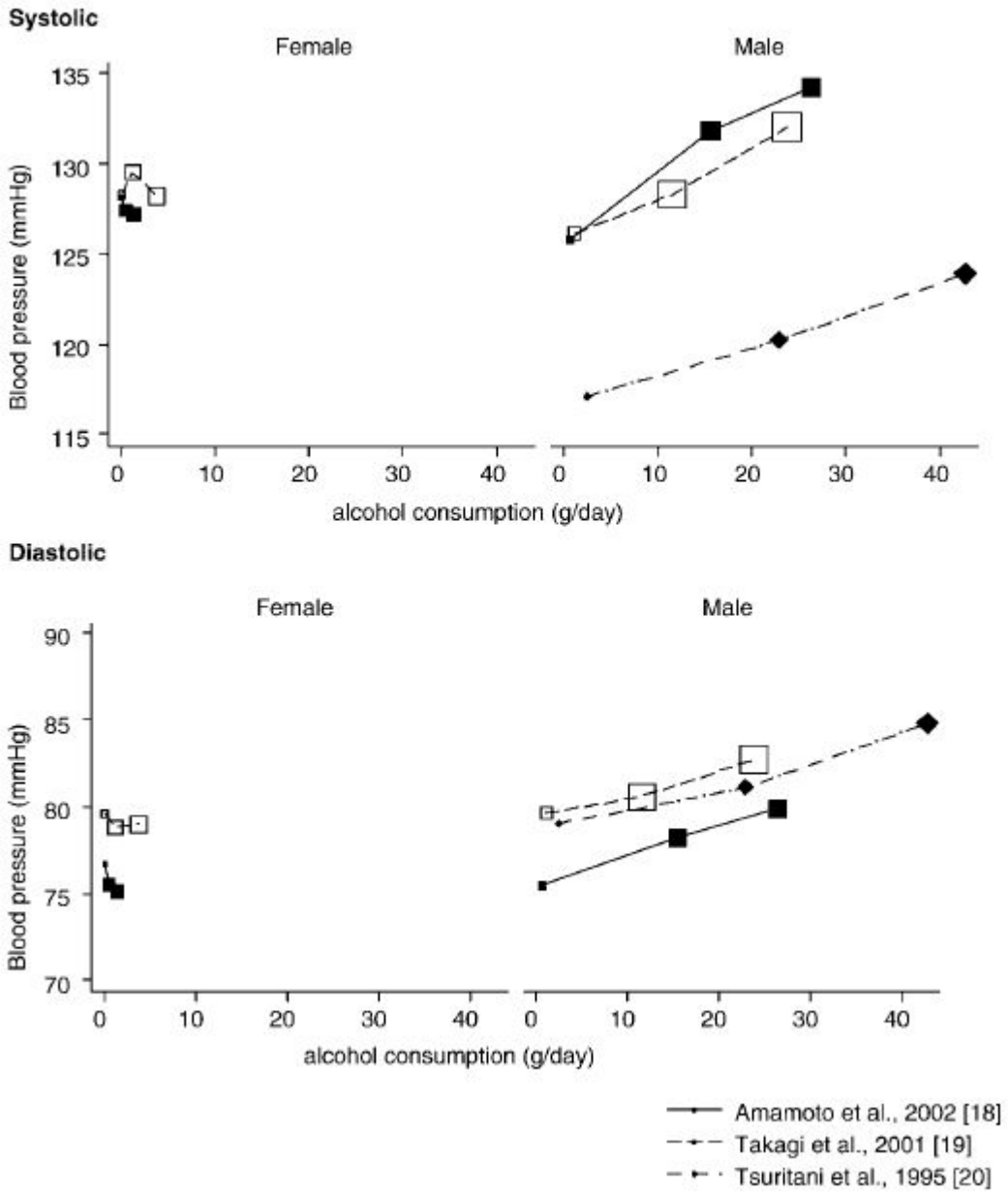
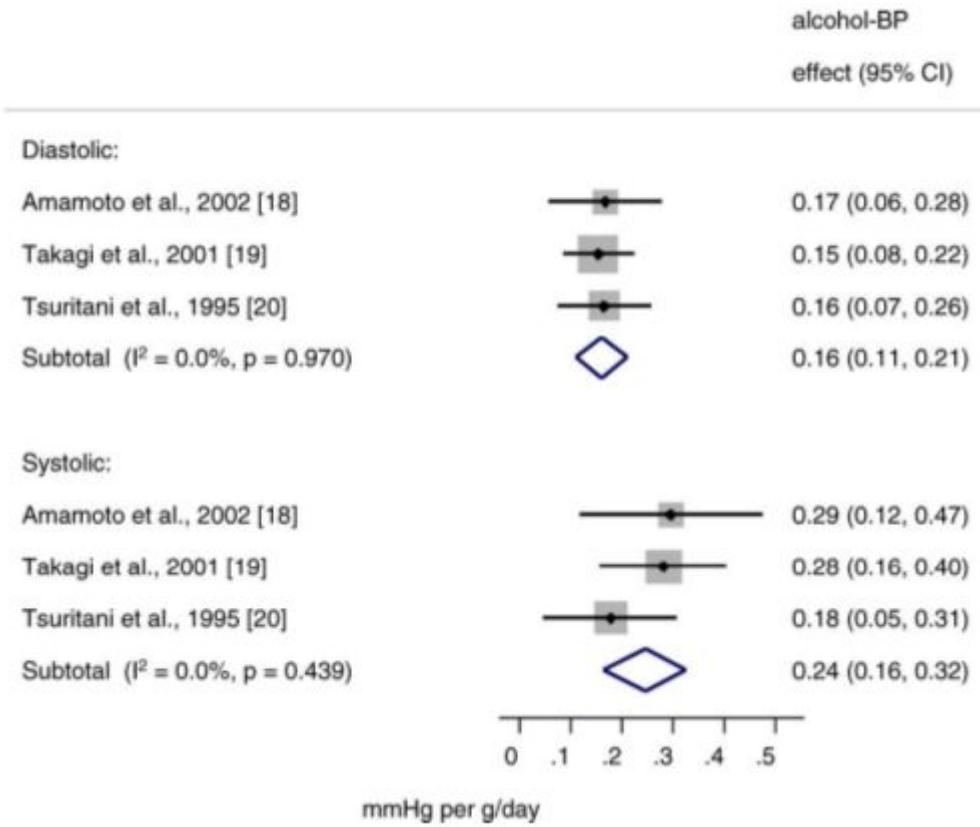

Systolic
Female
Male
135
130
125
120
115
Blood pressure (mmHg)
0
10
20
30
40
alcohol consumption (g/day)
Diastolic
Female
Male
90
85
80
75
70
Blood pressure (mmHg)
0
10
20
30
40
alcohol consumption (g/day)
Amamoto et al., 2002 [18]
Takagi et al., 2001 [19]
Tsuritani et al., 1995 [20]

**Table 3.** Distribution of Potential Confounding Factors by Genotype among Studies Included in the Meta-analysis

| Author | No. of Participants | Covariates | Category | Covariates by Genotypes[a] | | | Published *p*-Values[b] |
|---|---|---|---|---|---|---|---|
| | | | | *1*1 | *1*2 | *2*2 | |
| Amamoto et al., 2002 [18] | 2,035 | Sex | Male | 335 (44.7) | 351 (46·9) | 63 (8.4) | — |
| | | | Female | 644 (50.1) | 555 (43.2) | 87 (6.8) | — |
| | | Age | Male | 56.8 ± 15.7 | 58.1 ± 15.6 | 52.8 ± 17.7 | 0.05 |
| | | | Female | 54.8 ± 16.0 | 55.1 ± 15.4 | 58.5 ± 14.3 | 0.12 |
| | | BMI | Male | 22.9 ± 3.0 | 22.3 ± 2.8 | 21.9 ± 2.7 | 0.004 |
| | | | Female | 22.5 ± 3.1 | 22.4 ± 3.3 | 22.4 ± 3.1 | 0.89 |
| | | Current smoker | Male | 176 (52.5) | 173 (49.3) | 36 (56.5) | 0.49 |
| | | | Female | 39 (6.1) | 48 (8.6) | 9 (10.5) | 0.16 |
| | | Cigarette smoking, cigarettes per day | Male | 10.4 | 10.8 | 11.5 | 0.78 |
| | | | Female | 0.69 | 1.04 | 1.02 | 0.17 |
| Hashimoto et al., 2002 [13] | 133 | Age | — | 46.0 ± 7.1 | 46.2 ± 5.8 | — | >0.05 |
| | | BMI | — | 22.5 ± 2.0 | 22.2 ± 1.6 | — | >0.05 |
| | | Smoking, cigarettes per day | — | 14.2 ± 16.3 | 12.7 ± 13.8 | — | >0.05 |
| | | Exercise, times per month | — | 2.6 ± 3.1 | 2.6 ± 2.7 | — | >0.05 |
| Iwai et al., 2004 [31] | 1,849 | Sex | Male | 443 (51.6) | 348 (40.6) | 67 (7·8) | — |
| | | | Female | 520 (52.3) | 381 (38.3) | 93 (9·4) | — |
| Mackenzie et al., 2005 [12] | 28 | Sex | Male | 8 (47.1) | 9 (52.9) | — | — |
| | | Age | Female | 9 (81.8) | 2 (18.2) | — | — |
| | | | — | 30 ± 7 (12); 29 ± 5 (5) | 31 ± 10 | — | — |
| | | BMI | — | 23.8 ± 3.1(12); 22.0 ± 3.4 (5) | 23.0 ± 3.1 | — | — |
| Nishimura et al., 2002 [11] | 36 | Age | — | 22.1 ± 3.3 | 23.1 ± 3.4 | — | — |
| Okayama et al., 1994 [14] | 159 | Age | — | 45.7 ± 10.2 | 48.0 ± 9.4 | — | — |
| | | BMI | — | 22.9 ± 2.8 | 21.8 ± 2.7 | — | — |
| Saito et al., 2003 [28] | 335 | Age | — | 53.3 ± 8.6 | 53.4 ± 8.8 | 52.4 ± 8.4 | 0.89 |
| | | BMI | — | 23.5 ± 3.0 | 23.3 ± 3.1 | 23.0 ± 2.5 | 0.71 |
| | | Smoking | Current smokers | 100 (56.3) | 77 (55.9) | 14 (66..7) | 0.12 |
| | | | Never-smokers | 49 (27.8) | 25 (18.4) | 4 (19.1) | |
| | | | Ex-smokers | 28 (15.9) | 35 (25.7) | 3 (14.3) | |
| | | Physical activity | Inactive | 66 (37.5) | 52 (38.2) | 3 (14.3) | 0.32 |
| | | | Moderate | 50 (28.4) | 41 (30.2) | 8 (38.1) | |
| | | | Active | 60 (34.1) | 43 (31.6) | 10 (47.6) | |
| Takagi et al., 2001 [19] | 4,057 | Sex | Male | 924 (48.2) | 825 (43.0) | 170 (8.8) | — |
| | | | Female | 1,112 (52.0) | 838 (39.2) | 188 (8.8) | — |
| | | Age | Male | 60.2 ± 12.2 | 61.6 ± 11.5 | 61.3 ± 11.7 | 0.05 |
| | | | Female | 58.5 ± 13.3 | 59.6 ± 11.6 | 58.·4 ± 12·3 | 0.09 |
| | | BMI | Male | 23.2 ± 3.0 | 22.8 ± 2.9 | 23.1 ± 2.6 | 0.02 |
| | | | Female | 22.4 ± 3.3 | 22.1 ± 2.9 | 22.3 ± 2.7 | 0.11 |
| | | Current smoker | Male | 367 (39.7) | 326 (39.5) | 61 (35.9) | n.s. |
| | | | Female | 93 (8.4) | 68 (8.2) | 13 (7.5) | n.s. |
| Tsuritani et al., 1995 [20] | 403 | Age | — | 46.4 ± 5.8 | 45.7 ± 5.0 | 46.4 ± 4.4 | n.s. |
| | | BMI | — | 24.7 ± 2.8 | 24.3 ± 3.0 | 24.9 ± 2.4 | n.s. |



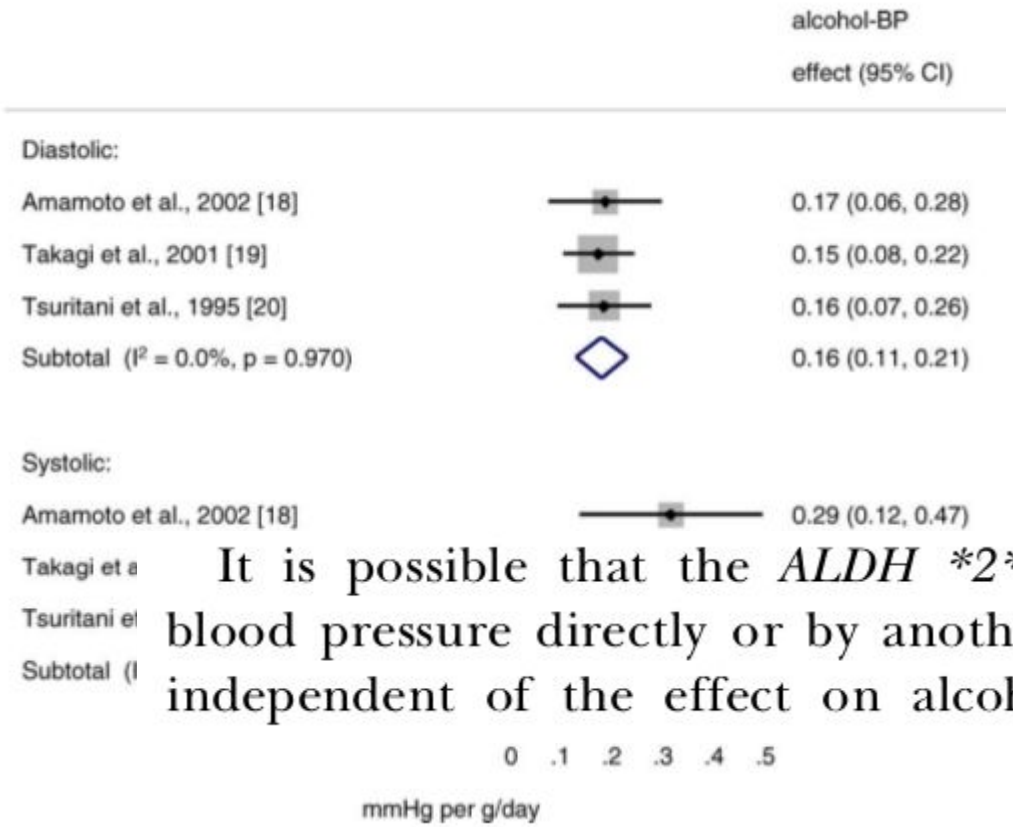


## Methods and Findings

We carried out fixed effect meta-analyses of the *ALDH2* genotype with blood pressure (five studies, $n = 7{,}658$) and hypertension (three studies, $n = 4{,}219$) using studies identified via systematic review. In males, we obtained an overall odds ratio of 2.42 (95% confidence interval [CI] 1.66–3.55, $p = 4.8 \times 10^{-6}$) for hypertension comparing **1*1* with **2*2* homozygotes and an odds ratio of 1.72 (95% CI 1.17–2.52, $p = 0.006$) comparing heterozygotes (surrogate for moderate drinkers) with **2*2* homozygotes. Systolic blood pressure was 7.44 mmHg (95% CI 5.39–9.49, $p = 1.1 \times 10^{-12}$) greater among **1*1* than among **2*2* homozygotes, and 4.24 mmHg (95% CI 2.18–6.31, $p = 0.00005$) greater among heterozygotes than among **2*2* homozygotes.

## Conclusions

These findings support the hypothesis that alcohol intake has a marked effect on blood pressure and the risk of hypertension.




It is possible that the *ALDH *2*2* genotype influenced blood pressure directly or by another mechanism that was independent of the effect on alcohol intake (pleiotropy).


## Methods and Findings

We carried out fixed effect meta-analyses of the *ALDH2* genotype with blood pressure (five studies, $n = 7{,}658$) and hypertension (three studies, $n = 4{,}219$) using studies identified via systematic review. In males, we obtained an overall odds ratio of 2.42 (95% confidence interval [CI] 1.66–3.55, $p = 4.8 \times 10^{-6}$) for hypertension comparing **1*1* with **2*2* homozygotes and an odds ratio of 1.72 (95% CI 1.17–2.52, $p = 0.006$) comparing heterozygotes (surrogate for moderate drinkers) with **2*2* homozygotes. Systolic blood pressure was 7.44 mmHg (95% CI 5.39–9.49, $p = 1.1 \times 10^{-12}$) greater among **1*1* than among **2*2* homozygotes, and 4.24 mmHg (95% CI 2.18–6.31, $p = 0.00005$) greater among heterozygotes than among **2*2* homozygotes.

## Conclusions

These findings support the hypothesis that alcohol intake has a marked effect on blood pressure and the risk of hypertension.

# Genes as modifiers

- Do specific biological mechanisms modify how we respond to “interventions”?
- Similarities to candidate gene type research, but hopefully based upon firmer foundation.
  - Does your genetic variant have robust evidence re its association with some outcome?

# Genetics of smoking

Genome-wide meta-analyses identify multiple loci associated with smoking behavior

The Tobacco and Genetics Consortium*

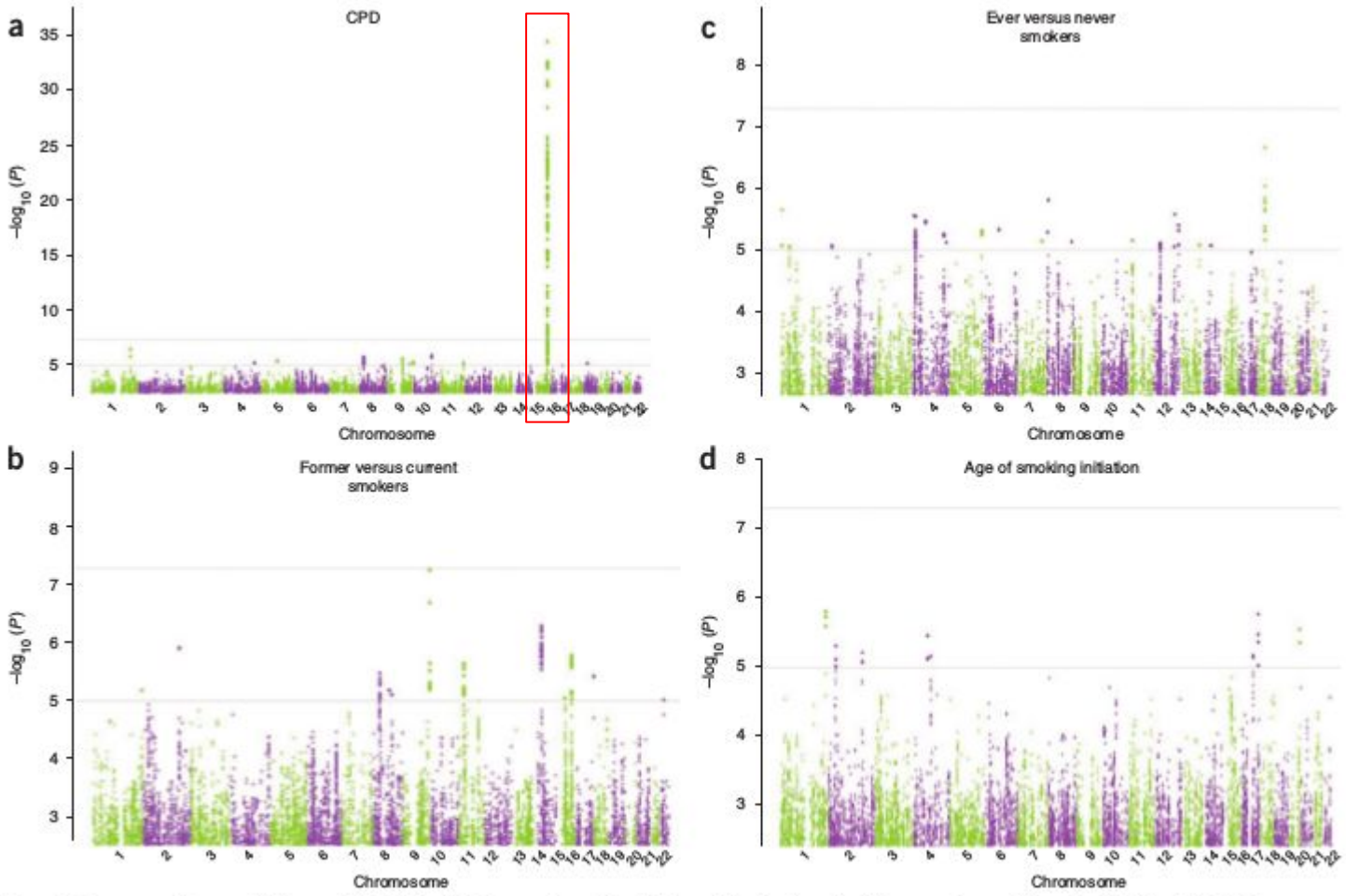



**Figure 1** Genome-wide association results for the TAG Consortium. Manhattan plots showing significance of association of all SNPs in the TAG Consortium meta-analyses for four smoking phenotypes. (**a**–**d**) Manhattan plots show SNPs plotted on the *x* axis according to their position on each chromosome against, on the *y* axis (shown as $-\log_{10}$ *P* value), the association with CPD (**a**), former versus current smoking (**b**), ever versus never smoking (**c**) and age of smoking initiation (**d**).

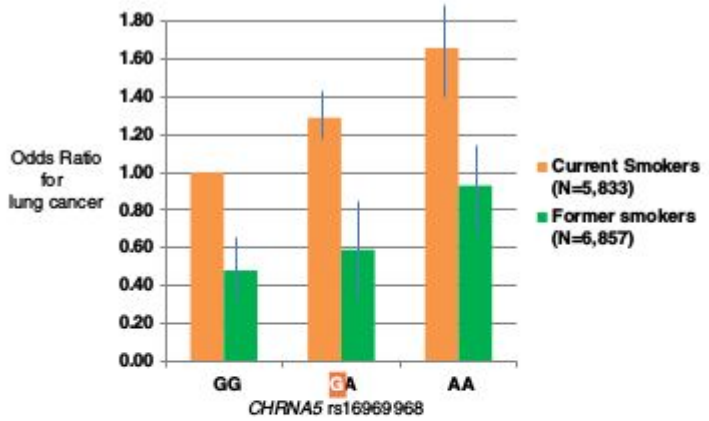
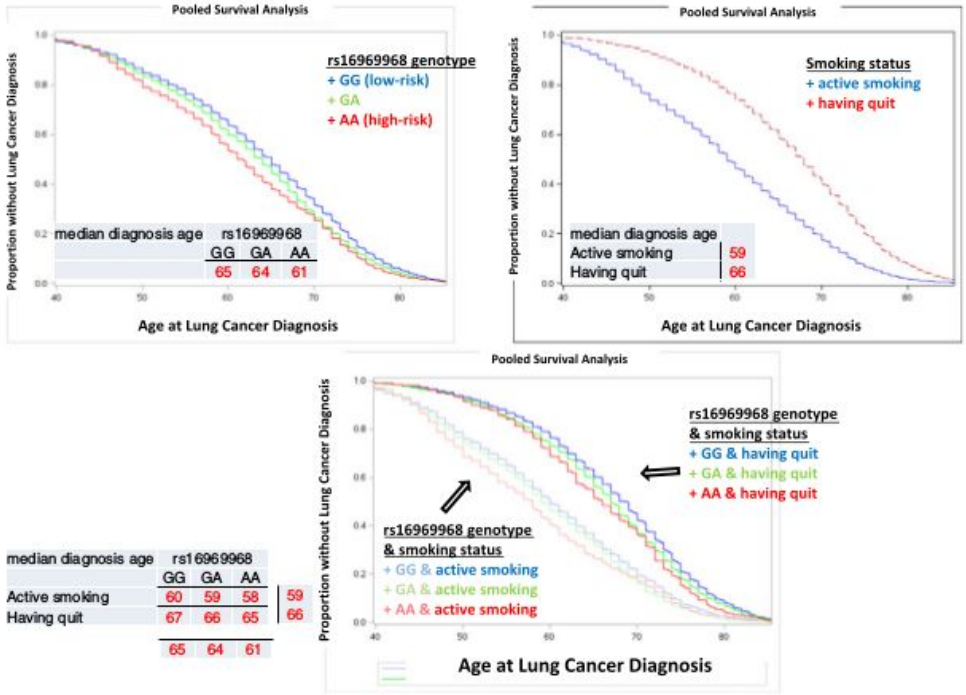


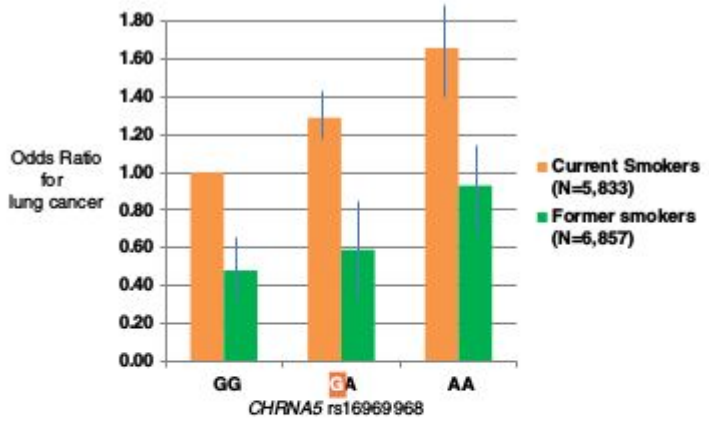

**Fig. 1.** *CHRNA5* rs16969968 predicts risk of lung cancer. Smoking cessation decreases probability of lung cancer regardless of *CHRNA5* rs16969968 genotype. Quit: individuals who report having quit smoking when ascertained as cases of lung cancer or controls. N = 12,690 (6988 cases of lung cancer and 5702 controls). All participants have smoked at least 100 cigarettes (5833 active smokers, 6857 smokers who have quit). Adjusted for age, sex and rs16969968 genotype.




Quit: individuals who report having quit smoking when ascertained as cases of lung cancer.
N=6,988 cases of lung cancer who are ever smokers. (3,471 current smokers, 3,517 former smokers who have quit.)
Adjusted for sex, study, and rs16969968.

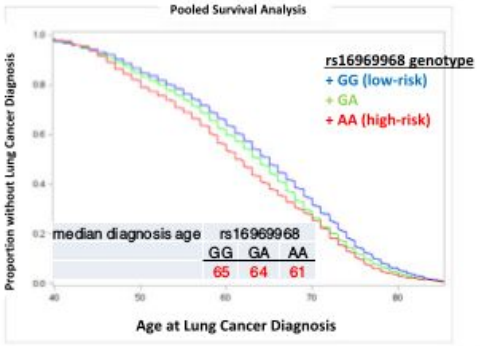
**Fig. 2.** *CHRNA5* rs16969968 predicts earlier lung cancer. Smoking cessation delays lung cancer, regardless of *CHRNA5* rs16969968 genotype. Quit: individuals who report having quit smoking when ascertained as cases of lung cancer. N = 6,988 cases of lung cancer who are ever smokers. (3,471 current smokers, 3,517 former Smokers. Adjusted for sex, study, and rs16969968 genotype.

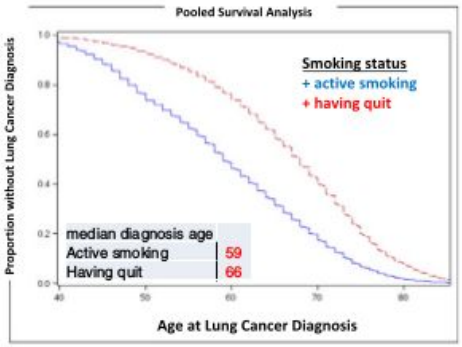
# Genetic Risk Can Be Decreased: Quitting Smoking Decreases and Delays Lung Cancer for Smokers With High and Low *CHRNA5* Risk Genotypes — A Meta-Analysis

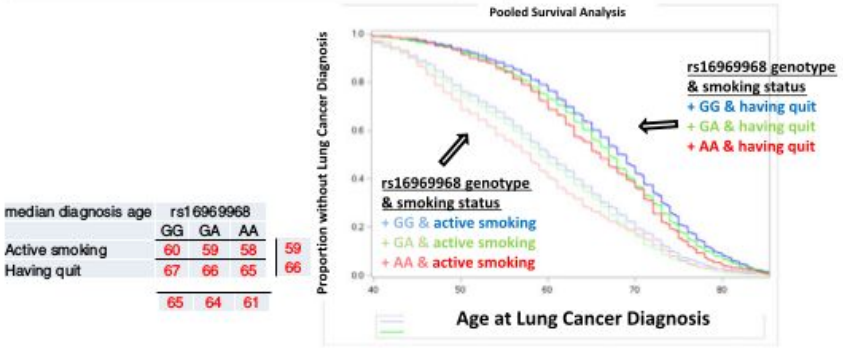



Quit: individuals who report having quit smoking when ascertained as cases of lung cancer or controls.
N=12,690 (6988 cases of lung cancer and 5702 controls)
All subjects have smoked at least 100 cigarettes (5833 active smokers, 6857 smokers who have quit)
Adjusted for age, sex and rs16969968.

**Fig. 1.** *CHRNA5* rs16969968 predicts risk of lung cancer. Smoking cessation decreases probability of lung cancer regardless of *CHRNA5* rs16969968 genotype. Quit: individuals who report having quit smoking when ascertained as cases of lung cancer or controls. N = 12,690 (6988 cases of lung cancer and 5702 controls). All participants have smoked at least 100 cigarettes (5833 active smokers, 6857 smokers who have quit). Adjusted for age, sex and rs16969968 genotype.




Quit: individuals who report having quit smoking when ascertained as cases of lung cancer.
N=6,988 cases of lung cancer who are ever smokers. (3,471 current smokers, 3,517 former smokers who have quit.)
Adjusted for sex, study, and rs16969968.

**Fig. 2.** *CHRNA5* rs16969968 predicts earlier lung cancer. Smoking cessation delays lung cancer, regardless of *CHRNA5* rs16969968 genotype. Quit: individuals who report having quit smoking when ascertained as cases of lung cancer. N = 6,988 cases of lung cancer who are ever smokers. (3,471 current smokers, 3,517 former Smokers. Adjusted for sex, study, and rs16969968 genotype.

But why?

## Genetic Risk Can Be Decreased: Quitting Smoking Decreases and Delays Lung Cancer for Smokers With High and Low *CHRNA5* Risk Genotypes — A Meta-Analysis

The mechanisms through which *CHRNA5* increases the risk for, and accelerates the onset age of, lung cancer likely involve multiple direct or indirect pathways (mediated through cigarette smoking (Liu et al., 2010; Saccone et al., 2010; TAG, 2010; Thorgeirsson et al., 2010), deeper inhalation of cigarettes leading to higher carcinogen exposure (Bloom et al., 2014; Le Marchand et al., 2008), and a delay in smoking cessation (Chen et al., 2015a). There has been inconsistent evidence on whether

# Implicated in cessation efforts

## Interplay of Genetic Risk Factors (*CHRNA5-CHRNA3-CHRNB4*) and Cessation Treatments in Smoking Cessation Success

Li-Shiun Chen, M.D., M.P.H., Sc.D.
Timothy B. Baker, Ph.D.
Megan E. Piper, Ph.D.
Naomi Breslau, Ph.D.
Dale S. Cannon, Ph.D.
Kimberly F. Doheny, Ph.D.
Stephanie M. Gogarten, Ph.D.
Eric O. Johnson, Ph.D.
Nancy L. Saccone, Ph.D.
Jen C. Wang, Ph.D.
Robert B. Weiss, Ph.D.
Alison M. Goate, D.Phil.
Laura Jean Bierut, M.D.

**Objective:** Smoking is highly intractable, and the genetic influences on cessation are unclear. Identifying the genetic factors affecting smoking cessation could elucidate the nature of tobacco dependence, enhance risk assessment, and support development of treatment algorithms. This study tested whether variants in the nicotinic receptor gene cluster *CHRNA5-CHRNA3-CHRNB4* predict age at smoking cessation and relapse after an attempt to quit smoking.

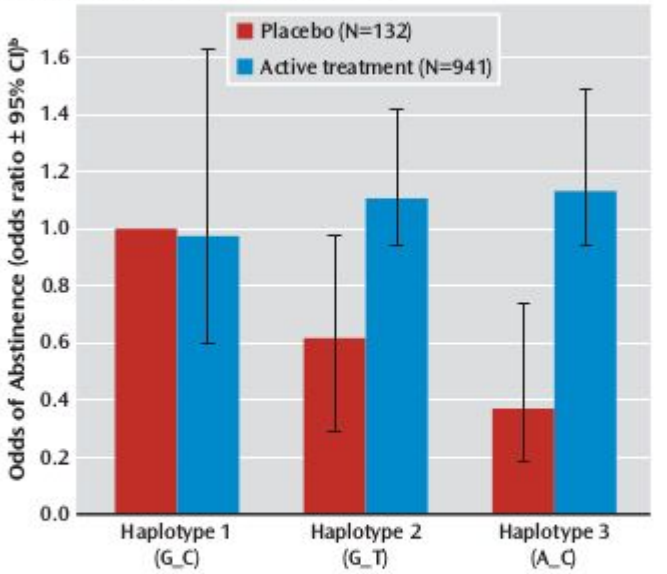
**Method:** In a community-based, cross-sectional study (N=5,216) and a randomized comparative effectiveness smoking cessation trial (N=1,073), the authors used Cox proportional hazard models and logistic regression to model the relationships of smoking cessation (self-reported quit age in the community study and point-prevalence abstinence at the end of treatment in the clinical trial) to three common haplotypes in the *CHRNA5-CHRNA3-CHRNB4* region defined by rs16969968 and rs680244.

**Results:** The genetic variants in the *CHRNA5-CHRNA3-CHRNB4* region that predict nicotine dependence also predicted a later age at smoking cessation in the community sample. In the smoking cessation trial, haplotype predicted abstinence at end of treatment in individuals receiving placebo but not among individuals receiving active medication. Haplotype interacted with treatment in affecting cessation success.

**Conclusions:** Smokers with the high-risk haplotype were three times as likely to respond to pharmacologic cessation treatments as were smokers with the low-risk haplotype. The high-risk haplotype increased the risk of cessation failure, and this increased risk was ameliorated by cessation pharmacotherapy. By identifying a high-risk genetic group with heightened response to smoking cessation pharmacotherapy, this work may support the development of personalized cessation treatments.

*(Am J Psychiatry 2012; 169:735–742)*

**FIGURE 1. Effect on Endpoint Abstinence of Interaction Between Treatment and Haplotype in Nicotinic Receptor Gene Cluster *CHRNA5-CHRNA3-CHRNB4* in an 8-Week Smoking Cessation Trial (N=1,073)[a]**



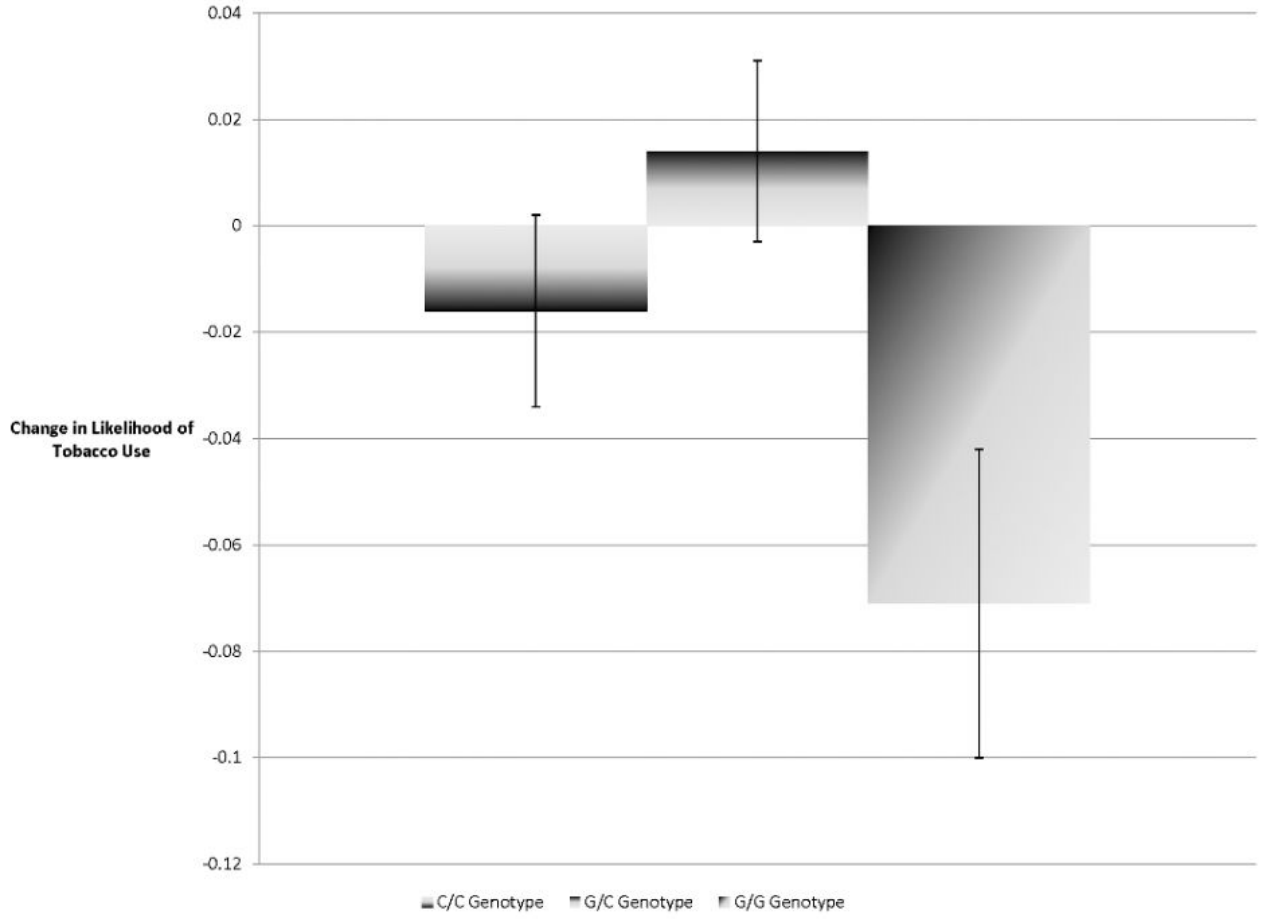

[a] Haplotypes were defined by the rs16969968 to rs680244 single-nucleotide polymorphisms. The frequencies of haplotypes 1, 2, and 3 were 20.8%, 43.7%, and 35.5%, respectively. The interaction of haplotype and treatment was significant ($\chi^2$=8.97, df=2, p=0.011).
[b] Adjusted for age and gender.

# Smoking & Taxes

## Why Have Tobacco Control Policies Stalled? Using Genetic Moderation to Examine Policy Impacts

Jason M. Fletcher

Published: December 5, 2012 • https://doi.org/10.1371/journal.pone.0050576




Notes: Author's calculation from NHANES Data. **Results from three separate regression analyses estimating the association between state level tobacco tax rates and tobacco use based on *CHRNA6* genotype (C/C, G/C, G/G).** Among the 1,278 individuals with C/C genotype, the estimated tax effect was not significant (b=−0.016, SE=0.018, z=.89, P=0.37). Among the 2,328 individuals with the G/C genotype, the estimated tax effect was not significant (b=0.014, SE=0.017, z=.82, P=.41). However, among the 2,572 individuals with the G/G genotype, the estimated tax effect was statistically significant (b=−0.071, SE=0.029, z=2.44, P=0.0148). These effects were statistically different based on genotype (see full results in Table 2). **This interaction showed that only adults with the G/G genotype respond to tobacco taxes in the manner predicted by economic theory.** Robust standard error bars clustered at the state level. Sample weights were used in order to produce nationally representative estimates.

# Biologically Agnostic

- When we use more powerful genetic predictors, we generally have less of a sense for the specific biological mechanisms.
  - E.g., pleiotropy becomes a major problem.
- We can now ask about the role of individual-level genetic endowments (that are fixed at birth) in a variety of lifecourse processes.

# Biologically Agnostic

- When we use more powerful genetic
  pre ense
  for
  – I
- We
  ind at are
  fixed at birth) in a variety of lifecourse processes.

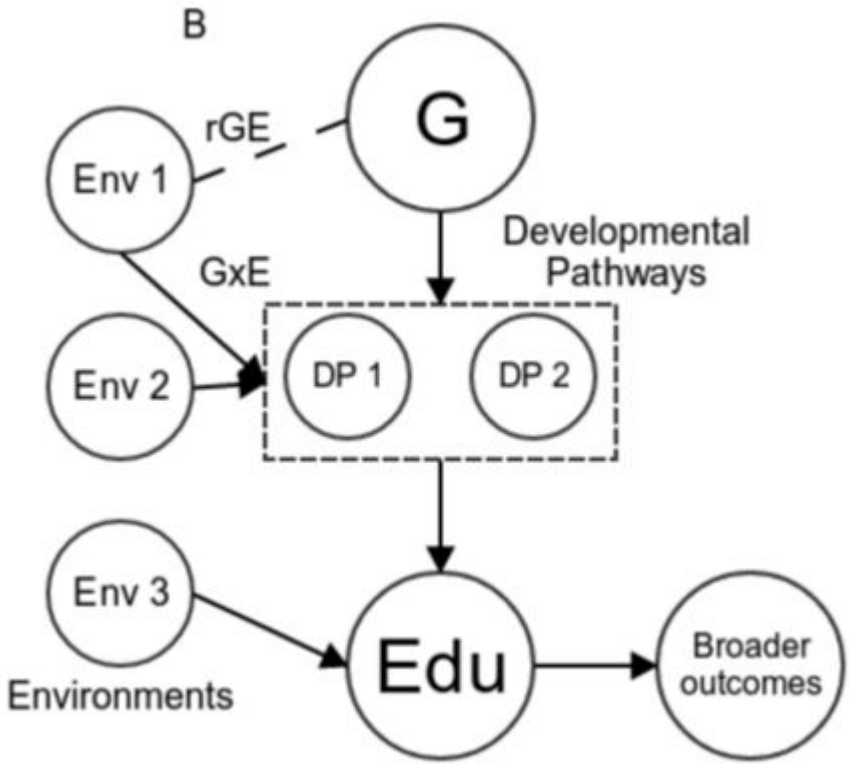
An Expanded View of Complex Traits: From Polygenic to Omnigenic

15 June 2017

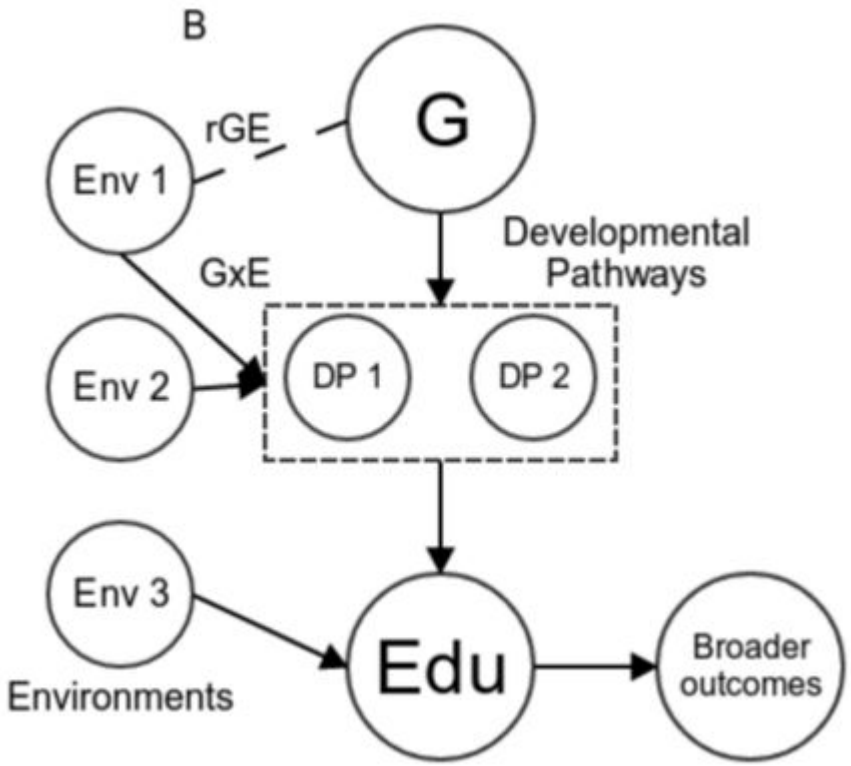
Evan A. Boyle | Yang I. Li | Jonathan K. Pritchard

Many complex genetic traits arise from large numbers of variants, each with small effects. This Perspective argues that risk is ultimately driven by an even larger number of genes with no direct impact on the phenotype or disease whose effects are propagated through regulatory networks.

# How might we study such individual-level endowments



# How might we study such individual-level endowments

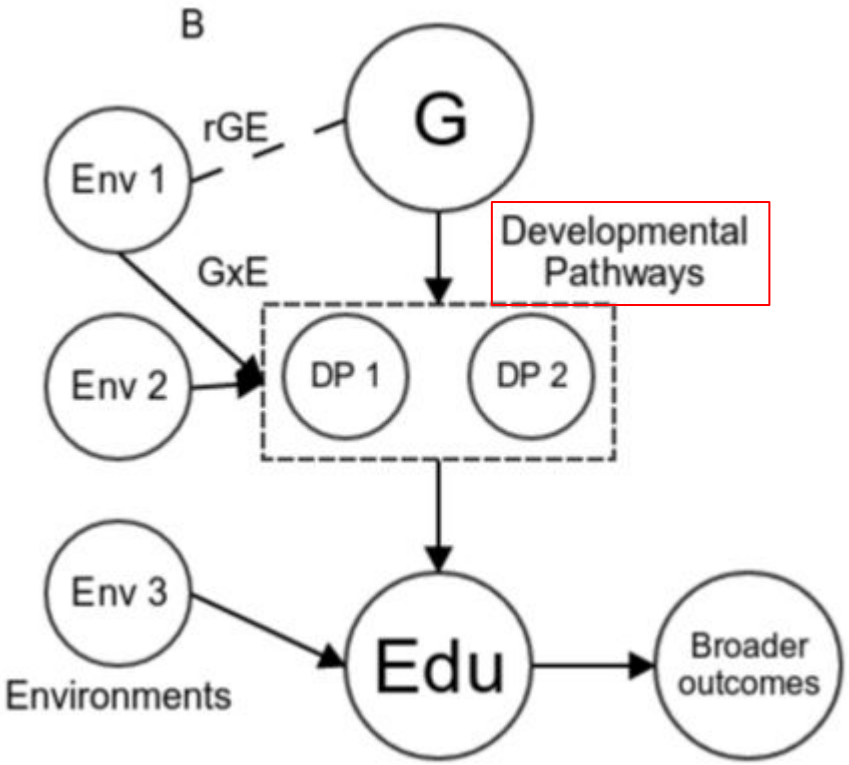



Recall that G is

- Fixed across lifecourse
- Not due to reverse causality

Not many other predictors have such properties

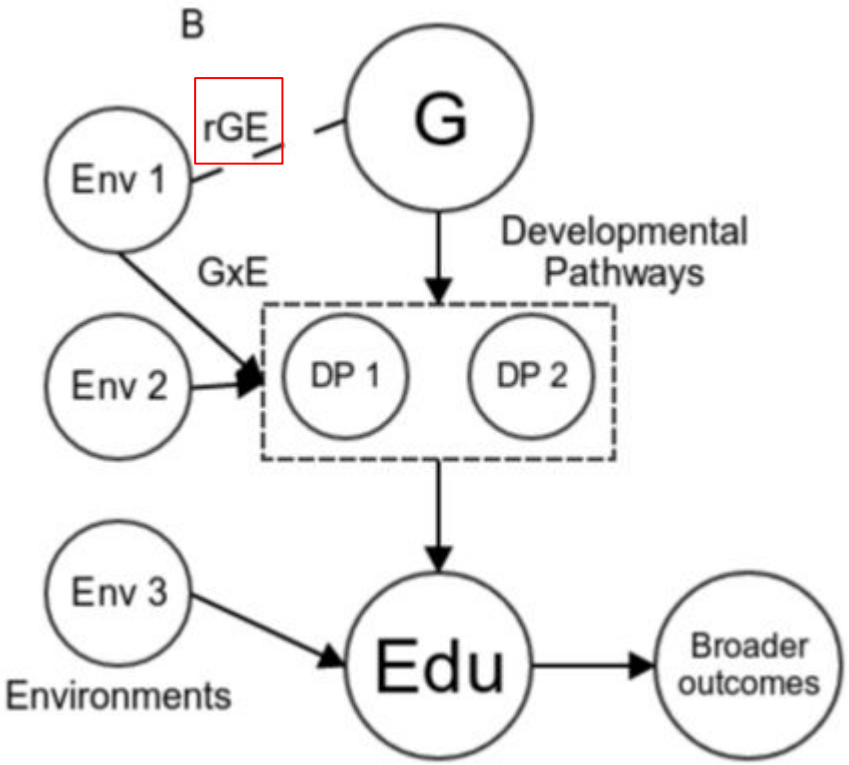
# How might we study such individual-level endowments




How do genes influence development of some phenotype?

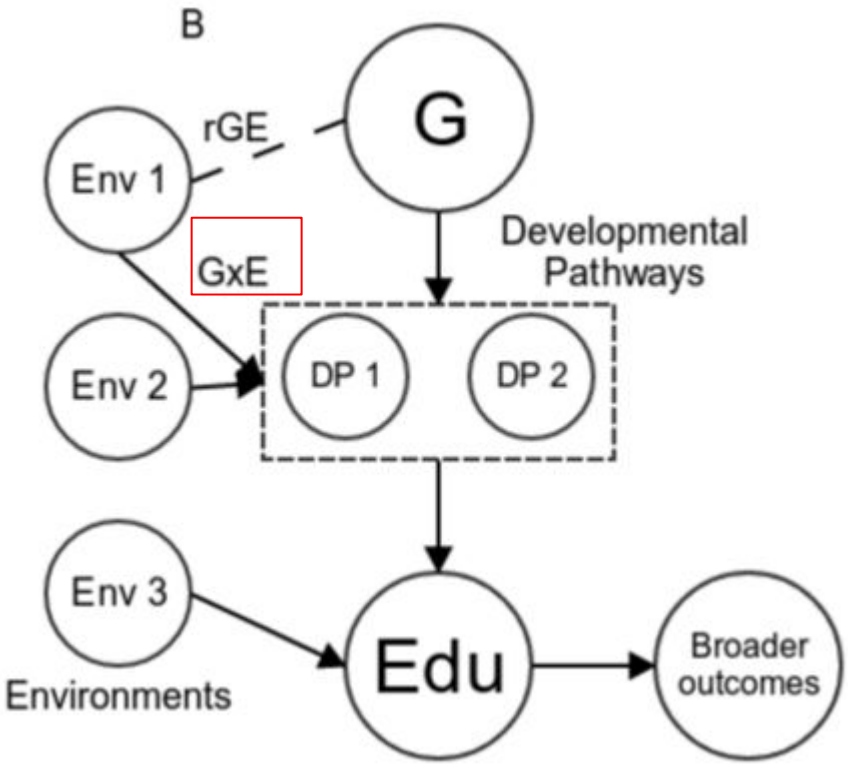
# How might we study such individual-level endowments

How are genes associated with various “environments”?

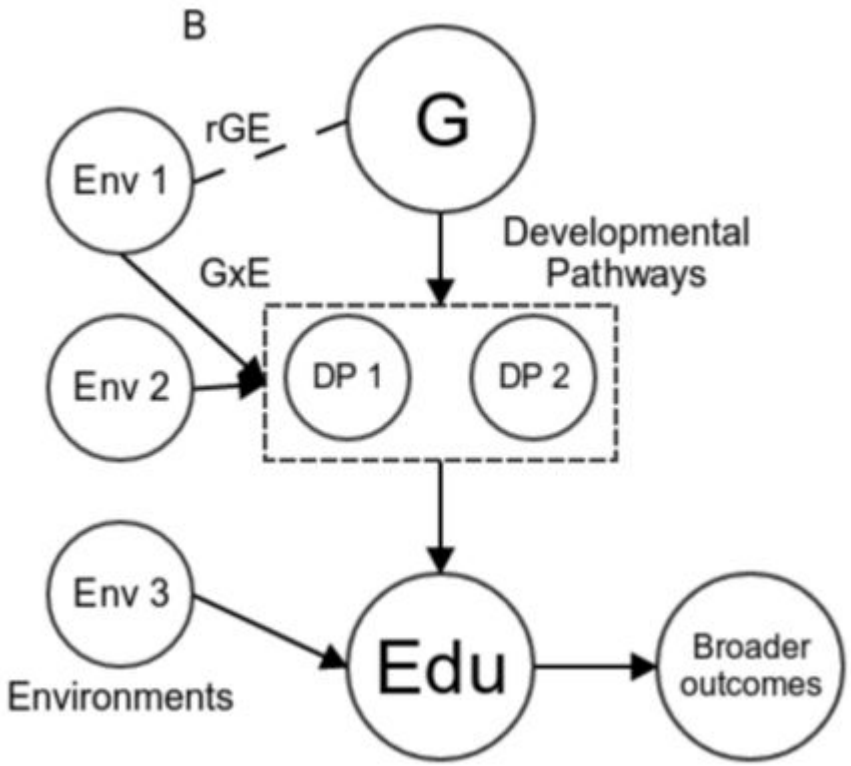


# How might we study such individual-level endowments

How might environments moderate genetic influences?



# How might we study such individual-level endowments




- Social surveys with rich measures are useful for fleshing out GWAS findings.
- Nothing need be constant across time and place.
- Questions about studying humans are area of expertise for social scientists.
- Sibling designs can still be useful.

# Illustrations

- Development
- Time
  - Mostly GxE
- Place
  - A little rGE, a little GxE
- Social Relationships

# Development

What is the developmental process through individual-level genetic endowments come to manifest as phenotypes?

Dan Belksy

# Obesity

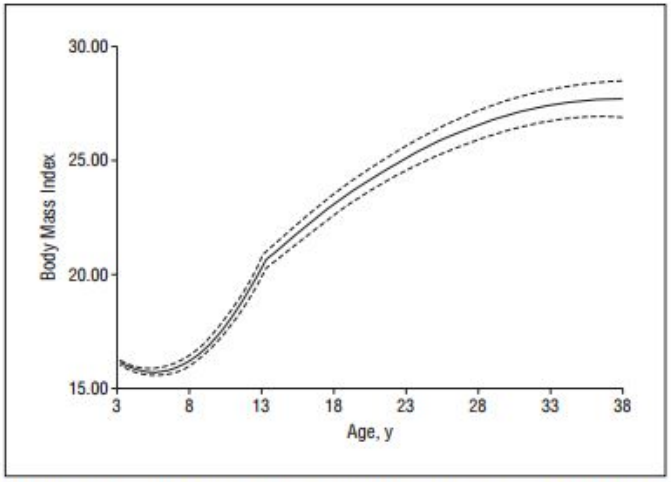



**Figure 2.** Life-course growth curves for children with high, low, and average genetic risk scores (GRSs). Individuals with higher-obesity GRSs were larger and grew more rapidly as children and adults. The solid line represents the population mean trajectory (average genetic risk). Dashed lines are for subgroups within 1 SD of the GRS (high and low genetic risk). Trajectories were derived from the life-course growth model (intercept fitted at 13 years of age; linear and quadratic slopes fitted during ages 3-13 years and 13-38 years), including intercept and linear slope effects for the GRS. Analyses included 856 individuals of European descent. Body mass index is calculated as weight in kilograms divided by height in meters squared.

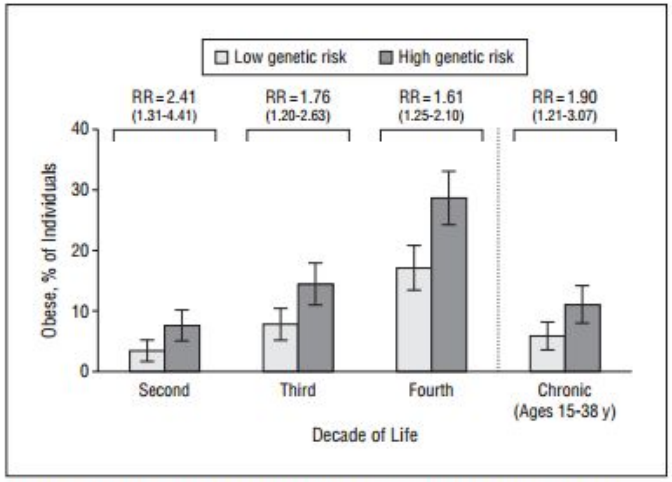


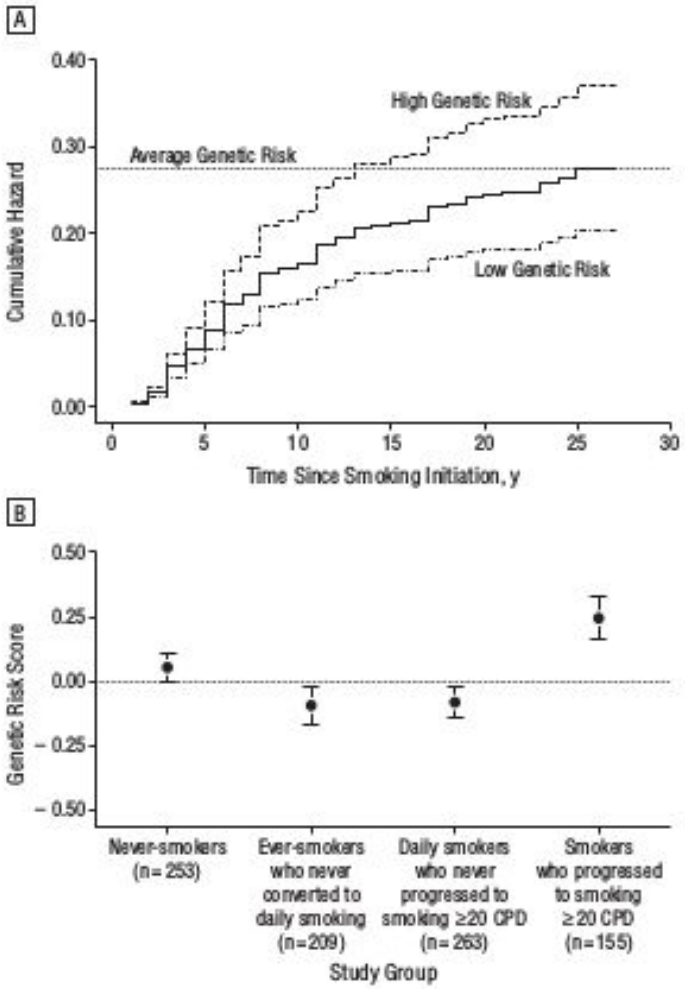

**Figure 3.** Obesity prevalence among low and high genetic risk cohort members in their second, third, and fourth decades of life and chronically across ages 15 to 38 years. Individuals with higher genetic risk scores (GRSs) were more likely to be obese across 2 decades of adult follow-up. Error bars and numbers in parentheses reflect 95% CIs. The GRS was dichotomized at the sample mean to create low and high genetic risk categories. Relative risks (RRs) (95% CIs) are reported from Poisson regression models adjusted for sex that included the 856 individuals of European descent in the analysis sample.

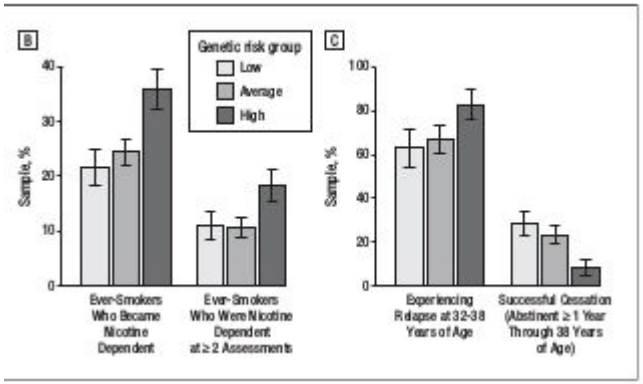
## Polygenic Risk, Rapid Childhood Growth, and the Development of Obesity

*Evidence From a 4-Decade Longitudinal Study*

*Daniel W. Belsky, PhD; Terrie E. Moffitt, PhD; Renate Houts, PhD; Gary G. Bennett, PhD; Andrea K. Biddle, PhD; James A. Blumenthal, PhD; James P. Evans, MD, PhD; HonaLee Harrington, BA; Karen Sugden, PhD; Benjamin Williams, BS; Richie Poulton, PhD; Avshalom Caspi, PhD*

# Smoking




ORIGINAL ARTICLE

## Polygenic Risk and the Developmental Progression to Heavy, Persistent Smoking and Nicotine Dependence

*Evidence From a 4-Decade Longitudinal Study*

*Daniel W. Belsky, PhD; Terrie E. Moffitt, PhD; Timothy B. Baker, PhD; Andrea K. Biddle, PhD; James P. Evans, MD, PhD; HonaLee Harrington, BA; Renate Houts, PhD; Madeline Meier, PhD; Karen Sugden, PhD; Benjamin Williams, BS; Richie Poulton, PhD; Avshalom Caspi, PhD*

# Other examples

THE LANCET
Respiratory Medicine
Volume 1, Issue 6, August 2013, Pages 453-461

Articles

Polygenic risk and the development and course of asthı
analysis of data from a four-decade longitudinal study

Dr Daniel W Belsky PhD[a, c], Prof Malcolm R Sears MB[d], Robert J Hancox MD[e], HonaLee Har
Renate Houts PhD[b, c, f], Prof Terrie E Moffitt PhD[b, c, f, g], Karen Sugden PhD[c, f, g], Benjamin Willia

Findings

Of the 880 cohort members included in our analysis, those at higher genetic risk developed asthma earlier in life than did those with lower genetic risk (hazard ratio [HR] 1·12, 95% CI 1·01–1·26). Of cohort members with childhood-onset asthma, those with higher genetic risk were more likely to develop life-course-persistent asthma than were those with a lower genetic risk (relative risk [RR] 1·36, 95% CI 1·14–1·63). Participants with asthma at higher genetic risk more often had atopy (RR 1·07, 1·01–1·14), airway hyper-responsiveness (RR 1·16, 1·03–1·32), and incompletely reversible airflow obstruction (RR 1·28, 1·04–1·57) than did those with a lower genetic risk. They were also more likely to miss school or work (incident rate ratio 1·38, 1·02–1·86) and be admitted to hospital (HR 1·38, 1·07–1·79) because of asthma. Genotypic information about asthma risk was independent of and additive to information derived from cohort members' family histories of asthma.

**The Genetics of Success**

*How Single-Nucleotide Polymorphisms Associated With Educational Attainment Relate to Life-Course Development*

Daniel W. Belsky, Terrie E. Moffitt, David L. Corcoran, ... Show all authors

First Published June 1, 2016 | Research Article

# Time [charts via CDC]

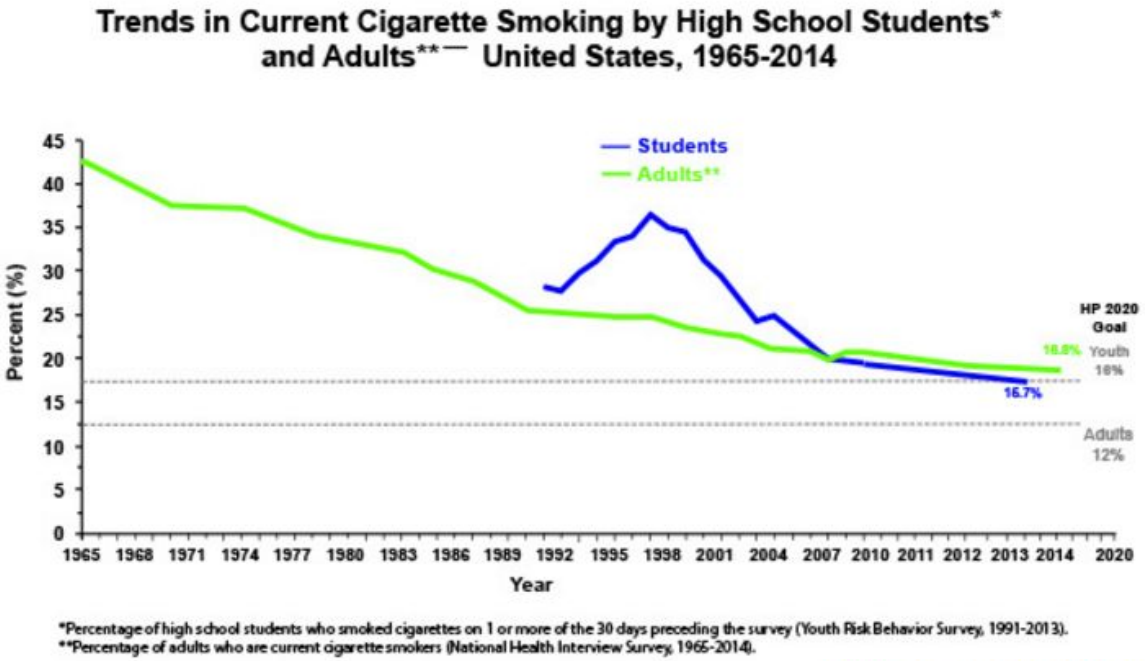
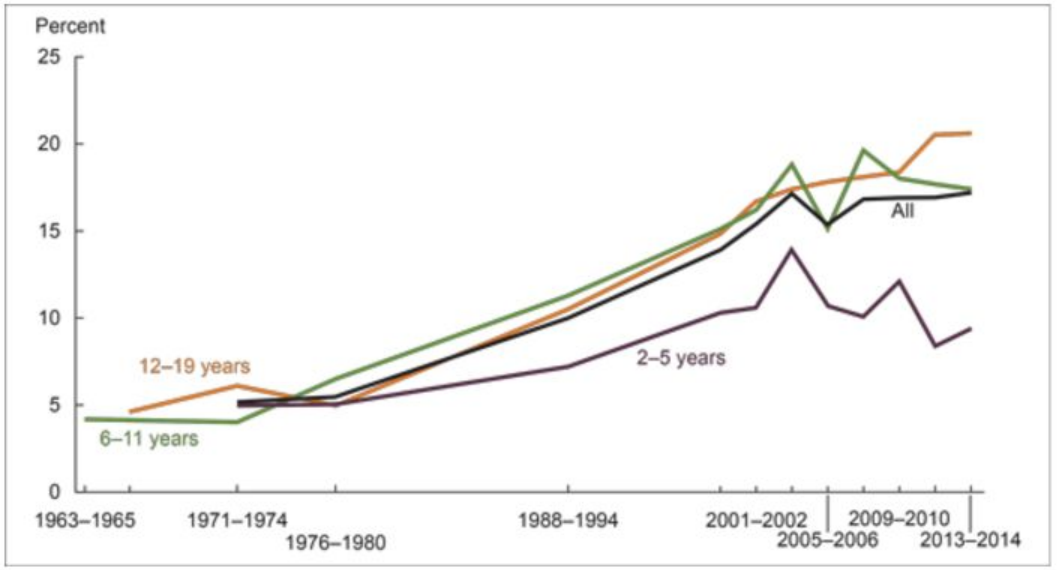
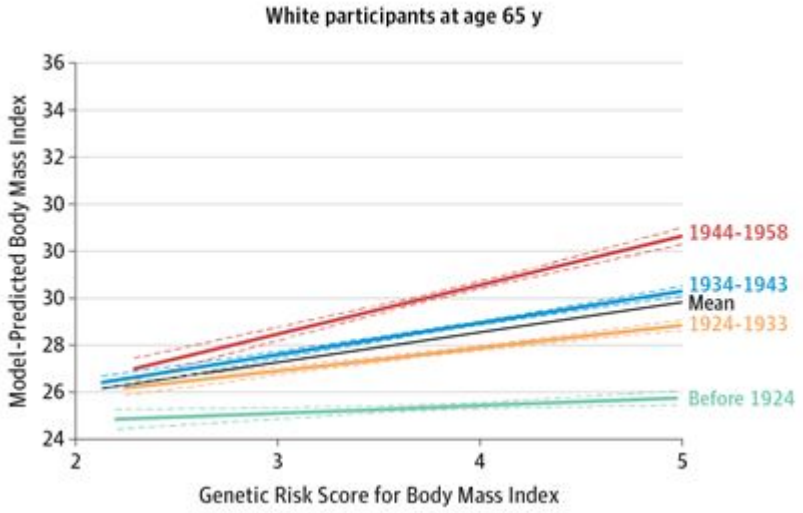



Figure. Trends in obesity among children and adolescents aged 2–19 years, by age: United States, 1963–1965 through 2013–2014




Given manifest changes in environment, how has influence of genes changed?

# BMI



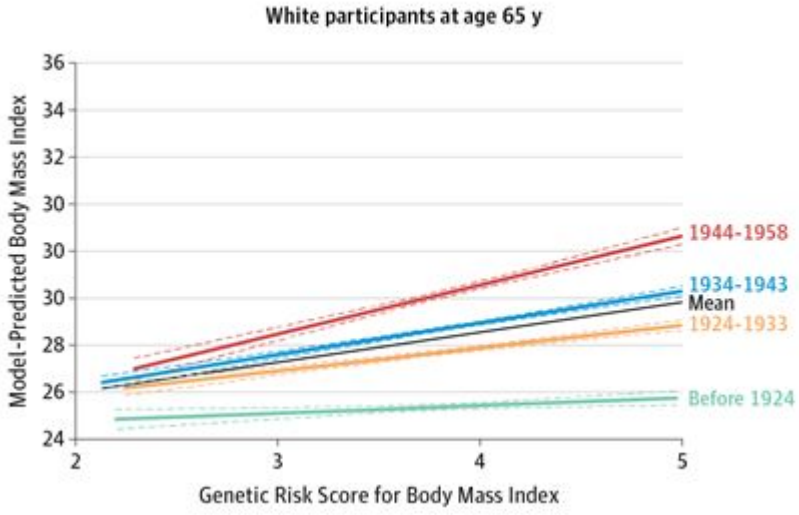

**Association of a Genetic Risk Score With Body Mass Index Across Different Birth Cohorts**

Stefan Walter, PhD[1]; Iván Mejía-Guevara, PhD[2]; Karol Estrada, PhD[3]; et al

Results based on top hits BMI PGS.

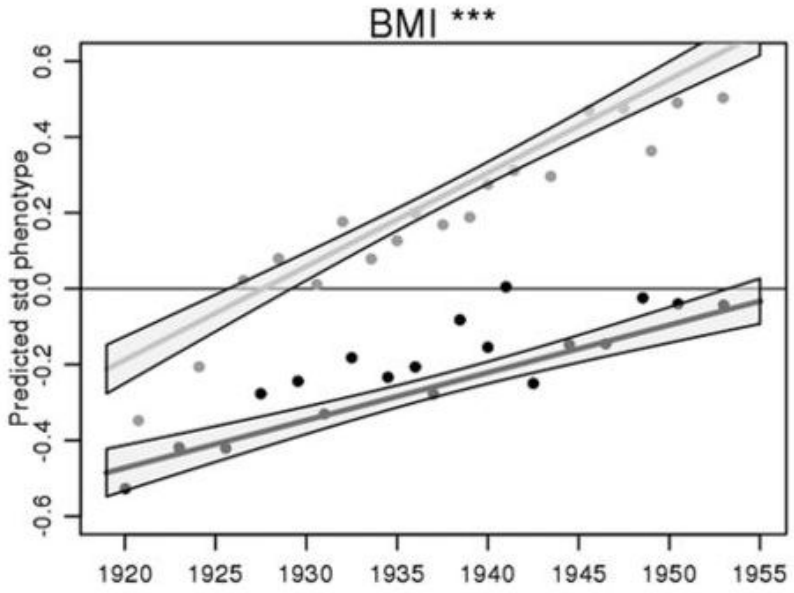
- Similar findings in Liu & Guo, AJS, 2015.

# BMI




**Association of a Genetic Risk Score With Body Mass Index Across Different Birth Cohorts**

Stefan Walter, PhD[1]; Iván Mejía-Guevara, PhD[2]; Karol Estrada, PhD[3]; et al

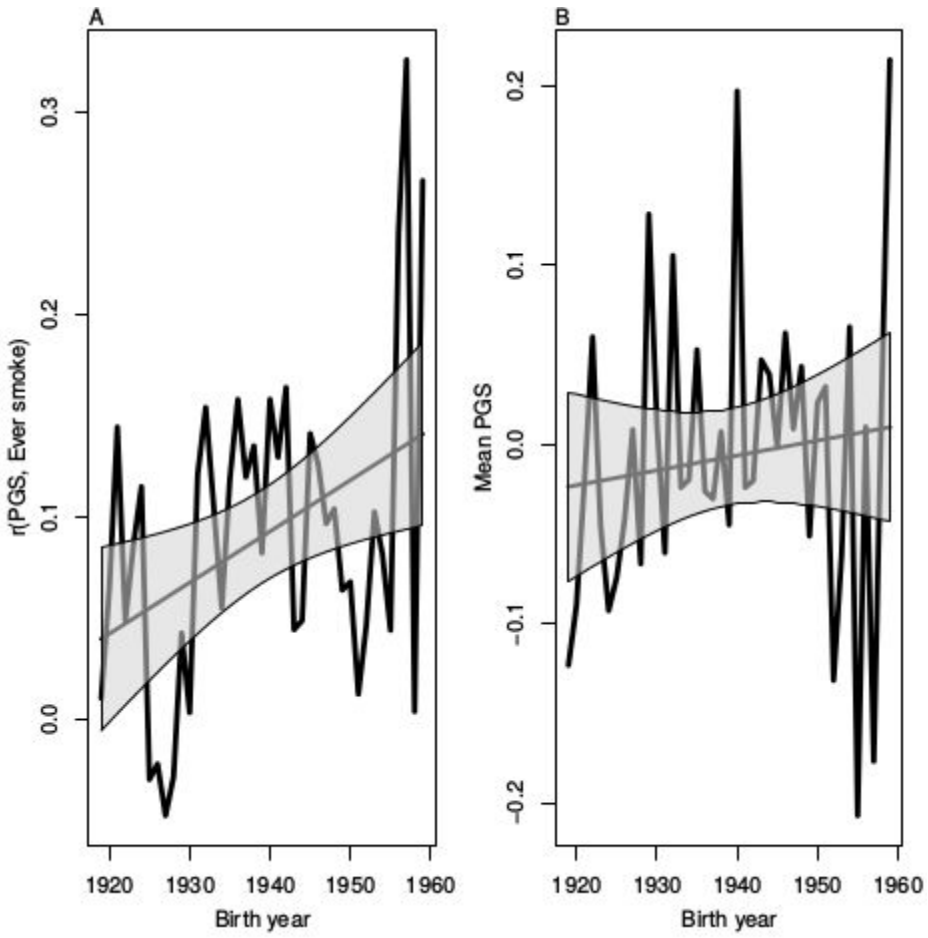



Changing Polygenic Penetrance on Phenotypes in the 20th Century Among Adults in the US Population

Dalton Conley, Thomas M. Laidley, Jason D. Boardman & Benjamin W. Domingue

Results based on top hits BMI PGS.

- Similar findings in Liu & Guo, AJS, 2015.




# Smoking

Behav Genet
DOI 10.1007/s10519-015-9731-9

ORIGINAL RESEARCH

## Cohort Effects in the Genetic Influence on Smoking

Benjamin W. Domingue[1] · Dalton Conley[2] · Jason Fletcher[3] · Jason D. Boardman[4]

# A note about time

**Table 1 Descriptive characteristics of the 16 studies participating in the TAG Consortium**

| Study | *n* (% female) | Age[a], mean (s.d.) | Ever smokers (%) | CPD, mean (s.d.)[b] | Age of initiation of smoking[a], mean (s.d.)[b] | Former smokers (%)[b] |
|---|---|---|---|---|---|---|
| **Population-based cohort studies** | | | | | | |
| Atherosclerosis Risk in Communities (ARIC) | 8,330 (52.9) | 54.3 (5.7) | 60.4 | 21.0 (11.7) | 18.6 (5.1) | 57.4 |
| Baltimore Longitudinal Study of Aging (BLSA) | 856 (46.0) | 48.1 (17.8) | 54.0 | NA | 19.3 (5.9) | NA |
| Cardiovascular Health Study (CHS) | 3,236 (60.8) | 72.3 (5.4) | 52.3 | 17.8 (11.8) | 19.6 (6.6) | 77.8 |
| Invecchiare in Chianti (InCHIANTI) | 1,200 (55.2) | 68.4 (15.5) | 43.9 | 14.8 (9.4) | 32.2 (16.7) | 57.0 |
| Rotterdam Study | 5,610 (60.3) | 69.1 (8.9) | 59.2 | 15.8 (11.7) | 20.4 (8.2) | 62.6 |
| Framingham Heart Study (FHS) | 7,257 (53.7) | 45.4 (10.9) | 54.2 | 15.5 (10.8) | 17.9 (4.2) | 61.7 |
| Women's Genome Health Study (WGHS) | 22,037 (100) | 54.7 (7.1) | 49.2 | 16.0 (11.0) | NA | 75.2 |
| **Case-control studies** | | | | | | |
| Atherosclerotic Disease Vascular Function and Genetic Epidemiology (ADVANCE) | 585 (58.8) | 45.8 (5.9) | 47.7 | 13.1 (14.2) | 17.0 (4.6) | 65.2 |
| Atherosclerosis, Thrombosis and Vascular Biology Italian Study Group (ATVB) | 3,260 (11.6) | 39.6 (4.9) | 68.1 | 23.4 (14.7) | 17.4 (4.0) | 21.3 |
| Diabetes Genetic Initiative (DGI) | 2,504 (50.0) | 61.6 (10.6) | 37.7 | NA | 19.0 (5.5) | NA |
| Finland-United States Investigation of NIDDM Genetics (FUSION) | 1,055 (52.8) | 64.0 (7.5) | 46.8 | 16.3 (12.4) | 21.0 (7.0) | 65.0 |
| International Agency for Research on Cancer (IARC) | 8,381 (24.7) | 59.6 (10.1) | 75.2 | 19.3 (12.9) | 18.7 (5.6) | 31.4 |
| Myocardial Infarction Genetics Consortium (MIGen) | 2,647 (38.5) | 48.8 (8.2) | 64.3 | NA | NA | 41.1 |
| Nurses' Health Study (NHS) | 2,249 (100) | 70.5 (6.4) | 53.8 | 18.5 (10.5) | 19.6 (3.6) | 88.7 |
| Netherlands Twin Registry-Netherlands Study of Depression and Anxiety (NTR/NESDA) | 3,438 (66.9) | 43.8 (13.4) | 64.9 | 14.5 (9.8) | 16.4 (4.2) | 52.6 |
| MGS (GAIN):controls | 1,390 (54.1) | 51.1 (17) | 55.9 | 19.3 (16.4) | NA | 62.9 |

[a]Age in years. [b]Calculated among ever regular smokers. NA, not available.

Genome-wide meta-analyses identify multiple loci associated with smoking behavior

The Tobacco and Genetics Consortium*

# A note about time

**Table 1 Descriptive characteristics of the 16 studies participating in the TAG Consortium**

| Study | *n* (% female) | Age[a], mean (s.d.) | Ever smokers (%) | CPD, mean (s.d.)[b] | Age of initiation of smoking[a], mean (s.d.)[b] | Former smokers (%)[b] |
|---|---|---|---|---|---|---|
| **Population-based cohort studies** | | | | | | |
| Atherosclerosis Risk in Communities (ARIC) | 8,330 (52.9) | 54.3 (5.7) | 60.4 | 21.0 (11.7) | 18.6 (5.1) | 57.4 |
| Baltimore Longitudinal Study of Aging (BLSA) | 856 (46.0) | 48.1 (17.8) | 54.0 | NA | 19.3 (5.9) | NA |
| Cardiovascular Health Study (CHS) | 3,236 (60.8) | 72.3 (5.4) | 52.3 | 17.8 (11.8) | 19.6 (6.6) | 77.8 |
| Invecchiare in Chianti (InCHIANTI) | 1,200 (55.2) | 68.4 (15.5) | 43.9 | 14.8 (9.4) | 32.2 (16.7) | 57.0 |
| Rotterdam Study | 5,610 (60.3) | 69.1 (8.9) | 59.2 | 15.8 (11.7) | 20.4 (8.2) | 62.6 |
| Framingham Heart Study (FHS) | 7,257 (53.7) | 45.4 (10.9) | 54.2 | 15.5 (10.8) | 17.9 (4.2) | 61.7 |
| Women's Genome Health Study (WGHS) | 22,037 (100) | 54.7 (7.1) | 49.2 | 16.0 (11.0) | NA | 75.2 |
| **Case-control studies** | | | | | | |
| Atherosclerotic Disease Vascular Function and Genetic Epidemiology (ADVANCE) | 585 (58.8) | 45.8 (5.9) | 47.7 | 13.1 (14.2) | 17.0 (4.6) | 65.2 |
| Atherosclerosis, Thrombosis and Vascular Biology Italian Study Group (ATVB) | 3,260 (11.6) | 39.6 (4.9) | 68.1 | 23.4 (14.7) | 17.4 (4.0) | 21.3 |
| Diabetes Genetic Initiative (DGI) | 2,504 (50.0) | 61.6 (10.6) | 37.7 | NA | 19.0 (5.5) | NA |
| Finland-United States Investigation of NIDDM Genetics (FUSION) | 1,055 (52.8) | 64.0 (7.5) | 46.8 | 16.3 (12.4) | 21.0 (7.0) | 65.0 |
| International Agency for Research on Cancer (IARC) | 8,381 (24.7) | 59.6 (10.1) | 75.2 | 19.3 (12.9) | 18.7 (5.6) | 31.4 |
| Myocardial Infarction Genetics Consortium (MIGen) | 2,647 (38.5) | 48.8 (8.2) | 64.3 | NA | NA | 41.1 |
| Nurses' Health Study (NHS) | 2,249 (100) | 70.5 (6.4) | 53.8 | 18.5 (10.5) | 19.6 (3.6) | 88.7 |
| Netherlands Twin Registry-Netherlands Study of Depression and Anxiety (NTR/NESDA) | 3,438 (66.9) | 43.8 (13.4) | 64.9 | 14.5 (9.8) | 16.4 (4.2) | 52.6 |
| MGS (GAIN):controls | 1,390 (54.1) | 51.1 (17) | 55.9 | 19.3 (16.4) | NA | 62.9 |

[a]Age in years. [b]Calculated among ever regular smokers. NA, not available.

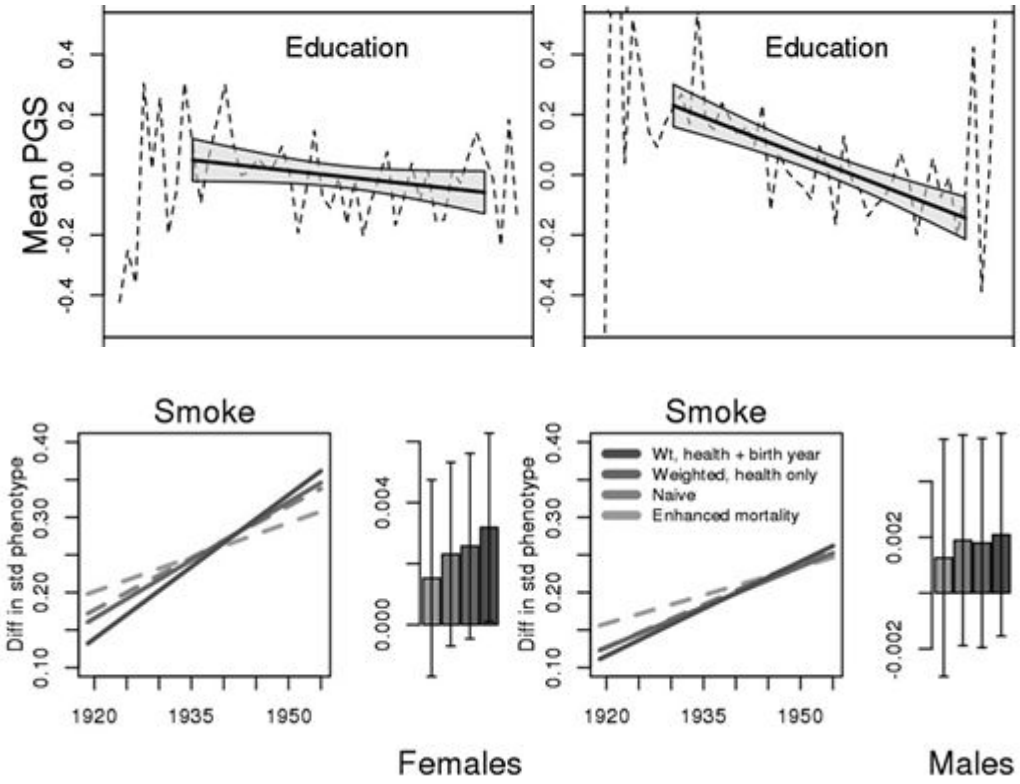
If I give you a sample of people with ages centered around 65, what might you infer about them?

Genome-wide meta-analyses identify multiple loci associated with smoking behavior

The Tobacco and Genetics Consortium*

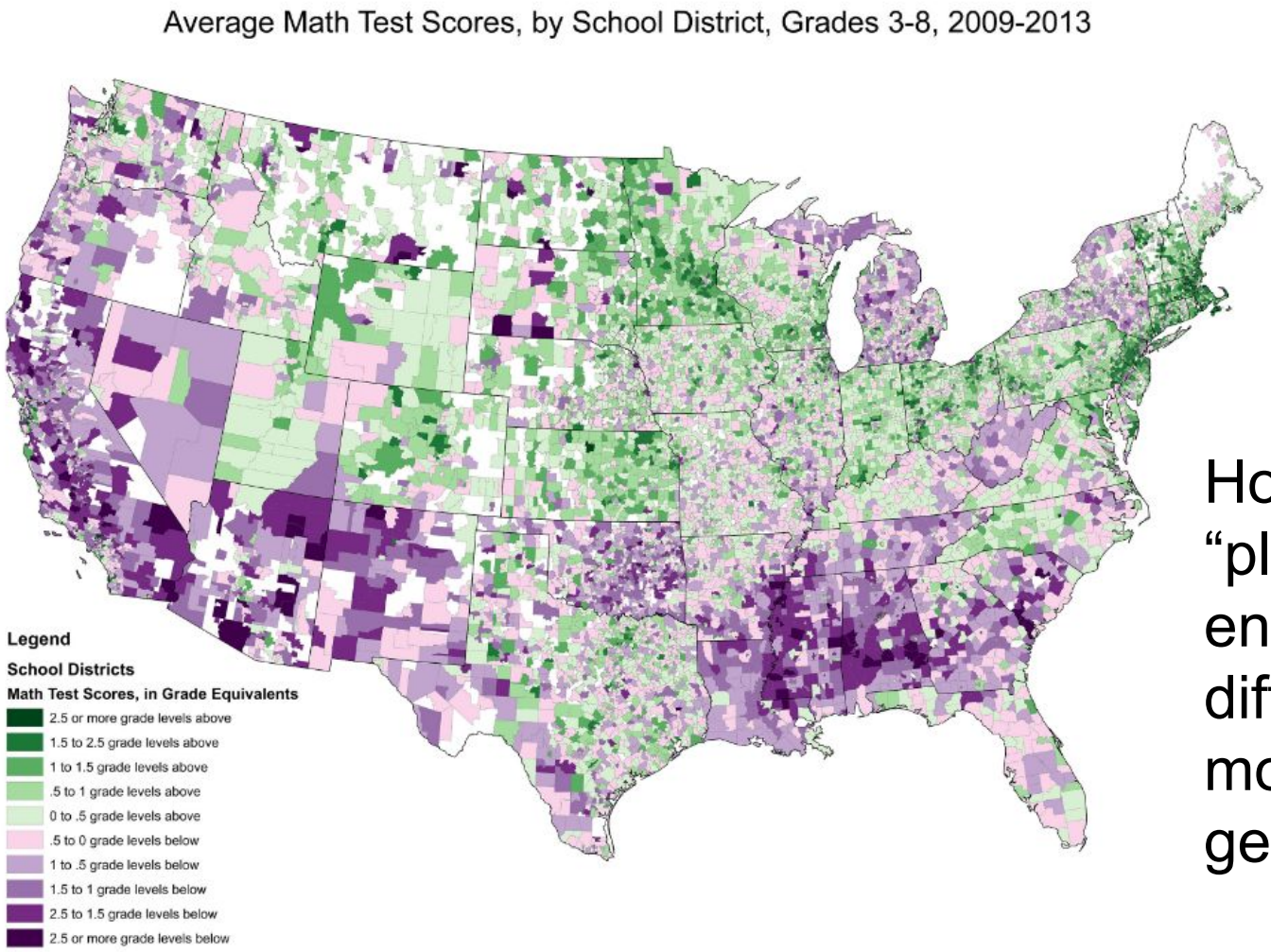
**Mortality selection in a genetic sample and implications for association studies**

Benjamin W. Domingue ✉, Daniel W. Belsky, Amal Harrati, Dalton Conley, David R. Weir, Jason D. Boardman




- Earliest born respondents in surveys are genetically distinct.
- Some evidence for bias in study of cohort-specific effects

# Place






How do "place"-based environmental differences moderate genetic effects?

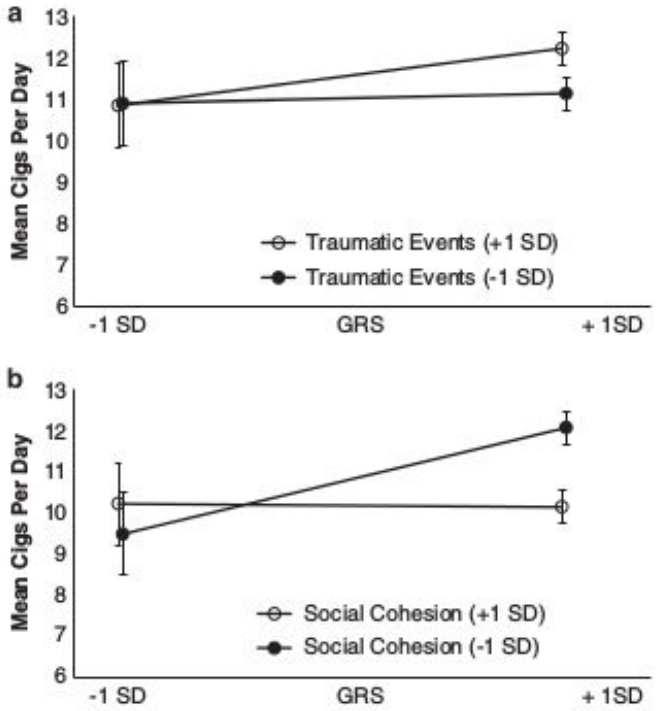



**Figure 1** The Interaction between genetic and environmental predictors of cigarette use in 399 individuals from the Detroit Neighborhood Health Study. (a) Genetic risk for smoking (GRS +1 s.d.) was greater for individuals who had experienced an increased number of traumatic events in their lifetimes. (b) Genetic risk for smoking (GRS +1 s.d.) was greater for individuals who lived in a neighborhood characterized by less social cohesion.

traumatic events,
neighborhood social cohesion and
neighborhood physical disorder.

Citation: Transl Psychiatry (2013) 3, e290; doi:10.1038/tp.2013.63

www.nature.com/tp

# Interaction between polygenic risk for cigarette use and environmental exposures in the Detroit neighborhood health study

JL Meyers[1], M Cerdá[1], S Galea[1], KM Keyes[1], AE Aiello[2], M Uddin[3,4], DE Wildman[3] and KC Koenen[1]

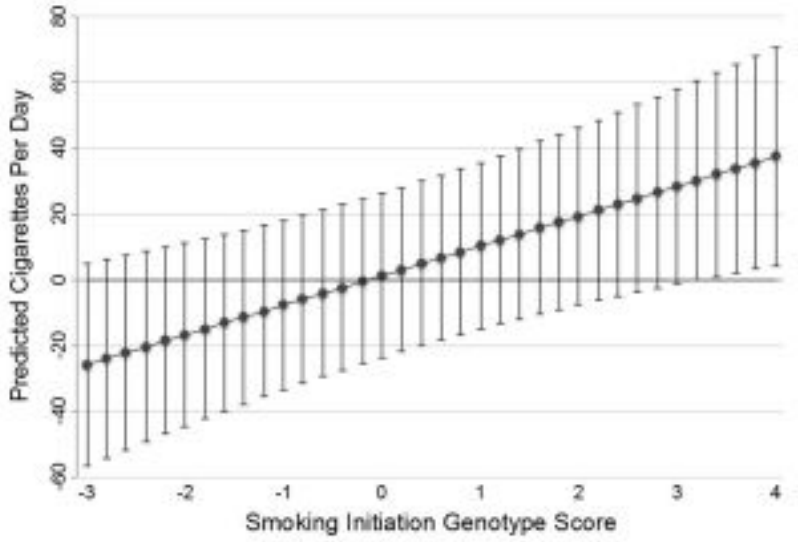



**Fig. 2** Difference in predicted cigarettes per day, veterans versus non-veterans, college educated

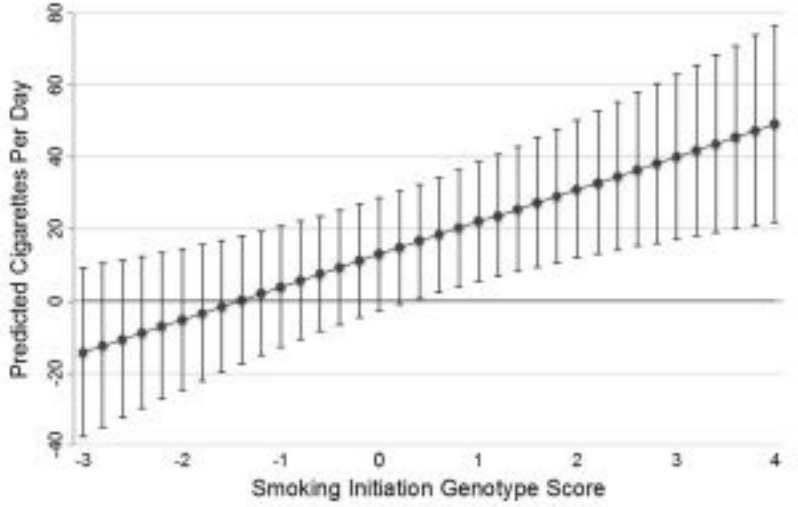



**Fig. 3** Difference in predicted cigarettes per day, veterans versus non-veterans, not college educated

## The Long-Term Consequences of Vietnam-Era Conscription and Genotype on Smoking Behavior and Health

**Lauren Schmitz[1] · Dalton Conley[2]**

# Genetics & Geography

**The Geographic Distribution of Genetic Risk as Compared to Social Risk for Chronic Diseases in the United States**

DAVID H. REHKOPF,[1] BENJAMIN W. DOMINGUE,[2] AND MARK R. CULLEN[1]

[1]School of Medicine, Division of General Medical Disciplines, Stanford University, Stanford, California, USA
[2]Graduate School of Education, Stanford University, Stanford, California, USA

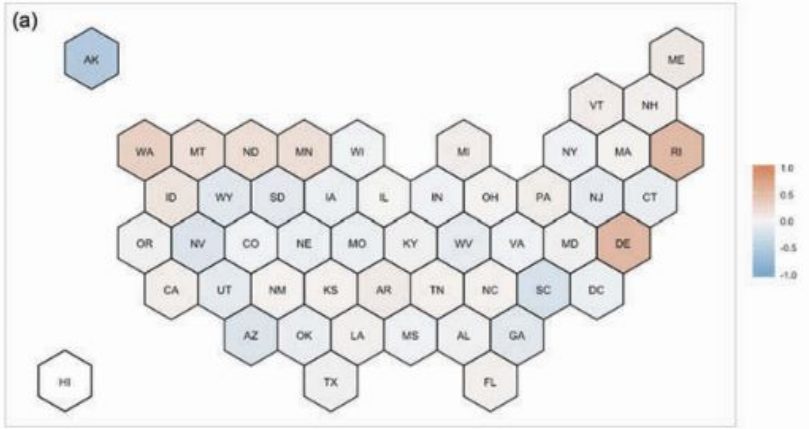



**Figure 4a.** Hexbin state map of the mean coronary artery disease (CAD) polygenic risk score by state of birth, non-Hispanic whites, Health and Retirement Study, 2006–2012.

# Genetics & Geography

**The Geographic Distribution of Genetic Risk as Compared to Social Risk for Chronic Diseases in the United States**

DAVID H. REHKOPF,[1] BENJAMIN W. DOMINGUE,[2] AND MARK R. CULLEN[1]

[1]School of Medicine, Division of General Medical Disciplines, Stanford University, Stanford, California, USA
[2]Graduate School of Education, Stanford University, Stanford, California, USA

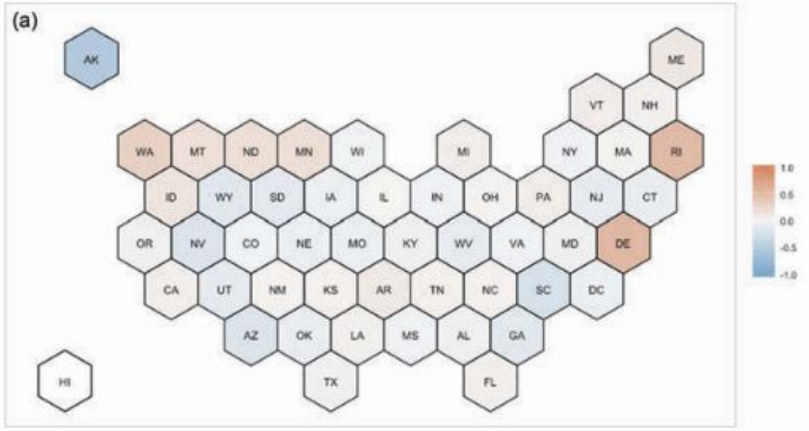
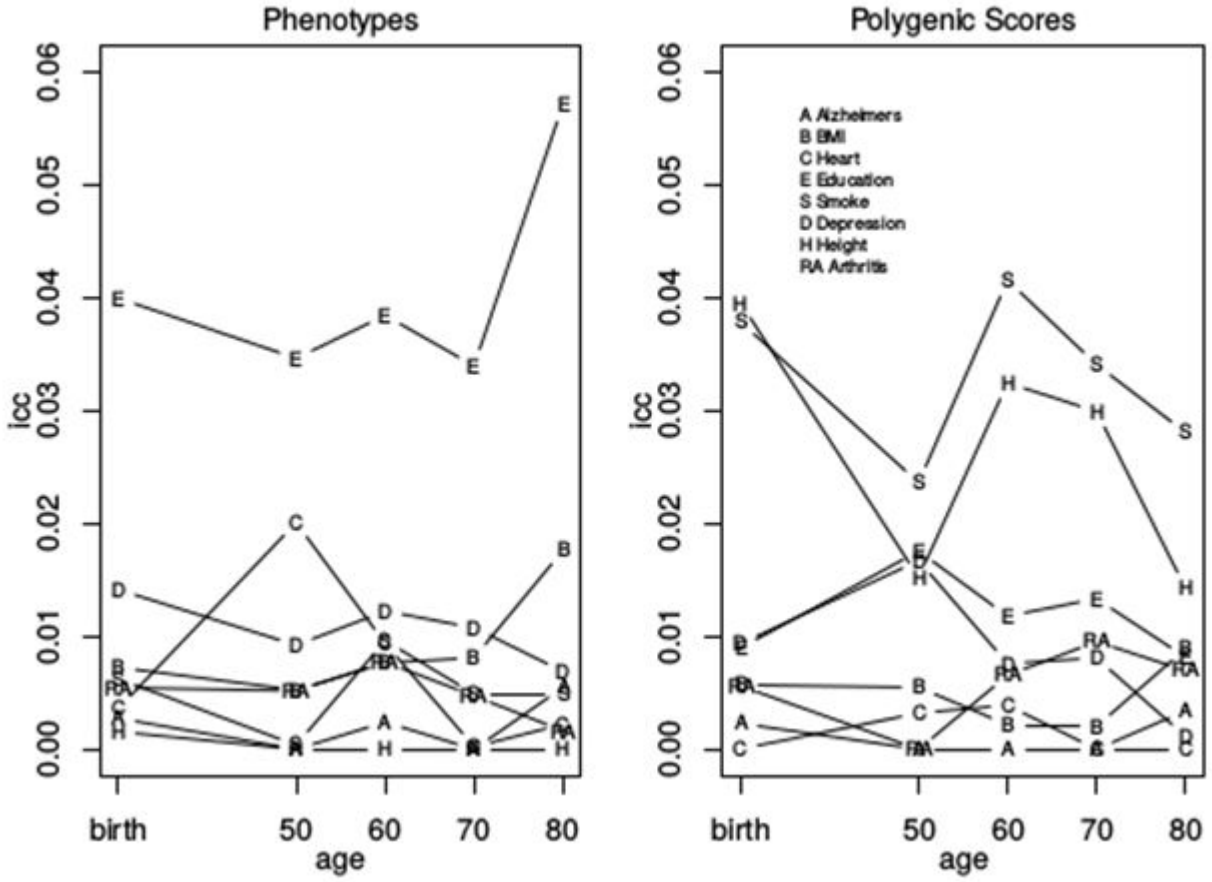
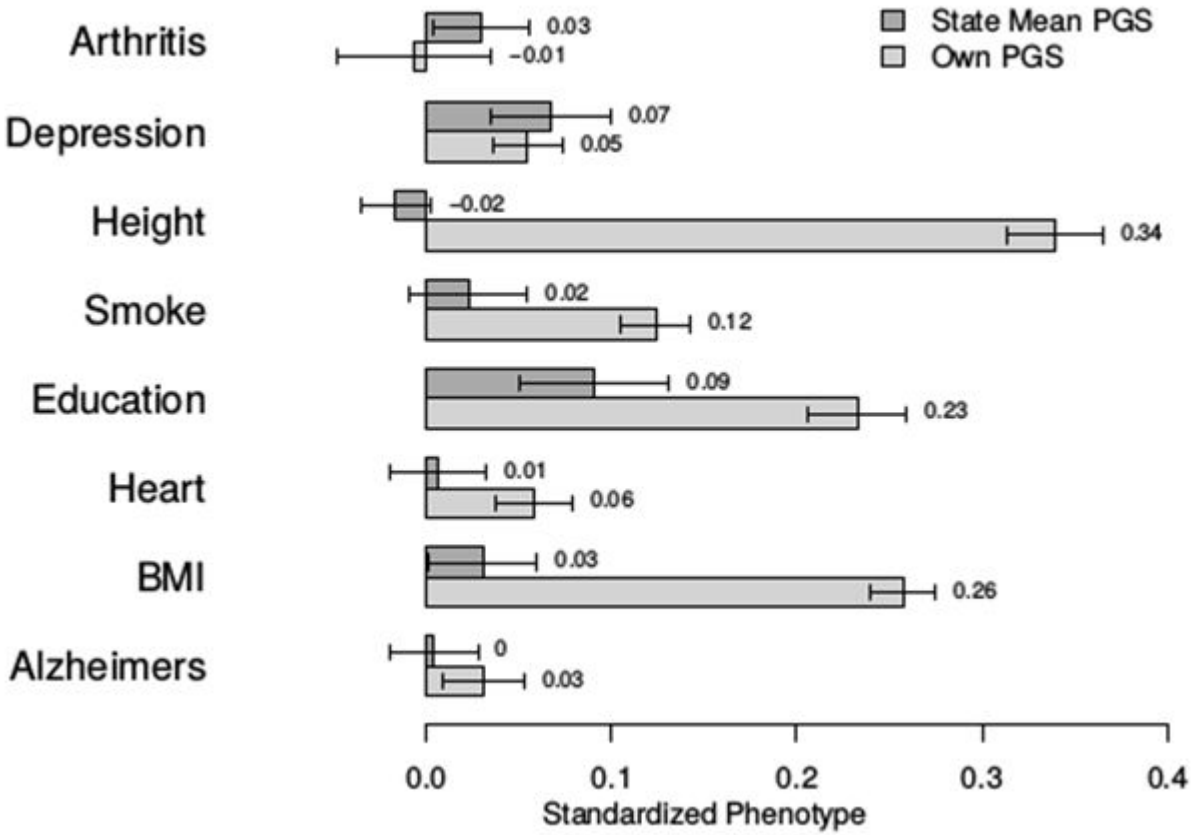


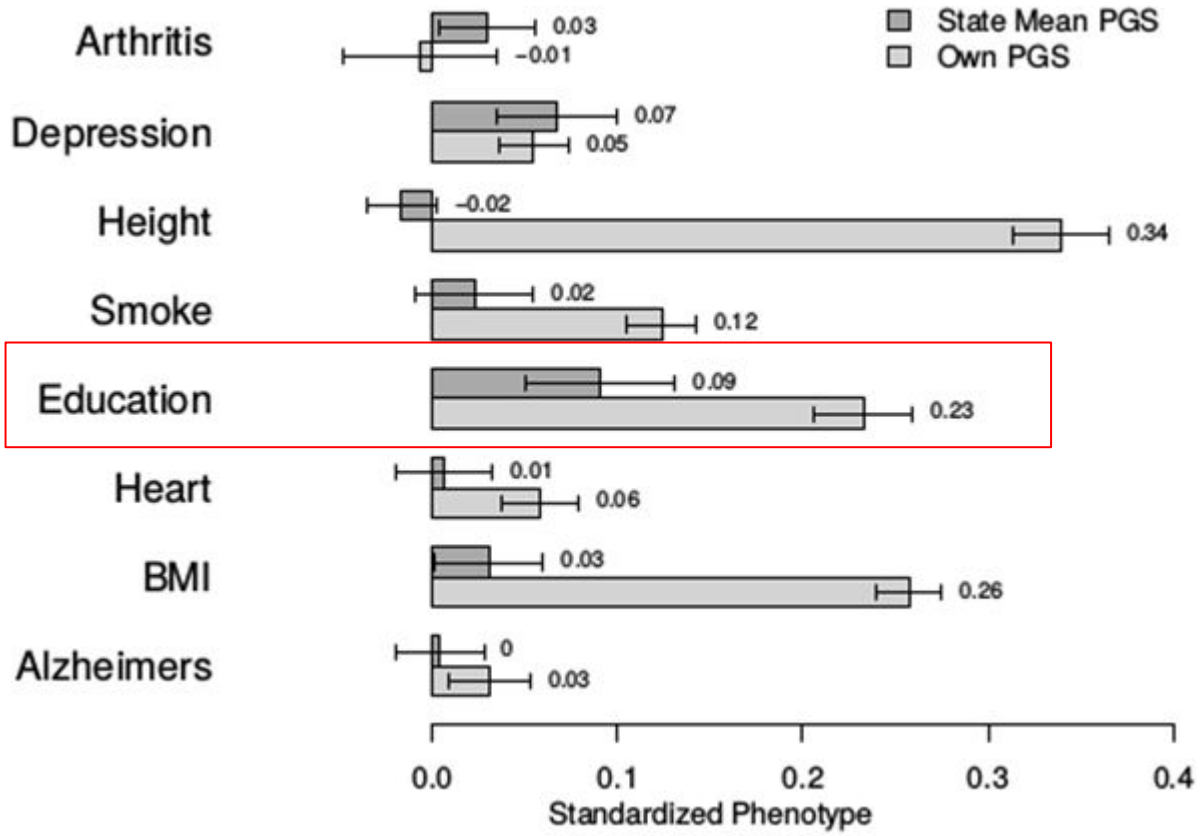

**Figure 4a.** Hexbin state map of the mean coronary artery disease (CAD) polygenic risk score by state of birth, non-Hispanic whites, Health and Retirement Study, 2006–2012.

Limited set of variants over limited set of diseases




- Perhaps some clustering of interest.
- What is its role?


Arthritis
Depression
Height
Smoke
Education
Heart
BMI
Alzheimers
State Mean PGS
Own PGS
0.03
−0.01
0.07
0.05
−0.02
0.34
0.02
0.12
0.09
0.23
0.01
0.06
0.03
0.26
0
0.03
0.0
0.1
0.2
0.3
0.4
Standardized Phenotype

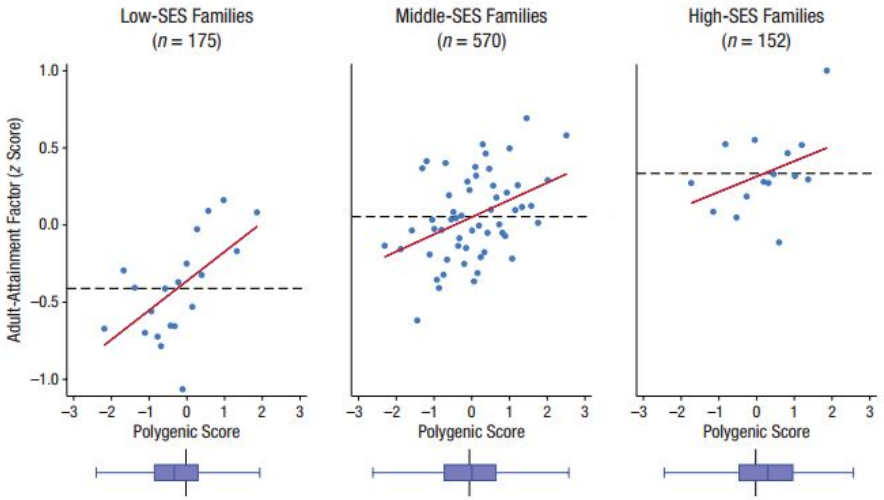



What do we make of this? We'll come back to an interpretation, but first let's talk a little about the educational attainment GWAS results.

# Educational Attainment GWAS

On the one hand:

- Stratified across environment. Associated with
  - Parental endowments
  - Family background
  - Neighborhood

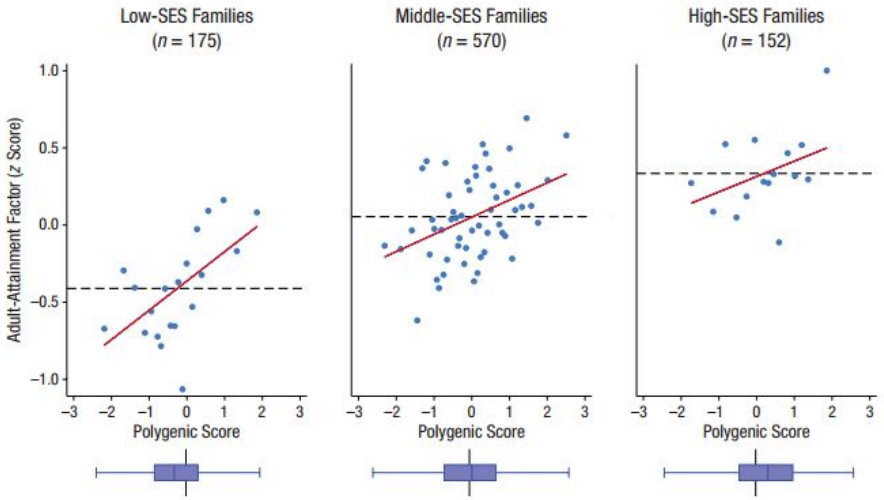



**The Genetics of Success: How Single-Nucleotide Polymorphisms Associated With Educational Attainment Relate to Life-Course Development**

# Educational Attainment GWAS

On the one hand:

- Stratified across environment. Associated with
  - Parental endowments
  - Family background
  - Neighborhood




**The Genetics of Success: How Single-Nucleotide Polymorphisms Associated With Educational Attainment Relate to Life-Course Development**

Is the Effect of Parental Education on Offspring Biased or Moderated by Genotype?

Dalton Conley,[a] Benjamin W. Domingue,[b] David Cesarini,[a] Christopher Dawes,[a] Cornelius A. Rietveld,[c] Jason D. Boardman[b]

**Polygenic Influence on Educational Attainment**

*New Evidence From the National Longitudinal Study of Adolescent to Adult Health*

Benjamin W. Domingue, Daniel W. Belsky, Dalton Conley, ... Show all authors

First Published August 19, 2015 | Research Article

On the other hand:

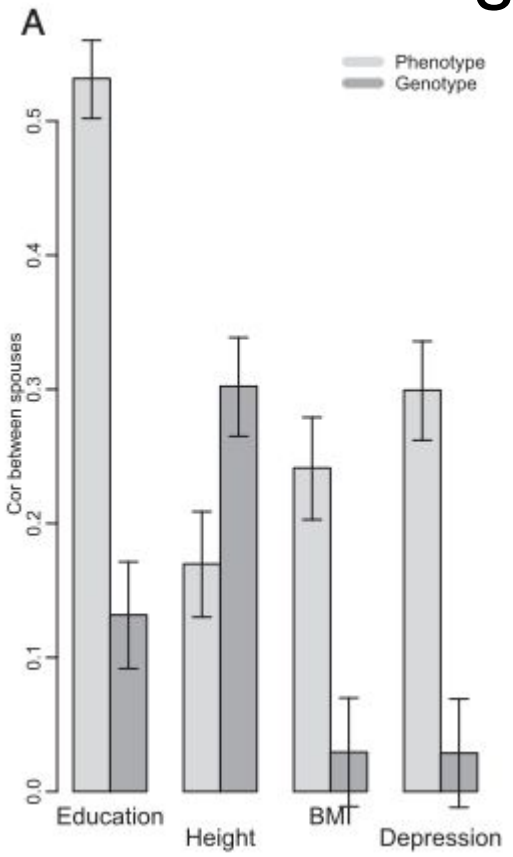
- Predictive net of family SES
- **and within family.**

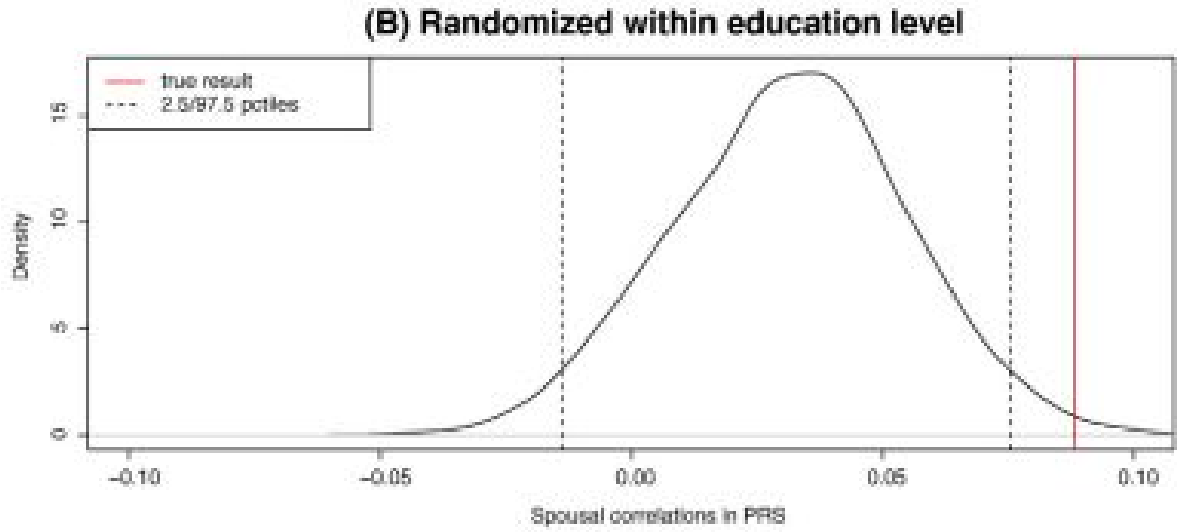
# Social Processes



What is role of genes in social life? How does social life complicate inferences re genetic effects?

# Genetic Associations between spouses




Assortative mating on educational attainment leads to genetic spousal resemblance for causal alleles

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, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands

**Assortative mating and differential fertility by phenotype and genotype across the 20th century**

Dalton Conley[a,b,1], Thomas Laidley[a], Daniel W. Belsky[c,d], Jason M. Fletcher[e], Jason D. Boardman[f], and Benjamin W. Domingue[g,1]

[a]Department of Sociology, New York University, New York, NY 10003; [b]School of Medicine, New York University, New York, NY 10003; [c]Duke University School of Medicine, Durham, NC 27708; [d]Social Science Research Institute, Durham, NC 27708; [e]LaFollette School of Public Affairs, University of Wisconsin, Madison, WI 53706; [f]Institute of Behavioral Science, University of Colorado, Boulder, CO 80309; and [g]Graduate School of Education, Stanford University, Stanford, CA 94305

# Implications for fitness?

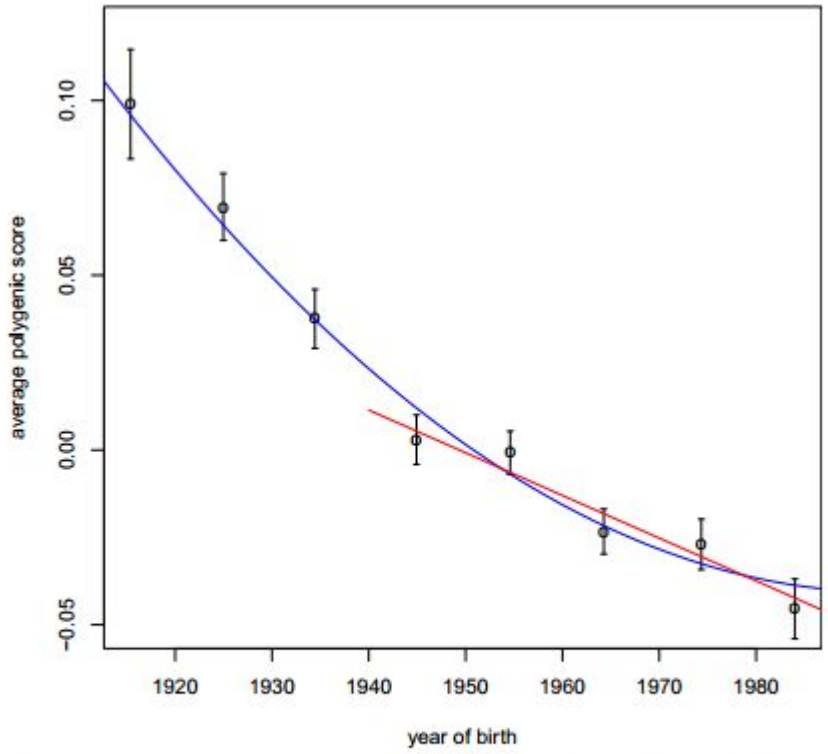



**Fig. 2.** Average educational attainment polygenic score and year of birth (yob). Results for 10-year bins are presented. Error bars indicate plus/minus 1 SE. The blue line is a quadratic fit for the full yob range indicated. The red line is a linear fit applied to individuals with yob ≥1940.

## Selection against variants in the genome associated with educational attainment

Augustine Kong[a,b,1], Michael L. Frigge[a], Gudmar Thorleifsson[a], Hreinn Stefansson[a], Alexander I. Young[c], Florian Zink[a], Gudrun A. Jonsdottir[a], Aysu Okbay[d,e], Patrick Sulem[a], Gisli Masson[a], Daniel F. Gudbjartsson[a,b], Agnar Helgason[a,f], Gyda Bjornsdottir[a], Unnur Thorsteinsdottir[a,g], and Kari Stefansson[a,g,1]

# Implications for fitness?

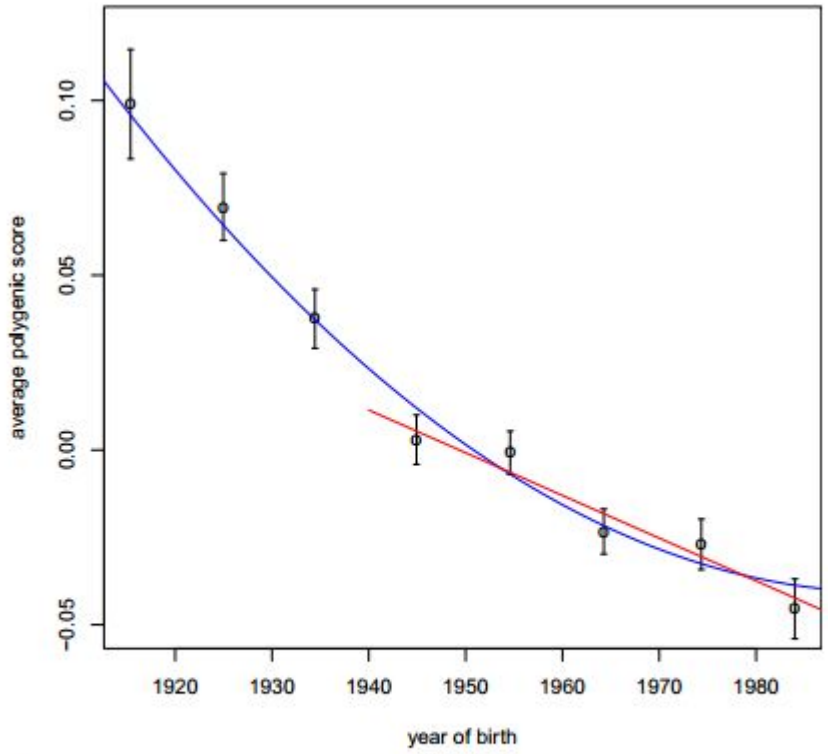



**Fig. 2.** Average educational attainment polygenic score and year of birth (yob). Results for 10-year bins are presented. Error bars indicate plus/minus 1 SE. The blue line is a quadratic fit for the full yob range indicated. The red line is a linear fit applied to individuals with yob ≥1940.

**Selection against variants in the genome associated with educational attainment**

Augustine Kong[a,b,1], Michael L. Frigge[a], Gudmar Thorleifsson[a], Hreinn Stefansson[a], Alexander I. Young[c], Florian Zink[a], Gudrun A. Jonsdottir[a], Aysu Okbay[d,e], Patrick Sulem[a], Gisli Masson[a], Daniel F. Gudbjartsson[a,b], Agnar Helgason[a,f], Gyda Bjornsdottir[a], Unnur Thorsteinsdottir[a,g], and Kari Stefansson[a,g,1]

Cautions:

I'm not a demographer.

When genes and environment disagree: Making sense of trends in recent human evolution

Alexandre Courtiol[a,1,2], Felix C. Tropf[b,c,1], and Melinda C. Mills[b,c,1,2]

I don't study evolution.

- Within-generation versus across generations.

# What about other implications of social clustering?

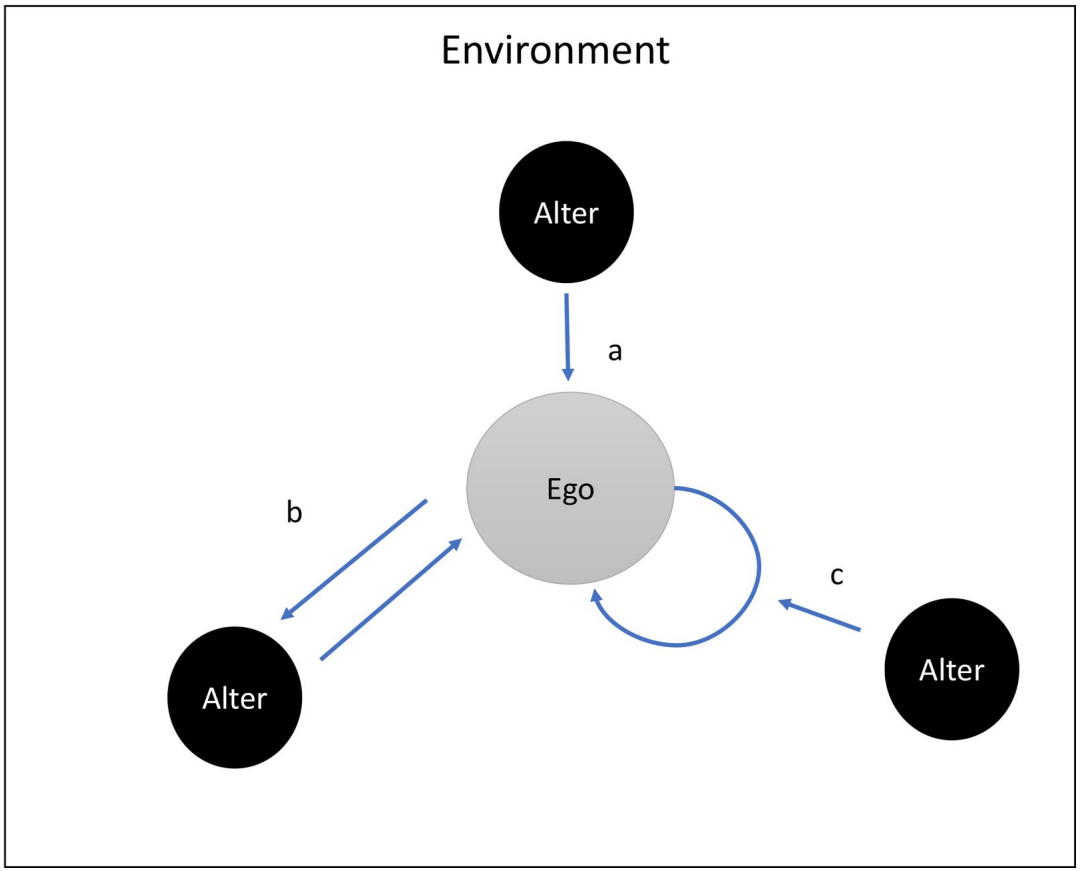
Focal Genome

Social Genome

a – **Social Genetic Effect**. An alter's genotype functions as an "environmental" influence on the ego's phenotype.

b – **Social Genetic Correlation**. Ego and alter share an environment together because shared genetic factors influence selection into that environment.

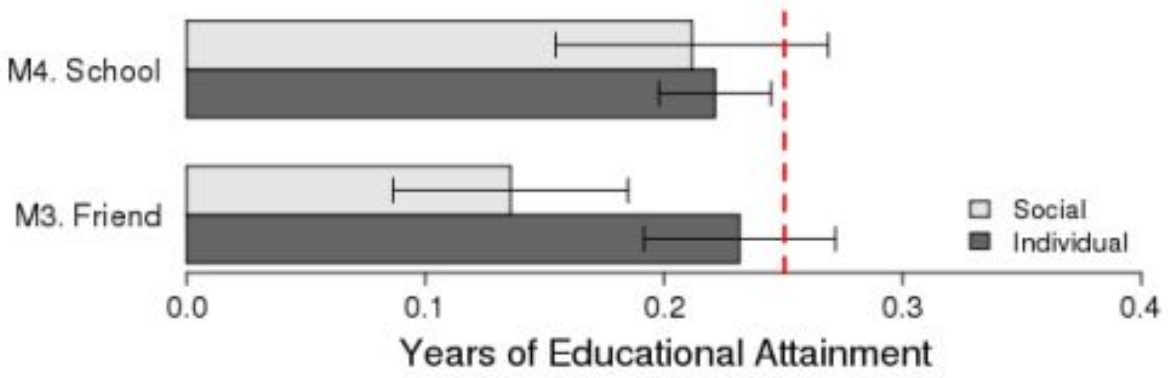
c – **Social Epistasis**. The effect of an ego's genotype on that ego's own phenotype is modified by the genotype of an alter in the same environment.

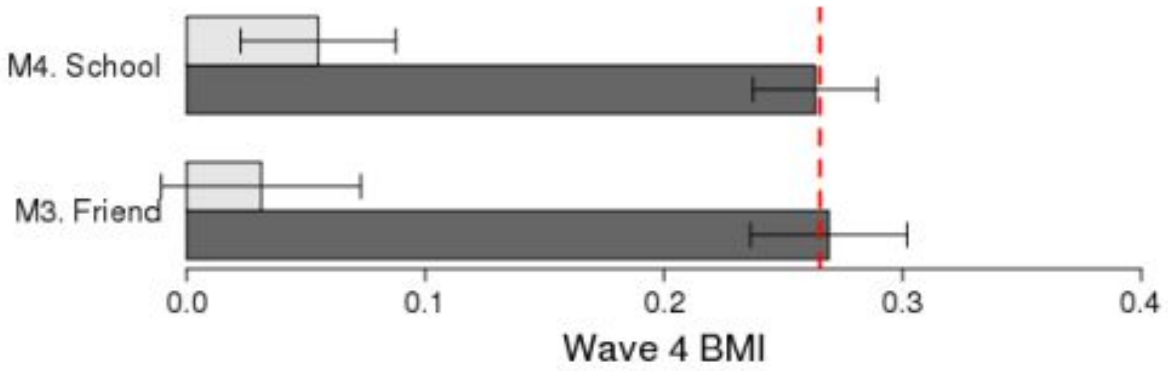


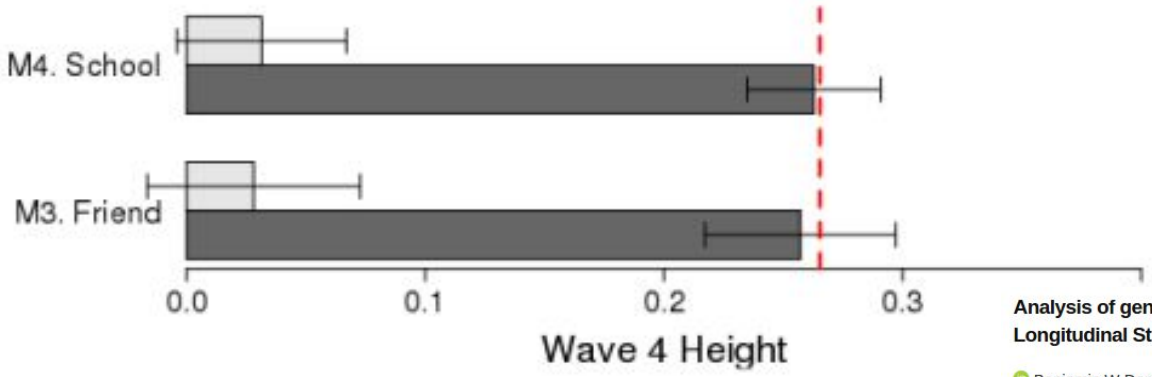

The social genome: Current findings and implications for the study of human genetics

Benjamin W. Domingue, Daniel W. Belsky

Published: March 16, 2017 • https://doi.org/10.1371/journal.pgen.1006615




Social genome, as measured by educational attainment PGS, **predicts** own educational attainment.

Social genome, as measured by height/BMI PGS, **irrelevant** for own phenotype.

**Analysis of genetic similarity among friends and schoolmates in the National Longitudinal Study of Adolescent to Adult Health (Add Health)**

Benjamin W Domingue, Daniel Belsky, Jason Fletcher, Dalton Conley, Jason D. Boardman, Kathleen Mullan Harris

**doi:** https://doi.org/10.1101/107045

# Some Closing Thoughts

# Gene-Environment Interplay

When “environment” is endogenous, interpretation is challenging at best.

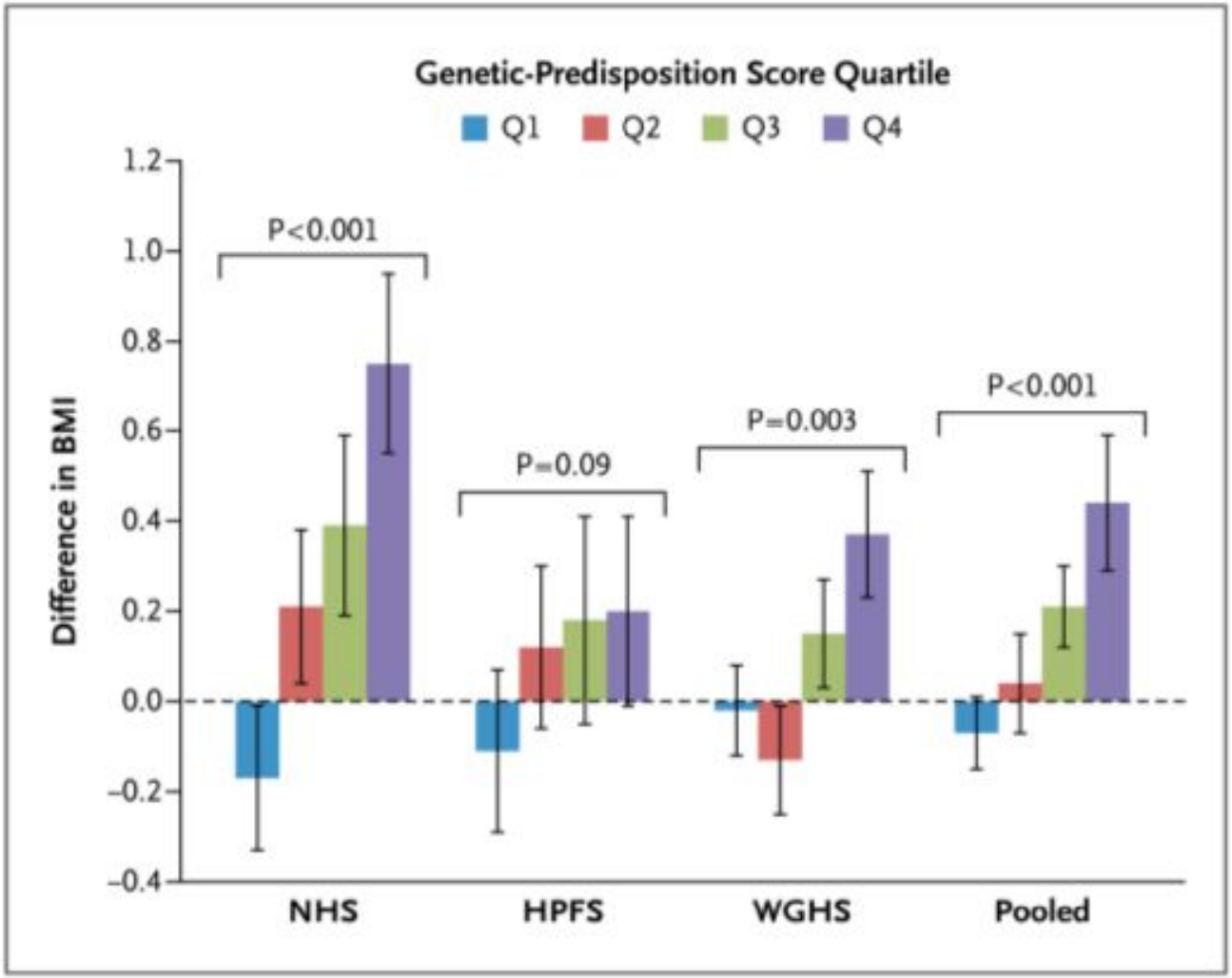
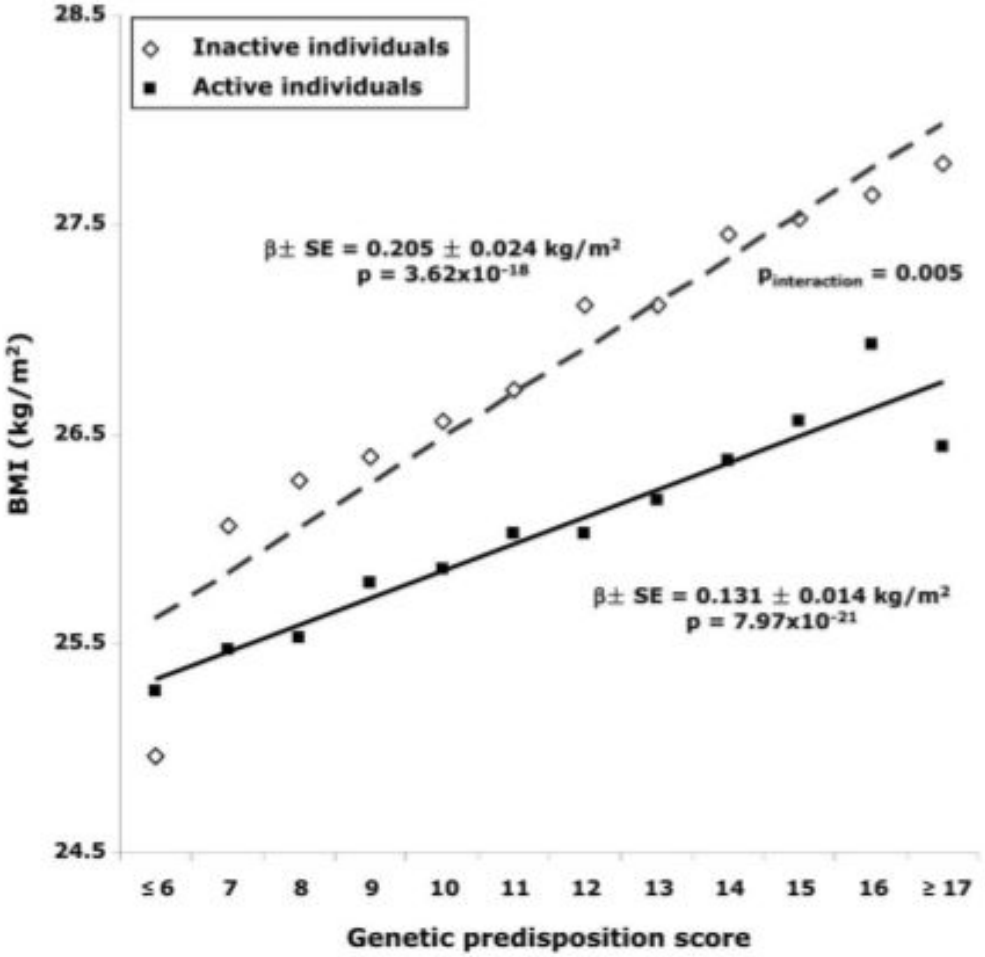



**Figure 2. Difference in BMI Associated with One Serving of a Sugar-Sweetened Beverage per Day, According to the Quartile of the Genetic-Predisposition Score**

## Sugar-Sweetened Beverages and Genetic Risk of Obesity

**Qibin Qi, Ph.D., Audrey Y. Chu, Ph.D., Jae H. Kang, Sc.D., Majken K. Jensen, Ph.D., Gary C. Curhan, M.D., Sc.D., Louis R. Pasquale, M.D., Paul M. Ridker, M.D., M.P.H., David J. Hunter, M.B., B.S., Sc.D., Walter C. Willett, M.D., Dr.P.H., Eric B. Rimm, Sc.D., Daniel I. Chasman, Ph.D., Frank B. Hu, M.D., Ph.D.,** and **Lu Qi, M.D., Ph.D.**




# Physical Activity Attenuates the Genetic Predisposition to Obesity in 20,000 Men and Women from EPIC-Norfolk Prospective Population Study

Shengxu Li[1], Jing Hua Zhao[1], Jian'an Luan[1], Ulf Ekelund[1], Robert N. Luben[2], Kay-Tee Khaw[2], Nicholas J. Wareham[1], Ruth J. F. Loos[1]*

Choices related to diet and exercise may not be orthogonal to the genetics of obesity as discovered in the GWAS.

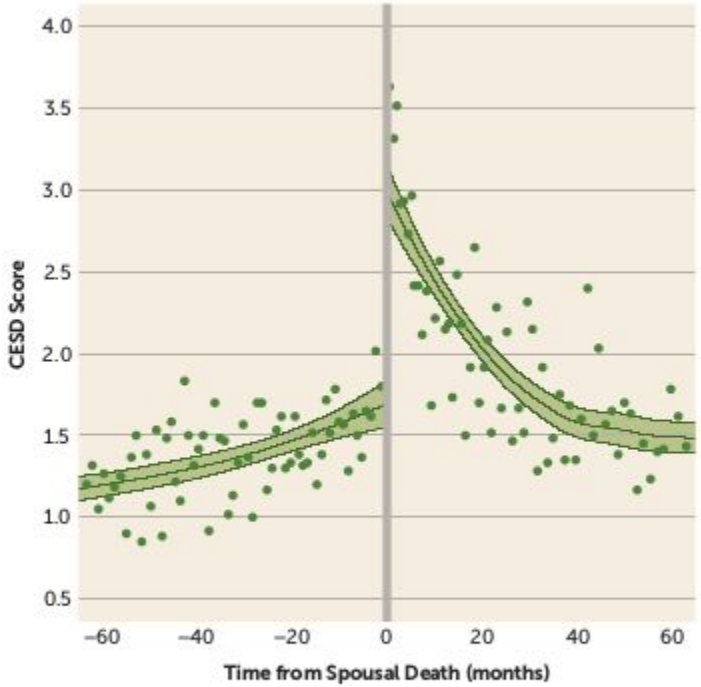
Articles

# Genetic Heterogeneity in Depressive Symptoms Following the Death of a Spouse: Polygenic Score Analysis of the U.S. Health and Retirement Study

Benjamin W. Domingue, Ph.D., Hexuan Liu, Ph.D., Aysu Okbay, Ph.D., Daniel W. Belsky, Ph.D.

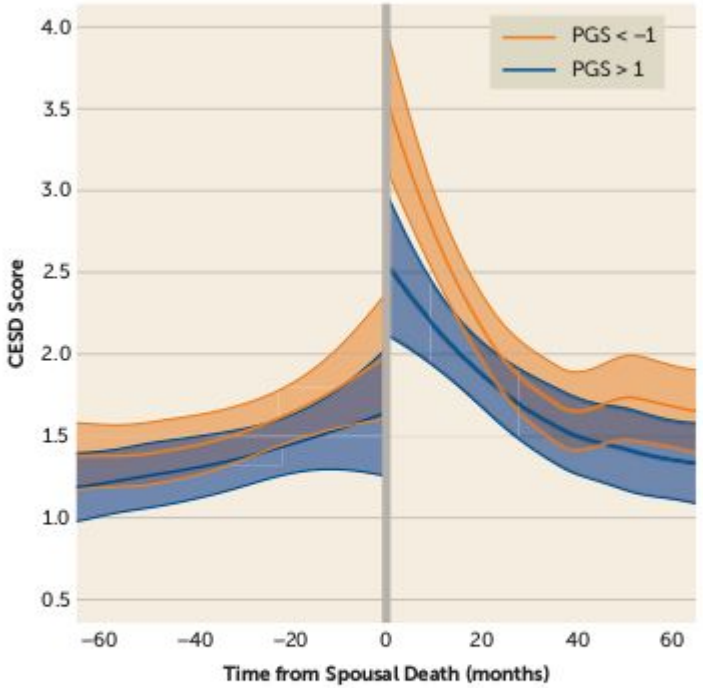
Received: November 02, 2016
Accepted: February 09, 2017
Published online: March 24, 2017 | http://dx.doi.org/10.1176/appi.ajp.2017.16111209

**FIGURE 2. Trajectories of Depressive Symptoms in the Months Surrounding the Death of a Spouse[a]**




[a] The scatterplot and local regression plot show depressive symptoms (Center for Epidemiological Studies Depression [CESD] scores) by month of measurement relative to the death of a spouse. The plot reflects data from Health and Retirement Study respondents who experienced spousal death (N=1,647; 14,309 total observations). Plotted points show mean x- and y- coordinates for bins of about 60 observations.

**of Depressive Symptoms for Health and Retirement Study (HRS) Respondents With Low Subjective Well-Being Polygenic Scores and High Subjective Well-Being Polygenic Scores[a]**




[a] The graph shows local regression plots of depressive symptoms (Center for Epidemiological Studies Depression [CESD] scores) by month of measurement relative to the death of a spouse. Trajectories for respondents with low subjective well-being polygenic scores are graphed in red (1 or more standard deviations below the mean, N=255 individuals with 2,164 observations). Trajectories for HRS respondents with high subjective well-being polygenic scores are graphed in blue (1 or more standard deviations above the mean, N=242 individuals with 2,048 observations). PGS=polygenic score.

# Repeating Myself

No matter what you are doing, make sure you're using a genetic predictor that has a robust association with the outcome of interest.

# What might the future hold?

# Ever more phenotypes (diseases, behaviors, personality traits, substance use, etc.)

## Figure 2: Results of SNP-based meta-analysis for intelligence based on 78,308 individuals.

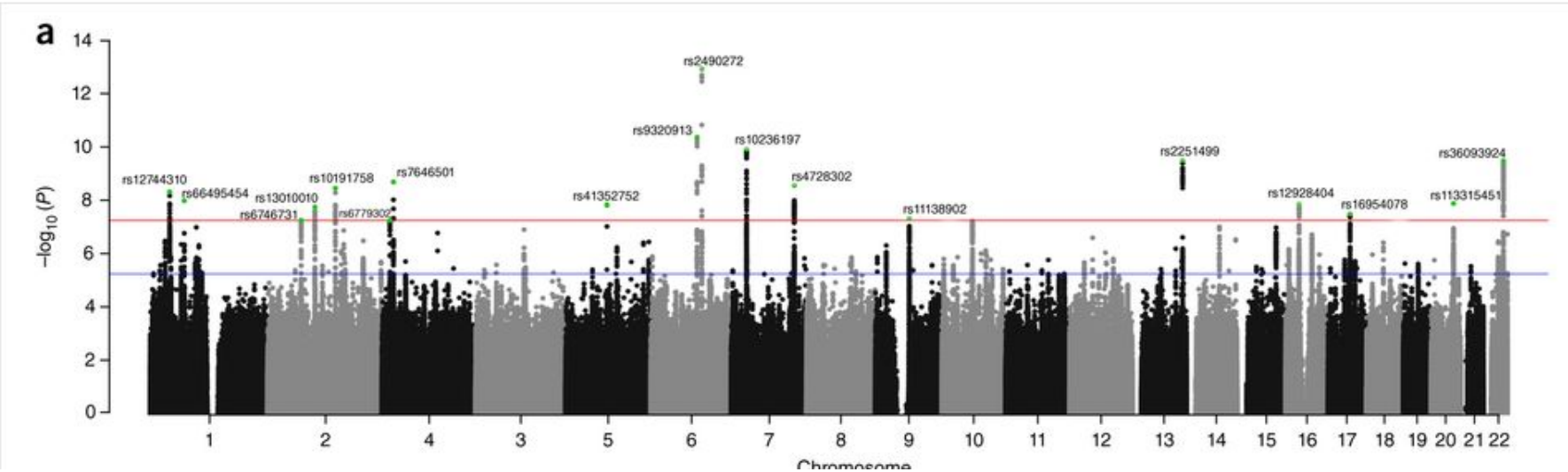
**From**

**Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence**

**Suzanne Sniekers, Sven Stringer, Kyoko Watanabe, Philip R Jansen, Jonathan R I Coleman, Eva Krapohl, Erdogan Taskesen, Anke R Hammerschlag, Aysu Okbay, Delilah Zabaneh, Najaf Amin, Gerome Breen, David Cesarini, Christopher F Chabris, William G Iacono, M Arfan Ikram, Magnus Johannesson, Philipp Koellinger, James J Lee, Patrik K E Magnusson, Matt McGue, Mike B Miller, William E R Ollier, Antony Payton, Neil Pendleton + *et al.***

*Nature Genetics* (2017) | doi:10.1038/ng.3869

Received 10 January 2017 | Accepted 24 April 2017 | Published online 22 May 2017 | Corrected online **31 May 2017**



## Genetics of the Research Domain Criteria (RDoC): genome-wide association study of delay discounting

Sandra Sanchez-Roige, Pierre Fontanillas, Sarah L. Elson, Anita Pandit, Ellen Schmidt, Johanna Forster, Goncalo R. Abecasis, Joshua C. Gray, Harriet de Wit, Lea K. Davis, James MacKillop, Abraham A. Palmer

## Genetic prediction of male pattern baldness

Saskia P. Hagenaars, W. David Hill, Sarah E. Harris, Stuart J. Ritchie, Gail Davies, David C. Liewald, Catharine R. Gale, David J. Porteous, Ian J. Deary, Riccardo E. Marioni

## Genome-wide association study of alcohol consumption and genetic overlap with other health-related traits in UK Biobank (N=112,117).

Toni-Kim Clarke[1]*, Mark J. Adams[1], Gail Davies[2], David M. Howard[1], Lynsey S. Hall[3], Sandosh Padmanabhan[4], Alison D. Murray[5], Blair H. Smith[6], Archie Campbell[7], Caroline Hayward[8], David J. Porteous[2,7], Ian J. Deary[2,9] & Andrew M. McIntosh[1,2].

## Genome-Wide Association Study Reveals First Locus for Anorexia Nervosa and Metabolic Correlations

Laramie Duncan, Zeynep Yilmaz, Raymond Walters, Jackie Goldstein, Verneri Antilla, Brendan Bulik-Sullivan, Stephan Ripke, Roger Adan, Lars Alfredsson, Tetsuya Ando,

## Genome-Wide Analysis Of 113,968 Individuals In UK Biobank Identifies Four Loci Associated With Mood Instability.

Joey Ward, Rona Strawbridge, Nicholas Graham, Mark Bailey, Amy Freguson, Donald Lyall, Breda Cullen, Laura Pidegon, Jonathan Cavanagh, Daniel Mackay, Jill Pell, Michael O'Donovan, Valentina Escott-Price, Daniel J Smith

## Genetic contributions to self-reported tiredness.

Deary V[1], Hagenaars SP[2,3,4], Harris SE[2,5], Hill WD[2,3], Davies G[2,3], Liewald DC[1,2]; International Consortium for Blood Pressure GWAS; CHARGE Consortium Aging and Longevity Group; CHARGE Consortium Inflammation Group, McIntosh AM[4], Gale CR[2,3,6], Deary IJ[2,3].

# "Clinical" Prediction

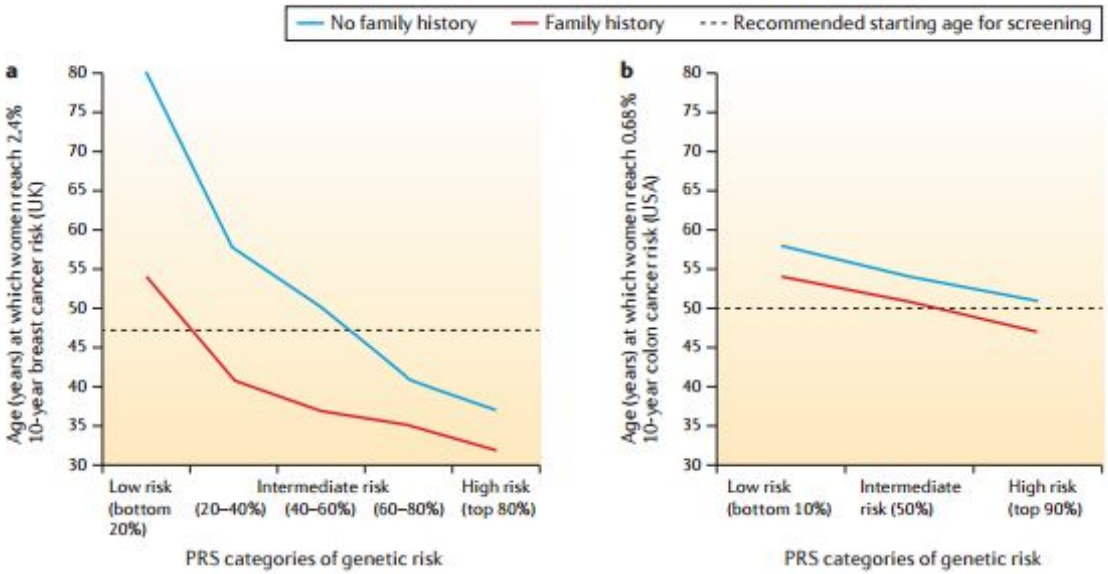



Figure 5 | **Role of polygenic risk in determining the optimal age of initiation for screening of breast and colon cancers.** Age at which the risk of developing breast cancer reaches 2.4% (panel **a**) or the risk of developing colon cancer reaches 0.68% (panel **b**) over the next 10 years, for women at different levels of polygenic risk, with and without a family history of the disease. The risk levels of 2.4% and 0.68% correspond to the average population 10-year risk of developing each disease for women at the currently recommended starting ages for screening in the countries where the original studies were conducted (that is, 47 years old for breast cancer in the United Kingdom, and 50 years old for colorectal cancer in the United States). The studies illustrate that the risk threshold for screening is reached at earlier ages for subjects with higher genetic risk, defined by the polygenic risk score (PRS) and a family history of the disease. Data in panel **a** from REF. 52. Panel **b** adapted with permission from REF. 53, Elsevier.

Developing and evaluating polygenic risk prediction models for stratified disease prevention

*Nilanjan Chatterjee[1–3], Jianxin Shi[3] and Montserrat García-Closas[3]*

- How might humans respond to genetic predictions?
- How might predictions exacerbate pre-existing disparities?

# Next wave GIV

**Genetic Instrumental Variable (GIV) Regression: Explaining Socioeconomic And Health Outcomes In Non-Experimental Data**

Thomas A. DiPrete, Casper Burik, Philipp Koellinger

**doi:** https://doi.org/10.1101/134197



Table 2: Regression of educational attainment on height in the Health and Retirement Study (HRS)

| | (1) OLS | (2) MR | (3) EMR | (4) GIV1 | (5) GIV2 |
|---|---|---|---|---|---|
| Height | 0.120*** | 0.130* | 0.0523 | 0.000594 | 0.0697 |
| | (0.0252) | (0.0599) | (0.0590) | (0.0609) | (0.0706) |
| PGS_EA_Total | 0.189*** | | 0.191*** | | |
| | (0.0165) | | (0.0166) | | |
| PGS_EA_SSGAC | | | | 0.364*** | |
| | | | | (0.0425) | |
| PGS_EA_UKB | | | | | 0.447*** |
| | | | | | (0.0438) |
| Birth year | 1.825 | 0.0763 | 2.834 | 8.066 | -2.086 |
| | (22.11) | (22.63) | (22.16) | (22.72) | (23.63) |
| Birth year, squared | -1.790 | -0.0439 | -2.800 | -8.030 | 2.115 |
| | (22.11) | (22.63) | (22.16) | (22.72) | (23.63) |
| Gender | 0.00707 | 0.0109 | -0.0435 | -0.0787 | -0.0297 |
| | (0.0250) | (0.0481) | (0.0473) | (0.0487) | (0.0559) |
| Mother's EA | 0.247*** | 0.256*** | 0.250*** | 0.231*** | 0.250*** |
| | (0.0201) | (0.0207) | (0.0203) | (0.0210) | (0.0217) |
| Father's EA | 0.198*** | 0.220*** | 0.199*** | 0.183*** | 0.176*** |
| | (0.0201) | (0.0205) | (0.0201) | (0.0210) | (0.0218) |
| $N$ | 2839 | 2839 | 2839 | 2839 | 2839 |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Standard errors in parentheses. All variables have been standardized. EA is measured in years of schooling needed to obtain the highest achieved educational degree according to ISCED classifications. The first 10 principal components in the genetic data were included as control variables. *PGS_EA_SSGAC*: PGS for EA using meta analysis from [27], excluding data from *23andMe*, UKB, and HRS; *PGS_EA_UKB*: PGS for EA using UKB data; *PGS_EA_Total*: PGS for EA using GWAS meta analysis of UKB + SSGAC[27], excluding data from *23andMe* and HRS. MR and EMR use *PGS_Height_UKB* as instrument for height. GIV1 uses *PGS_Height_UKB* and *PGS_EA_UKB* as instruments for height and *PGS_EA_SSGAC*. GIV2 uses *PGS_Height_GIANT* and *PGS_EA_SSGAC* as instruments for height and *PGS_EA_UKB*.

# THANKS!

bdomingue@stanford.edu