

Gene–environment interactions and the case of body mass index and obesity: How much do they matter?

Yiyue Huangfu ^a, Alberto Palloni ^{a,b,*}, Hiram Beltrán-Sánchez ^c and Mary C. McEniry^a

^aCenter for Demography and Health of Aging, University of Wisconsin-Madison, Madison, WI 53706, USA

^bInstituto de Economía, Geografía y Demografía (IEGD), Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Madrid 28006, Spain

^cFielding School of Public Health and California Center for Population Research, UCLA, CA 90095, USA

*To whom correspondence should be addressed: Email: palloni@ssc.wisc.edu

Edited By: Adelia Bovell-Benjamin

Abstract

We investigate the demographic and population health implications of gene–environment interactions (GxE) in the case of body mass index (BMI) and obesity. We seek to answer two questions: (a) what is the first-order impact of GxE effects on BMI and probability of obesity, e.g. the direct causal effect of G in different E's? and (b) how large is the impact of GxE effects on *second-order health outcomes associated with BMI* and obesity, such as type 2 diabetes (T2D) and disability? In contrast to most of the literature that focuses on estimating GxE effects, we study the implications of GxE effects for population health outcomes that are downstream of a causal chain that includes the target phenotype (in this case BMI) as the initial cause. To limit the scope of the paper, we focus on environments defined by birth cohorts. However, extensions to other environments (education, socioeconomic status (SES), early conditions, and physical settings) are straightforward.

Keywords: gene–environment interactions, obesity, type 2 diabetes

Significance Statement

Recent studies highlight the gene–environment interactions that may affect the risk of obesity in the US adult population. These studies show that individuals born after 1950 who have higher genetic propensities to develop higher body mass index (BMI) experience elevated risks of obesity than those with a similar genetic profile but born earlier. This paper assesses if and to what extent these excess risks of obesity translate into increases in the risks of type 2 diabetes (T2D) and disability. We find that the gene–cohort interaction effects exert direct impact on the youngest cohorts' obesity prevalence. Despite the strong relationship between obesity and T2D and disability, these additional increases in the risk of obesity do not translate into significant impact on T2D or disability.

Introduction

Increasing rates of obesity prevalence in the 20th century appear simultaneously in high-income countries around 1970–1980¹ and spread rapidly to low- and middle-income countries. Since 1975, the worldwide prevalence of obesity has trebled but varies widely across geographic regions (1). Between 2000 and 2018, the US population obesity prevalence grew from 30.5 to 42.4% continuing a trend that began in the middle-60s from a level of about 13%, clocking a doubling time of about 31 years (2).

¹ Throughout we use the WHO definition of obesity and use the term to refer to individuals with BMI, exceeding 30 (<https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>). As long as there is no equivocation, we use the expression “effects on obesity” to mean increased risk of obesity as well as “effects on BMI” (and vice versa). The same applies to expressions such as “obesity-related GxE effects” which we will use as equivalent to “BMI-related GxE effects.” We will refer to BMI and obesity as equivalent phenotypes by using the awkward “BMI–obesity” expression. Throughout we use “GxE effects” to mean gene–environment interaction effects.

More ominous is the rapid increase of obesity among children and adolescents. In the 40-year period between 1965–1969 to 2008, obesity prevalence among children and adolescents spiked from 5% to a level about twice as high while it trebled during the same period among those aged 6–19. Since then, prevalence rates increased to 15 and 20% in each age group respectively (3). By virtue of the association between child and parental obesity (4), on the one hand, and individuals' early childhood and adult obesity (5), on the other, these trends might lead to intergenerational “transmission” of the phenotype.

Sizeable increases in human girth are not by themselves of immediate concern.² What preoccupies scientists and health

² We do not belittle the personal consequences of obesity. In societies where the phenotype is stigmatized, it causes discrimination, maltreatment, isolation, and mental illness and imposes an incalculable psychological cost to individuals. Furthermore, an important fraction of the economic burden associated with obesity is borne by the individuals themselves.

Competing Interest: The authors declare no competing interest.

Received: April 15, 2023. **Revised:** June 7, 2022. **Accepted:** June 13, 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of National Academy of Sciences. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

policymakers alike is the evolution of phenotypes related to obesity. Obesity is associated with metabolic syndrome (6, 7), elevated risks of chronic conditions such as type 2 diabetes (T2D), cardiovascular disease (CVD), cancer, stroke, midlife cognitive performance, late-life cognitive decline, and increases in fragility and disability (8, 9). Although the direct impact of obesity on mortality is controversial, its indirect effects through chronic conditions and illnesses are strong and undisputable (10–12).

These relations between obesity, chronic illnesses, and disability are hugely consequential. It has been estimated, for example, that in 2010–2012, the US medical costs of obesity hovered around a staggering 150 billion per year (2014 US dollars) (10) and could have been as large as 210 billion. Most of this spending is associated with the treatment of T2D, other closely associated chronic conditions, and disability (11).

There is widespread consensus that the root of the post-1950 increase of obesity is environmental and associated with wholesale changes in diet, physical activity, sleeping patterns, and stress (12). These “obesogenic environments” are well entrenched and unlikely to be dismantled any time soon (13, 14).

Obesogenic environments, however, are not the only game in town. Other determinants may either reinforce (or weaken) future trends of obesity and associated chronic conditions, even if obesogenic environments remain unchanged. One of these is the genetic makeup of a population. Family- and kin-based studies estimate that additive genetic effects account for about .40 to .70 of the phenotype variance (15, 16). More recently, genome-wide association studies (GWAS) studies confirm that body mass index (BMI) and other obesity markers such as waist-hip ratio (WHR) and waist circumference (WC) are polygenic traits involving multiple allelic variants (17–20). Estimates of heritability from these studies are within a much lower range, between 5 and 15%. Barring shifts in assortative mating and sharp differentials in net reproduction rates by body size, it is improbable that these additive allelic effects could significantly shift future trends of BMI, obesity, or associated health outcomes.³

Enter gene–environment interactions (GxE), and we may have a different story. Not only can these have direct, first-order, impacts on BMI and obesity but may also have indirect, second-order, impacts on chronic illnesses, disability, and mortality. Were these second-order impacts of GxE associated with BMI to be important, their population health implications will be significant for current and subsequent generations. The goal of this paper is to examine jointly first- and second-order effects of GxE. We seek to answer two questions: (a) what is the likely range of first-order impacts of GxE effects on BMI and obesity, e.g. the direct effect of G in different E? and (b) how large is the second-order impact of GxE effects, e.g. the effect on health outcomes associated with BMI and obesity? Because they have attracted a great deal of recent attention, we focus on environments defined by birth cohorts. Birth cohort is one among many “environments” highlighted in social research on GxE (22–25). In the case of BMI and obesity, birth cohort is a surrogate for “timing of onset of widespread exposure to obesogenic environments.” The magnitude of GxE effects estimated with birth cohort as “environment” is among the largest involving obesity and BMI. Extensions of our arguments to other environments (education, socioeconomic status [SES], early conditions, physical settings, etc.) are straightforward.

³ While it is unlikely that vertical genetic transmission alone may become a driving force of the phenotype’s trajectory, it is possible that it, in combination with vertical cultural transmission, can have nonnegligible impacts (21). In addition, to potent vertical cultural transmission, and in the absence of mutations, the key drivers of genotype frequencies associated with BMI and obesity and the phenotype trajectories in large populations will be assortative mating and differential fertility.

GxE effects: How large are they?

Large GWAS studies have made possible the rapid growth of empirical studies that seek to identify the association between thousands of allelic variants and a growing array of phenotypes. An increasing fraction of these studies use polygenic risk scores (PRS) to estimate the additive effect of multiple allelic variants on a phenotype. Although in most cases the fraction of explained variance by PRS is quite small, these findings support the idea that knowledge about phenotypes of interest to social scientists could be much improved if, alongside other standard determinants, researchers consider the impact of genetic factors (26–32).

An important part of this new research program focuses on GxE, that is, variation of phenotypic response of a single genotype to changes in environments (33). In social sciences, GxE refer to situations in which the *additive causal genetic effect* on a trait or behavior is different among individuals belonging to well-characterized subgroups (gender, age, education, and SES) or social settings (normative vs. nonnormative and exposed vs. non-exposed to risks or interventions).⁴ GxE effects are also of interest in evolutionary biology and population genetics for they could influence the allelic composition of entire populations and drive the evolution of phenotypes under selection pressure (34–37).⁵

GxE have been at the center of heated debates regarding the relevance of heritability of phenotypes, including the relative importance of genes and environments in the production of social and economic inequalities (32, 33, 39, 40). Empirical evidence from GWAS-based studies that identify GxE in multiple phenotypes (education, IQ, noncognitive traits, depression, and onset of sexual activity) has recently been invoked to support the formulation of policy interventions that are better informed about the role of individuals’ genotypes (32). A common inference in these studies is that identification of GxE effects is not only relevant for theory building but might also benefit the design of interventions for they can guide identification of subgroups that, by virtue of their genetic makeup, are at higher risks of deleterious outcomes in some environments (28–30, 41, 42). Although this may be the correct inference, it is not always clear what the implications of detected GxE effects are, that is, the “so what question”: how large are they? How do they stack up against other determinants identified with similar or higher precision? If, for example, some phenotypes’ sensitivity to genetic risks has indeed increased across US birth cohorts (22), what does this mean for subpopulations that express them? How consequential are GxE effects involving obesity and relevant environments (educational attainment, SES, birth cohort, early conditions, and residential location)? And what do they imply for other phenotypes associated with them?

Inferences about GxE effects are usually made at three different analytic levels. The first demands quantification of the impact of GxE effects on shifts of the average phenotype across generations. This is the breeders’ concern and is the target of researchers interested in selection pressures under which a phenotype may be evolving.

The second level consists of identification of either environments that modify the contribution of individual genetic risks or

⁴ The literature distinguishes three main types of GxE, depending on the functional form of the relationship between genotype, environment, and outcomes. These are diathesis stress, differential susceptibility, and social push models (26, 30).

⁵ Throughout, we will focus on the causal effects of gene variants, namely, the slopes of phenotypes relative to variables measuring genetic variation. In standard linear models, these are not to be equated with heritability (h^2). The two metrics are indistinguishable only in classic path analysis, e.g. when all variables are standardized. When variables are in their natural scales, h^2 and slopes, though related, can behave differently (38).

of groups of individuals whose genetic profiles make them more (less) vulnerable to express a phenotype under well-defined E's. In studies of individual depression, for example, the phenotype under study is the target of interest. In these cases, GxE effects may matter a great deal because they could shed light on modifiable environmental conditions that exacerbate (attenuate) the role of genetic propensities.

These two analytic levels are both about first-order effects. The third level of analysis has been little studied. It is about second-order effects, e.g. those associated with traits further down a causal pathway in which the phenotype under study is either an initial condition or a mediator. For example, a researcher may investigate the genetic determinants of age at first birth (43) or age at first sexual intercourse (44) not because these phenotypes are of intrinsic interest (though they might be) but because they could be consequential for other outcomes such as higher fertility, lower female labor force participation, early high school dropout, and onset of criminal activity or substance abuse (45).⁶ The pressing concern may be to understand the mechanisms that produce these second-order outcomes as a response to changes in the phenotype under study. If age at first birth has a causal effect on females' subsequent labor force participation, a GxE associated with age at first birth could have large repercussions for aggregate female labor supply. For economists, the target quantity is not the fraction of total variance in age at first birth or age at first sexual encounter explained by G or GxE, but the magnitude of shifts in female labor supply expected under a set of genetic profiles situated in different environments.

A good illustration is the case of BMI and obesity. We know that increases in body size are just the beginning of a chain of physiological changes that lead to metabolic syndrome, prediabetes, T2D, circulatory dysfunctions, fragility, disability, and death. If risks of increased BMI (or probability of obesity) for individuals of a given genetic propensity are higher among those in the lowest educated groups (25, 46) or in younger birth cohorts (25), a relevant issue for population health scientists is what the impacts of these environmentally enhanced genetic risks are on T2D, kidney and heart disease, and stroke. The concern with the effect of GxE on BMI and obesity is ancillary and the more pressing problem is the impact of these GxE on chronic conditions, the phenotypes at the end of the causal chain.⁷

It is worth noting (see footnote 3) that GxE effects in one generation may have lasting impacts on subsequent generations. Because parents with a given genetic propensity for BMI (obesity) in one environment (in the absence of an obesogenic environment) may experience added risks of higher BMI (obesity) when exposed to a new environment (obesogenic), the genotypic composition of the obese population in the parental generation will change (its average genetic propensity will decrease). This has two consequences: first, due to phenotypic assortative mating, it will reduce the average genetic propensity that offspring inherit from parents who have higher BMI; second, it may increase the penetrance of parental shared environments that are characteristic of couples with higher BMI, even if parents in these couples have lower genetic propensity to be obese. Thus, even if the offspring generation does not experience additional GxE, its average BMI (and probabilities of obesity) may increase but now because of nongenetic inheritance.

⁶ Second-order phenotypes could be under the influence of additive allelic effects that may or may not affect the phenotype under immediate study.

⁷ Estimates of first-order GxE effects will, of course, always be relevant for those interested in the biology of obesity even if no effects are expected on second-order phenotypes.

To assess the magnitude of first- and second-order impacts of GxE on BMI, T2D, and disability, we employ a counterfactual approach. We ask, what would the BMI (or probability of obesity) be for the cohort that experienced the new post-1950 obesogenic environments if they had not been influenced by the estimated GxE? This quantity is the observed BMI for individuals belonging to the birth cohort that experienced obesogenic environment minus the estimated interaction effect (see Materials and methods section). We then use the counterfactual values of BMI and compute quantities related to T2D and disability in a hypothetical scenario in which those born after 1950 experience the impact of their enhanced genetic predisposition to have higher BMI.

Results

Results I: First-order effects on BMI and obesity

Table 1 and Fig. 1 display results after computing counterfactual (e.g. in the absence of GxE) BMI values using the Health and Retirement Study (HRS) estimate of GxE (see Table S6 and Fig. S2 for results including other counterfactuals) (47). Figure 1 displays the observed and counterfactual values of BMI by deciles of PRS. Table 1 displays the mean BMI and the share of obese individuals in the observed and counterfactual case. In the observed scenario, individuals in all deciles of PRS have higher BMI than in the counterfactual case and those with higher PRS experience the largest increases in BMI associated with GxE. On average, the observed BMI is 28.66, about 2.12 BMI units higher than in the counterfactual case. For a person of median height (175 cm), this is equivalent to an increase of about 14.3 lb. As expected, the increase in BMI is more pronounced for individuals with higher PRS. For those in the top decile, for example, the counterfactual BMI is 3 units lower than the observed, equivalent to 20.3 lb. Column (4) of Table 1 displays the impact of GxE on obesity prevalence. In the absence of GxE, the prevalence of obesity would be 23% or 13 percentage points lower than observed and for individuals in the top decile of BMI's PRS, the reduction is as large as 19 percentage points.

The question we pose next is about the magnitude of second-order effects on T2D and disability. To estimate these quantities,

Table 1. Mean BMI and share of obese individuals by decile of BMI PRS, cohorts 1945–1959, with and without GxE (GxE = 0.57).

Decile	With GxE (observed)		Without GxE (counterfactual, GxE = 0.57)	
	BMI (1)	% Obese (2)	BMI (3)	% Obese (4)
1	25.71	0.16	24.69	0.11
2	26.93	0.24	25.54	0.14
3	27.49	0.26	25.51	0.17
4	27.99	0.31	25.95	0.21
5	28.56	0.31	25.96	0.21
6	28.83	0.37	26.80	0.24
7	29.32	0.36	26.83	0.25
8	29.86	0.48	27.65	0.31
9	30.48	0.50	27.91	0.31
10	31.58	0.55	28.58	0.36
Total	28.66	0.36	26.54	0.23

Note: Counterfactual BMI/share of obese individuals refers to the BMI/share of obese individuals if there were no GxE; the value of GxE effect is as estimated using the HRS sample.

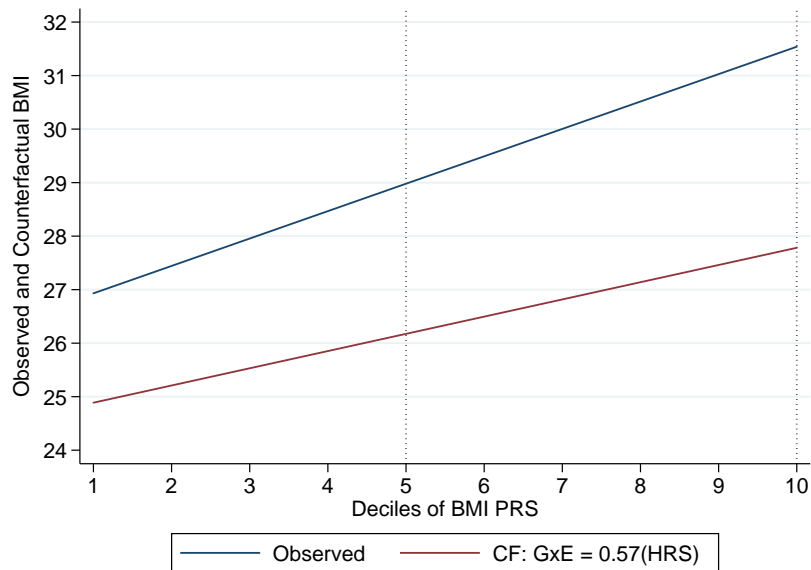
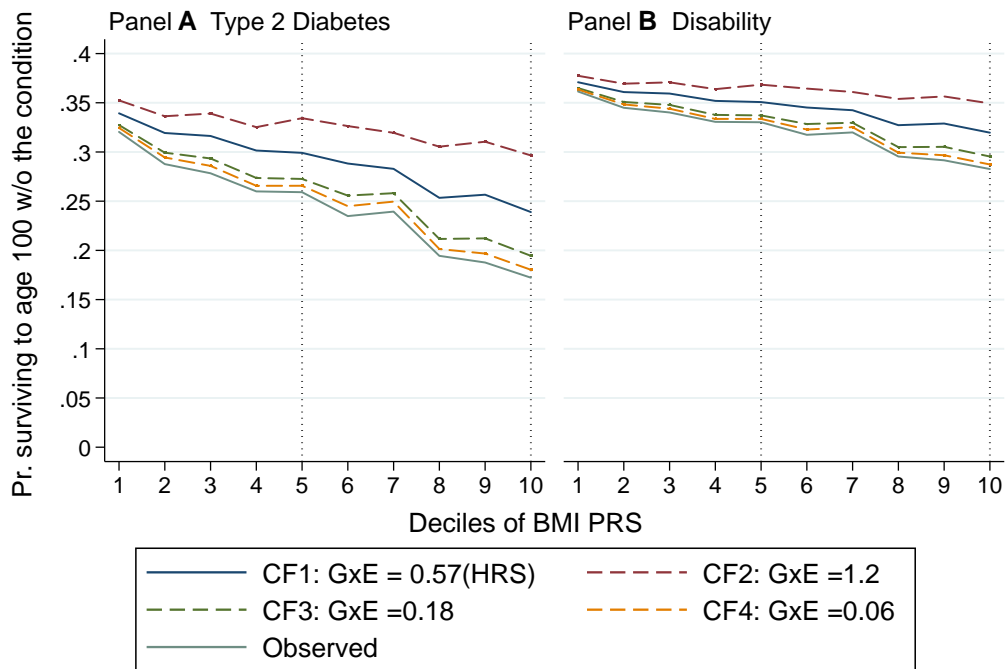


Fig. 1. BMI of cohort 1945–1959 by deciles of BMI PRS, as observed and in counterfactual. CF stands for counterfactual. Vertical dotted lines indicate the 5th and 10th deciles.



Graphs by condition

Fig. 2. A) Single decrement probability of never contracting T2D at age 100, cohort 1945–1959. B) Single decrement probability of never becoming disabled at age 100, cohort 1945–1959. Pr. stands for probability and CF stands for counterfactual. Vertical dotted lines indicate the 5th and 10th deciles.

we concentrate on impacts via obesity (rather than BMI) since the bulk of the literature on the subject uses obesity as the main predictor.

Results II: Second-order effects

Panels A and B in Fig. 2 display probabilities of never contracting T2D after age 50 and never becoming disabled, respectively,

predicted at each decile of the PRS distribution.⁸ In each figure, there are five lines: one of them (“observed”) plots BMI values that include the contribution of GxE. The line labeled CF1 plots

⁸ It is standard practice in demographic studies to gauge the importance of an event’s risk by employing the single decrement probability of never (ever) experiencing within the age or duration range within which the event has a non-zero risk.

counterfactual values that are obtained after eliminating the GxE effect estimated from HRS (CF1). Lines CF2–CF4 do the same for each of three remaining alternative estimates of GxE effects. Figure 2A shows that, except for the case when the GxE effect is as large as 1.2 (CF2), the impact of GxE on the probability of never contracting T2D is quite modest, except perhaps in the upper part of the PRS distribution. In fact, when the GxE effect is 1.2, the largest impact is for individuals in the top decile of PRS: if there was no GxE, the probability of never contracting T2D increases from .18 to .30. A lower increase, from .18 to .24, applies when the estimate of GxE is equivalent to that found in HRS. The magnitude of impacts is much smaller around the center of the PRS distribution, not larger than .03.

The estimated impacts on disability are even smaller. If the true GxE were as high as estimated in HRS (CF1) and we suppressed it, the probability of never becoming disabled would increase from .28 to .33 among those located in the top decile of PRS and from .33 to .35 for those in the 5th decile. If the true GxE is set to its maximum value (CF2), the change would be from .28 to .35 and .33 to .36 for those in the top and 5th decile, respectively.

We turn to effects on the expected number of years lived (from age 50) without T2D or disability using single decrement probabilities (ignoring mortality). Panel A of Fig. 3 shows that individuals in the 5th decile of PRS who avoid the impact of GxE are expected to live about 1.4 year longer with no T2D compared with those who experience GxE of the magnitude found in HRS (33.6 vs. 32.2). Among those in the top decile, the difference is about 3 years (31.5 vs. 28.5). When the GxE effect rises to the maximum we use here (CF2), the differences are 2.8 and 5.4 years of life for the 5th and top decile, respectively. In relative terms, these impacts fall in the range of 2–11% of the total lifespan after age 50. The bulk of these differences, however, is not larger than 5%. In the case of disability (Panel B), the contrasts between predicted and counterfactual are again smaller: the maximum value is 2.2 in the extreme case of the top decile of PRS and the most powerful GxE estimate. The range of relative differences is between 1 and 4% of total life expectancy after age 50, with most of the values falling below 2.5%.

We now focus on the quantity that, at least from a health policy perspective, is the most important: the expected number of years to be lived after age 50 *with T2D or with disability*. These are key numbers for estimating long-term costs associated with the burden of T2D and disability are, by far, the most relevant in the accounting of a nation's and households' budget. To compute these quantities, we construct multiple decrement tables accounting for the joint incidence of T2D and disability, on one hand, and mortality risks associated with them, on the other. We estimate hazard models for mortality in the same HRS sample and include T2D and disability as predictors (plus controls).⁹ We then combine the predicted hazards with the single decrement tables for T2D and disability estimated before and, finally, compute expected duration of life with T2D or with disability (see Section SII). Panels A and B of Fig. 4 display the average number of years of life after reaching the 50th birthday to be lived with T2D and disability, respectively. In both cases, and no matter how extreme a PRS decile and the magnitude of the GxE effect are, the differences are trivial, generally smaller than .5 years.¹⁰

⁹ The mortality hazard models do not include BMI or obesity since their effects are statistically insignificant.

¹⁰ A better assessment requires to introduce as actual yearly costs of the condition as weights. At least in the US health system, even a 5-year difference can be the equivalent of a large fraction of an average household's assets.

Discussion

Summary

Obesity is a phenotype strongly related to important health outcomes for one and, because of its potential reproduction via vertical genetic and cultural heritability, for several generations. The magnitude of GxE affecting the trait matters as it affects the reproduction, health status, and survival of the organism that bears it. Even if obesity is not influenced by assortative mating (which it is) and had no effects on humans' net reproduction rate (which, in some cases, may have), its association with modern human chronic conditions and disability is sufficiently tight to make GxE relevant from a public health standpoint. Thus, it is highly relevant to assess whether the direct effect of GxE on BMI also translates into significant impacts of these chronic conditions and disability. If that is not the case, then *while the original GxE effects (PRS for BMI/obesity and cohort) may be relevant as a tool to target groups for interventions to reduce obesity, the second-order effects will not provide additional helpful clues for interventions aimed at T2D or disability.*

Our empirical estimates indicate that the magnitude of known estimates of GxE has some impacts on the prevalence of obesity. However, its aggregate impact on two demographic outcomes of importance in population health is small. GxE effects are relevant only to the small fraction of individuals in the upper extreme of the genetic risk distribution and only when estimates of GxE are set to the maximum value we were able to retrieve from past studies. We also show that the impact on a key parameter for health policymaking, namely, the duration of life with a chronic condition or disability, is very small. If the magnitude of GxE effects is as documented in recent research (or less), then GxE effects are too small to influence health outcomes that demographers and population health scientists are interested in.

Limitations

These findings, however, should be interpreted with caution. It is possible that although commonly used in other research, the tools we employ are too blunt to detect first- and second-order GxE effects. First, the PRS, the "G" in our model, has well-known weaknesses as an indicator of genetic propensity. The issue of "missing heritability" looms large here. If it is due to imperfection of estimates of allelic effects, we will underestimate the influence of both additive and interaction effects. Furthermore, our estimates are potentially affected by confounding, sample selection, insufficient statistical power, model misspecification, and measurement error. These pitfalls may lead to over or underestimation of main and interaction effects.

Second, birth cohort defined by discrete periods, the "E" in the model, is at best a very coarse indicator. We use it to follow recent studies in which birth cohort is a proxy for a "treatment," namely, exposure to obesogenic environments. This may be justified on the grounds that it captures transformations (in physical, ideological, legal, judicial, and health conditions) that regulate the expression of the phenotype. But it is an extremely vague and crude construct for we know next to nothing about the mechanisms linking it to the phenotype and genotype of interest.

Third, we use a sample of non-Hispanic White (NHW) elderly residents of the United States. The rationale behind this stems from our observation that in recent studies consistent empirical support for the presence of GxE effects is only found among this specific demographic group. Consequently, our inference is confined to this demographic group. Additional research should

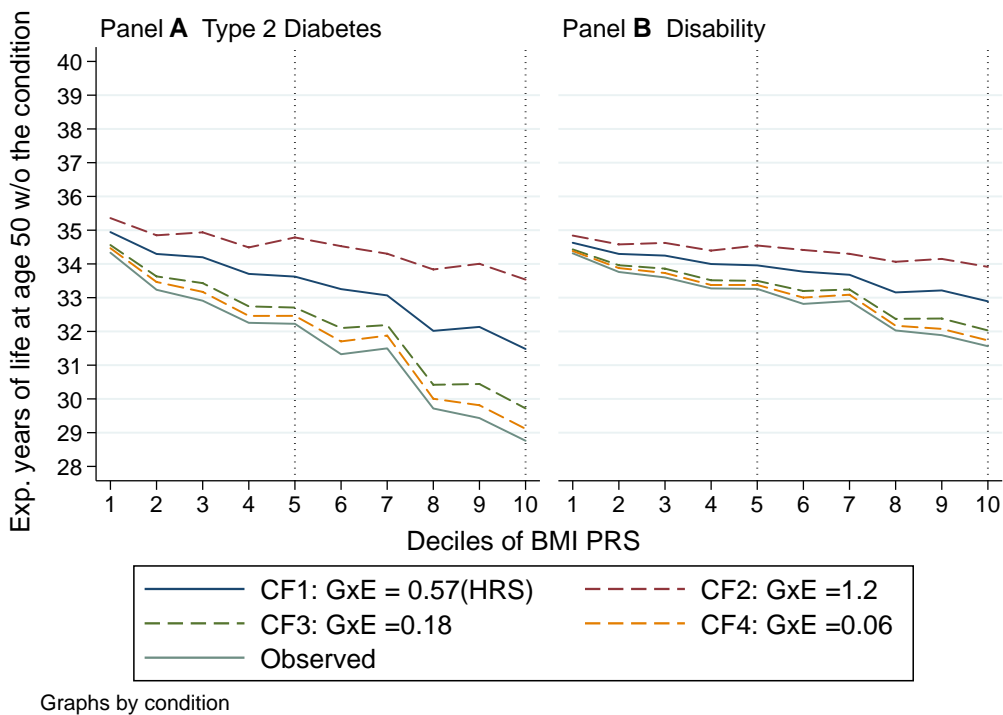


Fig. 3. A) Single decrement expected years of life at age 50 without T2D, cohort 1945–1959. B) Single decrement expected years of life at age 50 without disability, cohort 1945–1959. Exp. stands for expected and CF stands for counterfactual. Vertical dotted lines indicate the 5th and 10th deciles.

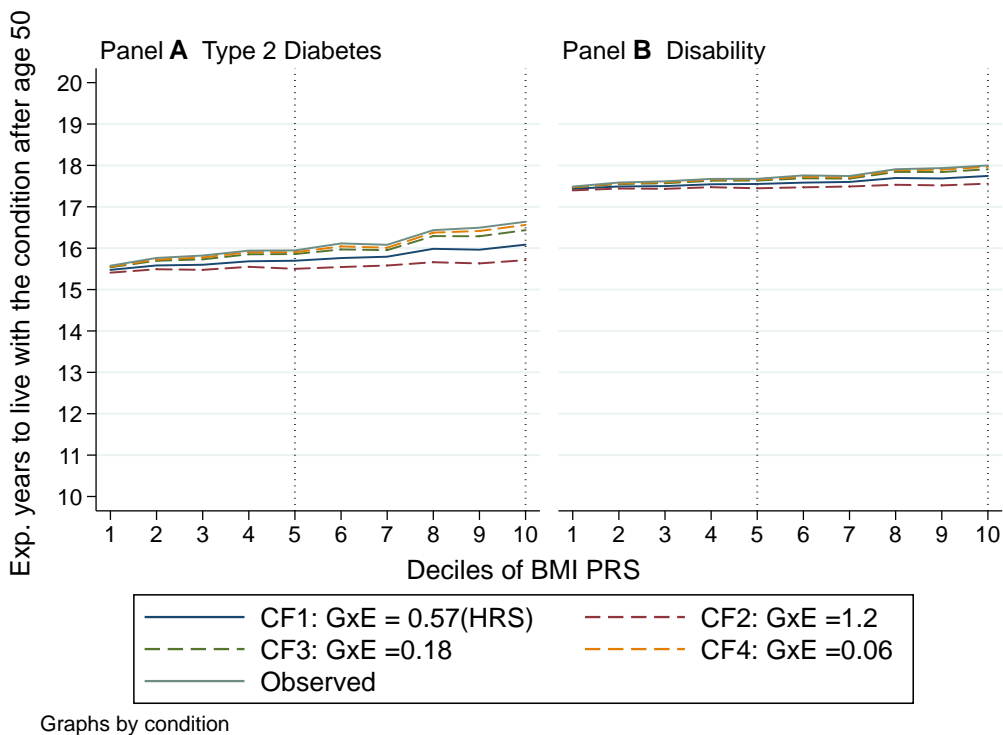


Fig. 4. A) Expected years of life to live with T2D after age 50, cohort 1945–1959. B) Expected years of life to live with disability after age 50, cohort 1945–1959. Exp. stands for expected and CF stands for counterfactual. Vertical dotted lines indicate the 5th and 10th deciles.

determine the magnitude of GxE effects in other groups and whether they have any significant second-order impacts.

Fourth, we study second-order effects manifested during a fraction of the total life span within which adult health conditions related to obesity may surface. This is especially the case for T2D, a disease whose incidence begins to increase at around ages 30 to 35 and peaks at ages 50 to 55. By starting life tables at age 50, we only capture the known GxE effects but only in oldest 70% of the relevant age range. However, if one assumes that GxE effects for the younger population are the same as for the population older than 50, the quantities estimated here would change minimally.

Fifth and finally, we hasten to emphasize that our inferences apply only to selected second-order phenotypes, T2D and disability, and refer only to effects induced by GxE. The *direct impact* of GxE on BMI and obesity is modest but not trivial and may explain, albeit partially, recent increases in rates of obesity prevalence and could have an impact on T2D and disability. However, the effects of GxE on T2D and disability mediated by BMI and obesity are inconsequential. It is of course possible that different results are obtained with other health outcomes related to BMI and obesity. This remains to be investigated but it is unlikely to be a deal breaker, as T2D is by far the strongest second-order outcome associated with obesity. By the same token, we study only one phenotype, BMI (obesity), and conclusions about the relevance of GxE effects should be confined only to it. The magnitude of GxE first- and second-order impacts could be more significant for other phenotypes. For example, there are important GxE effects involving smoking behavior and changing environments (birth cohort and social contexts) (22, 23, 48); might it not be case that second-order effects of GxE in the case of smoking (lung and other cancers, chronic obstructive pulmonary disease (COPD), and CVD) are highly significant? This is certainly possible but, to demonstrate it, researchers should abandon the common practice of halting their investigation after detecting and estimating only first-order impacts of GxE.

Conclusion

Population health scientists should routinely trace GxE effects on phenotypes located farther out in a causal chain that begins with a target phenotype located at the top of it. A comprehensive assessment of the relevance of GxE effects on human health ought not to be limited to first-order effects but should also include health-related second-order phenotypes. Until this becomes part of standard population research practice, studies of GxE will not have as much impact as they might deserve.

Materials and methods

We use HRS data including only the NHW population¹¹ and estimate the magnitude of GxE effects in models predicting BMI as a function of birth cohort, BMI's PRS, and controls (see Section S1 and Table S1 for further details about the HRS sample we used).¹² Estimates from the main model are displayed in Table S2. The predicted BMI by cohort and age is shown in

¹¹ We restrict analysis to NHW on account of the nature of population ancestry composition of the GWAS studies on which the HRS PRS is computed. We also estimated separate models for the African American (AA) samples in HRS and obtained lower (and statistically insignificant) effects. Furthermore, the additive allelic effects on BMI computed from GWAS studies that include representation of populations with African ancestry are about half the size as those for NHW in HRS (49).

¹² These are unadjusted estimates. Adjustments for survival selection and sample weights lead to virtually indistinguishable estimates for the parameters of interest to us.

Fig. S1. We then evaluate the impact of changes in BMI (and probability of obesity) induced by GxE effects on probabilities of T2D and disability, a secondary consequence of both obesity and T2D.

Estimation of GxE first-order effects on BMI and obesity

We estimate a growth curve model (GCM) for BMI:

$$BMI_{it} = \alpha_0 + \alpha_1 cohort_t + \alpha_2 zPRS_i + \alpha_3 cohort_t * zPRS_i + \beta_{4i} age_{it} + \beta_{5i} age_{it}^2 + \beta'_6 \mathbf{X}_i + \mu_i + \epsilon_{it}, \quad (1)$$

where BMI_{it} is an individual's BMI in year t , $zPRS_i$ is the z-score of the PRS, $cohort_t$ is a categorical variable for the cohort born between 1945 and 1959, 1935 and 1944, 1925 and 1934, and before 1925, age_{it} is median-centered age at year t , and \mathbf{X}_i is a vector of control variables including education, gender, principal components, and an indicator of early conditions. We also include two sets of interactions terms: between cohort and all control variables and between PRS, and all control variables. The coefficients of age_{it} and age_{it}^2 vary across individuals, μ_i is the random intercept for each respondent, and ϵ_{it} is the error term.

We are interested in the change in BMI attributable to the interaction between PRS and in the youngest HRS birth cohort, namely,

$$\Delta BMI_i^{GxE} = \alpha_3^T * zPRS_i, \quad (2)$$

where ΔBMI_i^{GxE} is the excess BMI attributable to GxE among those born between 1945 and 1959, α_3^T is the true interaction effect (estimated from the model), and $zPRS_i$ is the PRS' z-score. There is, however, a minor problem with (2) because the scale of the measure of genetic predisposition is relevant when evaluating the size of the impact of GxE. By construction, $zPRS$ has a mean of 0 and a standard deviation of 1 and using it in equation (2) implies to set the impact of GxE to be 0 for individuals with average PRS. Because $\alpha_3 > 0$, those with PRS below average would have higher BMI if they were not exposed to the obesogenic environment. This is counterintuitive and a result of an arbitrary decision that assigns null effects at the mean of PRS. It is also inconsistent with empirical evidence suggesting that the impact of PRS on BMI follows a pattern predicted by a diathesis-stress model (DSM) (22, 25). Under this model, increased genetic penetrance is expected only in obesogenic environments (26). To circumvent the problem, we use a rescaled value, $rsPRS_i$, with a lowest value of 0

$$rsPRS_i = zPRS_i - \min(zPRS_i),$$

where $\min(zPRS)$ is the minimum of the PRS' z-scores. This transformation sets the impact of gene to zero for individuals with the lowest $zPRS$ but does not change the coefficient on the interaction term, α_3 in equation (1). Its only advantage is interpretational as the rescaled variable has a "natural" zero, i.e. a value of the genetic risk score that is inconsequential for the phenotype.¹³

To assess the first-order impact of GxE effects on BMI, we compute the counterfactual value BMI_i^{-GxE}

$$BMI_i^{-GxE} = BMI_{it} - \Delta BMI_i^{GxE} = BMI_{it} - \alpha_3^T * rsPRS_i, \quad (3)$$

where BMI_{it} is the *observed* BMI, $rsPRS_i$ is the rescaled PRS, and α_3^T is the *true* (unobserved but estimated) interaction effect. BMI_i^{-GxE} is the value of BMI that individuals born in 1945–1959 would have

¹³ Although the rescaling is consistent with the DSM, it can also be consistent with other models. We use it here not as a consequence of assuming ex ante that the DSM is true but rather of pragmatic considerations to improve interpretability.

attained if GxE effects were suppressed (see description in Section S1). Note that the counterfactual value is a function of the observed value of an individual's BMI and the unobserved (but estimated) GxE effect, α_3^T . It does not depend on other parameters of the model and is not affected by the rescaling of PRS. Since the values we use for the interaction effects are all positive, the observed value of BMI will always be larger than the counterfactual values.

Because the parameter α_3^T is unknown, computations of first- and second-order impacts, must account, at least partially, for the uncertainty of the estimate we choose. With the HRS data, we could generate multiple estimates depending on variable definitions and model specification. For example, our model includes a control for interaction between the PRS and education. Estimates are different in a model that ignores such interactions. Similarly, different estimates are produced by different definitions of E, e.g. birth cohorts. For our purposes at least, a conservative solution is to select the *largest* estimate from a handful of justifiable birth cohort definitions and model specifications. We choose α_3^T , the estimate associated with the youngest birth cohort when contrasted with the oldest. This choice will yield an upper bound for first- and second-order effects.

There are, however, two additional sources of uncertainty. First, the HRS is one of many data sets that could generate estimates of GxE effects. Second, even if one uses the HRS data, the value of the PRS score employed depends on the model specification to retrieve single nucleotide polymorphism (SNP)'s weights. The number of SNPs included in the PRS we use here is significantly larger than the number of SNPs than enter the definition of PRS employed by other studies based on HRS.

To partially account for this class of uncertainty, we conducted a thorough search of recent empirical studies and identified a small set of estimates comparable with ours (see Section SIII and Table S5). We chose three of these values spanning an interval that includes our HRS-based estimate as its (approximately) midpoint. Altogether, we use four alternative estimates of values of α_3^T , α_3^j ($j = 1, \dots, 4$).

We use these alternative estimates to compute BMI_{it}^{GxE} and then modify an individual's obesity status according to the counterfactual value. This will only change the obesity status of those whose observed BMI is larger but relatively close to 30 and, therefore, the population's prevalence of obesity. The difference between the counterfactual and observed prevalence is the first-order impact of the GxE on obesity.¹⁴

Estimation of GxE second-order effects on T2D and disability

To assess the magnitude of second-order effects on T2D and disability, we proceed in two steps. First, we use the HRS data to estimate hazard models for T2D and disability for ages 50–99 as defined in equations (4a) and (4b). The models include a vector of controls, Z, and a dummy variable O valued 1 for BMI ≥ 30 using parameters calculated from equations (4a) and (4b). We then compute two sets of predicted hazards for T2D and two for disability for each decile i of the zPRS distribution. The first sets are values for T2D computed from expression (4a) by replacing the dummy variable for obesity with the *observed fraction* of individuals who are obese in the ith decile of the PRS distribution and

the average values of Z_i's in the corresponding decile. These are the *observed* T2D's hazards for an average individual in those deciles and, of course, include the impact of the GxE effect postulated by our model for BMI. The second sets of rates for T2D are the *counterfactual* values. These are computed from expression (4a) by inserting the fraction of individuals in the corresponding decile that would be obese *according to the counterfactual values*, BMI_{i}^{-GxE} . These rates are those we would observe at each PRS's decile if there were no GxE effects. Analogous calculations using equation (4b) lead to observed and counterfactual values for the disability hazards. The two expressions we use are as follows:

$$\mu_i^{T2D}(t, O_i, Z_i) = \mu_0^{T2D}(t) \exp(\kappa O_i + \tau Z_i), \quad (4a)$$

$$\mu_i^{Disab}(t, O_i, Z_i) = \mu_0^{Disab}(t) \exp(\kappa' O_i + \tau' Z_i), \quad (4b)$$

where $\mu_i^{T2D}(t, O_i, Z_i)$ and $\mu_i^{Disab}(t, O_i, Z_i)$ are the T2D and disability hazards at age t and decile i, $\mu_0^{T2D}(t)$ and $\mu_0^{Disab}(t)$ are the corresponding baseline hazards, Z_i is a vector of covariates, O_i is the observed or counterfactual fraction of obese individuals in decile i, and finally, κ , τ , κ' , and τ' are parameters. Estimates from these models are displayed in Table S3.

A key quantity for health policy is duration of life after 50 that individuals will live with (or without) T2D and disability. To estimate this parameter, we compute multiple decrement life tables that simultaneously consider both the hazard of the event of interest (T2D or disability) and the competing event, mortality. We first estimate a hazard model for the mortality experienced by the HRS sample. This model includes dummy variables for T2D in expression (5a) and disability in expression (5b). Once the parameters are estimated, we compute mortality hazards for individuals with and without T2D/disability by including and excluding dummies for T2D/disability. The hazard models to estimate mortality hazard are as follows:

$$\mu_j^{death}(t, T2D_j, Z_j) = \mu_0^{death}(t) \exp(\theta T2D_j + \gamma Z_j), \quad (5a)$$

$$\mu_j^{death}(t, disab_j, Z_j) = \mu_0^{death}(t) \exp(\theta' disab_j + \gamma' Z_j), \quad (5b)$$

where $\mu_j^{death}(t, T2D_j, Z_j)$ and $\mu_j^{death}(t, disab_j, Z_j)$ are the mortality hazard for individual j in the presence of T2D and disability, T2D_j and disab_j are 0/1 dummy variables, Z_j is a vector of covariates, and θ , γ , θ' , and γ' are parameters. We use quantities computed with expressions (4a) and (4b) and standard life table procedures to construct single decrement life tables for T2D/disability that yield single decrement probabilities of surviving to age x > 50 without contracting T2D/disability. Estimates from these models are displayed in Table S4.

Second, we combine the hazards computed with expressions (4a) and (4b) and those from expressions (5a) and (5b) to construct standard multiple decrement life tables. From these, we retrieve statistics such as the probability of never contracting T2D/disability in the presence of mortality, the mean number of years lived with noT2D/disability in the presence of mortality, and the mean number of years lived after contracting T2D or becoming disable.¹⁵

¹⁵ An important caveat is needed here. The total impact of BMI (obesity) on the risk of T2D may well implicate components we are not including here. For example, if obese individuals with a higher genetic propensity to be obese are exposed to higher (lower) T2D risks than individuals with the same genetic propensity who are not obese, then there is an added second-order component we are not accounting for. Some literatures suggest that the sign of this component is negative, not positive (50, 51) and, consequently, it would reduce the total second order effects of the original GxE effects. Furthermore, and at least in the HRS, its magnitude is close to 0 (we thank a reviewer for pointing this possibility to us).

¹⁴ To avoid cluttering and throughout presentation of results, we compute counterfactuals using effects associated with only the youngest birth cohort, the only one exhibiting a significant increase in effects at given values of PRS.

Supplementary material

Supplementary material is available at PNAS Nexus online.

Funding

Research for this paper was funded by a European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 788582). We also acknowledge research support from the National Institute on Aging (<https://www.nia.nih.gov/>) and National Institute of Child Health and Development (<https://www.nichd.nih.gov/>) via the following project grants R01-AG016209, R03-AG015673, R01-AG018016, R37-AG025216, R01-AG056608, and R01-AG052030 and D43-TW001586, R24 HD047873, P30-AG-017266, and P2CHD041022. This manuscript was posted on a preprint: <https://www.researchsquare.com/article/rs-2022298/v1>.

Author contributions

A.P. and Y.H. formulated models and conducted the analysis; Y.H. prepared the data; A.P. and Y.H. in collaboration with H.B.-S. and M.M. wrote the manuscript; and all authors edited and reviewed the manuscript.

Data availability

The data underlying this article are publicly available through the University of Michigan Health and Retirement Study and available at <https://hrsdata.isr.umich.edu/data-products/public-survey-data>. Code to process the data and reproduce the main results and the figures is publicly available on a GitHub repository: https://github.com/hfyy/G_E_interaction.

References

- Obesity and overweight. World Health Organization. <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>.
- Flegal KM, Carroll MD, Kuczmarski RJ, Johnson CL. 1998. Overweight and obesity in the United States: prevalence and trends, 1960–1994. *Int J Obes Relat Metab Disord*. 22:39–47.
- Childhood Obesity in the United States: CDC Grand Rounds. 2023.
- Lake JK, Power C, Cole TJ. 1997. Child to adult body mass index in the 1958 British birth cohort: associations with parental obesity. *Arch Dis Child*. 77:376–381.
- Whitaker RC, Wright JA, Pepe MS, Seidel KD, Dietz WH. 1997. Predicting obesity in young adulthood from childhood and parental obesity. *N Engl J Med*. 337:869–873.
- Després J-P, Lemieux I. 2006. Abdominal obesity and metabolic syndrome. *Nature* 444:881–887.
- Mehta NK, Chang VW. 2009. Mortality attributable to obesity among middle-aged adults in the United States. *Demography* 46: 851–872.
- Mehta NK, Chang VW. 2011. Secular declines in the association between obesity and mortality in the United States. *Popul Dev Rev*. 37:435–451.
- Stewart ST, Cutler DM, Rosen AB. 2009. Forecasting the effects of obesity and smoking on U.S. life expectancy. *N Engl J Med*. 361: 2252–2260.
- Kim DD, Basu A. 2016. Estimating the medical care costs of obesity in the United States: systematic review, meta-analysis, and empirical analysis. *Value Health*. 19:602–613.
- Finkelstein EA, Trogon JG, Cohen JW, Dietz W. 2009. Annual medical spending attributable to obesity: payer- and service-specific estimates: amid calls for health reform, real cost savings are more likely to be achieved through reducing obesity and related risk factors. *Health Aff*. 28:w822–w831.
- Swinburn BA, et al. 2019. The global syndemic of obesity, undernutrition, and climate change: the Lancet Commission report. *Lancet* 393:791–846.
- Popkin BM, Reardon T. 2018. Obesity and the food system transformation in Latin America: obesity and food system transformation. *Obes Rev*. 19:1028–1064.
- Popkin BM, Corvalan C, Grummer-Strawn LM. 2020. Dynamics of the double burden of malnutrition and the changing nutrition reality. *Lancet* 395:65–74.
- Farooqi IS. 2000. Recent advances: recent advances in the genetics of severe childhood obesity. *Arch Dis Child*. 83:31–34.
- Willyard C. 2014. The family roots of obesity. *Nature* 508:S58–S60.
- Cheng M, et al. 2018. Computational analyses of obesity associated loci generated by genome-wide association studies. *PLoS One* 13:e0199987.
- Wang K, et al. 2011. A genome-wide association study on obesity and obesity-related traits. *PLoS One* 6:e18939.
- Drong AW, Lindgren CM, McCarthy MI. 2012. The genetic and epigenetic basis of type 2 diabetes and obesity. *Clin Pharmacol Ther*. 92:707–715.
- Goodarzi MO. 2018. Genetics of obesity: what genetic association studies have taught us about the biology of obesity and its complications. *Lancet Diabetes Endocrinol*. 6:223–236.
- Daza S, Palloni A. 2022. Modeling the impact of heritability, assortative mating and fertility on population-level obesity trends. *SocArXiv bmjzk*. <https://doi.org/10.31235/osf.io/bmjzk>, preprint: not peer reviewed.
- Conley D, Laidley TM, Boardman JD, Domingue BW. 2016. Changing polygenic penetrance on phenotypes in the 20th century among adults in the US population. *Sci Rep*. 6:30348.
- Domingue BW, Conley D, Fletcher J, Boardman JD. 2016. Cohort effects in the genetic influence on smoking. *Behav Genet*. 46: 31–42.
- Guo G, Liu H, Wang L, Shen H, Hu W. 2015. The genome-wide influence on human BMI depends on physical activity, life course, and historical period. *Demography* 52:1651–1670.
- Walter S, Mejía-Guevara I, Estrada K, Liu SY, Glymour MM. 2016. Association of a genetic risk score with body mass index across different birth cohorts. *JAMA* 316:63.
- Belsky J, et al. 2009. Vulnerability genes or plasticity genes? *Mol Psychiatry*. 14:746–754.
- Belsky J, Beaver KM. 2011. Cumulative-genetic plasticity, parenting and adolescent self-regulation. *J Child Psychol Psychiatry*. 52: 619–626.
- Belsky J, Pluess M. 2009. Beyond diathesis stress: differential susceptibility to environmental influences. *Psychol Bull*. 135:885–908.
- Boardman JD, Daw J, Freese J. 2013. Defining the environment in gene–environment research: lessons from social epidemiology. *Am J Public Health*. 103:S64–S72.
- Boardman JD, et al. 2014. Is the gene–environment interaction paradigm relevant to genome-wide studies? The case of education and body mass index. *Demography* 51:119–139.
- Burt A. 2011. Some key issues in the study of gene–environment interplay: activation, deactivation, and the role of development. *Res Hum Dev*. 8:192–210.

- 32 Harden KP. 2021. *The genetic lottery: why DNA matters for social equality*. Princeton (NJ): Princeton University Press.
- 33 Lewontin RC. 2006. The analysis of variance and the analysis of causes. *Int J Epidemiol*. 35:520–525.
- 34 Coop G. 2020. Reading Tea leaves? Polygenic scores and differences in traits among groups. *arXiv* 00892. <https://doi.org/10.48550/arXiv.1909.00892>, preprint: not peer reviewed.
- 35 Fox RJ, Donelson JM, Schunter C, Ravasi T, Gaitán-Espitia JD. 2019. Beyond buying time: the role of plasticity in phenotypic adaptation to rapid environmental change. *Philos Trans R Soc Lond B Biol Sci*. 374:20180174.
- 36 Harpak A, Przeworski M. 2021. The evolution of group differences in changing environments. *PLoS Biol*. 19:e3001072.
- 37 Saltz JB, et al. 2018. Why does the magnitude of genotype-by-environment interaction vary? *Ecol Evol*. 8:6342–6353.
- 38 Daza S, Palloni A. 2022. Distinguishing between interaction and dispersion effects in GxE analysis: a review of strategies. Working Paper.
- 39 Feldman M, Lewontin R. 1975. The heritability hang-up. *Science* 190:1163–1168.
- 40 Manski CF. 2011. Genes, eyeglasses, and social policy. *J Econ Perspect*. 25:83–94.
- 41 Caspi A, et al. 2002. Role of genotype in the cycle of violence in maltreated children. *Science* 297:851–854.
- 42 Tuvblad C, Grann M, Lichtenstein P. 2006. Heritability for adolescent antisocial behavior differs with socioeconomic status: gene-environment interaction. *J Child Psychol Psychiatry*. 47:734–743.
- 43 Mills MC, et al. 2021. Identification of 371 genetic variants for age at first sex and birth linked to externalising behaviour. *Nat Hum Behav*. 5:1717–1730.
- 44 Harden KP, Mendle J, Hill JE, Turkheimer E, Emery RE. 2008. Rethinking timing of first sex and delinquency. *J Youth Adolescence*. 37:373–385.
- 45 Harden KP. 2014. Genetic influences on adolescent sexual behavior: why genes matter for environmentally oriented researchers. *Psychol Bull*. 140:434–465.
- 46 Tommerup K, Ajnakina O, Steptoe A. 2021. Genetic propensity for obesity, socioeconomic position, and trajectories of body mass index in older adults. *Sci Rep*. 11:20276.
- 47 Health and Retirement Study. 2023. Public use dataset and polygenic score data. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Ann Arbor (MI).
- 48 Boardman JD, Blalock CL, Pampel FC. 2010. Trends in the genetic influences on smoking. *J Health Soc Behav*. 51:108–123.
- 49 Duncan L, et al. 2019. Analysis of polygenic risk score usage and performance in diverse human populations. *Nat Commun*. 10:3328.
- 50 Davidson T, Vinneau-Palarino J, Goode JA, Boardman JD. 2021. Utilizing genome wide data to highlight the social behavioral pathways to health: the case of obesity and cardiovascular health among older adults. *Soc Sci Med*. 273:113766.
- 51 Vinneau JM, Huibregtse BM, Laidley TM, Goode JA, Boardman JD. 2021. Mortality and obesity among U.S. older adults: the role of polygenic risk. *J Gerontol B Psychol Sci Soc Sci*. 76:343–347.