Interpretation and potential biases of Mendelian randomization estimates with time-varying exposures

Jeremy A. Labrecque and Sonja A. Swanson

Correspondence to Dr. Jeremy A Labrecque, Department of Epidemiology, Erasmus MC, P. O. Box 2040, 3000 CA Rotterdam, The Netherlands (j.labrecque@erasmusmc.nl)

Author affiliations: Department of Epidemiology, Erasmus MC, P. O. Box 2040, 3000 CA Rotterdam, The Netherlands (Jeremy A. Labrecque, Sonja A. Swanson)

Financial support: SAS is supported by a NWO/ZonMW Veni grant [91617066]

Conflict of interest: None declared.

Running head: Mendelian randomization and time-varying exposures

ABSTRACT

Mendelian randomization (MR) is used to answer a variety of epidemiologic questions. One stated advantage of MR is that it estimates a "lifetime effect" of exposure though this term remains vaguely-defined. Instrumental variable analysis, on which MR is based, has focused on estimating the effects of point or time-fixed exposures rather than "lifetime effects". We use an empirical example with data from the Rotterdam Study to demonstrate how confusion can arise when estimating "lifetime effects". We provide one possible definition of a lifetime effect: the average change in outcome measured at time *t* when the entire exposure trajectory from conception to time *t* is shifted by one unit. We show that MR only estimates this type of lifetime effect under specific conditions, for example when the effect of the genetic variants used on exposure do not change over time (which many genetic variants commonly used in MR do). Lastly, we simulate the magnitude of bias that would result in realistic scenarios that use genetic variants with effects that change over time. We recommend future MR studies carefully consider the effect of interest and how genetic variants whose effects change with time may impact the interpretability and validity of their results.

Keywords: Mendelian randomization analysis, epidemiologic methods, bias, longitudinal studies

Abbreviations

MR: Mendelian randomization

IV: instrumental variable

BMI: body mass index

Introduction

Mendelian randomization (MR) uses genetic variation as a proposed instrumental variable to estimate the effect of an exposure on an outcome (1). The increasing availability of genetic data and the perception that the assumptions required for valid causal inference with MR are more plausible than those required for traditional methods have contributed to its popularity.

To estimate the effect of an exposure, MR relies on classical (2) or more robust versions (3) of instrumental variable analyses (IV). However, instrumental variable analyses have been developed to estimate the effect of exposure at one point in time while MR is often interpreted as a longitudinal or lifetime effect of a time-varying exposure (1, 4-6). The rationale behind this difference in interpretation is likely that genetic effects begin at conception and are present throughout the lifetime mimicking, to some extent, a longitudinal intervention. The validity of this interpretation of MR as a lifetime effect has, to our knowledge, never been statistically justified with time-varying exposures.

The lifetime effect interpretation commonly seen in MR studies also lacks clarity (7). To be a well-defined causal construct, a lifetime effect requires specification of a time frame and a comparison of treatment regimens (8). Not only is the lifetime effect interpretation used in MR ambiguous, but a recent review of MR contributions pointed out that even an ambiguous interpretation would have many exceptions depending on the biological context (9). A less ambiguous, well-defined causal definition of the MR parameter would make its clinical importance clearer and help identify when MR approaches would be valid.

Here, we explore issues related to MR with time-varying exposures and the estimation of lifetime effects. We provide a definition of a lifetime effect and derive the conditions under which common

approaches to MR estimates can be interpreted in this way. We begin with a simple empirical example to demonstrate a case where confusion arises in the interpretation of a MR study.

An empirical example of the difficulty of MR with time-varying exposures

Suppose our goal is to estimate "the" lifetime effect of body mass index (BMI) on systolic blood pressure using the rs9939609 variant within the FTO locus as a proposed genetic instrument. By proposing this genetic variant as an instrument, we (and other investigators who have proposed this instrument) are assuming that the genetic variant is associated with BMI, has no effect on blood pressure except through its effect on BMI, and shares no causes with blood pressure. Additionally, to obtain an average treatment effect, we further assume effect homogeneity (10).

To estimate this effect, we used data from the Rotterdam Study, a prospective cohort of people 55 years or older living in Rotterdam (11). The medical ethics committee of the Erasmus University of Rotterdam approved the study, and informed consent was obtained from all participants. We included all participants who contributed to follow-up visits between 2009 and 2013 ($n = 5123$). The MR estimate, estimated via IV analysis, can be obtained by dividing the per allele effect of FTO on systolic blood pressure by the per allele effect of FTO on BMI. Each high-risk allele was associated with an increase of 0.70mmHg (95% confidence interval (CI): -0.16, 1.57) in systolic blood pressure and an increase of 0.32kg/m^2 (95%CI: 0.15, 0.49) in BMI. Therefore, the IV estimate of the effect of BMI on systolic blood pressure is $0.70/0.32 = 2.19$ mmHg per kg/m^2 (95% CI: -0.65, 5.04).

Previous literature suggests, however, that the relationship between FTO and BMI changes with age (12–14). This time-varying relationship is also observed in the Rotterdam study (Figure 1). The per allele effect estimate at age 55 years is over 1kg/m^2 (with wide confidence intervals)

while the estimate at age 75 is nearly null. Therefore, measuring the exposure at different ages would result in very different IV estimates because of changes in the denominator of the IV estimate even when the numerator remains the same. If BMI was measured at age 55 in all participants the IV estimate would therefore be around $0.70/1 = 0.70\text{mmHg/kg/m}^2$. If BMI was measured at age 75, the denominator of the IV estimate would be very small, let us say 0.1 kg/m^2 and therefore the estimate would be $7.0\ \text{mmHg per kg/m}^2$ found previously. Given that very different estimates can be obtained by measuring the exposure at different times, several questions arise: what 'lifetime' effects are we actually estimating, and under what assumptions would one or both of these estimates be valid? From a public health perspective, addressing these questions is essential for these numeric estimates to be informative, as it would be important to discern whether changes in BMI would result in small or substantial changes in blood pressure.

This simple example brings up many important questions about using MR with time-varying exposures. In the following sections we will explore how this can complicate MR analyses, change the interpretation of MR estimates and under what specific conditions MR estimates can be validly interpreted as well defined effects.

A clearer definition of lifetime effects

A clearer definition of what is meant by *lifetime effects* in the context of MR is required in order to determine whether common MR analyses are in fact estimating lifetime effects. The most common definition in the MR literature is "the effects of long-term differences in exposures on disease risk,"(1) but this lacks the clarity of a properly defined causal parameter (8). Does *long-term* necessarily refer to the effect from conception or can it also refer to a shorter time period? Do long-term differences refer to differences in average lifetime exposure, differences induced by

intervening on exposure at conception or maintaining a fixed difference in exposure at all time points?

We will consider the canonical IV causal diagram with a single binary genetic variant G, continuous exposure A, continuous outcome Y and unmeasured confounders U (Figure 2A). We will use counterfactual notation such that Y_k^a represents the outcome that would have been observed at age k had A been set to a. Using this notation the effect of increasing a time-fixed exposure by one unit at age k would be $E[Y_k^{(a+1)}] - E[Y_k^a]$. In the context of MR with a time-fixed exposure (*i.e.* an exposure whose level is fixed at conception such as eye color or blood type) this can be interpreted as the effect of increasing A by one unit over the entire lifetime on Y at age k.

With a time-varying exposure, A is not a single value but a vector $\overline{A}_k = (A_0, A_1, \ldots, A_k)$ representing the value of A at each time point between conception ($k = 0$) and time k when the outcome is measured (Figure 2B). This vector can be thought of as the trajectory taken by A similar to the allele specific BMI trajectories depicted in Figure 1. Therefore, we propose the following definition for a lifetime effect of a time-varying exposure in the context of MR: the effect of shifting the entire exposure trajectory ($\overline{A}$) by one unit on Y at time k. Another way to state this is the effect of increasing the exposure by one unit at every point in time. This particular definition of the lifetime effect can be written in counterfactual notation as $E[Y_k^{\overline{a+1}}] - E[Y_k^{\overline{a}}]$.

Note the necessity of specifying the time at which the outcome is measured. For any exposure with cumulative effects starting at conception, the lifetime effects must be time or age-specific because older people will be exposed for longer periods of time. This means cumulative effects must always be heterogeneous across age. For some types of non-cumulative exposures or cumulative

exposures over shorter time periods, it is possible for lifetime effects to be constant across different age groups.

A simple example of MR with a time-varying exposure

Applying the above-defined definition of lifetime effects to a simple example with two time points will clarifies what is happening in the previous empirical example. We assume the causal structure in Figure 3 with binary genetic variant G, continuous exposure A_k which can take different values at $k = 0$ and $k = 1$, and outcome Y. The γ above each edge represents the causal effect of one variable on another, which, for simplicity, we assume is linear and constant for all causal relationships depicted. We have chosen two time points for simplicity but in Web Appendix 1 we extend this logic to continuous time.

The lifetime effect of increasing the exposure trajectory $\overline{A}$ by one unit in this example is simply the sum of the effects of A_0 and A_1 on Y_k: $\gamma_4 + \gamma_5$. The effect of G on Y, the numerator of the IV estimator, is the sum of the three possible pathways between them: $\gamma_1 * \gamma_4 + \gamma_1 * \gamma_3 * \gamma_5 + \gamma_2 * \gamma_5$. The effect of G on A can take two values in this example depending on when A was measured. At time 0, the effect of G on A_0 is γ_1 and at time 1, the effect of G on A_1 is $\gamma_1 * \gamma_3 + \gamma_2$. As such, there are at least two different IV estimates we may consider computing, by dividing the effect of G on Y by the effect of G on A at each time point:

$$
\begin{aligned}
MR_0 &= \frac{\gamma_1 * \gamma_4 + \gamma_1 * \gamma_3 * \gamma_5 + \gamma_2 * \gamma_5}{\gamma_1} \\
&= \gamma_4 + \gamma_5(\gamma_3 + \frac{\gamma_2}{\gamma_1}) \\
MR_1 &= \frac{\gamma_1 * \gamma_4 + \gamma_1 * \gamma_3 * \gamma_5 + \gamma_2 * \gamma_5}{\gamma_1 * \gamma_3 + \gamma_2} \\
&= \gamma_4(\frac{\gamma_1}{\gamma_1 * \gamma_3 + \gamma_2}) + \gamma_5
\end{aligned}
$$

For both MR_0 and MR_1, the terms inside the parentheses bias the estimates away from the true value. If we modify our example so the genetic effect on exposure is constant over time, the bias is eliminated. More specifically, if the genetic effect is constant over time then the effect of G on A_0 (γ_1) is equal to the effect of G on A_1 ($\gamma_1 * \gamma_3 + \gamma_2$). This equality can be rewritten as $\gamma_3 = 1 - \frac{\gamma_2}{\gamma_1}$ and substituted into the equation for MR_0:

$$
\begin{aligned}
MR_0 &= \gamma_4 + \gamma_5(1 - \frac{\gamma_2}{\gamma_1} + \frac{\gamma_2}{\gamma_1}) \\
&= \gamma_4 + \gamma_5
\end{aligned}
$$

We can do the same with MR_1 substituting γ_1 for $\gamma_1 * \gamma_3 + \gamma_2$:

$$
\begin{aligned}
MR_1 &= (\frac{\gamma_1}{\gamma_1}) * \gamma_4 + \gamma_5 \\
&= \gamma_4 + \gamma_5
\end{aligned}
$$

Therefore, even when A is time-varying, the IV estimate using either time point could potentially be a valid estimate of the lifetime effect of A on Y when the relationship between G and A is constant through time. However, when the effect of G on A changes over time, the IV estimate will be a biased estimate of the lifetime effect. The intuition behind this result is simple: if the relationship between G and A changes over time, it cannot be adequately summarized by measuring it at one time point.

MR and time-varying exposures: a simulation

To learn about the magnitude and direction of bias with realistic parameters, we performed analytic derivations for the following series of simulated scenarios. We simulate a longitudinal relationship between a binary genetic variant G, an exposure A and an outcome Y running the simulations to age 30 and age 50, with relationships between variables informed in part by prior literature on FTO and BMI (14). e used four different gene-exposure relationships depicted in each column of Figure 4 where the solid line is the per allele effect at different points in time. The first column represents a constant genetic effect where the per allele effect of G is constant over time. The second and third columns represent genetic variants whose effects increase and decrease over time respectively. The last column roughly emulates the relationship between FTO and BMI where the largest per allele difference occurs around age 20 and decreases thereafter. Where the per allele effect varies with time, the maximum change was $0.5\ kg/m^2$ to match a conservative estimate of the greatest change observed in the effect of FTO on BMI (15).

Each row in Figure 4 represents one of the four different exposure window scenarios we used. Here, an exposure window is defined as the time period during which the exposure is relevant to the etiology of the outcome (16). In other words, only changes in exposure during the exposure window will affect the outcome holding exposure at other time points constant. The dashed lines represent the instantaneous effect of exposure at different points in time on Y_K. The first row represents a pure cumulative effect where exposure at any time point has the same effect on Y_K. In the second row, the effect of recent exposure is stronger and is roughly limited to thirty years before K though most of the effect comes from the last ten years. Note that in the recent exposure scenario, the exposure window is relative to the age when the outcome is measured. The dashed gray line, therefore, represents the exposure window when the outcome is measured at age 30. The

third row represents a critical exposure window where only exposure around age 13 (i.e., slightly before the available initial measurement of exposure) affects Y_K. The last row is a simple increasing exposure window where later exposure has a larger effect on Y_K. All exposure windows were scaled to make the lifetime effect at age 50 equal to 2 and the lifetime effect at 30 was calculated by taking the area under the solid line in Figure 4 from birth to age 30. Altogether, we selected these scenarios to represent some biologically plausible relationships between FTO, BMI through mid-life, and blood pressure (15) while nonetheless relevant for an arbitrary MR genetic variant, exposure, and outcome combination.

The IV numerator, the effect of G on Y_K, was calculated by multiplying the genetic effect on exposure (the dashed line) by the effect of exposure on the outcome (the solid line) and taking the area under the resulting curve from birth to age 30 or 50. The denominator of the IV estimate was the value of the difference in BMI at age 30 or 50 (the dashed line in Figure 4). The final IV estimate was obtained by dividing the IV numerator by the denominator. The code for these derivations can be found in Web Appendix 2.

When the instrument strength was constant over time, the estimates were unbiased with respect to the lifetime effect at both ages regardless of the shape of the exposure window (Table 1). The estimate is biased in all other scenarios and the bias was sensitive to not only the type of relationship between the genetic variant and the exposure but also the exposure window and age when the outcome was measured. For example, in the FTO scenario, the IV analysis underestimates the true effect at age 30 but overestimates the effect at age 50.

When the G-A relationship changes with time, the bias is minimized to the degree that the measured G-A relationship is a good summary of the average G-A relationship within the exposure window.

For example, in the scenario with the lowest bias, the FTO G-A scenario with the recent exposure window at age 30, the G-A relationship at age 30 is relatively close to the average G-A difference in the most important part of the exposure window. In the scenarios with the highest bias, such as the FTO G-A scenario and a critical exposure window at age 50, the G-A relationship when measured at age 50 is much smaller than the level in the exposure window.

How common is it for genetic effects to vary with age?

Despite the widespread use of genome wide association studies, few have investigated how genetic effects change with age. This is likely because genome wide association studies require large sample sizes to achieve adequate power and investigating effect modification by age would require even larger sample sizes. One recent, large meta-analysis of genome wide association studies of BMI studies found 15 loci that demonstrated different effects by age including FTO and MC4R which are commonly used in MR studies of the effect of BMI (17). A variety of other studies have also found effect modification by age for the effect of genetic variants on BMI (13–15). The same phenomenon has been observed with other phenotypes such as LDL cholesterol (18,19), Alzheimer's Disease (20) and blood pressure (21–23).

For many genetic variants used in MR studies, there has been no investigation into whether age modifies the relationship between the genetic variant and the exposure. For some complex phenotypes we can infer effect modification by age. For example, consider ALDH2 and alcohol consumption. The ALHD2 genetic variant cannot have any effect on alcohol consumption until a person initiates alcohol consumption. Therefore, the effect of ALHD2 on alcohol consumption changes from zero before a person initiates alcohol consumption to a non-null effect afterward. If the effect of ALHD2 on alcohol consumption were constant with age after initiation, the MR estimate would correspond with a lifetime effect starting at initiation rather than conception.

However, people homozygous for the *2 allele are almost all never drinkers (24),therefore, for the effect of ALDH2 to be constant over time, people with at least one *1 allele would have to drink the same amount regardless of their age. Alcohol consumption peaks around age 20 and declines with age (25) meaning that the effect of ALDH2 on drinking necessarily decreases with age. In Web Appendix 3, we also demonstrate that genetic effects are almost guaranteed to change with time when there are bidirectional effects between two variables. Therefore, when there are bidirectional effects between two variables, including the more general and common situation of treatment-confounder feedback. Therefore, when there are bidirectional effects between two variables, the effect estimate will be biased but the test of whether one variable causes in the other will be valid in both directions. .

Discussion

MR is still a relatively new method having risen in popularity only in the last 20 years (1). The method has spread rapidly from exposures that are direct gene products (e.g., proteins) to complex phenotypes (e.g., BMI, alcohol consumption). While MR continues to develop new robust methods to address certain violations of the instrumental variable assumptions, little attention has been paid to issues that differentiate MR from classical IV analysis such as the estimation of lifetime effects and the (often unacknowledged) use of time-varying exposures.

We have demonstrated that genetic variants whose genetic effect on the exposure of interest changes over time cannot typically be used to validly estimate lifetime effects with MR. We show this by deriving a bias formula in a simple case, deriving bias in several hypothetical scenarios, and computing inconsistent estimates in an empirical example. Indeed, when the genetic effect on exposure changes with time, the MR estimate does not intuitively correspond with any causal parameter and certainly is not a valid estimate of the lifetime effect defined here. Our results, and

the potential ubiquity of such genetic variants, calls into question the numeric estimates from many MR studies. Moreover, our results corroborate the intuition of prior skeptics that variations over the life-course complicate if not invalidate MR analyses (26). Although not discussed here, one silver lining to this conclusion is that, even if the numeric results are not interpretable due to time-varying genetic effects, such MR studies can still potentially provide a valid test of certain causal null hypotheses (27).

Whenever the goal of an MR study is to estimate a lifetime effect, investigators should ideally limit their analyses to genetic variants whose effects on exposure are constant over time or carefully consider how the time-varying genetic effect might impact the effect estimate. In many contexts this is verifiable and can be achieved by looking for a statistical interaction between the genetic variant, age and the exposure or by plotting the relationship between exposure and age stratified by allele in samples with sufficient variation in age. One possible sensitivity analysis for the importance of variation in effect of a genetic variant would be to calculate the MR estimate with more than one of the possible denominator estimates, as we did in our empirical example based on the Rotterdam Study data. If the MR estimate is not sensitive to the differences, the investigators may perhaps argue their conclusions are unlikely to be impacted. Note, of course, such sensitivity analyses depend upon rich longitudinal data. Investigators may also consider adapting our provided code to investigate other hypothetical scenarios that map onto their beliefs about their particular genetic variant, exposure, and outcome relationships as a means for investigating possible magnitudes of bias.

We have proposed one definition of a lifetime effect in MR but certainly other interpretations could be given. One possible alternative is the effect of a point intervention at conception. Our interpretation and this interpretation are equivalent with time-fixed exposures or genetic variants

whose relationship to the exposure does not change with time. We demonstrate in Web Appendix 4 that estimating this effect usually requires that the genetic variant affect the exposure only at conception and at no future time point, which is biologically implausible in most MR settings. Another possible reading of the prior literature is that MR estimates the effect of an *average* of a one-unit change in exposure across the lifetime, as opposed to the exact one-unit change at all time points proposed in the current paper. In addition to the reasons our proposed definition can be problematic, however, this definition is further complicated by multiple versions of treatment: except in the case when the exposure's effect on the outcome is purely cumulative, the various versions of possible interventions would not necessarily result in the same effect size. More generally, this and recent work (7) suggests the importance of MR investigators providing unambiguous interpretations (and the assumptions upon which such interpretations rest) of their numeric estimates: if the investigators suggest they are not trying to estimate the lifetime effect as defined in the current paper, they should clarify what causal parameter they indeed are trying to estimate instead.

In conclusion, we have provided a definition of lifetime effects that can be used in MR studies. In so doing, we also demonstrate that many genetic variants used in MR studies cannot be leveraged to validly estimate this definition of lifetime effect. We recommend future MR studies carefully consider the effect of interest and how genetic variants whose effects change with time may impact the interpretability and validity of their results.

Acknowledgements

Author affiliations: Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands (Jeremy A Labrecque and Sonja A Swanson). Dr. Swanson is supported by a NWO/ZonMW Veni grant [91617066]. Conflicts of interest: none declared.

References

1. Smith GD, Ebrahim S. 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease? *International Journal of Epidemiology*. 2003;32(1):1–22.

2. Lawlor DA, Harbord RM, Sterne JAC, et al. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Statistics in Medicine*. 2008;27(8):1133–1163.

3. Bowden J, Smith GD, Burgess S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *International Journal of Epidemiology*. 2015;44(2):512–525.

4. Nitsch D, Molokhia M, Smeeth L, DeStavola BL, Whittaker JC, Leon DA. Limits to Causal Inference based on Mendelian Randomization: A Comparison with Randomized Controlled Trials. *American Journal of Epidemiology.* 2006;163(5):397-403.

5. Lawlor DA. Commentary: Two-sample Mendelian randomization: opportunities and challenges. *International Journal of Epidemiology.* 2016;45(3):908-915.

6. Ference BA. How to use Mendelian randomization to anticipate the results of randomized trials. *European Heart Journal*. 2017;39(5):360-362.

7. Swanson SA, Tiemeier H, Ikram MA, et al. Nature as a Trialist? *Epidemiology*. 2017;28(5):653–659.

8. Hernán MA, Robins JM. Causal inference. Boca Raton, FL: Chapman &Hall/CRC; 2018.

9. Holmes MV, Ala-Korpela M, Smith GD. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. *Nature Reviews Cardiology*. 2017;14(10):577–590.

10. Hernán MA, Robins JM. Instruments for causal inference: An epidemiologist's dream? *Epidemiology*. 2006;17(4):360–72.

11. Ikram MA, Brusselle GGO, Murad SD, et al. The Rotterdam Study: 2018 update on objectives, design and main results. *European Journal of Epidemiology*. 2017;(2017):1–44.

12. Rosenquist JN, Lehrer SF, O'Malley AJ, et al. Cohort of birth modifies the association between FTO genotype and BMI. *Proceedings of the National Academy of Sciences*. 2015;112(2):354–359.

13. Qi L, Kang K, Zhang C, et al. Fat Mass-and Obesity-Associated (FTO) Gene Variant Is Associated With Obesity: Longitudinal Analyses in Two Cohort Studies and Functional Test. *Diabetes*. 2008;57(11):3145–3151.

14. Sovio U, Mook-Kanamori DO, Warrington NM, et al. Association between common variation at the FTO locus and changes in body mass index from infancy to late childhood: The complex nature of genetic association through growth and development. *PLoS Genetics*. 2011;7(2).

15. Hardy R, Wills AK, Wong A, et al. Life course variations in the associations between FTO and MC4R gene variants and body size. *Human Molecular Genetics*. 2009;19(3):545–552.

16. Savitz DA. Interpreting Epidemiologic Evidence: Strategy for Study Design and Analysis. Oxford, UK: Oxford University Press; 2003.

17. Winkler TW, Justice AE, Graff M, et al. The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. *PLOS Genetics*. 2015;11(10):e1005378.

18. Dumitrescu L, Brown-Gentry K, Goodloe R, et al. Evidence for Age As a Modifier of Genetic Associations for Lipid Levels. *Annals of Human Genetics*. 2011;75(5):589–597.

19. Shirts B, Hasstedt S, Hopkins P, et al. Evaluation of the gene-age interactions in HDL cholesterol, LDL cholesterol, and triglyceride levels: The impact of the SORT1 polymorphism on LDL cholesterol levels is age dependent. *Atherosclerosis*. 2011;217(1):139–141.

20. Neu SC, Pa J, Kukull W, et al. Apolipoprotein E Genotype and Sex Risk Factors for Alzheimer Disease. *JAMA Neurology*. 2017;74(10):E1–E12.

21. Simino J, Shi G, Bis JC, et al. Gene-age interactions in blood pressure regulation: A large-scale investigation with the CHARGE, global BPgen, and ICBP consortia. *American Journal of Human Genetics*. 2014;95(1):24–38.

22. Tambs K, Eaves L, Mourn T, et al. Age-specific genetic effects for blood pressure. *Hypertension*. 1993;22(5):789–795.

23. Shi G, Gu CC, Kraja AT. Genetic Effect on Blood Pressure is Modulated by Age: The HyperGEN Study. 2010;23(2):83–88.

24. Chen L, Smith GD, Harbord RM, et al. Alcohol intake and blood pressure: A systematic review implementing a mendelian randomization approach. *PLoS Medicine*. 2008;5(3):0461–0471.

25. Substance Abuse and Mental Health Services Administration. Results from the 2010 National Survey on Drug Use and Health: Summary of National Findings. NSDUH Series H-41, HHS Publication No. (SMA) 11-4658. Rockville, MD: Substance Abuse; Mental Health Services Administration: Substance Abuse; Mental Health Services Administration; 2011. https://www.samhsa.gov/data/sites/default/files/NSDUHNationalFindingsResults2010-web/2k10ResultsRev/NSDUHresultsRev2010.pdf (accessed November, 2017).

26. Schooling CM, Yeung SL, Freeman G. Mendelian randomization estimates may be inflated. *Journal of the American College of Cardiology*. 2013;61(18):1931.

27. Swanson SA, Labrecque JA, Hernan M. Causal null hypotheses of sustained treatment strategies: What can be tested with an instrumental variable? *European Journal of Epidemiology*.2018 (in press).

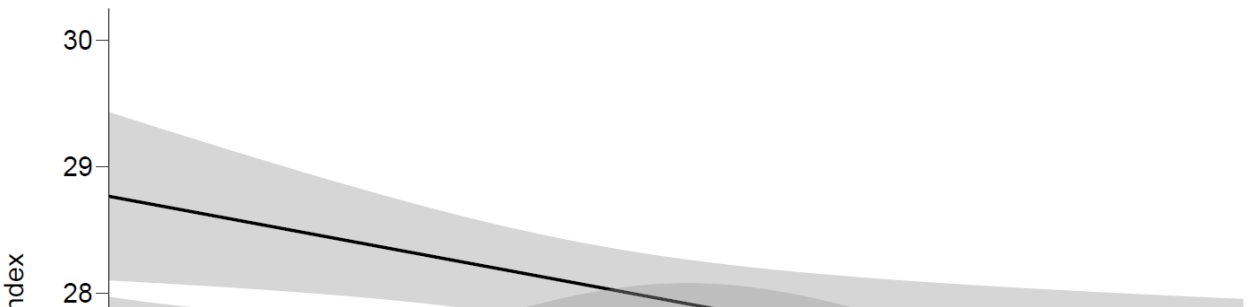

Figure 1—The relationship between body mass index and age by FTO allele in participants from the Rotterdam Study followed-up between 2009 and 2013 (n=5123) estimated using splines with five knots. The shaded regions are 95% confidence bands.

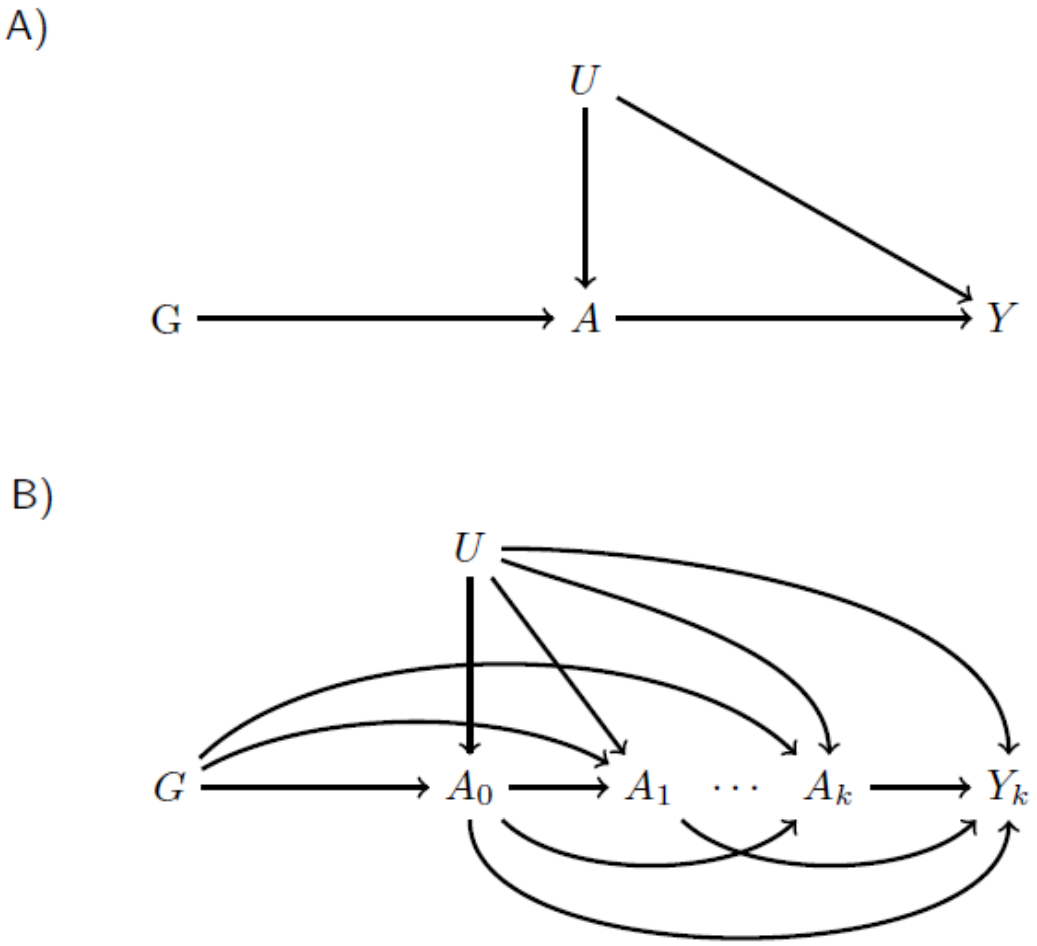

Figure 2—Causal graphs depicting the relationship between a genetic variant (G), an exposure (A) and an outcome (Y) when A) the exposure is time-fixed and B) the exposure is time-varying.

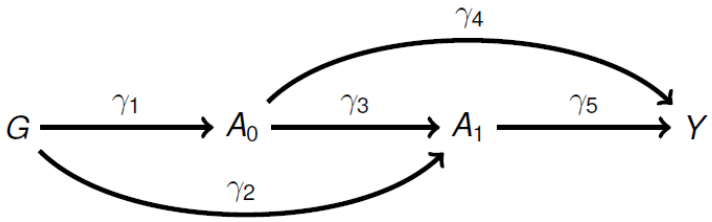

Figure 3—A causal graph depicting the relationship between a binary genetic variant (G) , an exposure (A) measured at two time points and an outcome (Y). The γ_n above the edges represent the parametric relationship between the variables joined by the edge.

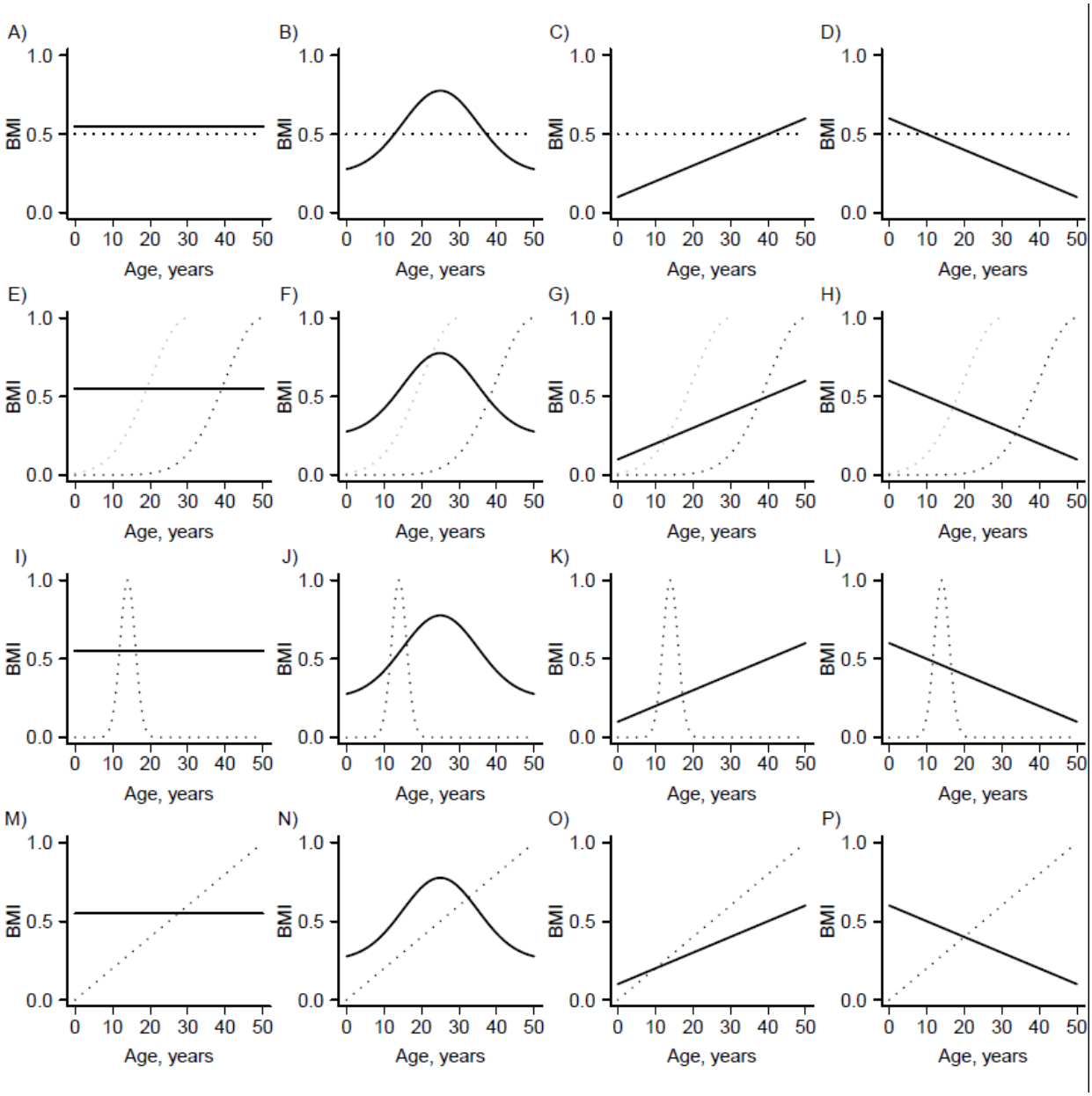

Figure 4—16 hypothetical scenarios used to investigate possible magnitudes of bias in IV analyses. The solid line shows the relationship between the genetic variant and body mass index (BMI) and the dashed line shows the relevant exposure window. For the recent exposure scenario, the gray dashed line shows the relevant exposure window at age 30. The dashed lines are only shown to demonstrate the shape of the exposure window and are not to scale.

Table 1. Results from the 16 hypothetical scenarios described in Figure 4 comparing the true lifetime effect with the MR estimate when the instrument strength varies over time.

	Age 30				Age 50			
Exposure window	**True effect**	**MR estimate**	**Absolute bias**	**Relative bias (%)**	**True effect**	**MR estimate**	**Absolute bias**	**Relative bias (%)**
Constant genetic scenario								
Uniform	1.2	1.2	0.0	0	2.0	2.0	0.0	0
Recent	2.0	2.0	0.0	0	2.0	2.0	0.0	0
Critical	2.0	2.0	0.0	0	2.0	2.0	0.0	0
Increasing	0.7	0.7	0.0	0	2.0	2.0	0.0	0
Increasing genetic scenario								
Uniform	1.2	1.0	-0.2	−18	2.0	1.5	−0.5	−25
Recent	2.0	1.8	-0.2	−10	2.0	1.8	−0.2	−8
Critical	2.0	1.6	-0.4	−20	2.0	1.3	−0.7	−36
Increasing	0.7	0.6	-0.1	−12	2.0	1.7	−0.3	−16
Decreasing genetic scenario								
Uniform	1.2	1.5	0.3	22	2.0	3.0	1.0	50
Recent	2.0	2.2	0.2	11	2.0	2.3	0.3	16
Critical	2.0	2.5	0.5	23	2.0	3.4	1.4	72
Increasing	0.7	0.8	0.1	14	2.0	2.7	0.7	34
FTO genetic scenario								
Uniform	1.2	0.9	−0.3	−22	2.0	3.7	1.7	85
Recent	2.0	2.0	0.0	−2	2.0	2.9	0.9	46
Critical	2.0	1.5	−0.5	−24	2.0	3.9	1.9	95
Increasing	0.7	0.7	−0.1	−8	2.0	3.7	1.7	85

MR: Mendelian randomization