



# Handling time varying confounding in observational research

Mohammad Ali Mansournia,<sup>1</sup> Mahyar Etminan,<sup>2</sup> Goodarz Danaei,<sup>3</sup> Jay S Kaufman,<sup>4</sup> Gary Collins<sup>5</sup>

<sup>1</sup>Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, PO Box 14155-6446, Tehran, Iran

<sup>2</sup>Department of Ophthalmology and Visual Sciences, University of British Columbia, Vancouver, British Columbia, Canada

<sup>3</sup>Department of Global Health and Population, Harvard TH Chan School of Public Health, Boston, MA, USA; Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA, USA

<sup>4</sup>Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Canada

<sup>5</sup>Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK

Correspondence to:

M A Mansournia  
mansournia\_m@sina.tums.ac.ir

Cite this as: *BMJ* 2017;359:j4587  
<http://dx.doi.org/10.1136/bmj.j4587>

Accepted: 28 August 2017

Many exposures of epidemiological interest are time varying, and the values of potential confounders may change over time leading to time varying confounding. The aim of many longitudinal studies is to estimate the causal effect of a time varying exposure on an outcome that requires adjusting for time varying confounding. Time varying confounding affected by previous exposure often occurs in practice, but it is usually adjusted for by using conventional analytical methods such as time dependent Cox regression, random effects models, or generalised estimating equations, which are known to provide biased effect estimates in this setting. This article explains time varying confounding affected by previous exposure and outlines three causal methods proposed to appropriately adjust for this potential bias: inverse-probability-of-treatment weighting, the parametric G formula, and G estimation.

Imagine you are a clinician scientist who is interested in examining the effect of testosterone treatment on risk of acute myocardial infarction. Recent studies have alluded to a possible harmful effect, although results have been contradictory.<sup>1</sup> One reason might be inadequate adjustment for the time varying confounding that may occur by the change in serum

testosterone levels over time, which in turn could affect treatment patterns in the future. After talking to a colleague, you are convinced that adjusting for serum testosterone levels as a time varying confounder is necessary for a valid estimate of the causal effect of interest. Here we discuss the concept of time varying confounding in observational studies and introduce methods that must be used to appropriately adjust for time varying confounders that are affected by previous treatment.

## Time varying confounding affected by past exposure

The goal of many observational studies is to estimate the valid causal effect of a treatment or exposure on a certain outcome. The value of exposure can depend on external factors (called confounders) that also affect the outcome. This may result in a “mixing” effect brought on by the confounders, commonly referred to as confounding.<sup>2-6</sup> Careful control of confounding remains one of the most challenging but important steps in the conduct and analysis of these types of observational studies.

Time varying confounding occurs when confounders have values that change over time. It often occurs with time varying exposures. Many exposures of epidemiological interest are time varying—for example, treatment dose, body mass index, smoking status, blood pressure, depression, air pollution, socioeconomic status.

Many longitudinal studies aim to estimate the overall causal effect of a time varying exposure on the outcome, which requires adjustment for time varying confounding. However, in certain clinical scenarios one or more time varying confounders are affected by past exposure. In our clinical example, serum testosterone level is the time varying confounder because doctors are likely to use the value of testosterone to titrate treatment. The level of testosterone after baseline is, however, affected by baseline testosterone treatment. This pattern, for lack of a shorter term, is referred to as “time varying confounding affected by past exposure.”

## The problem: controlling for time varying confounding affected by past exposure

Several methods can control confounding in observational studies at the design and analysis stages, including restriction, stratification, regression modelling, matching, and propensity scoring.<sup>2</sup> Nevertheless, these methods as well as more advanced conventional statistical methods such as time dependent Cox proportional hazard models, random effects models, or generalised estimating equations have been shown to be inadequate in controlling time varying confounding affected by past exposure.<sup>7-8</sup> In fact, conventional statistical methods can introduce bias in the presence of time varying confounding affected by past exposure.<sup>8-9</sup> This can happen for

## SUMMARY POINTS

- Many exposures of epidemiological interest are time varying, and time varying confounding affected by past exposure often occurs in practice
- Time varying confounding affected by past exposure is often adjusted for by using conventional analytical methods such as time dependent Cox regression, random effects models, or generalised estimating equations in clinical research, which are known to provide biased effect estimates in this setting
- Three causal methods have been proposed to appropriately adjust for time varying confounders that are affected by past exposure: inverse-probability-of-treatment weighting, parametric G formula, and G estimation

one or two reasons. Firstly, over-adjustment bias,<sup>10</sup> which occurs as a result of blocking the effect of past exposure on outcome, mediated through later confounders, leading to a downward bias. This dilemma occurs because the time varying confounder is affected by past exposure and therefore has a dual role as both mediator and confounder. In our clinical example, serum testosterone levels may mediate the impact of past testosterone treatment on myocardial infarction. Secondly, selection (also known as collider stratification) bias,<sup>11</sup> which occurs by inappropriately adjusting for a time varying confounder that may share a common cause with the outcome. In our example, restricting the study population to those who have lower serum testosterone levels post-baseline will create a (negative) association between baseline testosterone treatment and other causes of acute myocardial infarction, and this will lead to a false association between baseline treatment and the outcome.

The causal diagram in figure 1 represents our clinical scenario. Causal diagrams, represented as directed acyclic graphs, comprise variables and arrows.<sup>2 12-14</sup> The absence of an arrow pointing from one variable to another indicates that the former does not have a direct causal effect on the latter. To avoid clutter, the causal diagram in figure 1 only shows the exposure testosterone treatment at baseline ( $E_0$ ) and the second visit ( $E_1$ ), the time varying confounder (serum testosterone level) at the second visit ( $C_1$ ), and the outcome (myocardial infarction) between the second and third visits ( $D_2$ ); for the sake of simplicity we have omitted confounder at baseline ( $C_0$ ) and outcome between first and second visits ( $D_1$ ) as well as some arrows (eg, from  $E_0$  to  $E_1$ ). Figure 1 represents time varying confounding affected by past exposure as the post-baseline value of serum testosterone level  $C_1$  is a common cause of  $E_1$  and  $D_2$  (as there are arrows from  $C_1$  to  $E_1$  and  $D_2$ ) and thus is a time varying confounder. Also,  $C_1$  is affected by the previous exposure  $E_0$  as there is an arrow from  $E_0$  to  $C_1$ . Box 1 provides a graphical representation of over-adjustment and selection biases for the effect of testosterone treatment on risk of acute myocardial infarction.

### Occurrence of time varying confounding affected by past exposure

The issue of time varying confounding affected by past exposure is quite common and of concern in clinical epidemiology when drug indication can often depend on the value of a prognostic test, or a

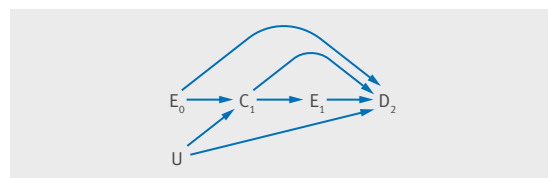


Fig 1 | Causal diagram showing time varying confounding affected by past exposure

biomarker that is itself affected by past treatment. Some examples include CD4 lymphocyte count for the effect of zidovudine therapy on AIDS,<sup>8</sup> asthma rescue drugs on pulmonary function measures,<sup>15</sup> and serum cholesterol level for the effect of statin treatment on coronary heart disease.<sup>16</sup>

Time varying confounding affected by past exposure also exists when there is feedback between exposure and outcome. For example, in a study on the effect of physical activity on knee pain in patients with osteoarthritis, previous knee pain can act as a strong confounder for the effect of current physical activity as it can limit current physical activity and at the same time may be affected by a patient's earlier physical activity.<sup>9</sup>

### The solution: G methods and longitudinal data

The resolution for the dilemma of time varying confounding affected by past exposure requires the longitudinal data on both the exposure and the confounders affected by the previous exposure. It also requires use of a statistical method that adjusts for the confounding effect of the covariate, but not for the effect of exposure on the confounder. Three causal methods have been suggested for this purpose: inverse-probability-of-treatment weighting,<sup>2 7-9 17 18</sup> parametric G formula,<sup>2 19 20</sup> and G estimation.<sup>2 21</sup> These methods are collectively known as G methods as they can be used for any generalised treatment, whether time fixed or time varying. The time varying treatment plan can be static (eg, always treat versus never treat) or dynamically dependent on the values of time varying confounders (eg, start treatment if CD4 is <200 versus start treatment if CD4 is <500). These three G methods can properly estimate the effect of a time varying exposure in the presence of time varying confounding affected by past exposure if longitudinal data are available.

The success of G methods for appropriately adjusting time varying confounders affected by past exposure is because unlike conventional methods such as stratification or regression modelling, they do not fix the value of confounders to adjust for them so they do not introduce over-adjustment or selection bias. Inverse-probability-of-treatment weighting creates a "pseudo-cohort," within which the time varying confounders are not associated with future exposure. Therefore, there is no confounding in the pseudo-cohort and so a crude (unadjusted) analysis in the pseudo-cohort (which is done using a weighted analysis in the original cohort) provides an unbiased effect of exposure, assuming there is no unmeasured confounding.

The parametric G formula effectively simulates the values of confounders, exposure, and outcome that would have been observed in a hypothetical study where every participant received the exposure regimen of interest (eg, always exposed, never exposed). Therefore, the risk for the exposure regimen of interest (eg, always exposed) adjusted for time varying confounders can be calculated by standardising the

**Box 1: Graphical representation of bias of standard methods for time varying confounding**

The bias introduced by conventional statistical models can be illustrated using the causal diagram in figure 1 and graphical rules.  $C_1$  is a common cause of  $E_1$  and  $D_2$ , so it is a time varying confounder and should be adjusted for in the analysis. However, the conventional adjustment (through restriction, matching, or regression) for  $C_1$  may introduce two biases if a time varying confounder is affected by past exposure, as the arrow from  $E_0$  to  $C_1$  in figure 1 suggests: (i) over-adjustment bias: it removes part of the effect of previous testosterone treatment ( $E_0$ ) on myocardial infarction that is mediated by serum testosterone level at the second visit ( $C_1$ ), as the causal path  $E_0 \rightarrow C_1 \rightarrow D_2$  will be blocked owing to adjusting for  $C_1$ , and (ii) selection (also known as collider stratification) bias:  $C_1$  is a collider on the path  $E_0 \rightarrow C_1 \leftarrow U \rightarrow D_2$ , where unmeasured risk factor  $U$  represents stress as two arrowheads from  $E_0$  and  $U$  converge on it. This path is blocked as it contains a collider ( $C_1$ ) that is not controlled. However, conventional adjustment for  $C_1$  opens this path and makes previous testosterone treatment ( $E_0$ ) a non-causal risk factor for myocardial infarction ( $D_2$ )

risk to the joint distribution of confounders across the entire follow-up.

G estimation adjusts for time varying confounders by separately examining the association between exposure and the counterfactual outcome under no exposure during the follow-up (ie, the outcome that would have been observed under no exposure during the follow-up) at each visit and only adjusting for the values of time varying confounders in the previous visits. The counterfactual outcome under no exposure during the follow-up is a pre-exposure variable that contains information about the participant's underlying predisposition for the outcome and is assumed to be independent of exposure given the measured confounders (no unmeasured confounding

assumption). Box 2 provides a brief description of the technical details of G methods.

As an example of inverse-probability-of-treatment weighting, the Multicenter AIDS Cohort Study estimated the hazard ratio for always use compared with never use of zidovudine on mortality of men who were positive for antibodies to HIV as 0.7 (95% confidence interval 0.6 to 1.0) using a weighted Cox model, which included the exposure and baseline confounders with inverse-probability-of-treatment weighting to adjust for time varying confounders (eg, CD4 count).<sup>8</sup> Under the assumption of no uncontrolled confounders, this inverse-probability-of-treatment weighted estimate represents the overall effect of zidovudine on mortality. The authors also reported

**Box 2: G methods**

G methods are causal methods for the estimation of the causal effect of a time varying exposure that can appropriately adjust for time varying confounders that are affected by past exposure: inverse-probability-of-treatment weighting, parametric G formula, and G estimation

In inverse-probability-of-treatment weighting, confounding is adjusted by weighting with weights equal to the inverse of the participant's probability of receiving their exposure history at each visit, given the value of previous confounders. The weights can be estimated from a logistic regression model for predicting exposure at each visit based on the previous exposure and covariate histories. Inverse-probability-of-treatment weighting in a cohort creates a "pseudo-cohort" with two important properties: (i) the measured confounders are not confounders in the pseudo-cohort because they do not predict the future exposure—that is, there is no arrow from  $C_1$  to  $E_1$  in figure 1 in the pseudo-cohort, and (ii) the causal effect of the exposure in the pseudo-cohort is the same as in the cohort. Therefore, the causal effect of the exposure can be unbiasedly estimated by a crude analysis in the pseudo-cohort, which is the same as a weighted analysis in the original cohort

To improve the efficiency of the causal effect estimate, it is usually only the post-baseline values of time varying confounders that are adjusted by inverse-probability-of-treatment weighting, while their baseline values along with the time fixed confounders (eg, sex) are adjusted traditionally using conventional statistical models. In other words, the standard outcome regression model includes the exposure and baseline confounders and is weighted by the inverse-probability-of-treatment. The ordinary 95% confidence interval for inverse-probability-of-treatment weighted estimates will generally not provide at least 95% coverage and thus should be avoided. Instead, robust or sandwich variance estimators or non-parametric bootstrapping (with resampling of participants) should be used to provide valid confidence intervals

The parametric G formula is a time dependent and model based generalisation of standardisation: a method, based on weighted averaging and a standard population, used to remove the effects of differences in age or other confounding variables (which yields the standardised mean outcome, eg, risk) under exposure regimens of interest (eg, always treat and never treat) by averaging the confounder specific mean outcomes under that regimen over the joint distribution of confounders across the entire follow-up. It effectively simulates the distribution of the exposure, outcome, and confounders that would have been observed in a hypothetical study where every participant received the exposure regimen of interest. Using the parametric G formula, two sets of regression models are constructed: (i) classic outcome regression with exposure and confounder histories as predictors, and (ii) confounder regressions with previous exposure and confounder histories as predictors. Then the standardised risk resulting from the exposure regimen of interest can be derived using Monte Carlo draws of confounders from the confounder regression model and substituting them in the outcome regression model along with the exposure regimen of interest. In many real applications, the exposure regimen is dynamic and therefore an exposure model is also needed. Confidence intervals can be constructed using non-parametric bootstrapping

G estimation is a two step procedure that uses two models to estimate the causal effect. The first is a causal model that includes the causal variable of interest and links the counterfactual outcome under no exposure during the follow-up (ie, the outcome that would have been observed under no exposure during the follow-up) to the weighted sum of time spent in a given exposure status. The second is a logistic regression for predicting exposure at each visit based on the previous exposure and covariate histories and the counterfactual outcome. Then G estimation iteratively searches for a causal variable value in the first (causal) model, which when used in the second (treatment) model, makes the exposure independent of the estimated counterfactual outcome given previous treatment and confounder history. G estimation succeeds in adjusting for time varying confounders that are affected by previous exposure by separately examining the association between exposure and counterfactual outcome at each visit and adjusting only for the time varying confounder values in past visits. Confidence intervals can be constructed using either a test based procedure or non-parametric bootstrapping

## Box 3: Glossary

- *Time fixed exposure*: any exposure that only occurs at the start of follow-up (eg, a one dose vaccine) or does not change over time (eg, genotype)
- *Time varying exposure*: any exposure that is not fixed. Many exposures of epidemiological interest are time varying
- *Time varying confounding*: confounding by confounders the values of which change over time. It often occurs with time varying exposures
- *Time varying confounding affected by past exposure*: time varying confounding in which one or more time varying confounders are affected by past exposure
- *Conventional methods for analysis of longitudinal data*: statistical methods such as time dependent Cox proportional hazard models, random effects models, or generalised estimating equations, which accounts for time varying confounders but at the expense of introducing bias in the presence of time varying confounding affected by past exposure
- *G methods*: causal methods proposed to appropriately adjust for time varying confounders that are affected by past exposure including inverse-probability-of-treatment weighting, parametric G formula, and G estimation. Unlike conventional methods, they do not fix the value of time varying confounders to adjust for them

the biased hazard ratio of 0.4 (95% confidence interval 0.3 to 0.5) from the standard time dependent Cox regression model, which included time varying confounders in the model; this was considerably less than inverse-probability-of-treatment weighted estimate of 0.7. As time varying confounders such as CD4 is affected by earlier zidovudine treatment, this estimate from a conventional model cannot be causally interpreted as either the overall zidovudine effect or the direct effect of zidovudine mediated by pathways not through time varying confounders.

As an example of G estimation, a recent study aimed to quantify the causal relation between obesity and coronary heart disease by appropriately adjusting for time varying confounders, including hypertension, diabetes mellitus, cholesterol, high density lipoprotein cholesterol, triglyceride, smoking, drinking status, and occurrence of stroke. The hazard ratios for abdominal obesity estimated by G estimation were 1.65 (95% confidence interval 1.35 to 1.92) based on waist circumference and 1.38 (1.13 to 1.99) based on waist-to-hip ratio, suggesting that abdominal obesity increased the risk of coronary heart disease. However, the standard time dependent Cox model adjusting for baseline time varying confounders yielded hazard ratios for abdominal obesity of 0.93 (95% confidence interval 0.77 to 1.12) based on waist circumference and 0.89 (0.69 to 1.15) based on waist-to-hip ratio.

Table 1 summarises the advantages and disadvantages of the three G methods. The important point is that while inverse-probability-of-treatment weighting and G estimation rely on correct exposure modelling (ie, regression modelling with exposure as the response variable and confounders as predictors), the parametric G formula is based on correct outcome and confounder modelling. In the case of time varying confounding affected by past exposure in pharmacoepidemiology, it can be argued that exposure models are more easily constructed based on the treatment indications and therefore inverse-probability-of-treatment weighting and G estimation may be preferred to the parametric G formula. Also inverse-probability-of-treatment weighting and G estimation can be more accurate (less subject to sparse data bias) than the parametric G formula in cohort studies if the exposure or treatment is common

but the outcome is rare.<sup>23</sup> The parametric G formula is the method of choice for estimating the effects of multiple risk factors (joint interventions) and dynamic interventions. For example, the Nurses' Health Study in 1982–2002 estimated the 20 year risk of coronary heart disease under a joint intervention of no smoking, increased exercise, improved diet, moderate alcohol consumption, and reduced body mass index using a parametric G formula as 1.9% (95% confidence interval 1.5% to 2.4%), whereas the observed risk was 3.5%.<sup>17</sup>

Inverse-probability-of-treatment weighting is the simplest method to apply and does not require special software, but it is sensitive to large weights and unlike the two other methods cannot be easily used to study the interaction between exposure and time varying confounders (ie, time varying effect modification).

### Conclusion

Time varying confounding affected by past exposure often occurs in longitudinal studies, but in practice it is usually either ignored or adjusted for by using conventional methods that are known to provide biased effect estimates in this setting. G methods are the methods of choice for adjustment of time varying confounders affected by the past exposure when cohort data with repeated measurements on the exposure and confounders are available. The cohort either can be clinical with a dynamic observation plan in which the interval between visits depends on the clinical evolution of the patients or can arise from a study with prespecified regular intervals between visits. In principle, G methods can also be used with case-control studies nested in longitudinal data using inverse probability weighting, with weights inversely proportional to sampling fractions in cases and controls required to properly account for the oversampling of cases, although further research is needed in this area.

Hybrid causal methods also enjoy the benefits of different G methods simultaneously. For example, the so-called double robust methods<sup>24</sup> are a group of causal methods that combine inverse-probability-of-treatment weighting and parametric G formula procedures and give valid causal effect estimate if either inverse-probability-of-treatment weighting or parametric G formula estimation is valid.

Advantages and disadvantages of three G methods		
G method	Advantages	Disadvantages
Inverse-probability-of-treatment weighting	They resemble standard statistical procedures and are simple to understand; available in almost all statistical software; useful when the reasons for exposure assignment are known (eg, in the presence of confounding by indication)	Cannot be used if there is a confounder level for which all participants are exposed or not exposed (eg, those who leave their occupation cannot be exposed to occupational exposures); unstable in the presence of extreme weights; less useful for studying the interaction between exposure and time varying confounders
Parametric G formula	Ideal for studies that examine interventions on multiple risk factors (joint interventions) and interventions dependent on evolving risk factor values (dynamic interventions); computes causal measures of interest such as risk ratios and risk differences	Computationally intensive, requires extra programming, and can lead to fitting problems; requires models for confounders as well as outcomes; subject to the “G null paradox”: the method rejects the causal null hypothesis, even when true, in sufficiently large samples, so it can only be used for interventions when the null is believed to be untrue
G estimation	Can be used even if there is a confounder level for which all participants are exposed or not exposed; useful for studying the interaction between exposure and time varying confounders	Computationally intensive, requires extra programming, and can lead to fitting problems; the methodology and resulting effect variable are somewhat difficult to understand

G methods rely on some assumptions for unbiased estimation of causal effects. In particular, the model represented in the causal diagram should be correct as we identify confounders based on causal diagrams. Other assumptions include well defined exposure, no measurement error, and correct specification of the statistical models. Inverse-probability-of-treatment weighting also needs positivity assumption (ie, all exposure levels are observed within each joint stratum of confounders), and violations of positivity will lead to huge or infinite weights and subsequently biased effect estimates. We note that standard statistical methods such as time dependent Cox regression also require these assumptions (except positivity), but still result in estimates of effect that may fail to have a causal interpretation even if all assumptions mentioned previously are met but there is a time varying confounder affected by past exposure.

We hope that this introduction to time varying confounding and appropriate handling of time varying confounding affected by past exposure will encourage researchers to use G methods for the analysis of longitudinal data in practice. As box 2 and table 1 suggest, however, application of G methods (especially parametric G formula and G estimation) involves many technicalities, so expert statistical advice is strongly recommended. Interested readers looking for the next step in understanding G methods are encouraged to read a recent tutorial written for epidemiologists.<sup>25</sup>

We thank Stephen Cole for helpful comments on earlier drafts of this paper.

**Contributors:** All authors were involved in the concept and design of the study. MAM wrote the first draft of the manuscript, and all authors contributed to subsequent revisions of the manuscript and approved the final version. MAM is the guarantor.

**Competing interests:** We have read and understood BMJ policy on declaration of interests and declare the following: none.

**Provenance and peer review:** Not commissioned; externally peer reviewed.

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

1 Hwang K, Miner M. Controversies in testosterone replacement therapy: testosterone and cardiovascular disease. *Asian J Androl* 2015;17:187-91.doi:10.4103/1008-682X.146968

2 Hernán MA, Robins JM. *Causal Inference*. Chapman & Hall/ CRC, 2017.

3 Greenland S, Robins JM, Pearl J. Confounding and collapsibility in causal inference. *Stat Sci* 1999;14:29-46.doi:10.1214/ss/1009211805.

4 Suzuki E, Tsuda T, Mitsuhashi T, Mansournia MA, Yamamoto E. Errors in causal inference: an organizational schema for systematic error and random error. *Ann Epidemiol* 2016;26:788-793. e1.doi:10.1016/j.annepidem.2016.09.008

5 Mansournia MA, Greenland S. The relation of collapsibility and confounding to faithfulness and stability. *Epidemiology* 2015;26:466-72.doi:10.1097/EDE.0000000000000291

6 Greenland S, Mansournia MA. Limitations of individual causal models, causal graphs, and ignorability assumptions, as illustrated by random confounding and design unfaithfulness. *Eur J Epidemiol* 2015;30:1101-10.doi:10.1007/s10654-015-9995-7

7 Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11:550-60. doi:10.1097/00001648-200009000-00011

8 Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 2000;11:561-70. doi:10.1097/00001648-200009000-00012

9 Mansournia MA, Danaei G, Forouzanfar MH, et al. Effect of physical activity on functional performance and knee pain in patients with osteoarthritis: analysis with marginal structural models. *Epidemiology* 2012;23:631-40.doi:10.1097/EDE.0b013e31824cc1c3

10 Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology* 2009;20:488-95.doi:10.1097/EDE.0b013e3181a819a1

11 Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology* 2003;14:300-6.doi:10.1097/01.EDE.0000042804.12056.6C

12 Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999;10:37-48.doi:10.1097/00001648-199901000-00008

13 Mansournia MA, Hernán MA, Greenland S. Matched designs and causal diagrams. *Int J Epidemiol* 2013;42:860-9.doi:10.1093/ije/dyt083

14 Mansournia MA, Higgins JPT, Sterne JAC, Hernán MA. Biases in randomized trials: a conversation between trialists and epidemiologists. *Epidemiology* 2017;28:54-9.doi:10.1097/EDE.0000000000000564

15 Mortimer KM, Neugebauer R, van der Laan M, Tager IB. An application of model-fitting procedures for marginal structural models. *Am J Epidemiol* 2005;162:382-8.doi:10.1093/aje/kwi208

16 Danaei G, Rodríguez LA, Cantero OF, Logan R, Hernán MA. Observational data for comparative effectiveness research: an emulation of randomised trials of statins and primary prevention of coronary heart disease. *Stat Methods Med Res* 2013;22:70-96. doi:10.1177/0962280211403603

17 Mansournia MA, Altman DG. Inverse probability weighting. *BMJ* 2016;352:i189.doi:10.1136/bmj.i189

18 Mohammad K, Hashemi Nazari SS, Mansournia N, Mansournia MA. Marginal versus conditional causal effects. *J BiostatEpidemiol* 2015;1:121-8.

19 Taubman SL, Robins JM, Mittleman MA, Hernán MA. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *Int J Epidemiol* 2009;38:1599-611. doi:10.1093/ije/dyp192

20 Gharibzadeh S, Mohammad K, Rahimiforushani A, Amouzegar A, Mansournia MA. Standardization as a Tool for Causal Inference in Medical Research. *Arch Iran Med* 2016;19:666-70.

- 21 Hernán MA, Cole SR, Margolick J, Cohen M, Robins JM. Structural accelerated failure time models for survival analysis in studies with time-varying treatments. *Pharmacoepidemiol Drug Saf* 2005;14:477-91.doi:10.1002/pds.1064
- 22 Shakiba M, Mansournia MA, Salari A, Soori H, Mansournia N, Kaufman JS. Accounting for time-varying confounding in the relation between obesity and coronary heart disease: Analysis with G-estimation, the Atherosclerosis Risk in Communities (ARIC) study. *Am J Epidemiol* [In press].
- 23 Greenland S, Mansournia MA, Altman DG. Sparse data bias: a problem hiding in plain sight. *BMJ* 2016;352:i1981.doi:10.1136/bmj.i1981
- 24 Funk MJ, Westreich D, Wiesen C, Stürmer T, Brookhart MA, Davidian M. Doubly robust estimation of causal effects. *Am J Epidemiol* 2011;173:761-7.doi:10.1093/aje/kwq439
- 25 Naimi AL, Cole SR, Kennedy EH. An introduction to g methods. *Int J Epidemiol* 2017;46:756-62.