## Education Corner

# An introduction to g methods

**Ashley I Naimi[1]\*, Stephen R Cole[2] and Edward H Kennedy[3]**

[1]Department of Epidemiology, University of Pittsburgh, [2]Department of Epidemiology, University of North Carolina at Chapel Hill and [3]Department of Statistics, Carnegie Mellon University

*Corresponding author. Department of Epidemiology University of Pittsburgh 130 DeSoto Street 503 Parran Hall Pittsburgh, PA 15261 ashley.naimi@pitt.edu

## Abstract

Robins' generalized methods (g methods) provide consistent estimates of contrasts (e.g. differences, ratios) of potential outcomes under a less restrictive set of identification conditions than do standard regression methods (e.g. linear, logistic, Cox regression). Uptake of g methods by epidemiologists has been hampered by limitations in understanding both conceptual and technical details. We present a simple worked example that illustrates basic concepts, while minimizing technical complications.

**Key words**: G Methods; Marginal Structural Model; Structural Nested Model; G Formula; Inverse Probability Weighting; G Estimation; Monte Carlo Estimation

---

**Key Messages**

- G methods include inverse probability weighted marginal structural models, g estimation of a structural nested model, and the g formula.
- G methods estimate contrasts of potential outcomes under a less restrictive set of assumptions than standard regression methods.
- Inverse probability weighting generates a pseudo-population in which exposures are independent of confounders, enabling estimation of marginal structural model parameters.
- G estimation exploits the conditional independence between the exposure and potential outcomes to estimate structural nested model parameters.
- The g formula models the joint density of the observed data to generate potential outcomes under different exposure scenarios.

---

Robins' g methods enable the identification and estimation of the effects of generalized treatment, exposure, or intervention plans. G methods are a family of methods that include the g formula, marginal structural models, and structural nested models. They provide consistent estimates of contrasts (e.g. differences, ratios) of average potential outcomes under a less restrictive set of identification conditions than standard regression methods (e.g. linear, logistic, Cox regression).[1] Specifically, standard regression requires no feedback between time-varying treatments and time-varying confounders, while g methods do not. Robins and Hernán[1] have provided a technically comprehensive

worked example of each of the three g methods. Here, we present a corresponding worked example that illustrates the need for and use of g methods, while minimizing technical details.

## Example

Our research question concerns the effect of treatment for HIV on CD4 count. Table 1 presents data from a hypothetical observational cohort study ($A = 1$ for treated, $A = 0$ otherwise). Treatment is measured at baseline ($A_0$) and once during follow up ($A_1$). The sole covariate is elevated HIV viral load ($Z = 1$ for those with $> 200$ copies/ml, $Z = 0$ otherwise), which is constant by design at baseline ($Z_0 = 1$) and measured once during follow up just prior to the second treatment ($Z_1$). The outcome is CD4 count measured at the end of follow up in units of cells/mm[3]. The CD4 outcome in Table 1 is summarized (averaged) over the participants at each level of the treatments and covariate. The number of participants is provided in the rightmost column of Table 1. In this hypothetical study of one million participants we ignore random error and focus on identifying the parameters defining our causal effect of interest, which we describe next.

Based on Figure 1, the average outcome in our simple data generating structure may be composed of several parts: the effects of $A_0$, $Z_1$, and $A_1$; the two-way interactions between $A_0$ and $Z_1$, $A_0$ and $A_1$, and $A_1$ and $Z_1$; and the three-way interaction between $A_0$, $Z_1$, and $A_1$. These components (some whose magnitudes may be zero) can be used to "build up" a contrast of substantive interest. Here, we focus on the average causal effect of always taking treatment ($a_0 = 1, a_1 = 1$) compared to never taking treatment ($a_0 = 0, a_1 = 0$),

**Table 1.** Prospective study data illustrating the number of subjects ($N$) within each possible combination of treatment at time 0 ($A_0$), HIV viral load just prior to the second round of treatment ($Z_1$), and treatment status for the 2nd round of treatment ($A_1$). The outcome column ($Y$) corresponds to the mean of $Y$ within levels of $A_0$, $Z_1$, $A_1$. Note that HIV viral load at baseline is high ($Z_0 = 1$) for everyone by design

| $A_0$ | $Z_1$ | $A_1$ | $Y$ | $N$ |
|---|---|---|---|---|
| 0 | 0 | 0 | 87.29 | 209,271 |
| 0 | 0 | 1 | 112.11 | 93,779 |
| 0 | 1 | 0 | 119.65 | 60,654 |
| 0 | 1 | 1 | 144.84 | 136,293 |
| 1 | 0 | 0 | 105.28 | 134,781 |
| 1 | 0 | 1 | 130.18 | 60,789 |
| 1 | 1 | 0 | 137.72 | 93,903 |
| 1 | 1 | 1 | 162.83 | 210,527 |

$$\psi = E(Y^{a_0=1,a_1=1}) - E(Y^{a_0=0,a_1=0})$$
$$= E(Y^{a_0=1,a_1=1} - Y^{a_0=0,a_1=0}),$$

where expectations $E(\cdot)$ are taken with respect to the target population from which our sample is a random draw. This average causal effect consists of the joint effect of $A_0$ and $A_1$ on $Y$.[2] Here, $Y^{a_0,a_1}$ represents a potential outcome value that would have been observed had the exposures been set to specific levels $a_0$ and $a_1$. This potential outcome is distinct from the observed (or actual) outcome.

This average causal effect $\psi = E(Y^{a_0,a_1} - Y^{0,0})$ is a *marginal* effect because it averages (or marginalizes) over all individual-level effects in the population. We can write this effect as $E(Y^{a_0,a_1} - Y^{0,0}) = \psi_0 a_0 + \psi_1 a_1 + \psi_2 a_0 a_1$, which states that our average causal effect $\psi$ may be composed of two exposure main effects (e.g., $\psi_0$ and $\psi_1$) and their two-way interaction ($\psi_2$). This marginal effect $\psi$ is indifferent to whether the $A_1$ component ($\psi_1 + \psi_2$) is modified by $Z_1$: whether such effect modification is present or absent, the marginal effect represents a meaningful answer to the question: what is the effect of $A_0$ and $A_1$ in the entire population?

Alternatively, we may wish to estimate this effect *conditional* on certain values of another covariate. A conditional effect would arise if, for example, one was specifically interested in effect measure modification by $Z_1$. When properly modeled, this conditional effect represents a meaningful answer to the question: what is the effect of $A_0$ and $A_1$ in those who receive $Z_1 = 1$ versus those who receive $Z_1 = 0$? Modeling such effect measure modification by time-varying covariates is the fundamental issue that distinguishes marginal structural from structural nested models. We thus return to this issue later. For simplicity, we define our effect of interest as $\psi = \psi_0 + \psi_1 + \psi_2$, and we explore a data example with no effect modification by time-varying confounders.

## Assumptions

Our average causal effect is defined as a function of two averages that would be observed if everybody in the population were exposed (or unexposed) at both time points.
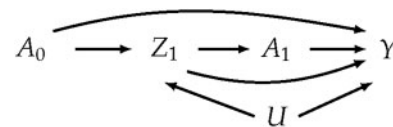


**Figure 1.** Causal diagram representing the relation between anti-retroviral treatment at time 0 ($A_0$), HIV viral load just prior to the second round of treatment ($Z_1$), anti-retroviral treatment status at time 1 ($A_1$), the CD4 count measured at the end of follow-up ($Y$), and an unmeasured common cause ($U$) of HIV viral load and CD4.

Yet we cannot directly acquire information on these averages because in any given sample, some individuals will be unexposed (or exposed). Part of our task therefore involves justifying use of averages among subsets of the population as what would be observed in the whole population. This is accomplished by making three main assumptions.

Counterfactual consistency[3] allows us to equate observed outcomes among those who received a certain exposure value to the potential outcomes that would be observed under the same exposure value:

$$E(Y|A_0 = a_0, A_1 = a_1) = E(Y^{a_0,a_1}|A_0 = a_0, A_1 = a_1)$$

The status of this assumption remains unaffected by the choice of analytic method (e.g., standard regression versus g methods). Rather, this assumption's validity depends on the nature of the exposure assignment mechanism.[4] Under counterfactual consistency, we partially identify our average causal effect.

Next, we assume exchangeability.[5] Exchangeability implies that the potential outcomes under exposures $a_0$ and $a_1$ (denoted $Y^{a_0,a_1}$) are independent of the actual (or observed) exposures $A_0$ and $A_1$. We make this exchangeability assumption within levels of past covariate values (conditional) and at each time point separately (sequential):

$$E(Y^{a_0,a_1}|A_1, Z_1, A_0) = E(Y^{a_0,a_1}|Z_1, A_0), \text{ and} \quad (1)$$
$$E(Y^{a_0,a_1}|A_0) = E(Y^{a_0,a_1}). \quad (2)$$

This sequential conditional exchangeability assumption would hold if there were no uncontrolled confounding and no selection bias. Equation 1 says that, within levels of prior viral load ($Z_1$) and a given treatment level $A_0$, $Y^{a_0,a_1}$ does not depend on the assigned values of $A_1$. Equation 2 says that $Y^{a_0,a_1}$ does not depend on the assigned values of $A_0$. Note the correspondence between these two equations and the causal diagram: because in Figure 1, $Z_1$ is a common cause of $A_1$ and $Y$, the assumption in equation 1 must be made conditional on $Z_1$. Failing to condition for $Z_1$ will result in uncontrolled confounding of the effect of $A_1$, and thus a dependence between the actual $A_1$ value and the potential outcome. However, adjusting for $Z_1$ using standard methods (restriction, stratification, matching, or conditioning in a linear regression model) would block part of the effect from $A_0$ through $Z_1$, and potentially lead to a collider bias of the effect of $A_0$ through $U$.[6] This is the central challenge that g methods were developed to address.

The third assumption, known as positivity,[7] requires $0 < P(A_1 = 1|Z_1 = z_1, A_0 = a_0) < 1$ and $0 < P(A_0 = 1) < 1$. Furthermore, this assumption must hold for all values of $a_0$ and $z_1$ where $P(A_0 = a_0, Z_1 = z_1) > 0$. This latter condition is required so that effects are not defined in strata of $a_0$ and $z_1$ that do not exist. Positivity is met when there are exposed and unexposed individuals within all confounder and prior exposure levels, which can be evaluated empirically.

Under these three assumptions, our hypothetical observational study can be likened to a sequentially randomized trial in which the exposure was randomized at baseline, and randomized again at time 1 with a probability that depends on $Z_1$. Under these assumptions, g methods can be used to estimate counterfactual quantities with observational data. In the Supplementary Material, we provide SAS code (SAS Institue, Cary, NC) in which standard regression and all three g methods are fit to the hypothetical data in Table 1.

## Results

### Standard Methods

Table 2 presents results from fitting a number of standard linear regression models to the data in Table 1. In the first model, $\hat{\beta} = 60.9$ cells/mm³ is the crude difference in mean CD4 count for the always treated compared to the never treated. In model two, $\hat{\beta} = 42.6$ cells/mm³ is the $Z_1$-adjusted difference in mean CD4 count for the same contrast. Other model results are provided in Table 2, and more could be entertained.

Table 3 presents the results from fitting all three g methods to the data in Table 1. The marginal structural model resulted in $\hat{\psi} = 50.0$ cells/mm³. The g formula resulted in $\hat{\psi} = 50.0$ cells/mm³. Finally, the structural nested model resulted in $\hat{\psi} = 50.0$ cells/mm³. Next we discuss how we obtained these results.

## G Methods

The g formula can be used to estimate the average CD4 level that would be observed in the population under a given treatment plan. To implement the approach, we start with a mathematical representation of the data generating

**Table 2.** A selection of regression models fit to the data in Table 1, and parameter estimates for various exposure contrasts

| Model Parameters | Estimate ($\hat{\beta}_1$) |
| --- | --- |
| $\beta_0 + \beta_1(A_0 + A_1)/2$ | 60.9 |
| $\beta_0 + \beta_1(A_0 + A_1)/2 + \beta_2 Z_1$ | 42.6 |
| $\beta_0 + \beta_1 A_0$ | 27.1 |
| $\beta_0 + \beta_1 A_0 + \beta_2 Z_1$ | 18.0 |
| $\beta_0 + \beta_1 A_1$ | 38.9 |
| $\beta_0 + \beta_1 A_1 + \beta_2 Z_1$ | 25.0 |

**Table 3.** G methods and corresponding estimates comparing contrasts quantifying always exposed versus never exposed scenarios fit to data in Table 1

| G Method | $\hat{\psi}^{a}$ |
|---|---|
| G Formula | 50.0 |
| IP-weighted marginal structural model | 50.0 |
| G Estimated Structural Nested Model | 50.0 |

$^{a}$ $\psi = E(Y^{1,1} - Y^{0,0})$

mechanism for all variables in Table 1. We refer to this as the joint density of the observed data. We factor the joint density in a way that respects the temporal ordering of the data by conditioning each variable on its history. For example, if $f(\cdot)$ represents the probability density function, then by the definition of conditional probabilities[8(p36)] we can factor this joint density as

$$f(y, a_1, z_1, a_0) = f(y|a_1, z_1, a_0)P(A_1 = a_1|Z_1 = z_1, A_0 = a_0)$$
$$P(Z_1 = z_1|A_0 = a_0)P(A_0 = a_0).$$

Our interest lies in the marginal mean of $Y$ that would be observed if $A_0$ and $A_1$ were set to some values $a_0$ and $a_1$, respectively. To obtain this expectation, we perform two mathematical operations on the factored joint density. The first is the well-known expectation operator,[8(p47)] which allows us to write the conditional mean of $Y$ in terms of its conditional density. The second is the law of total probability,[8(p12)] which allows us to marginalize over the distribution of $A_1$, $Z_1$ and $A_0$, yielding the marginal mean of $Y$:

$$E(Y) = \sum_{a_1, z_1, a_0} E(Y|A_1 = a_1, \quad Z_1 = z_1, A_0 = a_0)$$
$$P(A_1 = a_1|Z_1 = z_1, A_0 = a_0)$$
$$P(Z_1 = z_1|A_0 = a_0)P(A_0 = a_0).$$

We can now modify this equation to yield the average of potential outcomes that would be observed after intervening on the exposure [enabling us to drop out the terms for $P(A_1 = a_1|Z_1 = z_1, A_0 = a_0)$ and $P(A_0 = a_0)$], yielding

$$E(Y^{a_0, a_1}) = \sum_{z_1} E(Y|A_1 = a_1, Z_1 = z_1, A_0 = a_0)$$
$$P(Z_1 = z_1|A_0 = a_0).$$

This equation is the g formula; its proof, given in the Supplementary Material, follows from the three identifying assumptions. In our simple scenario, the expectation $E(Y^{0,0})$ can be calculated by summing the mean CD4 count in the never treated with $Z_1 = 1$ (weighted by the proportion of people with $Z_1 = 1$ in the $A_0 = 0$ stratum)

and the mean CD4 count in the never treated with $Z_1 = 0$ (weighted by the proportion of people with $Z_1 = 0$ in the $A_0 = 0$ stratum). Weighting the observed outcome's conditional expectation by the conditional probability that $Z_1 = z_1$ enables us to account for the fact that $Z_1$ is affected by $A_0$, but also confounds the effect of $A_1$ on $Y$. Computing this expectation's value yields a result of $\hat{E}(Y^{0,0}) = 100.0$, where we use $\hat{E}$ to denote a sample, rather than a population average, and with the understanding that $\hat{E}(Y^{0,0})$ is equal to the g formula with $A_0 = A_1 = 0$ (since the potential outcomes $Y^{0,0}$ are not directly observed). We repeat the process to obtain the corresponding value for treated at time 0 only: $\hat{E}(Y^{1,0}) = 125.0$; treated at time 1 only: $\hat{E}(Y^{0,1}) = 125.0$; and always treated: $\hat{E}(Y^{1,1}) = 150.0$. Thus, $\hat{\psi}_{GF} = 150.0 - 100.0 = 50.0$, which is the average causal effect of treatment on CD4 cell count.

This approach to computing the value of the g formula is referred to as nonparametric maximum likelihood estimation. Several authors[9–13] demonstrate how simulation from parametric regression models can yield a g formula estimator, which is often required in typical population-health studies with many covariates.

Modeling each component of the joint density of the observed data (including the probability that $Z_1 = z_1$) can lead to bias if any of these models are mis-specified. To compute the expectations of interest, we can instead specify a single model that targets our average causal effect, and avoid unnecessary modeling. Marginal structural models map a *marginal summary* (e.g., average) of potential outcomes to the treatment and parameter of interest $\psi$. Unlike the g formula, they do not require a model for $P(Z_1 = z_1|A_0 = a_0)$. Additionally, as we show in the Supplementary Material, while they cannot model it directly, they are indifferent to whether time-varying effect modification is present or absent. Because our interest lies in the marginal contrast of outcomes under always versus never treated conditions, our marginal structural model for the effect of $A$ can be written as $E(Y^{a_0, a_1}) = \beta_0 + \psi_0 a_0 + \psi_1 a_1 + \psi_2 a_0 a_1$, where $\beta_0 = E(Y^{0,0})$ is a (nuisance) intercept parameter, and $\psi = E(Y^{1,1} - Y^{0,0}) = (\psi_0 + \psi_1 + \psi_2)$ is the effect of interest.

Inverse probability weighting can be used estimate marginal structural model parameters (proofs are provided in the Supplementary Material). To estimate $\psi$ using inverse probability weighted regression, we first obtain the predicted probabilities of the observed treatments. In our example data, there are two possible $A_1$ values (exposed, unexposed) for each of the four levels in $Z_1$ and $A_0$. Additionally, there are two possible $A_0$ values (exposed, unexposed) overall. This leads to four possible exposure regimes: never treat, treat early only, treat late only, and

always treat. For each $Z_1$ value, we require the predicted probability of the exposure that was actually received. These probabilities are computed by calculating the appropriate proportions of subjects in Table 1. Because there are no variables that affect $A_0$, this probability is 0.5 for all individuals in the sample. Furthermore, in our example $A_1$ is not affected by $A_0$ (Figure 1). Thus, the $Z_1$ specific probabilities of $A_1$ are constant across levels of $A_0$. In settings where $A_0$ affects $A_1$, the $Z_1$ specific probabilities of $A_1$ would vary across levels of $A_0$.

In the stratum defined by $Z_1 = 1$, the predicted probabilities of $A_1 = 0$ and $A_1 = 1$ are 0.308 and 0.692, respectively. For example, $(210,527 + 136,293)/(210,527 + 136,293 + 93,903 + 60,654) = 0.692$. Thus, the probabilities for each treatment combination are: $0.5 \times 0.308 = 0.155$ (never treated), $0.5 \times 0.308 = 0.155$ (treated early only), $0.5 \times 0.692 = 0.346$ (treated late only), and $0.5 \times 0.692 = 0.346$ (always treated). Dividing the marginal probability of each exposure category (not stratified by $Z_1$) by these stratum specific probabilities gives stabilized weights of 1.617, 1.617, 0.725, and 0.725, respectively. For example, the never treated weight is $(0.5 \times 0.501)/(0.5 \times 0.308) = 1.617$. The same approach is taken to obtain predicted probabilities and stabilized weights in the stratum defined by $Z_1 = 0$. The weights and weighted data are provided in Table 4.

Fitting this model in the weighted data given in Table 4 provides the inverse-probability weighted estimates $[\hat{\psi}_{0_{IP}} = 25.0, \hat{\psi}_{1_{IP}} = 25.0, \hat{\psi}_{2_{IP}} = 0.0]$, thus yielding $\hat{\psi}_{IP} = 50.0$.

Weighting the observed data by the inverse of the probability of the observed exposure yields a "pseudo-population" (Table 4) in which treatment at the second time point ($A_1$) is no longer related to (and is thus no longer confounded by) viral load just prior to the second time point ($Z_1$). Thus, weighting a conditional regression model for the outcome by the inverse probability of treatment enables us to account for the fact that $Z_1$ both confounds $A_1$ and is affected by $A_0$.

**Table 4.** Stabilized inverse probability weights and Pseudo-population obtained by using inverse probability weights

| $A_0$ | $Z_1$ | $A_1$ | $Y$ | $sw$ | Pseudo $N$ |
|---|---|---|---|---|---|
| 0 | 0 | 0 | 87.23 | 0.72 | 151222.84 |
| 0 | 0 | 1 | 112.23 | 1.62 | 151680.46 |
| 0 | 1 | 0 | 119.79 | 1.62 | 98110.06 |
| 0 | 1 | 1 | 144.78 | 0.72 | 98789.40 |
| 1 | 0 | 0 | 105.25 | 0.72 | 97395.08 |
| 1 | 0 | 1 | 130.25 | 1.62 | 98321.62 |
| 1 | 1 | 0 | 137.80 | 1.62 | 151884.02 |
| 1 | 1 | 1 | 162.80 | 0.72 | 152596.51 |

Structural nested models map a *conditional contrast* of potential outcomes to the treatment, within nested sub-groups of individuals defined by levels of $A_1$, $Z_1$, and $A_0$. Our structural nested model can be written with two equations as

$$E(Y^{a_0,a_1} - Y^{a_0,0}|A_0 = a_0, Z_1 = z_1, A_1 = a_1)$$
$$= a_1(\psi_1 + \psi_2 a_0 + \psi_3 z_1 + \psi_4 a_0 z_1)$$
$$E(Y^{a_0,0} - Y^{0,0}|A_0 = a_0) = \psi_0 a_0$$

Note this model introduces two additional parameters: $\psi_3$ for the two-way interaction between $a_1$ and $z_1$, and $\psi_4$ for the three-way interaction between $a_1$, $z_1$, and $a_0$. Indeed, the ability to explicitly quantify interactions between time-varying exposures and time-varying covariates (which cannot be modeled via standard marginal structural models) is a major strength of structural nested models when effect modification is of interest.[1] To simplify our exposition, we set $(\psi_3, \psi_4) = (0, 0)$ in our data example, allowing us to drop the $\psi_3 z_1$ and $\psi_4 a_0 z_1$ terms from the model. In effect, this renders our structural nested mean model equivalent to a semi-parametric marginal structural model. In the Supplementary Material, we explain how marginal structural and structural nested models each relate to time-varying interactions in more detail.

We can now use gestimation to estimate $(\psi_0, \psi_1, \psi_2)$ in the above structural nested model. Gestimation is based on solving equations that directly result from the sequential conditional exchangeability assumptions in (1) and (2), combined with assumptions implied by the structural nested model. If, at each time point, the exposure is conditionally independent of the potential outcomes (sequential exchangeability) then the conditional covariance between the exposure and potential outcomes is zero.[14] Formally, these conditional independence relations can be written as:

$$0 = \text{Cov}(Y^{a_0,0}, A_1|Z_1, A_0)$$
$$= \text{Cov}(Y^{0,0}, A_0)$$

where $\text{Cov}(\cdot)$ is the well-known covariance formula.[8(p52)] These equalities are of little direct use for estimation, though, as they contain unobserved potential outcomes and are not yet functions of the parameters of interest. However, by counterfactual consistency and the structural nested model, we can replace these unknowns with quantities estimable from the data.

Specifically, as we prove in the Supplementary Material, the structural nested model, together with exchangeability and counterfactual consistency imply that we can replace the potential outcomes $Y^{a_0,0}$ and $Y^{0,0}$ in the above covariance formulas with their values derived from the structural nested model, yielding:

$$0 = \text{Cov}\{Y - A_1(\psi_1 + \psi_2 A_0), A_1 | Z_1, A_0\}$$
$$= \text{Cov}\{Y - A_1(\psi_1 + \psi_2 A_0) - \psi_0 A_0, A_0\}.$$

We provide an intuitive explanation for this substitution in the Supplementary Material. We also show how these covariance relations yield three equations that can be used to solve each of the unknowns in the above structural nested model ($\psi_0, \psi_1, \psi_2$). Two of the three equations yield the following g estimators:

$$\hat{\psi}_{1_{GE}} = \frac{\hat{E}[(1 - A_0)Y\{A_1 - \hat{E}(A_1|Z_1, A_0)\}]}{\hat{E}[(1 - A_0)A_1\{A_1 - \hat{E}(A_1|Z_1, A_0)\}]}$$

$$\hat{\psi}_{1_{GE}} + \hat{\psi}_{2_{GE}} = \frac{\hat{E}[A_0 Y\{A_1 - \hat{E}(A_1|Z_1, A_0)\}]}{\hat{E}[A_0 A_1\{A_1 - \hat{E}(A_1|Z_1, A_0)\}]}$$

Note that to solve these equations we need to model $E(A_1|Z_1, A_0)$, which in practice we might assume can be correctly specified as the predicted values from a logistic model for $A_1$. In our simple setting, the correctness of this model is guaranteed by saturating it (i.e., conditioning the model on $Z_1, A_0$ and their interaction).

As we show in the Supplementary Material, implementing these equations in software can be easily done using either an instrumental variables (i.e., two-stage least squares) estimator, or ordinary least squares.

Once the above parameters are estimated, the next step is to subtract the effect of $A_1$ and $A_1 A_0$ from $Y$ to obtain $\tilde{Y} = Y - \hat{\psi}_{1_{GE}} A_1 - \hat{\psi}_{2_{GE}} A_1 A_0$. We can then solve for the last parameter using a sample version of the third g estimation equality, yielding our final estimator and completing the procedure:

$$\hat{\psi}_{0_{GE}} = \frac{\hat{E}[\tilde{Y}\{A_0 - \hat{E}(A_0)\}]}{\hat{E}[A_0\{A_0 - \hat{E}(A_0)\}]}.$$

Again the above estimator can be implemented using an instrumental variable or ordinary least squares estimator. Implementing this procedure in our example data, we obtain [$\psi_{0_{GE}} = 25.0, \psi_{1_{GE}} = 25.0, \psi_{2_{GE}} = 0.0$], thus yielding $\psi_{GE} = 50.0$.

The potential outcome under no treatment can be thought of as a given subject's baseline prognosis: in our setting, individuals with poor baseline prognosis will have low CD4 levels, no matter what their treatment status may be. In the absence of confounding or selection bias, one expects this baseline prognosis to be independent of treatment status. G estimation exploits this independence by assuming no uncontrolled confounding (conditional on measured confounders), and assigning values to $\hat{\psi}_{GE}$ that render the potential outcomes independent of the exposure. However, assigning the correct values to $\hat{\psi}_{GE}$ depends on there being no confounding or selection bias.

## Discussion

Having constructed these data using the causal diagram shown in Figure 1, we know the true effect of combined treatment is indeed 50 cells/mm$^3$ (25 cells/mm$^3$ for each exposure main effect) as well approximated by all three g methods, but not by any of the standard regression models we fit, with one exception. The final standard result presented in Table 2 correctly estimates the effect of the second treatment (an effect of 25 cells/mm$^3$), as would be expected from the causal diagram.

For the past several years, we have used the foregoing simple example to initiate epidemiologists to g methods with some success. Once having studied this simple example in detail, we recommend working through more comprehensive examples by Robins and Hernán[1] and Hernán and Robins.[16] A recent tutorial[2] may then be of further use. G methods are becoming more common in epidemiologic research.[17] We hope this commentary facilitates the process of better understanding these useful methods.

## References

1. Robins J and Hernán M. Estimation of the causal effects of time-varying exposures. In: Fitzmaurice G, Davidian M, Verbeke G, and Molenberghs G (Eds.) *Advances in Longitudinal Data Analysis*. Boca Raton, FL: Chapman & Hall. 2009; 553–599.
2. Daniel R, Cousens S, De Stavola B, Kenward MG, and Sterne JAC. Methods for dealing with time-dependent confounding. *Stat Med* 2013; **32**:1584–618.
3. Cole SR and Frangakis CE. The consistency statement in causal inference: a definition or an assumption? *Epidemiol* 2009; **20**:3–5.
4. VanderWeele TJ and Hernán MA. Causal inference under multiple versions of treatment. *Journal of Causal Inference* 2013; **1**:1–20.
5. Greenland S and Robins JM. Identifiability, exchangeability, and epidemiological confounding. *Int J Epidemiol* 1986; **15**(3):413–19.
6. Cole SR, Platt RW, Schisterman EF *et al*. Illustrating bias due to conditioning on a collider. *Int J Epidemiol* 2010; **39**:417–420.
7. Westreich D and Cole SR. Invited commentary: Positivity in practice. *Am J Epidemiol* 2010; **171**:674–677.

8.  Wasserman L. *All of Statistics: A Concise Course in Statistical Inference*. New York, NY: Springer, 2005.

9.  Taubman SL, Robins JM, Mittleman MA, and Hernán MA. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *Int J Epidemiol* 2009; **38**:1599–611.

10. Westreich D, Cole SR, Young JG *et al*. The parametric g-formula to estimate the effect of highly active antiretroviral therapy on incident aids or death. *Stat Med* 2012; **31**:2000–2009.

11. Cole SR, Richardson DB, Chu H, and Naimi AI. Analysis of occupational asbestos exposure and lung cancer mortality using the g formula. *Am J Epidemiol* 2013; **177**:989–996.

12. Keil A, Edwards JK, Richardson DB, Naimi AI, and Cole SR. The parametric g-formula for time-to-event data: towards intuition with a worked example. *Epidemiol* 2014; **25**:889–97.

13. Edwards JK, McGrath L, Buckley JP and Schubauer-Berigan, MK *et al*. Occupational radon exposure and lung cancer mortality: Estimating intervention effects using the parametric g-formula. *Epidemiol* 2014; **25**:829–34.

14. Vansteelandt S and Joffe M. Structural nested models and g-estimation: The partially realized promise. *Statist Sci* 2014; **29**:707–731.

15. Robins JM, Mark SD, and Newey WK. Estimating exposure effects by modelling the expectation of exposure conditional on confounders. *Biometrics* 1992; **48**:479–95.

16. Hernán MA and Robins J. *Causal Inference*. Forthcoming. Chapman/Hall, http://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/, accessed 14 Oct 2016.

17. Suarez D, Borras R, and Basagana X. Differences between marginal structural models and conventional models in their exposure effect estimates: a systematic review. *Epidemiol* 2011; **22**:586–588.