



Methods

G-computation for policy-relevant effects of interventions on time-to-event outcomes

Alexander Breskin ,^{1,2*} Andrew Edmonds,² Stephen R Cole,² Daniel Westreich,² Jennifer Cocohoba,³ Mardge H Cohen,⁴ Seble G Kassaye,⁵ Lisa R Metsch,⁶ Anjali Sharma,⁷ Michelle S Williams⁸ and Adaora A Adimora^{2,9}

¹NoviSci Inc, Durham, NC, USA, ²Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, ³Department of Clinical Pharmacy, University of California, San Francisco, CA, USA, ⁴Stroger Hospital, Cook County Bureau of Health Services, Chicago, IL, USA, ⁵Department of Medicine, Georgetown University Medical Center, Washington, DC, USA, ⁶Department of Sociomedical Sciences, Mailman School of Public Health, Columbia University, New York, NY, USA, ⁷Department of Medicine, Albert Einstein College of Medicine, Bronx, NY, USA, ⁸Department of Population Health Science, and the Cancer Center and Research Institute, University of Mississippi Medical Center, Jackson, MS, USA and ⁹Institute for Global Health and Infectious Diseases, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

*Corresponding author. NoviSci Inc PMB 218 201 W Main St, Ste 200 Durham, NC 27701, USA.

E-mail: abreskin@novisci.com

Editorial decision 14 July 2020; Accepted 27 July 2020

Abstract

Background: Parametric g-computation is an analytic technique that can be used to estimate the effects of exposures, treatments and interventions; it relies on a different set of assumptions than more commonly used inverse probability weighted estimators. Whereas prior work has demonstrated implementations for binary exposures and continuous outcomes, use of parametric g-computation has been limited due to difficulty in implementation in more typical complex scenarios.

Methods: We provide an easy-to-implement algorithm for parametric g-computation in the setting of a dynamic baseline intervention of a baseline exposure and a time-to-event outcome. To demonstrate the use of our algorithm, we apply it to estimate the effects of interventions to reduce area deprivation on the cumulative incidence of sexually transmitted infections (STIs: gonorrhoea, chlamydia or trichomoniasis) among women living with HIV in the Women's Interagency HIV Study.

Results: We found that reducing area deprivation by a maximum of 1 tertile for all women would lead to a 2.7% [95% confidence interval (CI): 0.1%, 4.3%] reduction in 4-year STI incidence, and reducing deprivation by a maximum of 2 tertiles would lead to a 4.3% (95% CI: 1.9%, 6.4%) reduction.

Conclusions: As analytic methods such as parametric g-computation become more accessible, epidemiologists will be able to estimate policy-relevant effects of interventions to better inform clinical and public health practice and policy.

Key words: Causal inference, G-computation, survival analysis, HIV, sexually transmitted infections, area deprivation

Key Messages

- Being able to estimate the policy-relevant effects of interventions is increasingly important for informing public health decisions.
- Parametric g-computation can be easily implemented to estimate the effects of interventions on baseline treatments and exposures.

Introduction

The g-methods are a class of causal inference techniques that are increasingly being used for epidemiologic research. These methods can be used to estimate the effects of exposures, treatments and interventions under less restrictive conditions than typical regression-based methods.¹ The three canonical g-methods include parametric g-computation,² inverse probability-weighted estimation³ and g-estimation of structural nested models.⁴ Of these, inverse probability-weighted estimation has been most widely used.⁵

The relatively widespread adoption of inverse probability-weighted estimation may be due to the ease of implementation using standard software.⁶ However, inverse probability weighting is only valid under statistical modelling assumptions, e.g. properly specified models for treatment and censoring. Parametric g-computation requires a different set of statistical modelling assumptions for the outcome and covariates and therefore may be used in cases where investigators have better knowledge of the outcome and covariate processes than the treatment and censoring processes. Ideally, practitioners would use both of these distinct methods in parallel to yield robust results.

Though often perceived as difficult to implement, in many settings parametric g-computation can be implemented with a similar level of ease as inverse probability-weighted estimation. For instance, Snowden *et al.* provided a simple algorithm for parametric g-computation to estimate the effect of a fixed baseline exposure on a continuous outcome.⁷ For more complicated scenarios involving time-varying treatments and confounding, most examples

require far more complex algorithms that require fitting multiple models plus Monte Carlo simulation^{8–11} or estimating a set of iterated conditional expectations.¹²

Here, we extend Snowden *et al.*'s algorithm for cases of intermediate complexity—the effects of dynamic baseline interventions on time-to-event outcomes. Our algorithm can easily be implemented using standard statistical software. In the following sections, we introduce parametric g-computation and dynamic baseline interventions, provide an algorithm for implementing parametric g-computation and illustrate the use of our proposed algorithm to estimate the effects of an intervention to reduce area-level deprivation on sexually transmitted infection (STI) incidence among women with HIV.

Statistical methods

Notation

We index individuals in the sample with i , taking values $1, 2, \dots, n$, where n is the size of the study sample. The exposure for individual i , which may be continuous or categorical, is denoted by X_i . The time that individual i experiences the outcome is denoted T_i , and the end of follow-up is τ . The quantity of interest is the risk of experiencing the outcome by time t , which we denote $Y_{t,i} = I(T_i \leq t)$, where $I(x)$ is the indicator function taking value 1 when x is true. Throughout, capital letters represent random variables, whereas lowercase letters represent their possible realizations.

Some individuals may be lost to follow-up before the end of the study. The time of loss to follow-up is denoted C_i . The observed time that an individual exits the study is

denoted $T_i^* = \min(C_i, T_i, \tau)$, i.e. the first of the following events: loss to follow-up, the outcome or the end of study (τ). Loss to follow-up at time t is denoted by $L_{t,i} = I(C_i \leq t)$. Individuals for whom $T_i^* = C_i$ are missing values for $Y_{t,i}$. Here, we only consider a baseline exposure or treatment, denoted X_i . We also observe a set of baseline covariates, denoted W_i . We assume individuals are independent and identically distributed, and therefore drop the subscript i hereafter for notational simplicity.

Dynamic baseline interventions

The interventions we consider here are deterministic and dynamic,^{13,14} meaning they involve changing the level of the exposure for individuals by a certain known amount (deterministic), with the amount depending on an individual's baseline exposure and covariates (dynamic). In general, the intervention takes the form of a function mapping the observed exposure and covariates to a new level of exposure (we note that, more precisely, such interventions can be described as dynamic interventions that depend on the natural value of treatment¹⁵). For instance, we may be interested in estimating the effect of shifting an ordinal exposure down 1 level from that observed for each individual with covariate value $W = 0$ and 2 levels for those with covariate value $W = 1$. In that case, those in the lowest exposure category have their level of exposure unchanged, and the exposure cannot be changed to be below its lowest possible level. The intervention function for the previously described dynamic baseline intervention can be expressed as $g(x, w) = \max[1, x - I(w = 0) - 2 \times I(w = 1)]$. With this framework, many policy-relevant interventions can be expressed.

Estimating the effects of dynamic baseline interventions

To estimate the effects of dynamic baseline interventions, we must introduce additional notation. Using notation similar to Young *et al.*,¹⁵ we define T^g as the value T would take if, possibly counter to fact, an individual experienced exposure consistent with the intervention represented by g . We similarly define Y_t^g and X^g as the values Y_t and X would take under the intervention represented by g . The effect of a dynamic baseline intervention on the outcome, compared with no intervention, can be expressed as $\psi_t = E[Y_t] - E[Y_t^g]$.

Because the values of T^g (and thus Y_t^g) cannot be observed, a set of identification conditions must be met to estimate ψ_t . One set of sufficient conditions begins with a graphical criterion described in detail by Young *et al.*¹⁵ The criterion requires that, conditional on baseline covariates, all backdoor paths between observed treatment and

the potential outcome are blocked on the single world intervention graph¹⁶ drawn for the intervention of interest. This condition is generally violated when there is unmeasured confounding present.¹⁷ In addition to the graphical criterion, additional conditions that are together sufficient to identify the effect include causal consistency,¹⁸ conditionally uninformative censoring, positivity,¹⁹ and, if parametric models are used, correct model specification. Causal consistency means that the outcome under an observed level of exposure is equal to the outcome that would be observed had exposure been set to that level by an intervention, e.g. $T = T^g$ if $X = X^g$. Causal consistency is typically violated if there are multiple versions of the intervention, referred to as treatment variation relevance,²⁰ or if the outcome for an individual depends on the exposure for other individuals, referred to as interference.²¹ Conditionally uninformative censoring, denoted $T^g \perp\!\!\!\perp C | X, W$, means the potential event time is independent of time of loss to follow-up conditional on observed exposure and the covariates. Finally, positivity means that, within every combination of covariate values, there are individuals observed with every level of exposure that occurs under the intervention and individuals who are not censored at every time, meaning that if $\Pr(W = w) > 0$, then $\Pr(X = x | W = w) > 0$ and $\Pr(L_t = 0 | X = x, W = w) > 0$ for all w and x that occur under the intervention. Violations of positivity occur when individuals with certain covariate values cannot experience all levels of exposure (e.g. men cannot have hysterectomies).

When the aforementioned conditions are met, the risk under the intervention can be expressed with the extended g-formula²² as

$$E[Y_t^g] = \sum_{w,x} \left[\left(1 - \prod_{t=0}^{\tau} \Pr(Y_t = 0 | X = x, W = w, \bar{L}_t = 0) \right) \Pr(g(X, w) = x | W = w) \Pr(W = w) \right]$$

where \bar{L}_t denotes the entire history of censoring through time t . The intervened-on distribution of X is identified by

$$\Pr(g(X, w) = x | W = w) = \sum_k I(g(k, w) = x) \Pr(X = k | W = w).$$

With $E[Y_t^g]$ expressed in this form, all the necessary quantities can be estimated from the observed data. In our proposed algorithm, $\Pr(g(X, w) = x | W = w)$ and $\Pr(W = w)$ are implicitly estimated nonparametrically by averaging the predicted outcomes across observed individuals with each individual's X set to $g(X, W)$. The remaining term, $\Pr(Y_t = 0 | X = x, W = w, \bar{L}_t = 0)$, can be estimated with any properly specified model, for instance, a model for the discrete-time hazard, e.g. with pooled

 Parametric g-computation for a dynamic baseline exposure and time-to-event outcome

- 1: Select the baseline record for each individual.
 - 2: Set X to $g(X, W)$ based on the observed values of X and W for each record.
 - 3: **for** Every time point $k \leq t$ **do**
 Copy the intervened-on baseline record (including the vector of covariates) for each individual.
 Set the time for the copied record to k .
 Set the outcome for the copied record to missing.
 - 4: Fit a model, e.g. pooled logistic regression, to the unaltered observed data, for the discrete-time hazard of the outcome conditional on exposure, covariates, and time.
 - 5: Using the fitted model, compute the predicted probability that $Y_{k,i} = 0$ for each copied record.
 - 6: Compute the cumulative product of the predicted probabilities at each time for each individual using the copied records, $S_i = \prod_{k=1}^t Pr(Y_{k,i} = 0)$.
 - 7: Compute the average, over all n individuals, of the complement of the cumulative products from the previous step, $\frac{1}{n} \sum_{i=1}^n (1 - S_i)$. This is the estimated risk of the outcome by time t under the intervention.
 - 8: Repeat steps 1 – 7 for the other intervention of interest (including, possibly, no intervention).
 - 9: The differences of the risks estimated in steps 1 – 7 under each interventions are the estimates of ψ_t .
 - 10: Estimate 95% confidence intervals with a nonparametric bootstrap.
-

Figure 1. Algorithm for estimating the effects of a dynamic baseline intervention on a time-to-event outcome using g-computation.

logistic regression (for discrete-time data),²³ or a Cox model with the Breslow estimator for the baseline hazard function (for continuous-time data).²⁴ When parametric models are used, the estimation procedure is referred to as parametric g-computation. Standard errors and confidence intervals (CIs) can be estimated using the nonparametric bootstrap.²⁵

Here, to be consistent with the following illustrative example, we consider the discrete-time data case. In practice, ψ_t can be estimated with the procedure displayed in Figure 1, an extension of the parametric g-computation procedure described by Snowden *et al.*⁷ to handle dynamic interventions and time-to-event outcomes. We assume the data are set up so that each row represents a person-visit.

Illustrative example

In this illustrative example, we estimate the effects of an intervention on area deprivation on the incidence of an STI in the Women's Interagency HIV Study (WIHS). The WIHS is a long-running interval cohort study of women living with HIV and women at high risk of HIV.²⁶ Beginning with the biannual WIHS visit that occurred between 1 October 2013 and 31 March 2014, the addresses of WIHS participants from 10 locations were geocoded (Bronx, NY; Brooklyn, NY; Washington, DC; San Francisco, CA; Chicago, IL; Chapel Hill, NC; Atlanta, GA; Miami, FL; Birmingham, AL; Jackson, MS). We linked each woman's census block group at the time of geocoding to 2015 Area Deprivation Index (ADI)²⁷ data, which represented percentile of deprivation across all census block groups in the USA. The ADI is designed to characterize area deprivation based on four domains, including income, education, employment and housing quality. Women with HIV were included in this analysis if (i) at their first

geocoded visit they lived in a block group classified in the 2015 ADI data (i.e. not with <100 persons, <30 housing units, or >33% of the population living in group quarters), and (ii) had at least one visit in which they did not self-report an STI diagnosis visit (here, gonorrhea, chlamydia or trichomoniasis) at or after their geocoding visit. Included women were followed until the first of: one or more self-reported incident STIs (gonorrhea, chlamydia or trichomoniasis) diagnosed since last WIHS visit, loss to follow-up (the projected date of the second consecutive missed study visit), death, or their last attended visit between October 2017 and March 2018.

The intervention considered was an improvement in ADI by k tertiles for each individual, for $k \in \{1, 2\}$. For instance, with $k = 1$, individuals in the bottom tertile of ADI stayed in the bottom tertile, individuals in the middle tertile moved to the bottom tertile, and individuals in the top tertile moved to the middle tertile. Here, the intervened-on value of ADI was $ADI^* = \max(1, ADI - 1)$, where ADI is categorized into 3 levels. The intervention was compared with the natural course, which is the modelled incidence of STI under no intervention.

We modelled the discrete-time hazard of STI with pooled logistic regression. We modelled ADI as a 3-level categorical variable and time with a 3-knot restricted cubic spline. We included several baseline covariates in the model including: age (3-knot restricted cubic spline), race/ethnicity (non-Hispanic White, non-Hispanic Black or any other race/ethnicity combination), detectable viral load (binary, lower limit of detection 20 copies/mL), highly active antiretroviral therapy (HAART) since last visit (binary), CD4 count (3-knot restricted cubic spline), history of AIDS (binary), health insurance (binary), injection drug use (binary), depression (3-knot restricted cubic spline for Center for Epidemiologic Studies Depression Scale score²⁸) and

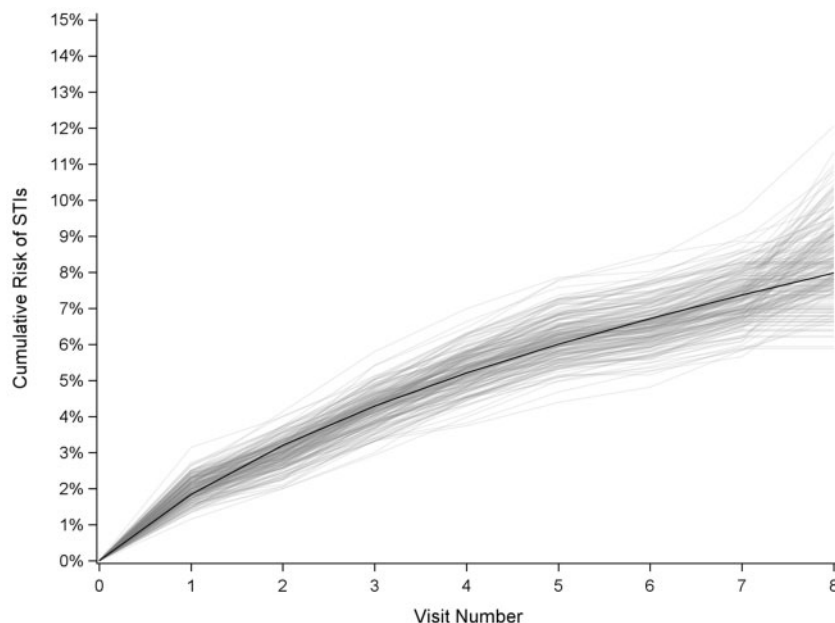


Figure 2. Comparison of the observed and modelled cumulative incidence of STIs among women with HIV in the Women's Interagency HIV Study, 2013–2018. Black line, cumulative incidence function estimated using parametric models; grey lines, cumulative incidence functions estimated non-parametrically from 200 bootstrap replicates from the original data.

sexual risk factors [any vaginal sex (binary), any anal sex (binary), unprotected vaginal sex (always, sometimes, never), unprotected anal sex (always, sometimes, never) and multiple sexual partners (binary)]. We also included product terms between the ADI and time. Women with missing data were excluded from the analysis. We estimated 95% CIs as the 2.5th and 97.5th percentiles of 200 bootstrapped analyses. A SAS macro implementing this analytic strategy is provided in the [Supplementary Material](#), available as [Supplementary data](#) at *IJE* online.

Of 1605 women who had a geocoded address in a classified block group and did not report having been diagnosed with an STI since the last WIHS visit, 1572 women (98%) had complete data. The distribution of ADI was: 743 (47.3%) in the highest tertile (most deprivation), 325 (20.7%) in the middle tertile and 504 (32.1%) in the lowest tertile (least deprivation). Over the course of follow-up, 105 women had an incident STI and 152 were censored due to loss to follow-up ($n = 85$) or death ($n = 67$).

To check our modelling assumptions, we compared the observed risk of survival with the risk from our models under no intervention. The observed estimate of 4-year STI incidence was 8.1% (95% CI: 6.4, 10.8), which was very similar to our modelled 4-year incidence of STI of 8.0% (95% CI: 6.7, 9.6, [Figure 2](#)).

The estimated 4-year STI incidence after intervening to improve ADI by 1 tertile for all individuals was 5.3% (95% CI: 3.5, 7.5, [Figure 3](#)), corresponding to a risk difference of -2.7% (95% CI: -4.3 , -0.1 , [Figure 4](#)). The

estimated 4-year STI incidence after intervening to improve ADI by 2 tertiles for all individuals was 3.7% (95% CI: 1.8, 6.4, [Figure 3](#)), yielding a risk difference of -4.3% (95% CI: -6.4 , -1.9 , [Figure 4](#)). The 4-year risk difference comparing a 1-tertile ADI improvement with a 2-tertile ADI improvement was -1.6% (95% CI: -3.5 , 0.3 , [Figure 4](#)).

Discussion

We present a simple algorithm for implementing parametric *g*-computation to estimate risks under dynamic baseline interventions. Our algorithm can be described succinctly in 6 steps: (i) fit a model for the discrete-time hazard of the outcome conditional on baseline treatment or exposure, baseline covariates, and time, (ii) set the treatment or exposure of each baseline record according to the intervention of interest, (iii) copy the intervened-on baseline records for each time point, (iv) predict the discrete-time hazard for each copied record, (v) compute the cumulative product of the complement of the hazards over time, (vi) average the complement of the cumulative products across individuals. This procedure is intuitively appealing, as it involves directly setting each individual's treatment or exposure based on the intervention and estimating each individual's risk under that intervention. Recently developed SAS²⁹ and R^{30,31} packages that implement parametric *g*-computations and regression standardization can greatly ease the implementation of this approach.

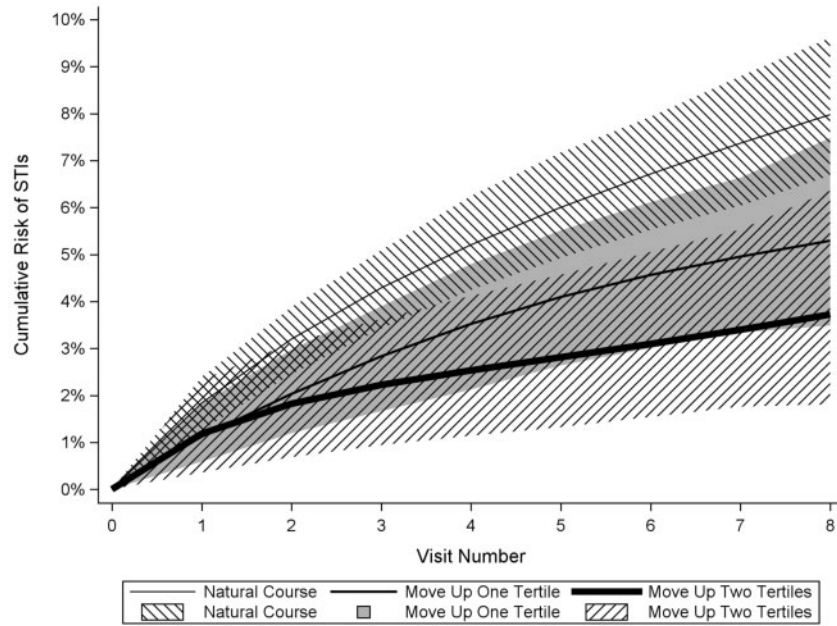


Figure 3. Cumulative incidence of STIs under interventions to reduce area-level deprivation among women with HIV in the Women's Interagency HIV Study, 2013–2018. Thin line, no intervention; medium line, reduce area-level deprivation by 1 tertile; thick line, reduce area-level deprivation by 2 tertiles. Lines are point estimates, shaded regions are 95% CIs.

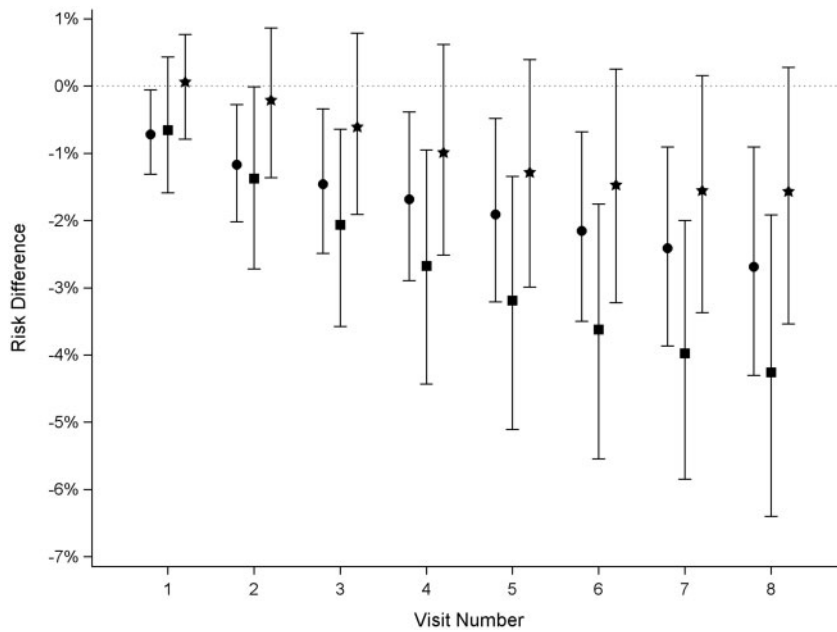


Figure 4. Risk differences comparing the risk of STIs under interventions to reduce area-level deprivation among women with HIV in the Women's Interagency HIV Study, 2013–2018. Circle, reduce area-level deprivation by 1 tertile vs no intervention; square, reduce area-level deprivation by 2 tertiles vs no intervention; star, reduce area-level deprivation by 1 tertile vs reduce area-level deprivation by 2 tertiles. Symbols are point estimates, bars are 95% CIs.

Care should be taken when choosing a method for estimating effects, as each requires distinct conditions to be met. When estimating the effect of a baseline intervention, in addition to the identification assumptions of causal consistency, conditional exchangeability, positivity and no measurement error, the parametric g-formula also requires the discrete-time hazard model be properly specified.

For instance, continuous variables must be modelled flexibly enough to capture their true relationship with the outcome, and effect measure modification must be accounted for, e.g. by including interaction terms in the model. In our example, we used pooled logistic regression, but more flexible modelling approaches may be used as well (we note, however, that inference for g-computation may be invalid when data-

adaptive methods are used³²). Similarly, for inverse probability-weighted estimation, the models for the probabilities of remaining uncensored due to loss to follow-up and due to not following the intervention of interest must be properly specified. The analyst's knowledge of treatment and outcome processes will thus inform the choice of the most appropriate method. Recently, doubly-robust methods that (i) remain valid if either set of modelling conditions, but not necessarily both, are met, and (ii) allow for the use of extremely flexible, data-adaptive machine learning to fit the necessary models, have been proposed,^{33,34} but these methods are beyond the scope of this paper.

One often cited reason for choosing inverse probability-weighted estimation methods is that parametric g-computation may suffer from the g-null paradox.³⁵ Briefly, the g-null paradox states that model misspecification is guaranteed due to the incompatibility of the outcome and treatment models used for estimating effects in the presence of time-varying confounding affected by past exposure. When the g-null paradox applies, then asymptotically one is guaranteed to reject the null hypothesis of no effect of the intervention. However, in the case of a baseline intervention as discussed here, there is no time-varying confounding and only one model is used, so the g-null paradox does not apply.

In our example, we used observational data to estimate the effects of an intervention on area deprivation on the incidence of the combined outcome of at least one of three specific STIs. We note that these represent idealized interventions that may be difficult to implement in practice, and we emphasize that the example is for illustrative purposes and the results should be interpreted with caution. In particular, our example highlights the need to carefully consider the causal consistency assumption, an assumption that is often of concern in health disparities research.^{36,37} The validity of our estimates relies on the strong assumption that individuals set to a given ADI by the intervention will experience social mixing patterns, and thus STI incidence, similar to those of individuals already living in areas with that ADI. However, interventions to improve area deprivation can take at least two distinct forms—moving individuals to better areas or improving the areas individuals live in. Each form may have different effects on a given outcome, and concerns about treatment-variation relevance and interference between individuals must be carefully considered. When estimating the effects of hypothetical interventions, considering the exact form of the intervention and its potential side effects and consequences is necessary to provide unbiased estimates of the effects of future policy decisions.

The algorithm we present can be used to estimate the effects of dynamic baseline interventions, which encompass a large class of interventions that may be of interest

for informing public health practice and policy. Though we did not present more advanced cases, our proposed algorithm can be extended to handle many such situations. These may include the effects of stochastic interventions,³⁸ in which the exact level the treatment takes under the intervention for each individual is random, and the effects of interventions generalized or transported to target populations of interest.^{39–41}

With epidemiology increasingly focusing on consequentialism⁴² and impact,⁴³ being able to estimate the policy-relevant effects of interventions is increasingly important for informing public health decisions. With an understanding of causal inference and the ability to implement modern methods, epidemiologists at all levels will be able to evaluate the public health impacts of implementing policy-oriented interventions within specific populations.

Supplementary data

Supplementary data are available at *IJE* online

Funding

This work was supported by the National Institutes of Health (DP2HD084070 to D.W. and A.B.) and a UNC Gillings Innovation Laboratory Award (S.R.C.). Data in this manuscript were collected by the MACS/WIHS Combined Cohort Study (MWCCS). The contents of this publication are solely the responsibility of the authors and do not represent the official views of the National Institutes of Health (NIH). MWCCS (Principal Investigators): Atlanta CRS (Ighowwerha Oforokun, Anandi Sheth, and Gina Wingood), U01-HL146241-01; Baltimore CRS (Todd Brown and Joseph Margolick), U01-HL146201-01; Bronx CRS (Kathryn Anastos and Anjali Sharma), U01-HL146204-01; Brooklyn CRS (Deborah Gustafson and Tracey Wilson), U01-HL146202-01; Data Analysis and Coordination Center (Gypsyamber D'Souza, Stephen Gange and Elizabeth Golub), U01-HL146193-01; Chicago-Cook County CRS (Mardge Cohen and Audrey French), U01-HL146245-01; Chicago-Northwestern CRS (Steven Wolinsky), U01-HL146240-01; Connie Wofsy Women's HIV Study, Northern California CRS (Bradley Aouizerat and Phyllis Tien), U01-HL146242-01; Los Angeles CRS (Roger Detels and Otoniel Martinez-Maza), U01-HL146333-01; Metropolitan Washington CRS (Seble Kassaye and Daniel Merenstein), U01-HL146205-01; Miami CRS (Maria Alcaide, Margaret Fischl, and Deborah Jones), U01-HL146203-01; Pittsburgh CRS (Jeremy Martinson and Charles Rinaldo), U01-HL146208-01; UAB-MS CRS (Mirjam-Colette Kempf and Deborah Konkle-Parker), U01-HL146192-01; UNC CRS (Adaora Adimora), U01-HL146194-01. The MWCCS is funded primarily by the National Heart, Lung, and Blood Institute (NHLBI), with additional co-funding from the Eunice Kennedy Shriver National Institute Of Child Health & Human Development (NICHD), National Human Genome Research Institute (NHGRI), National Institute On Aging (NIA), National Institute Of Dental & Craniofacial Research (NIDCR), National Institute Of Allergy And Infectious Diseases (NIAID), National Institute Of Neurological Disorders And Stroke (NINDS), National Institute Of Mental

Health (NIMH), National Institute On Drug Abuse (NIDA), National Institute Of Nursing Research (NINR), National Cancer Institute (NCI), National Institute on Alcohol Abuse and Alcoholism (NIAAA), National Institute on Deafness and Other Communication Disorders (NIDCD), National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). MWCCS data collection is also supported by UL1-TR000004 (UCSF CTSA), P30-AI-050409 (Atlanta CFAR), P30-AI-050410 (UNC CFAR), and P30-AI-027767 (UAB CFAR).

Conflict of interest

None declared.

References

- Naimi AI, Cole SR, Kennedy EH. An introduction to g methods. *Int J Epidemiol* 2017;**46**:756–62.
- Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math Model* 1986;**7**:1393–512.
- Breskin A, Cole SR, Westreich D. Exploring the subtleties of inverse probability weighting and marginal structural models. *Epidemiology* 2018;**29**:352–55.
- Robins JM, Hernán MA. Longitudinal data analysis. In: Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G, (eds). *Estimation of the Causal Effects of Time-Varying Exposures*. New York: Chapman and Hall/CRC Press, 2009, pp. 553–99.
- Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol* 2008;**168**: 656–64.
- Hernán MÁ, 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.
- Snowden JM, Rose S, Mortimer KM. Implementation of G-computation on a simulated data set: Demonstration of a causal inference technique. *Am J Epidemiol* 2011;**173**:731–8.
- 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.
- Young JG, Cain LE, Robins JM, O'Reilly EJ, Hernán MA. Comparative effectiveness of dynamic treatment regimes: an application of the parametric g-formula. *Stat Biosci* 2011;**3**: 119–43.
- Keil AP, Edwards JK, Richardson DR, Naimi AI, Cole SR. The parametric G-formula for time-to-event data: towards intuition with a worked example. *Epidemiology* 2014;**25**:889–97.
- Breskin A, Westreich D, Hurt CB *et al*. The effects of hepatitis C treatment eligibility criteria on all-cause mortality among people with human immunodeficiency virus. *Clin Infect Dis* 2019;**69**: 1613–20.
- Neugebauer R, Laan MJ, van der G. computation estimation for causal inference with complex longitudinal data. *Comput Stat Data Anal* 2006;**51**:1676–97.
- Hernán MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. *Basic Clin Pharmacol Toxicol* 2006;**98**:237–42.
- Westreich D. From patients to policy. *Epidemiology* 2017;**28**: 525–28.
- Young JG, Hernán MA, Robins JM. Identification, estimation and approximation of risk und interventions that depend on the natural value of treatment using observational data. *Epidemiol Method* 2014;**3**:1–19.
- Richardson TS, Robins JM. *Single World Intervention Graphs (SWIGs): A Unification of the Counterfactual and Graphical Approaches to Causality*. Center for Statistics and the Social Sciences, Working Paper 128, University of Washington, 2013. Available at: <https://csss.uw.edu/research/working-papers/single-world-intervention-graphs-swigs-unification-counterfactual-and>.
- Greenland S, Robins JM. Identifiability, exchangeability, and epidemiological confounding. *Int J Epidemiol* 1986;**15**:413–9.
- Cole SR, Frangakis CE. The consistency statement in causal inference. *Epidemiology* 2009;**20**:3–5.
- Westreich D, Cole SR. Invited commentary: positivity in practice. *Am J Epidemiol* 2010;**171**:674–7.
- VanderWeele TJ. Concerning the consistency assumption in causal inference. *Epidemiology* 2009;**20**:880–83.
- Hudgens MG, Halloran ME. Toward causal inference with interference. *J Am Stat Assoc* 2008;**103**:832–42.
- Robins JM, Hernán MA, Siebert U. Comparative quantification of health risks: global and regional burden of disease attributable to selected major risk factors. In: Ezzati M, Lopez AD, Rodgers AA, Murray CJ, (eds). *Effects of Multiple Interventions*, Volume 2. Geneva: World Health Organization, 2004. p. 2191–230.
- D'Agostino RB, Lee M-L, Belanger AJ, Cupples LA, Anderson K, Kannel WB. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. *Statist Med* 1990;**9**:1501–15.
- Lin DY. On the Breslow estimator. *Lifetime Data Anal* 2007;**13**: 471–80.
- Efron B, Tibshirani R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Stat Sci* 1986;**1**:54–75.
- Adimora AA, Ramirez C, Benning L *et al*. Cohort profile: The Women's Interagency HIV Study (WIHS). *Int J Epidemiol* 2018;**47**:393–94.
- University of Wisconsin School of Medicine and Public Health. 2015 Area Deprivation Index v2.0. <https://www.neighborhoodatlas.medicine.wisc.edu/> (15 May 2020, date last accessed).
- Radloff LS. The CES-D Scale: A self-report depression scale for research in the general population. In: *Applied Psychological Measurement*, Vol. 1. Thousand Oaks: Sage Publications, 1977, 385–401.
- Logan RW, Young JG, Taubman S, *et al*. GFORMULA SAS MACRO version 3.0. <https://www.hsph.harvard.edu/causal/soft-ware/> (21 October 2019, date last accessed).
- McGrath S, Lin V, Zhang Z *et al*. gfoRmula: An R package for estimating the effects of sustained treatment strategies via the parametric g-formula. *Patterns (NY)* 2020;**1**:100008.
- Sjölander A. Regression standardization with the R package stdReg. *Eur J Epidemiol* 2016;**31**:563–74.
- Naimi AI, Kennedy EH. Nonparametric Double Robustness. arXiv, doi: 1711.07137v1, 20 November 2017, preprint: not peer reviewed.

33. Bang H, Robins JM. Doubly robust estimation in missing data and causal inference models. *Biometrics* 2005;61: 962–72.
34. Schuler MS, Rose S. Targeted maximum likelihood estimation for causal inference in observational studies. *Am J Epidemiol* 2017;185:65–73.
35. Robins JM, Wasserman L Estimation of effects of sequential treatments by reparameterizing directed acyclic graphs. Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence, 1997. p. 409–20.
36. VanderWeele TJ, Robinson WR. On the causal interpretation of race in regressions adjusting for confounding and mediating variables. *Epidemiology* 2014;25:473–84.
37. Glymour MM, Spiegelman D. Evaluating Public Health Interventions: 5. Causal inference in public health research—do sex, race, and biological factors cause health outcomes? *Am J Public Health* 2017;107:81–85.
38. Díaz I, van der Laan MJ. Assessing the causal effect of policies: An example using stochastic interventions. *Int J Biostat* 2013;9: 161–74.
39. Lesko CR, Buchanan AL, Westreich D, Edwards JK, Hudgens MG, Cole SR. Generalizing study results: a potential outcomes perspective. *Epidemiology* 2017;28:553–61.
40. Westreich D, Edwards JK, Lesko CR, Stuart E, Cole SR. Transportability of trial results using inverse odds of sampling weights. *Am J Epidemiol* 2017;186:1010–14.
41. Westreich D, Edwards JK, Lesko CR, Cole SR, Stuart EA. Target validity and the hierarchy of study designs. *Am J Epidemiol* 2019;188:438–43.
42. Galea S. An argument for a consequentialist epidemiology. *Am J Epidemiol* 2013;178:1185–91.
43. Westreich D, Edwards JK, Rogawski ET, Hudgens MG, Stuart EA, Cole SR. Causal impact: epidemiological approaches for a public health of consequence. *Am J Public Health* 2016;106:1011–2.