Harvard University
Harvard University Biostatistics Working Paper Series

Year 2015 *Paper* 190

A general framework for diagnosing confounding of time-varying and other joint exposures

John W. Jackson*

*Harvard University, john.jackson@mail.harvard.edu

A general framework for diagnosing confounding of time-varying and other joint exposures

John W. Jackson

ABSTRACT

Statistical methods for studying the causal effects of time-varying exposures and other joint effects in observational settings (g-methods) are increasingly diffusing into applied practice. Unfortunately, unlike methods for point exposures (e.g. propensity-score matching), there are few tools to help researchers understand the extent and nature of measured confounding for joint exposures, and also how well an applied method resolves it. There is a substantial need for tools that explicitly map to the exchangeability assumptions that undergird causal inference for effects of joint exposures while retaining intuitive interpretations for less technical audiences (e.g. reviewers for applied journals, policy makers, etc.).

Beginning from the implications of exchangeability conditions that hold in hypothetical sequentially randomized trials, we develop covariate balance metrics and trellised plots for time-varying exposures that (1) describe measured confounding (2) assess time-dependent confounding—the indication for g-methods (3) describe residual confounding after certain g-methods have been applied (e.g. inverse probability weighting). We extend this framework to cover distinct point and time-varying exposures, and also data where there is informative censoring.

Using a simulated dataset (n=100,000) of time-varying exposures and covariates (some affected by exposure) under scenarios of model specification, weight truncation, and censoring, we demonstrate that this conceptually grounded framework can aid the transparent design, analysis, and reporting of observational studies for joint exposures.

Correspondence: John W. Jackson, Yerby Postdoctoral Research Fellow, Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, john.jackson@mail.harvard.edu. Acknowledgements: The author would like to acknowledge helpful conversations with James M. Robins and Tyler J. VanderWeele. All errors belong to the author.

INTRODUCTION

Medicine and public health are often concerned with sustained or varying exposures over time, the co-action of multiple exposures, or how exposure effects are mediated. These can all be conceptualized as joint exposures. Whether studies of joint exposures yield valid results hinges upon the ability to measure and adjust for all outcome risk factors that influence each exposure under investigation (e.g. confounders—common causes of exposure and outcome). Typically authors pursuing causal effects will adjust for confounding by estimating exposure-outcome associations within levels of these risk factors or their measurable proxies (covariates). But now it is well known that more sophisticated adjustment methods—g-methods—are required when there is time-dependent confounding i.e. any covariate itself is affected by any exposure under study.[1] Inverse probability weighted estimation of marginal structural models is by far the most popular choice of g-methods and is increasingly appearing substantive and clinical journals.

Unfortunately, there are few practical tools available to diagnose the severity and nature of confounding for joint exposures. Most studies only report how covariates distribute at baseline or over person-time but these portrayals are incomplete. There are also few tools to inform whether g-methods are warranted. Most arguments for their use rest on external subject matter knowledge divorced from empirical arguments because existing tools (e.g. regression) to diagnose time-dependent confounding are tedious and liable to bias in censored data. G-methods involve layers of parametric decision-making to find the optimal tradeoff between validity and statistical precision. But the impact of these choices on residual confounding is expressed in summary statistics that lack intuitive meaning for less technical audiences, including those who review clinical research manuscripts and guide policy.

Here, we develop a general framework for covariate balance based on the implications of exchangeability assumptions for joint exposures, and use it to diagnose: (1) confounding in the study population (2) time-dependent confounding (3) residual confounding after adjustment by inverse probability weighting (see **Appendix** for a special case of the parametric g-formula). Beginning with the setting of a time-varying exposure, we first review the expected statistical relationships between exposures and covariates under sequential randomization. We present the diagnostic framework for a time-varying exposure (**Table 1**) and expand this to distinct point exposures (and also distinct time-varying exposures in the **Appendix**). We then discuss interpretations and issues that arise for censored data and present suitable extensions (**Table 2**). Throughout, we illustrate the approach with simulated data where the data-generating process is known, and address practical issues in the Discussion. We have developed flexible and easy-to-use R code (forthcoming).

NOTATION AND SIMULATED DATA EXAMPLE

For any random variable V, let $\bar{V}(t)$ be a historical vector of its realizations through time t i.e. $V(0), \ldots, V(t)$. Let k index the amount of time by which a variable precedes time t. Let V_x represent the counterfactual value V would have realized had X been set to value x. Finally, let the expression $V \coprod A \mid C$ represent statistical independence between V and A conditional on C. For illustration we simulated a baseline-randomized trial of a time-varying exposure $A(t)$ subject to confounding over follow-up by a vector of covariates $C(t) = (L(t), M(t), N(t), O(t), P(t))$, and also a censoring indicator $S(t)$ with no censoring at baseline (i.e. $P[S(0) = 0] = 1$). Realizations were drawn at times $t = (0,1,2)$ according to the structural equation model in **Figure 1**, except that covariates $N(t)$, $O(t)$, and $P(t)$ were independent of previous exposures $A(t-k-1)$ conditional on full exposure and covariate histories $\bar{A}(t-k-2)$ and $\bar{C}(t-k-1)$; the remaining covariates $L(t)$ and $M(t)$ were time-dependent confounders. Because our metrics ignore data on the outcome Y we did not simulate it. See the **Appendix** for details.

EXCHANGEABILITY AND STATISTICAL EXOGENEITY FOR A TIME-VARYING EXPOSURE

In our simulated baseline-randomized example trial, suppose that we want to estimate causal effects for people who follow different exposure regimes over time. This query compares counterfactual outcomes for the joint set of exposures over time e.g. $E[Y_{\bar{a}'}] - E[Y_{\bar{a}''}]$. Robins[1,2] proved that this contrast can be estimated under certain conditions. Among them, it must be the case for each time t that the exposed and unexposed are exchangeable given their exposure and covariate history up through time $t-1$ i.e. $Y_{\bar{a}} \coprod A(t) | \bar{A}(t-1), \bar{C}(t)$. This untestable assumption requires that for each time t, nature randomly assigned exposure within levels of exposure history and covariate history (a sequentially randomized trial).

As recognized by Hernán and Robins,[3,4] the concept of a sequential randomized trial illuminates notions of confounding for joint exposures. If we could imagine a study where at each time t exposure is randomly assigned given exposure history alone, there would be no confounding because longitudinal analyses condition upon full exposure history. In this sequentially randomized trial, a stronger form of exchangeability $Y_{\bar{a}} \coprod A(t) | \bar{A}(t-1)$ would hold which has clear implications for how data are structured under no time-varying confounding: we would have $\bar{C}(t) \coprod A(t) | \bar{A}(t-1)$, covariate balance within levels of exposure history. This phenomenon, known also as statistical exogeneity, cannot prove exchangeability because unmeasured covariates may be imbalanced across non-randomized exposures. But observing statistical exogeneity does provide reassuring evidence that exposure groups are comparable for measured confounders at least.

DIAGNOSTIFC FRAMEWORK

Overview

Our framework uses covariate balance to examine departures from statistical exogeneity (a metric for measured confounding) in the crude and inverse probability weighted populations (Diagnostics 1 and 3) and also to assess time dependent confounding (Diagnostic 2). In **Table 1** we present diagnostics for a time-varying exposure in uncensored data. Diagnostics for distinct point exposures are briefly covered in the text, and distinct time-varying exposures in the **Appendix**. Extensions for censored data are presented in **Table 2** (for a time-varying exposure) and **Appendix Table 2** (distinct exposures). We use the mean difference of covariates between any arbitrary level of exposure a' and a common referent a'' to measure covariate balance because it contributes to bias expressions for confounding.[5] The framework can use other metrics, even ones that incorporate data on the outcome (see **Appendix**).

The framework can describe confounding for inquiries involving all observed regimes (e.g. to find the optimal one) or ones comparing specific regimes (e.g. always versus never exposed). Our presentation considers all observed regimes for simplicity but we detail application to precise regime contrasts in the **Appendix**.

Diagnostics for a Time-Varying Exposure in the Absence of Censoring

Diagnostic 1: is time-varying confounding present, and how is it patterned over time?

In **Table 1** we present a diagnostic for time-varying confounding: to examine across levels of exposure at each time t, the balance of all preceding covariates among each pattern of exposure history through time $t-1$:

$$E[C(t-k)|A(t)=a', \bar{A}(t-1)] - E[C(t-k)|A(t)=a'', \bar{A}(t-1)]$$

for all $t \in \{0, \ldots, t\}$ and all $k \in \{0, \ldots, t\}$ where $k \leq t'$, and all observed $\bar{A}(t-1)$. The structure of this diagnostic is shown in **Appendix Figure 1a** for the association between $A(t)$ and $\bar{C}(t)$. In **Figure 2**, we applied Diagnostic 1 to our simulated data and present the results as a trellised[6] covariate-balance plot, which belies useful information: imbalance was largest for the most recent covariates and decayed with increasing distance between exposure and covariate times; there was no confounding at $t=0$ but the pattern and magnitude of confounding was the same for times $t=1$ and $t=2$; the least balanced covariates were L (higher among the exposed) and M (higher among the unexposed), and N was the least imbalanced.

Diagnostic 2: is time-dependent confounding present, and how is it patterned over time?

Time-dependent confounding occurs when a time-varying covariate $C(t)$ is affected by any earlier exposure. Robins showed that while one must adjust for such confounders to validly estimate the causal effect of an exposure regime, any estimator that merely stratifies the analysis within levels of these covariates will block exposure effects and suffer from selection-bias.[1,2] But the selection-bias would arise even if time-varying exposures were only associated with subsequent covariates through an unmeasured common cause. Any metric of time-dependent confounding, then, should capture associations between exposures and future covariates that are either causal or due to unmeasured common causes.

Diagnosing time-dependent confounding requires some form of adjustment. Consider **Figure 1**, if we compared the crude covariate prevalence at time $t = 2$ between levels of exposure at time $t = 1$, our metric would be confounded by exposure and measured covariates at time $t = 0$. These confounds do not contribute to the bias from time-dependent confounders that motivates the use of g-methods. One might examine statistical dependence between $C(t)$ and $A(t-k-1)$ within levels of exposure history $\bar{A}(t-k-2)$ and covariate history $\bar{C}(t-k-1)$ but this would be tedious even with regression-based tools. We propose to instead examine such dependence in populations where exposure $A(t-k-1)$ is statistically exogenous, either in a weighted population or within each stratum of the propensity score for $A(t-k-1)$ (where the model for $A(t-k-1)$ includes all measured common causes of $A(t-k-1)$ and Y).

In **Table 1** we propose a novel diagnostic for time-dependent confounding in the absence of censoring: to examine the weighted balance of a covariate at each time t, across levels of exposure at all previous times $t-k-1$, among each pattern of exposure history through time $t-k-2$ (Diagnostic 2a):

$$E\begin{bmatrix} W_{a(t-k-1)} \times I(A(t-k-1)=a'|\bar{A}(t-k-2)) \\ \times C(t) \end{bmatrix} - E\begin{bmatrix} W_{a(t-k-1)} \times I(A(t-k-1)=a''|\bar{A}(t-k-2)) \\ \times C(t) \end{bmatrix}$$

for all $t \in \{0,\dots,t\}$ and all $k \in \{0,\dots,t-1\}$ where $k < t'$, and all observed $\bar{A}(t-k-2)$. This is accomplished using a stabilized inverse probability weight for exposure $A(t-k-1)$

$$W_{a(t-k-1)} = \frac{P[A(t-k-1)=a|\bar{A}(t-k-2)]}{P[A(t-k-1)=a|\bar{A}(t-k-2),\bar{C}(t-k-1)]}$$

which requires a model for exposure $A(t-k-1)$ given exposure history $\bar{A}(t-k-2)$ and covariate history $\bar{C}(t-k-1)$. As depicted in **Appendix Figure 1b**, $A(t-k-1)$ and $\bar{C}(t-k-1)$ will not be

associated within levels of $\bar{A}(t-k-2)$ in the weighted population if the models and weights correctly specified.[7]

Time-dependent confounding can also be assessed using propensity-score stratification (Table 1). The approach examines the balance of a covariate at each time t across levels of exposure at all previous times $t-k-1$, among each level of the propensity score $e_{a(t-k-1)}$ (Diagnostic 2b):

$$E\left[C(t)|A(t-k-1)=a', e_{a(t-k-1)}\right] - E\left[C(t)|A(t-k-1)=a'', e_{a(t-k-1)}\right]$$

for all $t \in \{0,\ldots,t\}$ and all $k \in \{0,\ldots,t-1\}$ where $k < t'$. This metric is estimated for all propensity score levels which are the predicted values from a model for exposure $A(t-k-1)$ given exposure history $\bar{A}(t-k-2)$ and covariate history $\bar{C}(t-k-1)$

$$e_{a(t-k-1)} = P[A(t-k-1)=a'|\bar{A}(t-k-2), \bar{C}(t-k-1)]$$

, but in practice are coarsened into strata (usually quantiles) in which statistical exogeneity $\bar{C}(t-k-1)\coprod A(t-k-1)$ holds for each stratum. As shown in Appendix Figure 1c, $A(t-k-1)$ will not be associated with $\bar{C}(t-k-1)$ within levels of the propensity score $e_{a(t-k-1)}$ provided that the model is correctly specified.[8]

These balance metrics will often lack causal interpretations because investigators rarely identify and measure all common causes of $C(t)$ and $A(t-k-1)$; nevertheless, they capture precisely the associations that lead to selection-bias if time-dependent confounders are conditioned upon while estimating joint effects. In Figure 3, we applied Diagnostic 2a to the simulated data which we report as a trellised plot. The patterns observed in Figure 3 correctly suggest that only covariates L and M were associated with previous exposures: for L the association was strong and positive; for M it was weaker and negative. An investigator studying exposure regimes in this data should consider adjusting for L and M using g-methods.

Diagnostic 3: is residual time-varying confounding present, and how is it patterned over time?

When investigators estimate the average causal effect of a treatment regime using a marginal structural model, cumulative inverse probability weights are estimated for each exposure time, based on models for each exposures $A(t)$ given covariate history $\bar{C}(t)$ and exposure history $\bar{A}(t-1)$ (see Robins[9] and Cole[10] for detailed reviews). Hernán and Robins[7] proved if certain conditions are met, applying these weights to persons following treatment regimes through time t yields a pseudo-population resembling a sequentially randomized trial: at every time t, $A(t)$ is not associated with preceding covariates $\bar{C}(t)$ within levels of exposure history $\bar{A}(t-1)$, as shown in Appendix Figure 1d. The necessary conditions include correct specification of the exposure models and also positivity (non-zero probability of each

exposure level given observed covariate and exposure histories). Residual confounding in the weighted-population will occur when either condition fails or the weights are mis-specified. The standard practice of examining the mean of the stabilized weights captures residual imbalance for covariates that only appear in the denominator.[10]

In Table 1 we present a substantively meaningful diagnostic for residual time-varying confounding in the inverse probability weighted population: to examine for each time t across levels of exposure, the weighted balance of all preceding covariates among each pattern of exposure history through time $t-1$ (Diagnostic 3):

$$E\begin{bmatrix} W_{a(t)} \times I(A(t)=a'|\bar{A}(t-1)) \\ \times C(t-k) \end{bmatrix} - E\begin{bmatrix} W_{a(t)} \times I(A(t)=a''|\bar{A}(t-1)) \\ \times C(t-k) \end{bmatrix}$$

for all $t \in \{0,\dots,t\}$ and $k \in \{0,\dots,t\}$ where $k \leq t'$, and observed $\bar{A}(t-1)$, using an inverse probability weight, perhaps

$$W_{a(t)} = \prod_{k=0}^{k=t} \frac{P[A(t-k)=a|\bar{A}(t-k-1)]}{P[A(t-k)=a|\bar{A}(t-k-1),\bar{C}(t-k)]}$$

to target the average causal effect. This analogue of Diagnostic 1 summarizes residual imbalance in the weighted population. In Figure 4, we applied this diagnostic to the simulated data after constructing correctly specified cumulative inverse probability weights and present the results as a trellised plot. There was no measured confounding in the weighted population.

Residual imbalance in the weighted population can arise from many sources, including: incorrectly specified weights or models, positivity violations, and finite-sample bias. In Figure 5 we demonstrate that Diagnostic 3 captures imbalance from these sources. Noticeable imbalances were induced when the weights were truncated to the 10th and 90th percentiles, and imbalances from model-misspecification and conservative weight truncation were less severe.

Diagnostics for Distinct Point Exposures in the Absence of Censoring

Suppose in our example trial we measured an unrandomized co-exposure Z as shown in Figure 6. Any causal question involving both A and Z pertains to their joint effect (e.g. direct effects or causal interaction between A and Z). We can use exchangeability assumptions to estimate such causal effects under certain conditions.[11] If both A and Z were taken as point exposures, we would require that (a) levels of exposure A are marginally exchangeable $Y_{az} \coprod A$—or within levels of common causes had A not been randomized—and (b) levels of exposure Z are exchangeable within levels of A and L i.e. $Y_{az} \coprod Z|A,L$ (or $Y_{az} \coprod Z|A,L,C$ if A were not randomized). When there is no confounding for the joint

effect, all measured and unmeasured covariates preceding exposure A are independent of A, and those preceding exposure Z are independent of Z within levels of A. This forms the basis of examining measured confounding for joint exposures A and Z (Diagnostic 1):

$$E[C|A = a'] - E[C|A(t) = a'']$$

and

$$E[C|Z = z', A] - E[C|Z = z'', A]$$

for each level of A; one would also check for balance of L across Z within levels of A. (See Jackson et al[12,13] for applied examples). Diagnosing residual confounding after applying inverse probability weights for both A and Z also follows (Diagnostic 3):

$$E[W_{az} \times I(A = a') \times C] - E[W_{az} \times I(A = a'') \times C]$$

and

$$E[W_{az} \times I(Z = z'|A) \times C] - E[W_{az} \times I(Z = z''|A) \times C]$$

For each level of A, where W_{az} is an inverse probability weight for joint exposure to A and Z, perhaps

$$W_{az} = \frac{P[A = a]}{P[A = a|C]} \times \frac{P[Z = z|A]}{P[Z = z|A, C, L]}$$

(one would also check the weighted balance of L across levels of Z within levels of A). To assess time-dependent confounding (Diagnostic 2) one could fit a model for exposure A conditional upon covariates C and evaluate the inverse probability weighted or propensity-score stratified balance of L across levels of A. In the Appendix we provide diagnostics for distinct exposures that vary over time in uncensored and censored data.

Diagnostics for a Time-Varying Exposure in the Presence of Censoring

So far we have implicitly assumed complete follow-up for all subjects. In the real world participants drop out of studies or investigators administratively censor them. We now extend our framework to cover these settings (see Appendix for when censoring is used to focus on specific exposure regimes). Our comments concern a time-varying exposure but also apply to distinct exposures.

Diagnostic 1: is time-varying confounding present, and how is it patterned over time?

In **Table 2** we extend Diagnostic 1 for a time-varying exposure for the uncensored: to examine across levels of exposure the balance of all preceding covariates among each pattern of exposure through time $t-1$:

$$E[C(t-k)|A(t)=a', \bar{A}(t-1), \bar{S}(t)=0] - E[C(t-k)|A(t)=a'', \bar{A}(t-1), \bar{S}(t)=0]$$

for all $t \in \{0, \ldots, t\}$ and $k \in \{0, \ldots, t\}$ where $k \leq t'$ and observed $\bar{A}(t-1)$. In this case, imbalance of covariates $\bar{C}(t-k)$ across level of $A(t)$ still reflect associations between $\bar{C}(t-k)$ and $A(t)$ that are causal or driven by common causes. As seen in **Figure 1**, it also reflects associations that arise from stratifying upon a common effect (conditioning upon $\bar{S}(t)=0$). Regardless of the structure, any non-causal association between $C(t-k)$ and $A(t)$ will bias estimates for causal effects of A on Y so Diagnostic 1 retains a useful interpretation even in the presence of censoring.

Diagnostic 2: is time-dependent confounding present, and how is it patterned over time?

Informative censoring—that which is influenced by exposure and covariates—can bias assessments of time-dependent confounding even when the query concerns a randomized exposure's effect on a covariate (consider open path $A(0) \rightarrow \boxed{S(2)} \leftarrow C(1) \leftarrow U \leftarrow C(2)$ in **Figure 1** among the uncensored). In **Table 2** we propose a novel method for assessing time-dependent confounding in censored data that, when fairly strong assumptions hold, will recover the association between $C(t)$ and $A(t-k-1)$ arising from their causal relationship or unmeasured common causes. Specifically, we modify Diagnostics 2a and 2b to incorporate a cumulative inverse probability of censoring weight that removes the collider-stratification bias when all common causes of $S(t)$ and $C(t)$ are measured and the models and final weight are specified correctly:

(Diagnostic 2a)

$$E\left[\begin{matrix} W_{as(t-k-1)} \times I\big(A(t-k-1)=a'|\bar{A}(t-k-2)\big) \\ \times I(\bar{S}(t)=0) \times C(t) \end{matrix}\right] - E\left[\begin{matrix} W_{as(t-k-1)} \times I\big(A(t-k-1)=a''|\bar{A}(t-k-2)\big) \\ \times I(\bar{S}(t)=0) \times C(t) \end{matrix}\right]$$

for all $t \in \{0, \ldots, t\}$ and all $k \in \{0, \ldots, t-1\}$ where $k < t'$, and observed $\bar{A}(t-k-2)$, where the weight $W_{as(t-k-1)}$ is the product of an exposure weight

$$W_{a(t-k-1)} = \frac{P[A(t-k-1)=a|\bar{A}(t-k-2), \bar{S}(t-k-1)=0]}{P[A(t-k-1)=a|\bar{A}(t-k-2), \bar{C}(t-k-1), \bar{S}(t-k-1)=0]}$$

and a censoring weight

$$W_{s(t)} = \prod_{k=0}^{k=t} \frac{P[S(t-k)=0|\bar{S}(t-k-1)=0]}{P[S(t-k)=0|\bar{S}(t-k-1)=0, \bar{A}(t-k-1), \bar{C}(t-k-1)]}$$

(Diagnostic 2b)

$$E\begin{bmatrix} W_{s(t)} \times I\big(A(t-k-1)=a'|e_{a(t-k-1)}\big) \\ \times I(\bar{S}(t)=0) \times C(t) \end{bmatrix} - E\begin{bmatrix} W_{s(t)} \times I\big(A(t-k-1)=a''|e_{a(t-k-1)}\big) \\ \times I(\bar{S}(t)=0) \times C(t) \end{bmatrix}$$

for all $t \in \{0, \dots, t\}$ and all $k \in \{0, \dots, t-1\}$ where $k < t'$, and all propensity score levels $e_{a(t-k-1)}$ (coarsened into strata), where the propensity score is

$$e_{a(t-k-1)} = P[A(t-k-1)=a'|\bar{A}(t-k-2), \bar{C}(t-k-1), \bar{S}(t-k-1)=0],$$

and the censoring weight is

$$W_{s(t)} = \prod_{k=0}^{k=t} \frac{P[S(t-k)=0|\bar{S}(t-k-1)=0]}{P[S(t-k)=0|\bar{S}(t-k-1)=0, \bar{A}(t-k-1), \bar{C}(t-k-1)]}$$

The modified Diagnostics 2a and 2b for censored data incorporate the same cumulative inverse probability of censoring weight for time t, the product of inverse probability weights from baseline through time t. These require models for each censoring time $t-k$ given exposure history $\bar{A}(t-k-1)$ and covariate history $\bar{C}(t-k-1)$, fit among the uncensored at time $t-k-1$ (note that exposure history does not appear in the numerator). The weights used in Diagnostic 2a and the strata in Diagnostic 2b are based on models for exposure $A(t-k-1)$ fit among the uncensored at $t-k-1$. The final weighted balance metric is estimated among the uncensored at time t, when $C(t)$ is measured.

For Diagnostic 2a the weighted population exhibits: (a) statistical exogeneity for exposure $A(t-k-1)$ i.e. $\bar{C}(t-k-1) \coprod A(t-k-1)|\bar{A}(t-k-2), S(t-k-1)=0$ and (b) censoring as a random process from baseline up through time t i.e. $\bar{S}(t) \coprod \bar{A}(t-1), \bar{C}(t-1)$ (Appendix Figure 2a). For the weighted population in Diagnostic 2b (Appendix Figure 2b), censoring is random but statistical exogeneity for exposure $A(t-k-1)$ only holds among the uncensored within levels of the propensity score $e_{a(t-k-1)}$ (in practice, coarsened strata sufficiently narrow to preserve this property). Thus, Diagnostics 2a and 2b isolate associations between $A(t-k-1)$ and $C(t)$ that reflect their causal relationship or association from unmeasured common causes—these are the associations that will lead to bias when mere stratification is used to adjust joint effect estimates for time-dependent confounders even after accounting for censoring. In Appendix Figure 3, we applied Diagnostic 2a with and without

the censoring weight to examine whether, among those with complete data at time $t = 2$, exposures $A(0)$ and $A(1)$ affect covariates $C(2)$. Before the censoring weight is applied it appears that $O(2)$ might be a time-dependent confounder along with $L(2)$ and $M(2)$, but upon applying the censoring weight we correctly see that it is not.

Diagnostic 3: is residual time-varying confounding present, and how is it patterned over time?

In Table 2 we propose a diagnostic for residual time-varying confounding in weighted uncensored population: to examine for each time t across levels of exposure, the weighted balance of all preceding covariates among each pattern of exposure through time $t - 1$ (Diagnostic 3):

$$E\begin{bmatrix} W_{a(t)} \times I(A(t) = a' | \bar{A}(t-1)) \\ \times I(\bar{S}(t) = 0) \times C(t-k) \end{bmatrix} - E\begin{bmatrix} W_{a(t)} \times I(A(t) = a'' | \bar{A}(t-1)) \\ \times I(\bar{S}(t) = 0) \times C(t-k) \end{bmatrix}$$

for all $t \in \{0, \ldots, t\}$ and $k \in \{0, \ldots, t\}$ where $k \leq t'$, and all observed $A(t-1)$, using an inverse probability weight, perhaps

$$W_{a(t)} = \prod_{k=0}^{k=t} \frac{P[A(t-k) = a | \bar{A}(t-k-1), \bar{S}(t-k) = 0]}{P[A(t-k) = a | \bar{A}(t-k-1), \bar{C}(t-k), \bar{S}(t-k) = 0]}$$

When censoring is informative this metric stratifies upon a collider and bias ensues, as would happen with any naïve analysis that incorrectly assumed non-informative censoring. Thus Diagnostic 3 captures the imbalance from selection-bias induced by informative censoring along with any residual imbalance from weight or model-misspecifications or other sources.

Sometimes investigators who wish to remove the selection-bias from informative censoring will apply cumulative inverse probability weights that address both confounding and selection-bias.[9] This strategy abandons the causal effect of exposure on Y among the uncensored in favor of another estimand, the average causal effect of interventions for exposure and censoring on Y. One can appropriate Diagnostic 3 for distinct exposures in censored data (see Appendix) to explore residual confounding after adjusting for measured confounding and selection-bias in this way.

DISCUSSION

We have outlined a fairly general framework to use covariate balance to assess confounding for one or more point exposures or time-varying exposures in the presence or absence of censoring that examines: (1) the nature of confounding in the study population (2) whether confounders associate with previous exposures (3) the nature of residual confounding after applying certain g-methods. These metrics map precisely to the exchangeability conditions often used to define confounding for

complex exposures, and yet can be succinctly portrayed as a trellised plots that intuitively present a wealth of substantive information.

Diagnostic 1 can aid model parsimony by identifying risk factors most associated with exposure. Diagnostic 2 allows one to examine which covariates, if any, might require adjustment by g-methods. Diagnostic 3 describes provides a rich summary of how well inverse probability weights achieve their aim of emulating a sequentially randomized trial. Sometimes alternate weights are designed to control for a subset of confounders[9,12] and here Diagnostic 3 shows how well the weights balance the covariates the intended covariates while describing remaining confounding to be addressed in the model for the outcome. This framework is also compatible with sub-population estimands such as the average exposure effect among the exposed: Diagnostic 1 reports covariate balance in the full population which contributes to confounding bias for this estimand[5] and Diagnostic 3 shows how well so-called SMR weights[13] reweight the unexposed to mirror the exposed. In the **Appendix**, we consider a fourth diagnostic to examine residual confounding for a special case of the parametric g-formula[14]—an alternative method to adjust for time-dependent confounders.[12]

Some limitations deserve more attention. As often done in point exposure studies, Diagnostics 1 and 3 focuses on marginal balance for each covariate but the exchangeability assumptions that undergird estimation of causal effects imply a much stronger condition: balance of strata formed by covariates. While randomized studies imply both joint and marginal balance, using marginal balance as proxy for joint balance in observational designs (as done here) requires a leap of faith. Another general issue with covariate balance is that it omits data on the outcome (although some welcome this[15]) so inferring the magnitude and direction of confounding bias requires external knowledge. In the **Appendix** we propose a straightforward way to incorporate empirical data on the covariate-outcome relationships to yield bias metrics on the risk difference scale. As others emphasize,[15] finding balance for measured confounders says nothing of balance for unmeasured ones, and should interpreted with care. When Diagnostic 2 is applied to censored data using the censoring weight, residual selection-bias will remain if shared determinants of censoring and covariates go unmeasured, or if censoring is nearly deterministic for some strata and there is strong selection-bias.[16] Addressing some forms of censoring (e.g. mortality or a time-to-event outcome) imply subtle interpretations that we discuss in the **Appendix**. While the finite sample properties of these metrics have not been evaluated, we suspect that they will perform better when used to diagnose confounding for specific regimes of interest (see **Appendix**). One might also consider standardizing balance metrics over exposure history or propensity score strata when these are sparsely populated. When reporting many covariates or measurement times on media with fixed dimensions, it may be preferable to select a subset of times and report separately each row of the trellised plot (as in **Figure 3**) or the diagonal (as in **Figure 4**).

These metrics describe persons who follow exposure regimes in ways that relate to the potential direction, strength, and nature of confounding. In an era of "Big Data" and significant advances in causal modeling, epidemiologists must preserve intimacy with their data and use adaptive tools that portray its complexity in ways that are concise, intuitive, and conceptually grounded. Moreover, intimacy with data should inform our statistical analyses, however complex, and we are responsible for communicating their performance in ways that less technical readers might understand. We have developed tools that we hope embody these values and provide easy-to-use R functions (Austria, Vienna) for their implementation (forthcoming).

REFERENCES

1. Robins JM. A new approach to causal inference in mortality studies with sustained exposure periods - Application to control of the healthy worker survivor effect. *Mathematical Modelling* 1986;**7**:1393-1512.

2. Robins JM. Addendum to "A new approach to causal inference in mortality studies with sustained exposure periods - Application to the control of the healthy worker survivor effect. *Computers and Mathematics with Applications* 1987;**14**:923-945.

3. Robins JM. Marginal Structural Models versus Structural Nested Models as Tools for Causal Inference. In: Holloran M, Berry D, eds. *Statistical Models in Epidemiology: The Environment and Clinical Trials*. New York: Springer-Verlag, 1999;95-134.

4. Hernán MA, Robins JM. Observational Studies. *Causal Inference* Chapman & Hall/CRC Press, 2016.

5. VanderWeele TJ, Arah OA. Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. *Epidemiology* 2011;**22**(1):42-52.

6. Becker RA, Cleveland WS, Shyu M. The visual design and control of trellis display. *Journal of Computational and Graphical Statistics* 1996;**5**(2):123-155.

7. Hernán MA, Brumback B, Robins JA. Marginal structural models to estimate the joint causal effect of nonrandomized treatments. *Journal of the American Statistical Association* 2001;**96**(454):440-448.

8. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;**70**(1):41-55.

9. Robins JM, Hernán MA, Brumback BA. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;**11**(5):550-560.

10. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *American Journal of Epidemiology* 2008;**168**(6):656-664.

11. VanderWeele TJ. Technical Details and Proofs. *Explanation in Causal Inference: Methods for Mediation and Interaction*. New York, NY: Oxford University Press, 2015.

12. Robins JM. Estimation of the causal effects of time-varying exposures. In: Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G, eds. *Longitudinal Data Analysis*. New York, NY: Chapman and Hall/CRC Press, 2009;553-599.

13. Sato T, Matsuyama Y. Marginal structural models as a tool for standardization. *Epidemiology* 2003;**14**(6):680-686.

14. Achy-Brou AC, Frangakis CE, Griswold M. Estimating treatment effects of longitudinal designs using regression models on propensity scores. *Biometrics* 2010;**66**(3):824-833.

15. Rubin DB. On principles for modeling propensity scores in medical research. *Pharmacoepidemiology and Drug Safety* 2004;**13**(12):855-857.

16. Howe CJ, Cole SR, Chmiel JS, Muñoz A. Limitation of inverse probability-of-censoring weights in estimating survival in the presence of strong selection bias. *American Journal of Epidemiology* 2011;**173**(5):569-577.

17. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis.* Use R. New York, NY: Springer-Verlag, 2009.

18. Wickham, H. Tidy Data. *Journal of Statistical Software* 2014;**59**(10):1-23.

19. VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. *Epidemiology* 2009;**20**(1):6-13.

20. Shinohara RT, Narayan AK, Hong K, Kim HS, Coresh J, Streiff MB, Frangakis C. Estmiating parsimonious models of longitudinal causal effects using regressions on propensity scores. *Statistics in Medicine* 2013;**30**(22):3829-3837.

APPENDIX

Extensions to Distinct Time-Varying Exposures

To diagnose confounding for the joint effect of distinct time-varying exposures, we can adapt our framework by conditioning the metrics and requisite models on joint exposure history and assessing each exposure separately. We develop this result for the case of two time-varying exposures where at each time t, exposure $A(t)$ affects covaraites $C(t)$ which affect exposure $Z(t)$, and covariates $C(t)$ affect subsequent exposures $A(t+k)$ and $Z(t+k)$.

Exchangeability conditions for distinct time-varying exposures

To identify the average causal effect of a treatment regime involving time-varying $A(t)$ and $Z(t)$ in observational data, sequential ignorability (along with positivity and consistency) must hold for both $A(t)$ and $Z(t)$ for each time t, given joint exposure history and covariate history (i.e. $Y_{az} \coprod A(t) | \bar{A}(t-1), \bar{Z}(t-1), \bar{C}(t)$ and $Y_{az} \coprod Z(t) | \bar{Z}(t-1), \bar{A}(t), \bar{C}(t)$ for all t).[19] Suppose in a sequentially randomized trial, $A(t)$ and $Z(t)$ were randomly assigned each time t within levels of joint exposure history. Stronger versions of exchangeability $Y_{az} \coprod A(t) | \bar{A}(t-1), \bar{Z}(t-1)$ and $Y_{az} \coprod Z(t) | \bar{Z}(t-1), \bar{A}(t)$ would hold for all t and we would expect statistical exogneiety (covariate balance within joint exposure history) for both $A(t)$ and $Z(t)$ i.e. $\bar{C}(t) \coprod A(t) | \bar{A}(t-1), \bar{Z}(t-1)$ and $\bar{C}(t) \coprod Z(t) | \bar{A}(t), \bar{Z}(t-1)$ for all t. Thus, departures from statistical exogeneity for either exposure A or Z signals measured confounding for their joint effect.

Diagnostic 1: is time-varying confounding present, and how is it patterned over time?

In **Appendix Table 1** we present a simple diagnostic for confounding for two time-varying exposures in uncensored data: to examine at each time t across levels of exposure, the balance of all preceding covariates among each pattern of joint exposure history, first for $A(t)$ and then for $Z(t)$ (Diagnostic 1). This balance metric could be presented as two trellised plots, one for all $A(t)$ and another for all $Z(t)$, where each dot represents the balance for a particular joint exposure history (see **Figure 2** for a similar layout). An extension for censored data is presented in **Appendix Table 2**.

Diagnostic 2: is time-dependent confounding present, and how is it patterned over time?

A covariate in C associated with any preceding exposure A or Z, causally or through unmeasured common causes, is a time-dependent confounder for the average joint effect of $A(t)$ and $Z(t)$ on Y. In **Appendix Table 1**, we present diagnostics for time-dependent confounding when the co-action of two time-varying exposures is estimated in uncensored data. Diagnostic 2a compares the balance of $C(t)$ across levels of all $A(t-k-1)$ within levels of joint exposure history after applying stabilized inverse probability weights for $A(t-k-1)$,

and repeats this assessment for all $Z(t-k-1)$ using weights for $Z(t-k-1)$; note that the weights condition on joint exposure history in the numerator and denominator. The alternate Diagnostic 2b examines the balance of $C(t)$ across levels of all $A(t-k-1)$ within levels of the propensity score for $A(t-k-1)$, and repeats this assessment for $Z(t-k-1)$ within levels of propensity score for $Z(t-k-1)$. Either form of Diagnostic 2 could be presented as two-trellised plots, one for all $A(t-k-1)$ and the other for all $Z(t-k-1)$, where now each dot represents the balance for a particular joint exposure history (see **Figure 3** for a similar layout).

Both approaches isolate the relevant associations between $C(t)$ and $A(t-k-1)$, or between $C(t)$ and $Z(t-k-1)$, when the models for exposures $A(t-k-1)$ and $Z(t-k-1)$ and weights (or strata) are correctly specified (see **Appendix Table 1**). These models must condition on joint exposure history; all common causes of $A(t-k-1)$ and Y should be included in the model for $A(t-k-1)$; and all common causes of $Z(t-k-1)$ and Y should be included in the model for $Z(t-k-1)$. Extensions for censored data are presented in **Appendix Table 2**.

Diagnostic 3: is residual time-varying confounding present, and how is it patterned over time?

Inverse probability weights can be used to adjust for time dependent confounding of the average joint effect of regimes involving $A(t)$ and $Z(t)$; the theory and technical details are covered elsewhere.[19] The weight at each time t is the product of the cumulative inverse probability weights for $A(t)$ and $Z(t)$ that condition upon joint exposure history in both the numerator and denominator, but only covariate history in the denominator. When the requisite models are correctly specified and other conditions are met these weights will create a pseudo-population that resembles a trial where at each time t, both $A(t)$ and $Z(t)$ are randomized within levels of joint exposure history.[7] If the weights or the models for $A(t)$ and $Z(t)$ are incorrectly specified, or there are positivity violations for either $A(t)$ or $Z(t)$, there will be residual confounding in the weighted population.

In **Appendix Table 1** we present a diagnostic for residual confounding of the joint effect of $A(t)$ and $Z(t)$ in the inverse probability weighted population: to examine at each time t across levels of exposure, the weighted balance of all preceding covariates among each pattern of joint exposure history through time $t-1$, first for $A(t)$ and then for $Z(t)$ (Diagnostic 3). These metrics could be presented as two trellised plots, one for all $A(t)$ and the other for all $Z(t)$, where now each dot represents a particular exposure history stratum (see **Figure 4** for a similar layout). This analogue of Dialogue 1 examines residual confounding for joint exposures in the population weighted by the cumulative inverse probability of joint exposure to $\bar{A}$(t) and $\bar{Z}(t)$. As in the case with a single exposure, the diagnostic will capture residual confounding when alternate weight-specifications are used to improve efficiency or target other estimands. Extensions for censored data are presented in **Appendix Table 2**.

Applications to Causal Inquiries Contrasting Specific Regimes

For a single time-varying exposure $A(t)$, it is possible to forgo comparing all exposure regimes and focus on two or more specific ones. For an analogous sequentially randomized trial, this would involve randomizing $A(t)$ for persons with exposure histories that are compatible with at least one exposure regime of interest and terminating the trial for all others, repeating this process until the trial has ended. Persons who at time $t-1$ have followed a regime of interest but their assignment at time t places them in a regime of no interest would have their exposure $A(t)$ recorded, but would not undergo randomization at time $t+1$. They would be administratively censored after being assigned to regimes outside the study's purview.

In such a trial we would have would have exchangeability $Y_{\bar{a}^\dagger} \coprod A(t) | \bar{A}(t-1)^\dagger$ for all times t and all exposure histories $\bar{A}(t-1)^\dagger$, where $Y_{\bar{a}^\dagger}$ corresponds to the counterfactual outcome under an entire regime of interest $\bar{a}^\dagger$, and $\bar{A}(t-1)^\dagger$ represents a strata of exposure history up through time $t-1$ compatible with an entire regime of interest. There would be no confounding for this inquiry and we would observe $\bar{C}(t) \coprod A(t) | \bar{A}(t-1)^\dagger$ i.e. statistical exogeneity within levels of exposure histories through time $t-1$ that are compatible with an entire regime of interest.

Applying Diagnostics 1 and 3 in this setting requires formulae for censored data (**Table 2** and **Appendix Table 2**) where persons are censored (i.e. $S(t)=1$) once their history $\bar{A}(t-1)$ becomes incompatible with all exposure regimes of interest (and $S(t)=0$ otherwise). Diagnostics 1 and 3 then compare the distribution of covariates for those who remain on particular regimes of interest at time t vs. those who deviate from them. Applying Diagnostic 2 also requires formulae for censored data where persons are censored (i.e. $S(t-k-1)=1$) once their history $\bar{A}(t-k-2)$ becomes incompatible with all exposure regimes of interest (and $S(t-k-1)=0$ otherwise), regardless of their exposure trajectories after time $t-k-2$:

(Diagnostic 2a)

$$E\begin{bmatrix} W_{as(t-k-1)} \times I\big(A(t-k-1)=a' | \bar{A}(t-k-2)\big) \\ \times I(\bar{S}(t-k-1)=0) \times C(t) \end{bmatrix} - E\begin{bmatrix} W_{as(t-k-1)} \times I\big(A(t-k-1)=a'' | \bar{A}(t-k-2)\big) \\ \times I(\bar{S}(t-k-1)=0) \times C(t) \end{bmatrix}$$

for all $t \in \{0, \ldots, t\}$ and all $k \in \{0, \ldots, t-1\}$ where $k < t'$, and observed $\bar{A}(t-k-2)$ compatible with regimes of interest, where $W_{a(t-k-1)}$ is

$$W_{a(t-k-1)} = \frac{P[A(t-k-1)=a | \bar{A}(t-k-2), \bar{S}(t-k-1)=0]}{P[A(t-k-1)=a | \bar{A}(t-k-2), \bar{C}(t-k-1), \bar{S}(t-k-1)=0]}$$

(Diagnostic 2b)

$$E\begin{bmatrix} W_{s(t)} \times I\big(A(t-k-1)=a'|e_{a(t-k-1)}\big) \\ \times I(\bar{S}(t-k-1)=0) \times C(t) \end{bmatrix} - E\begin{bmatrix} W_{s(t)} \times I\big(A(t-k-1)=a''|e_{a(t-k-1)}\big) \\ \times I(\bar{S}(t-k-1)=0) \times C(t) \end{bmatrix}$$

for all $t \in \{0,\dots,t\}$ and all $k \in \{0,\dots,t-1\}$ where $k < t'$, and all propensity score levels $e_{a(t-k-1)}$, where the propensity score is

$$e_{a(t-k-1)} = P[A(t-k-1)=a''|\bar{A}(t-k-2),\bar{C}(t-k-1),\bar{S}(t-k-1)=0]$$

(the exposure weight and propensity score definitions are of the same form as proposed in Table 2 and Appendix Table 2; the key difference is how the censoring indicator is defined for the balance metric). These modified versions of Diagnostics 2a and 2b examine whether exposures $A(t-k-1)$ are associated with future covariates $C(t)$ among those whose exposure history $\bar{A}(t-k-2)$ is compatible with a regime of interest. Similar results apply in the setting of distinct time-varying exposures, say $A(t)$ and $Z(t)$. For Diagnostics 1 and 3 censoring $S(t)=0$ if exposure histories $\bar{A}(t-1)$ and $\bar{Z}(t-1)$ are both compatible with a joint exposure regime of interest and $S(t)=1$ otherwise. In the case of Diagnostic 2 censoring $S(t-k-1)=0$ if exposure histories ($\bar{A}(t-k-2)$ and $\bar{Z}(t-k-2)$ are both compatible with a joint exposure regime of interest and $S(t-k-1)=1$ otherwise.

If there are other censoring mechanisms at play (e.g. study dropout) one can simply estimate the these metrics using additional censoring indicators indexed at time t and applying cumulative inverse probability of censoring weights for these mechanisms, as shown in Table 2 and Appendix Table 2.

Extension to the Parametric G-formula

A special case of the parametric g-formula for longitudinal settings uses strata from models for longitudinal propensity score histories $\bar{e}_{a(t)}$ conditional on propensity score history $\bar{e}_{a(t-1)}$ and exposure history $\bar{A}(t-1)$ to replace models for covariate histories $\bar{C}(t)$.

Achy-Brou et al[14] proved that if $Y_{\bar{a}} \coprod A(t)|\bar{C}(t),\bar{A}(t-1)$ holds then $Y_{\bar{a}} \coprod A(t)|\bar{e}_{a(t)},\bar{A}(t-1)$ holds as well, implying statistical exogeneity $\bar{C}(t) \coprod A(t)$ within joint levels of $\bar{e}_{a(t)}$ and $\bar{A}(t-1)$. It would be infeasible to check for covariate balance within joint levels of all propensity scores $\bar{e}_{a(t)}$ and exposure history $\bar{A}(t-1)$, but such balance relies on balance within levels of each $e_{a(t)}$ and exposure history $\bar{A}(t-1)$ as a prerequisite.

Therefore, as proposed by Shinohara et al[20], diagnosing residual confounding for applications of the parametric g-formula that rely on models for propensity-score history in place of models for covariate history could proceed as follows: evaluate for each time t across levels of exposure, the balance of all preceding covariates $C(t-k)$ among each stratum of the longitudinal propensity score $e_{a(t)}$ and each pattern of exposure history

up through time $t-1$ (for each distance k where $k \leq t$). The results can be plotted in the same format as figure 4.

Bias Metrics for Measured Confounding

The framework outlined in the main text proposes to use covariate balance

$E[C(t-k)|A(t)=a', \bar{A}(t-1)] - E[C(t-k)|A(t)=a'', \bar{A}(t-1)]$ (Diagnostic 1)
or

$$E\left[W(t) \times I\left(A(t)=a' \middle| \bar{A}(t-1) \times C(t)\right)\right] - E\left[W(t) \times I\left(A(t)=a'' \middle| \bar{A}(t-1) \times C(t)\right)\right]$$

(Diagnostic 3) to diagnose confounding, which omit data on the outcome. Consequently the balance metrics do not describe the magnitude and direction of bias from failing to adjust for a covariate.

It follows from VanderWeele and Arah 2010[5] that if one failed to adjust for a dichotomous confounder $C(t-k)$ while estimating the additive average causal effect for exposure $A(t)$ on outcome Y, the bias would equal (if covariates did not interact with each other or A to cause Y):

$$\sum_{\bar{A}(t-1)} (E[C(t-k)|A(t)=a', \bar{A}(t-1)] - E[C(t-k)|A(t)=a'', \bar{A}(t-1)]) \times (E[Y|C(t-k)=1|A(t)=a'', \bar{A}(t-1)] - E[Y|C(t-k)=0|A(t)=a'', \bar{A}(t-1)])$$

Thus, we can incorporate data on the outcome for binary covariates (and indicator variables for categorical ones) in a principled way by multiplying our proposed balance metrics by the difference in mean outcome across levels of the covariate, conditional on the unexposed at time t and exposure history through time $t-1$:

$$E[Y|C(t-k)=1|A(t)=a'', \bar{A}(t-1)] - E[Y|C(t-k)=0|A(t)=a'', \bar{A}(t-1)]$$

in the case of Diagnostic 1, and

$$E\left[W(t) \times I\left(C(t)=1 \middle| A(t)=a'', \bar{A}(t-1) \times Y(t)\right)\right] - E\left[W(t) \times I\left(C(t)=0 \middle| A(t)=a'', \bar{A}(t-1) \times Y(t)\right)\right]$$

in the case of Diagnostic 3 where $W(t)$ represents the cumulative inverse probability of exposure weight at time t. For uncensored data one would simply estimate them among the uncensored at time t. This bias metric will equal zero when either (a) the covariate C is not associated with the exposure A or (b) the covariate C is not associated with the outcome Y. As with covariate balance, it is not an aggregate summary of bias. It does, however, allow for a fairer assessment of the strength and direction of confounding across covariates. The results of VanderWeele and Arah[5] provide more general bias expressions than the one applied here.

Subtle Interpretations for Diagnostic 2 in Censored Data

In the main text we presented our diagnostic framework in the context of an outcome Y measured at the end of follow-up and expanded our assessment of time-dependent confounding (Diagnostic 2) to the setting where censoring occurs, proposing the use of a cumulative inverse probability of censoring weight (IPCW) to remove the selection-bias from conditioning on a common effect of exposures and covariates. Standard implementations of IPCW[9] (used alongside weights for exposure) to estimate causal effects on Y suppose a joint intervention on exposure and co-intervention to enforce study retention, which may be easy to imagine for processes involving attrition but more challenging for those of mortality, competing risks, or medical time-to-event phenomena (**Appendix Figure 4**). Using IPCW in those settings targets our inference to a world where no one dies, or has a competing risk, or has an event. These populations, arguably, are less interesting for estimating policy-relevant estimates of causal effects. However, they remain relevant in our diagnostic setting where the goal of Diagnostic 2 is to isolate and describe associations between exposures and future covariates that are causal or from unmeasured common causes.

Details of Example Simulations

We simulated a baseline-randomized trial (n=100,000) of a time-varying exposure $A(t)$ subject to confounding over follow-up by a vector of covariates $C(t) = (L(t), M(t), N(t), O(t), P(t))$, and also a censoring indicator $S(t)$ with no censoring at baseline (all realized at times $t = (0,1,2)$) (**Figure 1**). Each covariate in $C(t)$ was specified as a function of its prior realizations and an unmeasured variable U, and also exposure $A(t)$ in the case of time-dependent covariates $L(t)$ and $M(t)$. For times $t > 0$ exposure $A(t)$ depended on entire exposure and covariate histories $\bar{A}(t)$ and $\bar{C}(t)$ and also the interactions $M(t-1) \times N(t-1)$ and $O(t-1) \times P(t-1)$. For time $t = 0$ there was no censoring but for times $t > 0$ censoring depended on full exposure and covariate histories $\bar{A}(t)$ and $\bar{C}(t)$.

Realizations for all variables $V(t)$ were simulated as random bernoulli draws with probability $p_v(t)$. For $A(0)$, $S(0)$, and also U, the probability was specified directly. For all other variables we used linear combinations on the log odds scale to calculate $p_v(t)$. This setup allowed for encoding of time-varying confounding by both immediate and distant covariates, temporally indexed by $t-k$. The expanded linear combinations we specified are listed below, where $\beta_x^v(t-k)$ is a vector of log odds ratios expressing effect of some variable X (measured at time $t-k$) on variable V (measured at time t), where $1 \leq k \leq t$ when $X = V$, and possibly $0 \leq k \leq t$ otherwise):

$$\text{logit}\, p_l(t) = \beta_{0l(t)} + \beta_u^l u + \sum_{k=1}^{k=t} L'(t-k)\beta_l^l(t-k) + \sum_{k=1}^{k=t} A'(t-k)\beta_a^l(t-k)$$

$$\text{logit}\, p_m(t) = \beta_{0m(t)} + \beta_u^m u + \sum_{k=1}^{k=t} M'(t-k)\beta_m^m(t-k) + \sum_{k=1}^{k=t} A'(t-k)\beta_a^m(t-k)$$

$$\text{logit}\, p_n(t) = \beta_{0n(t)} + \beta_u^n u + \sum_{k=1}^{k=t} N'(t-k)\beta_n^n(t-k)$$

$$\text{logit}\, p_o(t) = \beta_{0o(t)} + \beta_u^o u + \sum_{k=1}^{k=t} O'(t-k)\beta_o^o(t-k)$$

$$\text{logit}\, p_p(t) = \beta_{0p(t)} + \beta_u^p u + \sum_{k=1}^{k=t} P'(t-k)\beta_p^p(t-k)$$

$$\begin{aligned}\text{logit}\, p_a(t) = {} & \beta_{0a(t)} + \sum_{k=1}^{k=t} A'(t-k)\beta_a^a(t-k) + \sum_{k=0}^{k=t} L'(t-k)\beta_l^a(t-k) \\ & + \sum_{k=0}^{k=t} M'(t-k)\beta_m^a(t-k) + \sum_{k=0}^{k=t} N'(t-k)\beta_n^a(t-k) \\ & + \sum_{k=0}^{k=t} O'(t-k)\beta_o^a(t-k) + \sum_{k=0}^{k=t} P'(t-k)\beta_p^a(t-k) \\ & + \beta_{m\times n}^a(t) \times (M(t-k) \times N(t-k)) + \beta_{o\times p}^a(t) \times (O(t-k) \times P(t-k))\end{aligned}$$

$$\begin{aligned}\text{logit}\, p_s(t) = {} & \beta_{0s(t)} + \sum_{k=1}^{k=t} A'(t-k)\beta_a^s(t-k) + \sum_{k=1}^{k=t} L'(t-k)\beta_l^s(t-k) \\ & + \sum_{k=1}^{k=t} M'(t-k)\beta_m^s(t-k) + \sum_{k=1}^{k=t} N'(t-k)\beta_n^s(t-k) \\ & + \sum_{k=1}^{k=t} O'(t-k)\beta_o^s(t-k) + \sum_{k=1}^{k=t} P'(t-k)\beta_p^s(t-k)\end{aligned}$$

To obtain censored data used in **Appendix Figure 3** we repeated the simulation but censored records once $S(t) = 1$ was realized, and to obtain data with positivity violations used in Figure 5c, 5d, and 5e we set $p_a(t) = 1/10{,}000{,}000$ whenever $L(t) = 0$ and $O(t) = 1$. We report the specific parameter values for all $\beta_{0v(t)}$ and $\beta_x^v(t-k)$ on directed acyclic graphs in **Appendix Figure 5**.

TABLES

Table 1. Balance metrics for confounding of a time-varying exposure $A(t)$ by time-varying risk-factors $C(t)$ in the absence of censoring $S(t)$	
Diagnostic 1: Time-varying confounding i.e. $C(t-k)$ across levels of $A(t)$ (for all $t \in \{0, \dots, t\}$ and $k \in \{0, \dots, t\}$ where $k \leq t'$)	
Balance metric	Definitions
$E[C(t-k)\|A(t)=a', \bar{A}(t-1)]$ $-E[C(t-k)\|A(t)=a'', \bar{A}(t-1)]$	N/A
Diagnostic 2: Time-dependent confounding i.e. $C(t)$ across levels of $A(t-k-1)$ (for all $t \in \{0, \dots, t\}$ and all $k \in \{0, \dots, t-1\}$ where $k < t'$)	
Balance metric	
a. by inverse probability weighting	Weight $W_{a(t-k-1)}$
$E\left[\begin{matrix}W_{a(t-k-1)} \times I(A(t-k-1)=a'\|\bar{A}(t-k-2)) \\ \times C(t)\end{matrix}\right]$ $-E\left[\begin{matrix}W_{a(t-k-1)} \times I(A(t-k-1)=a''\|\bar{A}(t-k-2)) \\ \times C(t)\end{matrix}\right]$	$W_{a(t-k-1)}$ $= \frac{P[A(t-k-1)=a\|\bar{A}(t-k-2)]}{P[A(t-k-1)=a\|\bar{A}(t-k-2), \bar{C}(t-k-1)]}$
b. by propensity score stratification	Propensity score $e_{a(t-k-1)}$
$E\left[C(t)\|A(t-k-1)=a', e_{a(t-k-1)}\right]$ $-E\left[C(t)\|A(t-k-1)=a'', e_{a(t-k-1)}\right]$	$e_{a(t-k-1)}$ $= P[A(t-k-1)=a'\|\bar{A}(t-k-2), \bar{C}(t-k-1)]$
Diagnostic 3: Residual time-varying confounding i.e. $C(t-k)$ across levels of $A(t)$ (for all $t \in \{0, \dots, t\}$ and $k \in \{0, \dots, t\}$ where $k \leq t'$)	
Balance metric	Weight $W_{a(t)}$
$E\left[\begin{matrix}W_{a(t)} \times I(A(t)=a'\|\bar{A}(t-1)) \\ \times C(t-k)\end{matrix}\right]$ $-E\left[\begin{matrix}W_{a(t)} \times I(A(t)=a''\|\bar{A}(t-1)) \\ \times C(t-k)\end{matrix}\right]$	$W_{a(t)}$ $= \prod_{k=0}^{k=t} \frac{P[A(t-k)=a\|\bar{A}(t-k-1)]}{P[A(t-k)=a\|\bar{A}(t-k-1), \bar{C}(t-k)]}$

Table 2. Balance metrics for confounding of a time-varying exposure $A(t)$ by time-varying risk-factors $C(t)$ in the presence of censoring $S(t)$

Diagnostic 1: Time-varying confounding i.e. $C(t-k)$ across levels of $A(t)$ (for all $t \in \{0, \dots, t\}$ and $k \in \{0, \dots, t\}$ where $k \leq t'$)	
Balance metric	Definitions
$E[C(t-k)\|A(t)=a', \bar{A}(t-1), \bar{S}(t)=0]$ $- E[C(t-k)\|A(t)=a'', \bar{A}(t-1), \bar{S}(t)=0]$	N/A
Diagnostic 2: Time-dependent confounding i.e. $C(t)$ across levels of $A(t-k-1)$ (for all $t \in \{0, \dots, t\}$ and all $k \in \{0, \dots, t-1\}$ where $k < t'$)	
Balance metric	
a. by inverse probability weighting	Weight $W_{as(t-k-1)} = W_{a(t-k-1)} \times W_{s(t)}$
$E\begin{bmatrix} W_{as(t-k-1)} \times I(A(t-k-1)=a'\|\bar{A}(t-k-2)) \\ \times I(\bar{S}(t)=0) \times C(t) \end{bmatrix}$ $-E\begin{bmatrix} W_{as(t-k-1)} \times I(A(t-k-1)=a''\|\bar{A}(t-k-2)) \\ \times I(\bar{S}(t)=0) \times C(t) \end{bmatrix}$	$W_{a(t-k-1)} = \frac{P[A(t-k-1)=a\|\bar{A}(t-k-2), \bar{S}(t-k-1)=0]}{P[A(t-k-1)=a\|\bar{A}(t-k-2), \bar{C}(t-k-1), \bar{S}(t-k-1)=0]}$ $W_{s(t)} = \prod_{k=0}^{k=t} \frac{P[S(t-k)=0\|\bar{S}(t-k-1)=0]}{P[S(t-k)=0\|\bar{S}(t-k-1)=0, \bar{A}(t-k-1), \bar{C}(t-k-1)]}$
b. by propensity score stratification	Propensity score $e_{a(t-k-1)}$ and weight $W_{s(t)}$
$E\begin{bmatrix} W_{s(t)} \times I(A(t-k-1)=a'\|e_{a(t-k-1)}) \\ \times I(\bar{S}(t)=0) \times C(t) \end{bmatrix}$ $-E\begin{bmatrix} W_{s(t)} \times I(A(t-k-1)=a''\|e_{a(t-k-1)}) \\ \times I(\bar{S}(t)=0) \times C(t) \end{bmatrix}$	$e_{a(t-k-1)} = P[A(t-k-1)=a''\|\bar{A}(t-k-2), \bar{C}(t-k-1), \bar{S}(t-k-1)=0]$ $W_{s(t)} = \prod_{k=0}^{k=t} \frac{P[S(t-k)=0\|\bar{S}(t-k-1)=0]}{P[S(t-k)=0\|\bar{S}(t-k-1)=0, \bar{A}(t-k-1), \bar{C}(t-k-1)]}$
Diagnostic 3: Residual time-varying confounding i.e. $C(t-k)$ across levels of $A(t)$ (for all $t \in \{0, \dots, t\}$ and $k \in \{0, \dots, t\}$ where $k \leq t'$)	
Balance metric	Weight $W_{a(t)}$
$E\begin{bmatrix} W_{a(t)} \times I(A(t)=a'\|\bar{A}(t-1)) \\ \times I(\bar{S}(t)=0) \times C(t-k) \end{bmatrix}$ $- E\begin{bmatrix} W_{a(t)} \times I(A(t)=a''\|\bar{A}(t-1)) \\ \times I(\bar{S}(t)=0) \times C(t-k) \end{bmatrix}$	$W_{a(t)} = \prod_{k=0}^{k=t} \frac{P[A(t-k)=a\|\bar{A}(t-k-1), \bar{S}(t-k)=0]}{P[A(t-k)=a\|\bar{A}(t-k-1), \bar{C}(t-k), \bar{S}(t-k)=0]}$

Appendix Table 1. Balance metrics for confounding of distinct time-varying exposures $A(t)$ and $Z(t)$ by time-varying risk-factors $C(t)$ in the absence of censoring

Diagnostic 1: Time-varying confounding i.e. $C(t-k)$ across levels of $A(t)$ (for all $t \in \{0, \dots, t\}$ and $k \in \{0, \dots, t\}$ where $k \leq t'$)	
Balance metrics	Definitions
$E[C(t-k)\|A(t)=a', \bar{A}(t-1), \bar{Z}(t-1)]$ $- E[C(t-k)\|A(t)=a'', \bar{A}(t-1), \bar{Z}(t-1)]$	N/A
$E[C(t-k)\|Z(t)=z', \bar{Z}(t-1), \bar{A}(t)]$ $- E[C(t-k)\|Z(t)=z'', \bar{Z}(t-1), \bar{A}(t)]$	N/A
Diagnostic 2: Time-dependent confounding i.e. $C(t)$ across levels of $A(t-k-1)$ (for all $t \in \{0, \dots, t\}$ and all $k \in \{0, \dots, t-1\}$ where $k < t'$)	
Balance metrics	
a. by inverse probability weighting	Weight $W_{a(t-k-1)}$ and $W_{z(t-k-1)}$
$E\left[\begin{matrix}W_{a(t-k-1)} \times I(A(t-k-1)=a'\|\bar{A}(t-k-2)) \\ \times C(t)\end{matrix}\right]$ $- E\left[\begin{matrix}W_{a(t-k-1)} \times I(A(t-k-1)=a''\|\bar{A}(t-k-2)) \\ \times C(t)\end{matrix}\right]$	$W_{a(t-k-1)} = \frac{P[A(t-k-1)=a\|\bar{A}(t-k-2)]}{P[A(t-k-1)=a\|\bar{A}(t-k-2), \bar{Z}(t-k-2), \bar{C}(t-k-1)]}$
$E\left[\begin{matrix}W_{z(t-k-1)} \times I(Z(t-k-1)=a'\|\bar{Z}(t-k-2)) \\ \times C(t)\end{matrix}\right]$ $- E\left[\begin{matrix}W_{z(t-k-1)} \times I(Z(t-k-1)=z''\|\bar{Z}(t-k-2)) \\ \times C(t)\end{matrix}\right]$	$W_{z(t-k-1)} = \frac{P[Z(t-k-1)=z\|\bar{Z}(t-k-2)]}{P[Z(t-k-1)=z\|\bar{Z}(t-k-2), \bar{A}(t-k-1), \bar{C}(t-k-1)]}$
b. by propensity score stratification	Propensity score $e_{a(t-k-1)}$
$E\left[C(t)\|A(t-k-1)=a', e_{a(t-k-1)}\right]$ $- E\left[C(t)\|A(t-k-1)=a'', e_{a(t-k-1)}\right]$	$e_{a(t-k-1)}$ $= P[A(t-k-1)=a'\|\bar{A}(t-k-2), \bar{Z}(t-k-2), \bar{C}(t-k-1)]$
$E\left[C(t)\|Z(t-k-1)=z', e_{z(t-k-1)}\right]$ $- E\left[C(t)\|Z(t-k-1)=z'', e_{z(t-k-1)}\right]$	$e_{z(t-k-1)}$ $= P[Z(t-k-1)=z'\|\bar{Z}(t-k-2), \bar{A}(t-k-1), \bar{C}(t-k-1)]$
Diagnostic 3: Residual time-varying confounding i.e. $C(t-k)$ across levels of $A(t)$ (for all $t \in \{0, \dots, t\}$ and $k \in \{0, \dots, t\}$ where $k \leq t'$ for Z and $k < t'$ for A)	
Balance metrics	Weight $W_{az(t)} = W_{a(t)} \times W_{z(t)}$
$E\left[\begin{matrix}W_{az(t)} \times I(A(t)=a'\|\bar{A}(t-1)), \bar{Z}(t-2) \\ \times C(t-k)\end{matrix}\right]$ $- E\left[\begin{matrix}W_{az(t)} \times I(A(t)=a''\|\bar{A}(t-1)), \bar{Z}(t-2) \\ \times C(t-k)\end{matrix}\right]$	$W_{a(t)} = \prod_{k=0}^{k=t} \frac{P[A(t-k)=a\|\bar{A}(t-k-1), \bar{Z}(t-k-2)]}{P[A(t-k)=a\|\bar{A}(t-k-1), \bar{Z}(t-k-2), \bar{C}(t-k-1)]}$
$E\left[\begin{matrix}W_{az(t)} \times I(Z(t)=z'\|\bar{Z}(t-1)), \bar{A}(t-1) \\ \times C(t-k)\end{matrix}\right]$ $- E\left[\begin{matrix}W_{az(t)} \times I(Z(t)=z''\|\bar{Z}(t-1)), \bar{A}(t-1) \\ \times C(t-k)\end{matrix}\right]$	$W_{z(t)} = \prod_{k=0}^{k=t} \frac{P[Z(t-k)=z\|\bar{Z}(t-k-1), \bar{A}(t-k-1)]}{P[Z(t-k)=z\|\bar{Z}(t-k-1), \bar{A}(t-k-1), \bar{C}(t-k)]}$

Appendix Table 2. Balance metrics for confounding of distinct time-varying exposures $A(t)$ and $Z(t)$ by time-varying risk-factors $C(t)$ in the presence of censoring

Balance metrics	Definitions
Diagnostic 1: Time-varying confounding i.e. $C(t-k)$ across levels of $A(t)$ (for all $t \in \{0, \dots, t\}$ and $k \in \{0, \dots, t\}$ where $k \leq t'$)	
$E[C(t-k)\|A(t)=a', \bar{A}(t-1), \bar{Z}(t-1), \bar{S}(t)=0]$ $- E[C(t-k)\|A(t)=a'', \bar{A}(t-1), \bar{Z}(t-1), \bar{S}(t)=0]$	N/A
$E[C(t-k)\|Z(t)=z', \bar{Z}(t-1), \bar{A}(t), \bar{S}(t)=0]$ $- E[C(t-k)\|Z(t)=z'', \bar{Z}(t-1), \bar{A}(t), \bar{S}(t)=0]$	N/A
Diagnostic 2: Time-dependent confounding i.e. $C(t)$ across levels of $A(t-k-1)$ (for all $t \in \{0, \dots, t\}$ and all $k \in \{0, \dots, t-1\}$ where $k < t'$)	
Balance metrics	
a. by inverse probability weighting	Weight $W_{az(t-k-1)} = W_{a(t-k-1)} \times W_{s(t)}$ and $W_{zs(t-k-1)} = W_{z(t-k-1)} \times W_{s(t)}$
$E\begin{bmatrix} W_{as(t-k-1)} \times I(\bar{S}(t)=0) \\ \times I(A(t-k-1)=a'\|\bar{A}(t-k-2)) \times C(t) \end{bmatrix}$ $-\begin{bmatrix} W_{as(t-k-1)} \times I(\bar{S}(t)=0) \\ \times I(A(t-k-1)=a''\|\bar{A}(t-k-2)) \times C(t) \end{bmatrix}$	$W_{a(t-k-1)} = \frac{P[A(t-k-1)=a\|\bar{A}(t-k-2), \bar{S}(t-k-1)=0]}{P[A(t-k-1)=a\|\bar{A}(t-k-2), \bar{Z}(t-k-2), \bar{C}(t-k-1), \bar{S}(t-k-1)=0]}$
$E\begin{bmatrix} W_{zs(t-k-1)} \times I(\bar{S}(t)=0) \\ \times I(Z(t-k-1)=z'\|\bar{Z}(t-k-2)) \times C(t) \end{bmatrix}$ $-\begin{bmatrix} W_{zs(t-k-1)} \times I(\bar{S}(t)=0) \\ \times I(Z(t-k-1)=z''\|\bar{Z}(t-k-2)) \times C(t) \end{bmatrix}$	$W_{z(t-k-1)} = \frac{P[Z(t-k-1)=z\|\bar{Z}(t-k-2, \bar{S}(t-k-1)=0)]}{P[Z(t-k-1)=z\|\bar{Z}(t-k-2), \bar{A}(t-k-1), \bar{C}(t-k-1), \bar{S}(t-k-1)=0]}$ $W_{s(t)} = \prod_{k=0}^{k=t} \frac{P[S(t-k)=0\|\bar{S}(t-k-1)=0]}{P[S(t-k)=0\|\bar{S}(t-k-1)=0, \bar{Z}(t-k-1), \bar{A}(t-k-1), \bar{C}(t-k-1)]}$
b. by propensity score stratification	Propensity scores $e_{a(t-k-1)}$ and $e_{z(t-k-1)}$ and Weight $W_{s(t)}$
$E\begin{bmatrix} W_{s(t)} \times I\big(A(t-k-1)=a'\|e_{a(t-k-1)}\big) \\ \times I(\bar{S}(t)=0) \times C(t) \end{bmatrix}$ $-\begin{bmatrix} W_{s(t)} I\big(A(t-k-1)=a''\|e_{a(t-k-1)}\big) \\ \times I(\bar{S}(t)=0) \times C(t) \end{bmatrix}$	$e_{a(t-k-1)} = P[A(t-k-1)=a'\|\bar{A}(t-k-2), \bar{Z}(t-k-2), \bar{C}(t-k-1), \bar{S}(t-k-1)=0]$ $e_{z(t-k-1)} = P[Z(t-k-1)=z'\|\bar{Z}(t-k-2), \bar{A}(t-k-1), \bar{C}(t-k-1), \bar{S}(t-k-1)=0]$ $W_{s(t)} = \prod_{k=0}^{k=t} \frac{P[S(t-k)=0\|\bar{S}(t-k-1)=0]}{P[S(t-k)=0\|\bar{S}(t-k-1)=0, \bar{Z}(t-k-1), \bar{A}(t-k-1), \bar{C}(t-k-1)]}$
$E\begin{bmatrix} W_{s(t)} \times I\big(Z(t-k-1)=z'\|e_{z(t-k-1)}\big) \\ \times I(\bar{S}(t)=0) \times C(t) \end{bmatrix}$ $-\begin{bmatrix} W_{s(t)} \times I\big(Z(t-k-1)=z''\|e_{z(t-k-1)}\big) \\ \times I(\bar{S}(t)=0) \times C(t) \end{bmatrix}$	
Diagnostic 3: Residual time-varying confounding i.e. $C(t-k)$ across levels of $A(t)$ (for all $t \in \{0, \dots, t\}$ and $k \in \{0, \dots, t\}$ where $k \leq t'$ for Z and $k < t'$ for A)	
Balance metrics	Weight $W_{az(t)} = W_{a(t)} \times W_{z(t)}$
$E\begin{bmatrix} W_{az(t)} \times I(A(t)=a'\|\bar{A}(t-1)), \bar{Z}(t-2) \\ \times I(\bar{S}(t)=0) \times C(t-k) \end{bmatrix}$ $- E\begin{bmatrix} W_{az(t)} \times I(A(t)=a''\|\bar{A}(t-1)), \bar{Z}(t-2) \\ \times I(\bar{S}(t)=0) \times C(t-k) \end{bmatrix}$	$W_{a(t)} = \prod_{k=0}^{k=t} \frac{P[A(t-k)=a\|\bar{A}(t-k-1), \bar{Z}(t-k-2), \bar{S}(t-k)=0]}{P[A(t-k)=a\|\bar{A}(t-k-1), \bar{Z}(t-k-2), \bar{C}(t-k-1), \bar{S}(t-k)=0]}$
$E\begin{bmatrix} W_{az(t)} \times I(Z(t)=z'\|\bar{Z}(t-1)), \bar{A}(t-1) \\ \times I(\bar{S}(t)=0) \times C(t-k) \end{bmatrix}$ $- E\begin{bmatrix} W_{az(t)} \times I(Z(t)=z''\|\bar{Z}(t-1)), \bar{A}(t-1) \\ \times I(\bar{S}(t)=0) \times C(t-k) \end{bmatrix}$	$W_{z(t)} = \prod_{k=0}^{k=t} \frac{P[Z(t-k)=z\|\bar{Z}(t-k-1), \bar{A}(t-k-1), \bar{S}(t-k)=0]}{P[Z(t-k)=z\|\bar{Z}(t-k-1), \bar{A}(t-k-1), \bar{C}(t-k), \bar{S}(t-k)=0]}$

FIGURE LEGENDS

Figure 1. Directed Acyclic Graph for data generating mechanism of simulated data examples (excluding the outcome Y, see Appendix for details).

Figure 2. A trellised covariate balance plot for Diagnostic 1 in the uncensored simulated data. The plot is comprised of sub-plots indexed by specific exposure measurement times (rows) and covariate measurement times (columns). Each sub-plot thus presents the mean difference across of each covariate C (i.e. L, M, N, O, P) assessed at time $t-k$, comparing exposures $A=1$ and $A=0$ at time t, for each pattern of exposure history through time $t-1$ (dots). When all of the dots in a horizontal plane align at zero there is no confounding of exposure assessed at any time t by the confounder assessed at time $t-k$. When this holds for all covariates in each sub-plot, the exposures are statistically exogenous and there is no measured time-varying confounding. Note that the sub-plots along the trellis diagonal running from the lower left to upper right constitute an assessment of covariate balance for the most recent covariates i.e. $C(t)$, which many applied studies privilege for adjustment.

Figure 3. A row from the full trellised covariate balance plot for Diagnostic 2a to examine whether covariates are time-dependent confounders (here, those measured at time $t=2$). In the full plot, exposure measurement times would be indexed by columns and covariate measurement times by rows. Each subplot presents the mean difference of each covariate C (i.e. L, M, N, O, P) assessed at time t, comparing exposures $A=1$ and $A=0$ at time $t-k-1$, for each pattern of exposure history through time $t-k-2$ (dots). If for a given subpanel, a dot aligns at zero, then that covariate measurement is not associated with prior exposure assessed at time $t-k-1$, and if this holds for all previous exposure measurement times (and covariate measurement times) then that covariate is not a time-dependent confounder. If this holds for all subpanels, no measured covariates are time-dependent confounders. Note that in the full plot, sub-plots along the trellis diagonal running from lower left to upper right constitute an assessment of whether covariates are associated with the most recent exposure i.e. $A(t-1)$.

Figure 4. The diagonal from the full trellised covariate balance plot for Diagnostic 3 for residual time-varying confounding in the weighted population. Here the diagonal shows that the weights balanced covariates most proximal to exposure (and all other covariates as well, not shown). In the full plot, exposure measurement times would be indexed by rows and covariate measurement times by columns. The interpretation is similar to the one given for Diagnostic 1: each subplot presents the weighted balance of covariates C (i.e. L, M, N, O, P) assessed at times $t-k$, comparing the exposures $A=1$ and $A=0$ at time t, among each pattern of exposure history through time $t-1$ (dots). There is no confounding of $A(t)$ by a given covariate when all dots in a horizontal plane align at zero; when this holds for all covariates in all subplots in the full plot, there is no confounding by measured covariates in the weighted population. Note that in the full plot, the sub-plots along the trellis diagonal running from the lower left to upper right constitute an assessment of

covariate balance for the most recent covariates i.e. $C(t)$, which many applied studies privilege for adjustment.

Figure 5. Demonstration of Diagnostic 3 for residual imbalance in the weighted population under various scenarios. (a) correct specification of the exposure model. (b) mis-specification of the exposure model by only including recent covariates $C(t-1)$ and omitting covariate-covariate interactions (c) random non-positivity specified in the data-generating model for $A(t)$ such that $P[A(t)|\bar{A}(t-1), \bar{C}(t)] = 1/10{,}000{,}000$ when $L(t) = 0$ and $O(t) = 1$. (e) weights truncated to 1st and 99th percentiles under injected positivity violation. (f) weights truncated to 10th and 90th percentiles under injected positivity violation.

Figure 6. Directed Acyclic Graph for example of two distinct point exposures A (randomized) and Z (not randomized). Here, L represents a time-dependent confounder because it is an effect of A but cause of Z and Y.

Appendix Figure 1. Directed Acyclic Graphs for each diagnostic in the absence of censoring. (a) Diagnostic 1 for assessing $A(t)$'s association with $C(t)$. (b) Diagnostic 2a implemented by inverse probability weighting to assess whether $C(t)$ is affected by or associated with $A(t-1)$ not through measured common causes (c) Diagnostic 2b by propensity score stratification to assess whether $C(t)$ is affected by or associated with $A(t-1)$ not through measured common causes. (d) Diagnostic 3 to assess residual associations between $C(t)$ and $A(t)$ after applying cumulative inverse probability weights for $A(t)$. Bold arrows represent associations captured by the diagnostic and (for diagnostics that apply weights) dashed arrows represent associations that the weights are meant to remove. Note that in (a) only the causal path between $C(t)$ and $A(t)$ is bolded but many other non-causal paths also contribute to Diagnostic 1.

Appendix Figure 2. Directed Acyclic Graphs for each diagnostic in the presence of censoring. (a) Diagnostic 1 for assessing $A(t)$'s association with $C(t)$. (b) Diagnostic 2a implemented by inverse probability weighting to assess whether $C(t)$ is affected by or associated with $A(t-1)$ not through measured common causes or selection-bias (c) Diagnostic 2b to assess whether $C(t)$ is affected by or associated with $A(t-1)$ not through measured common causes or selection-bias. (d) Diagnostic 3 to assess residual associations between $C(t)$ and $A(t)$ after applying cumulative inverse probability weights for $A(t)$. Bold arrows represent associations captured by the diagnostic and (for diagnostics that apply weights) dashed arrows represent associations that the weights are meant to remove. Note that in (a) only the causal path between $C(t)$ and $A(t)$ is bolded but many other non-causal paths also contribute to Diagnostic 1.

Appendix Figure 3. Demonstration of Diagnostic 2a for time-dependent confounding for censored data using inverse probability weighting (a) omitting the cumulative inverse probability of censoring weight (b) incorporating the cumulative inverse probability of censoring weight.

Appendix Figure 4. Directed Acyclic Graph for a time-to-event or competing risk $Y(t)$ (e.g. death), time-varying exposure A, and time-dependent confounders C.

Appendix Figure 5. Directed Acyclic Graph for simulation parameters involving a vector of covariates $C(t)$. Values assigned to an edge represent conditional odds ratios given the child node's parents. $S(1)$ and $S(2)$ are not shown, and deterministic arrows from $S(t)$ are present when $S(t)$ is conditioned upon, representing whether the child node is observed. The base line odds were (Figure 5a) 0 for $S(0)$, 0.15 for $S(1)$, 0.15 for $S(2)$, 0.67 for U, 0.5 for $A(0)$, 0.4 for $A(1)$, 0.3 for $A(2)$; (Figure 5b) 0.3 for $L(0)$, $L(1)$, and $L(2)$; (Figure 5c) 0.75 for $M(0)$, $M(1)$, and $M(2)$; (Figure 5d) 0.5 for $N(0)$, $N(1)$, and $N(2)$; (Figure 5e) 0.6 for $O(0)$, $O(1)$, and $O(2)$; (Figure 5f) 0.4 for $P(0)$, $P(1)$, and $P(2)$.

FIGURES

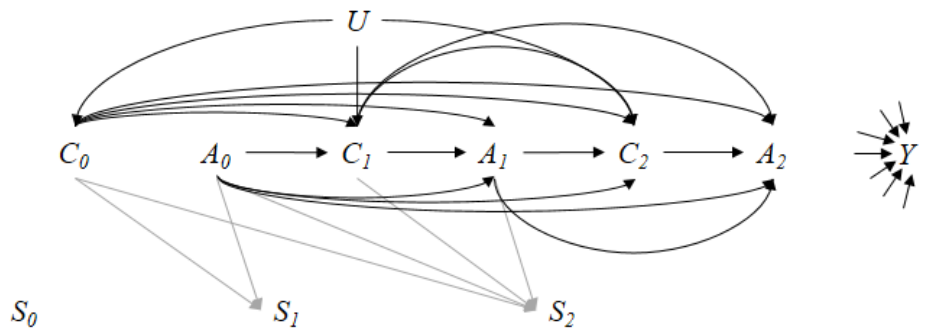


Figure 1.

Figure 2.

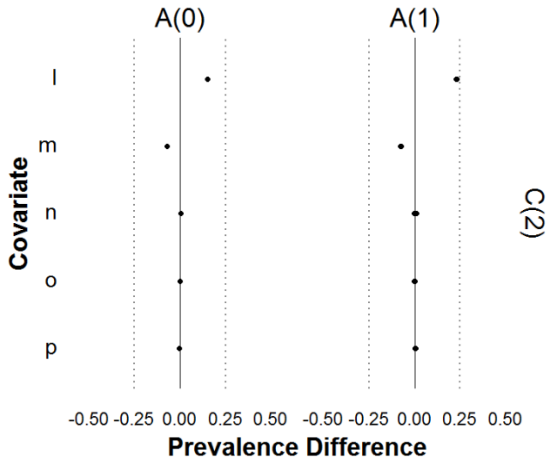


Figure 3.

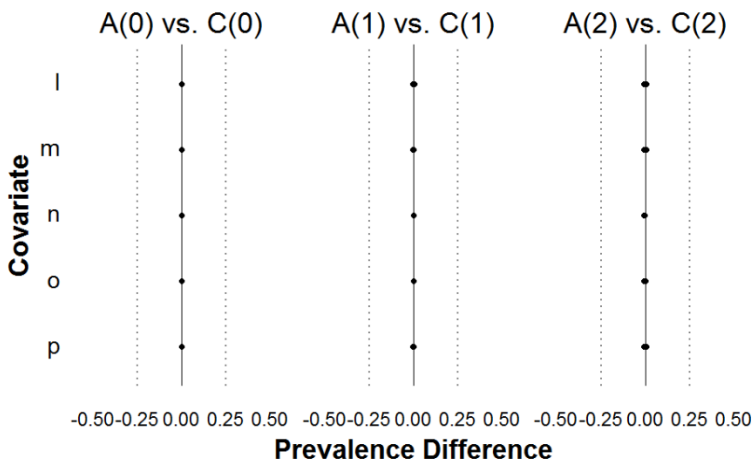


Figure 4.

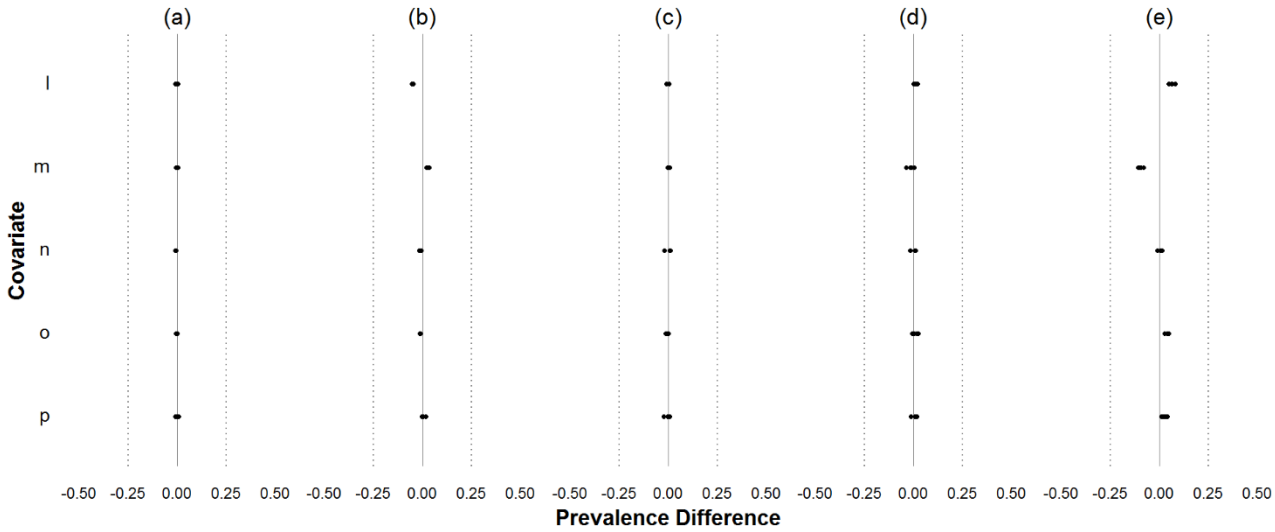


Figure 5.

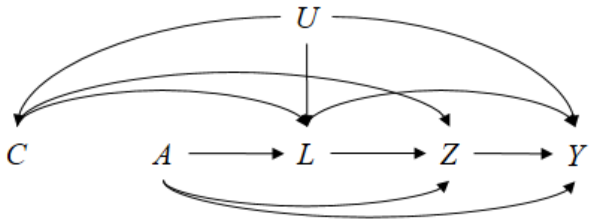
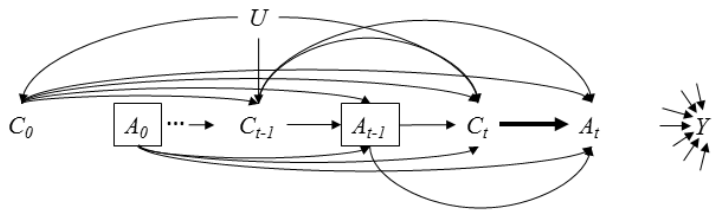
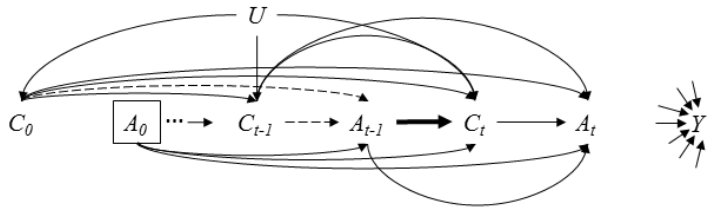


Figure 6.

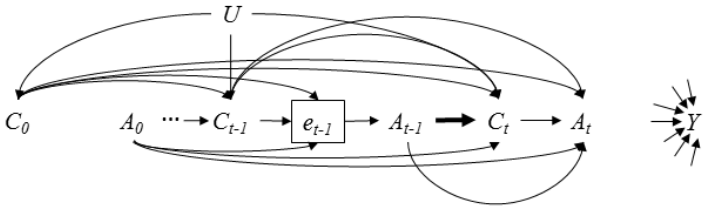
(a)



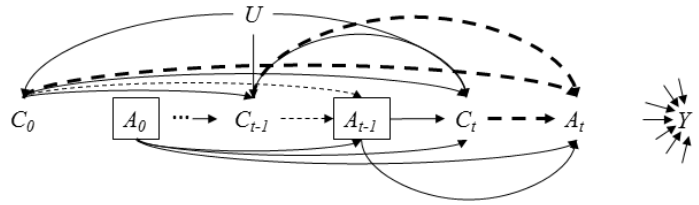
(b)



(c)

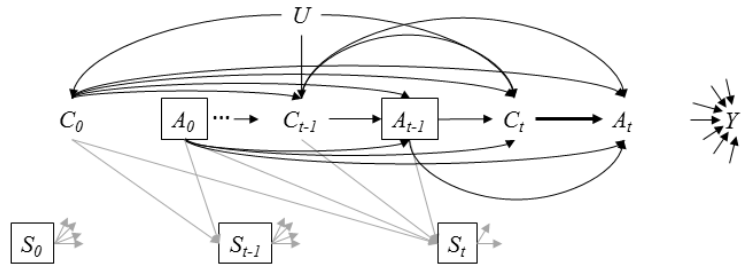


(d)

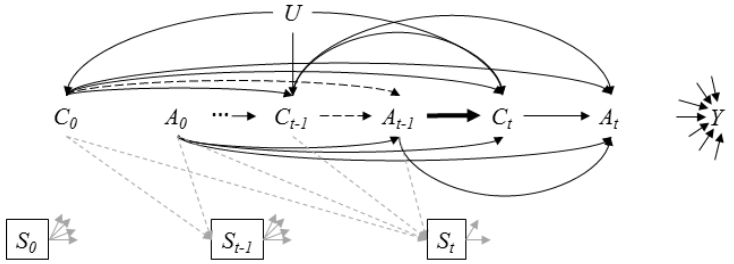


Appendix Figure 1.

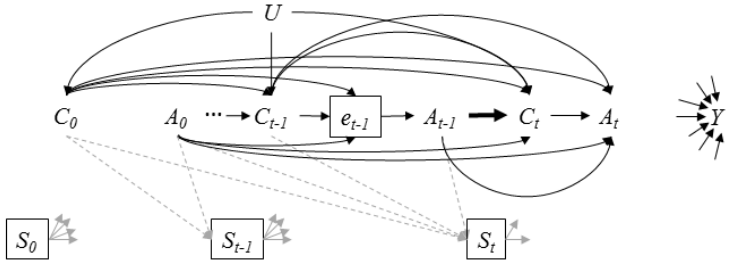
(a)



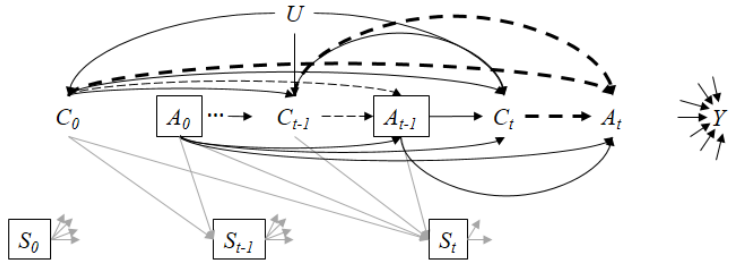
(b)



(c)

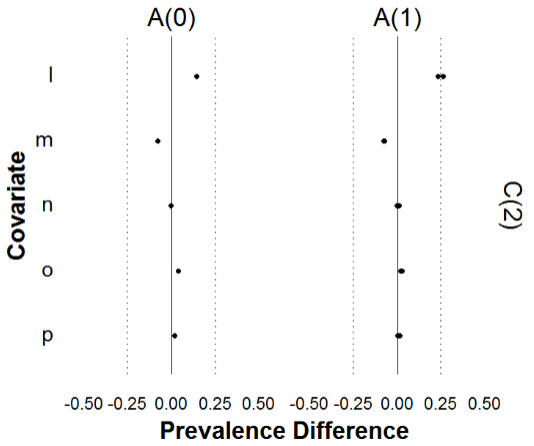


(d)

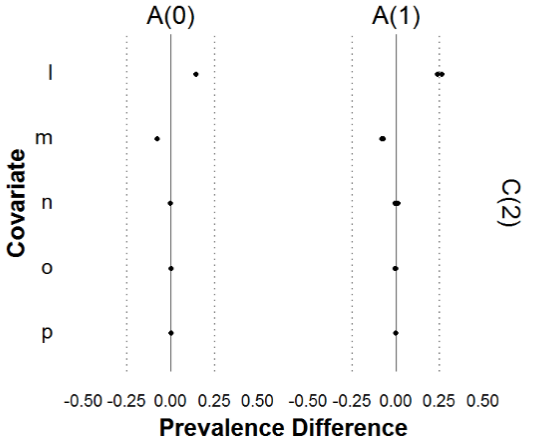


Appendix Figure 2.

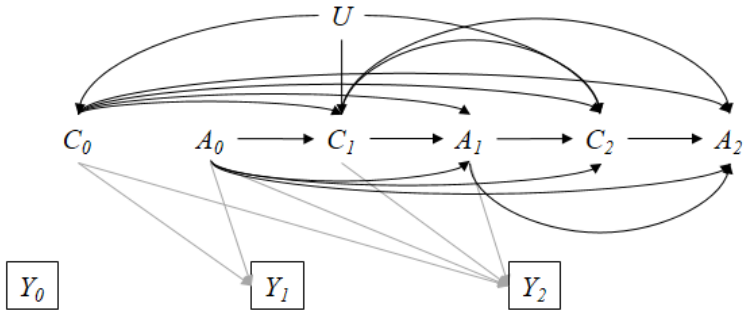
(a)



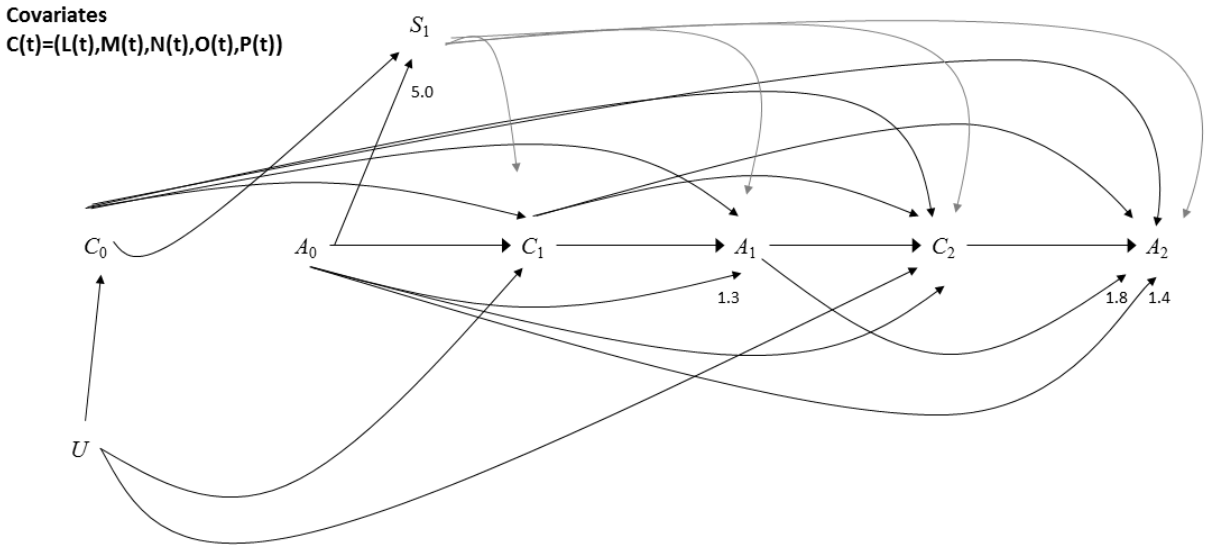
(b)



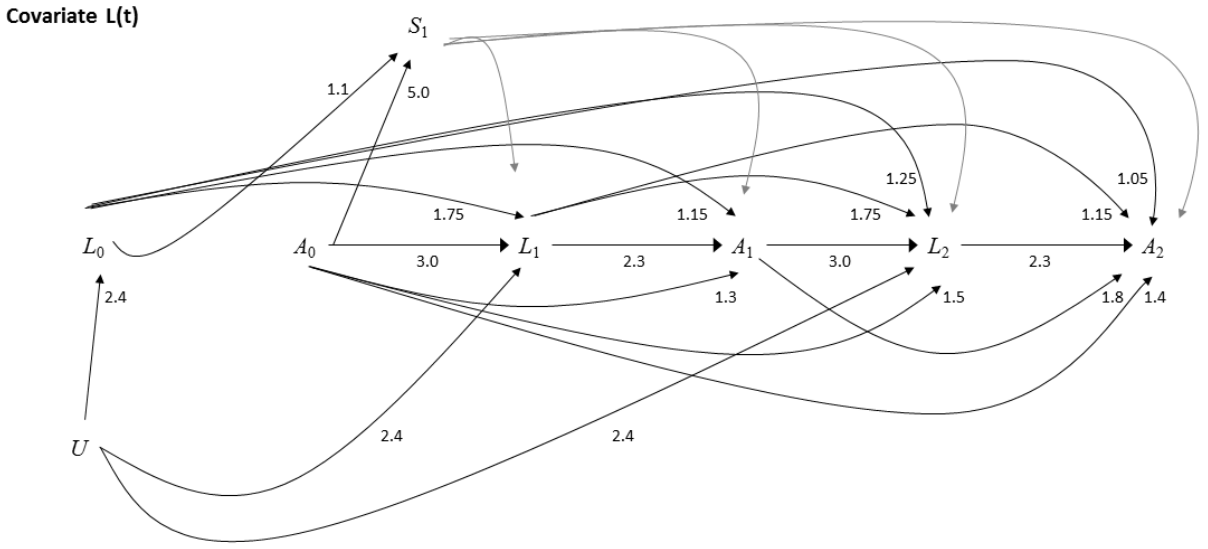
Appendix Figure 3.



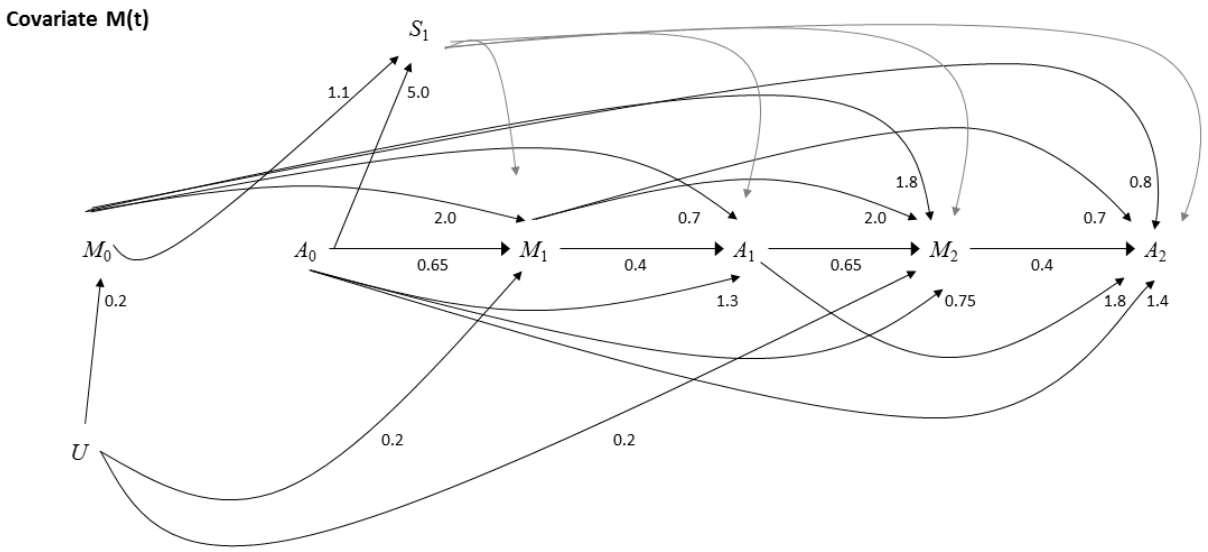
Appendix Figure 4.



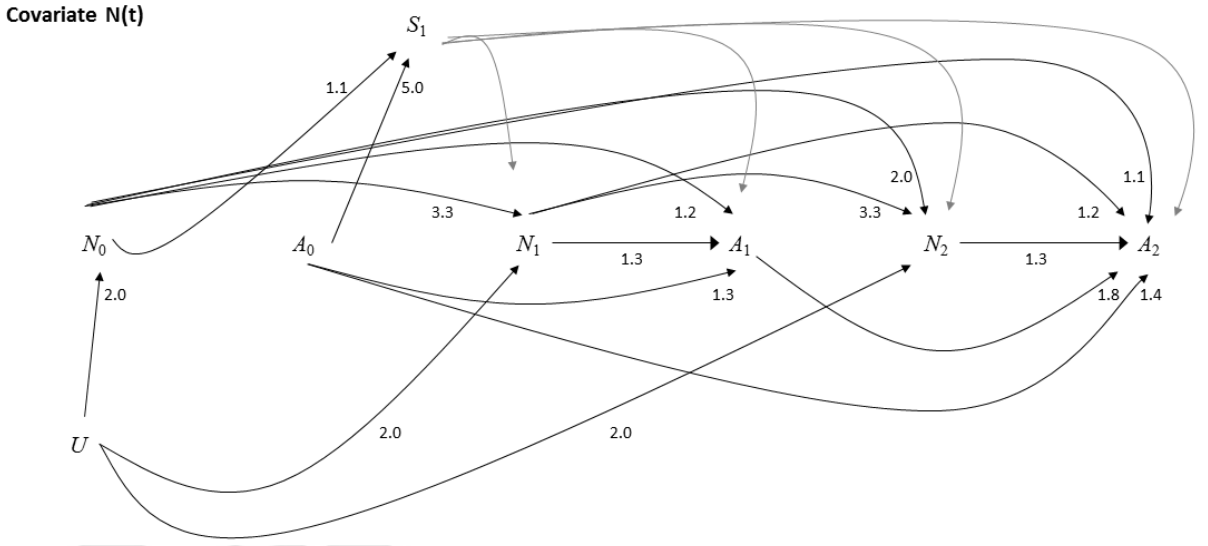
Appendix Figure 5a.



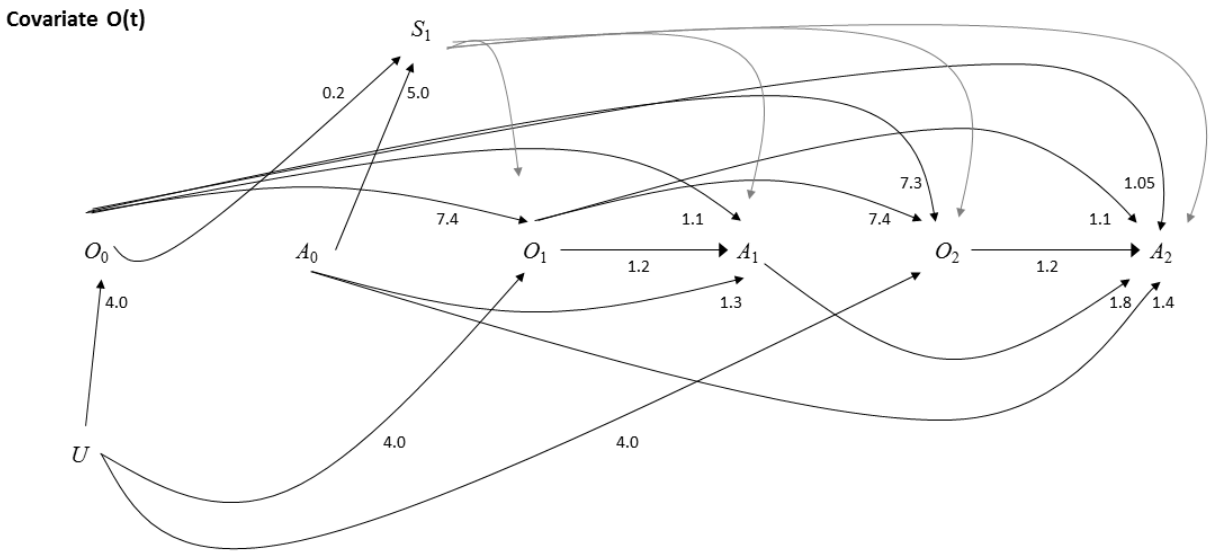
Appendix Figure 5b.



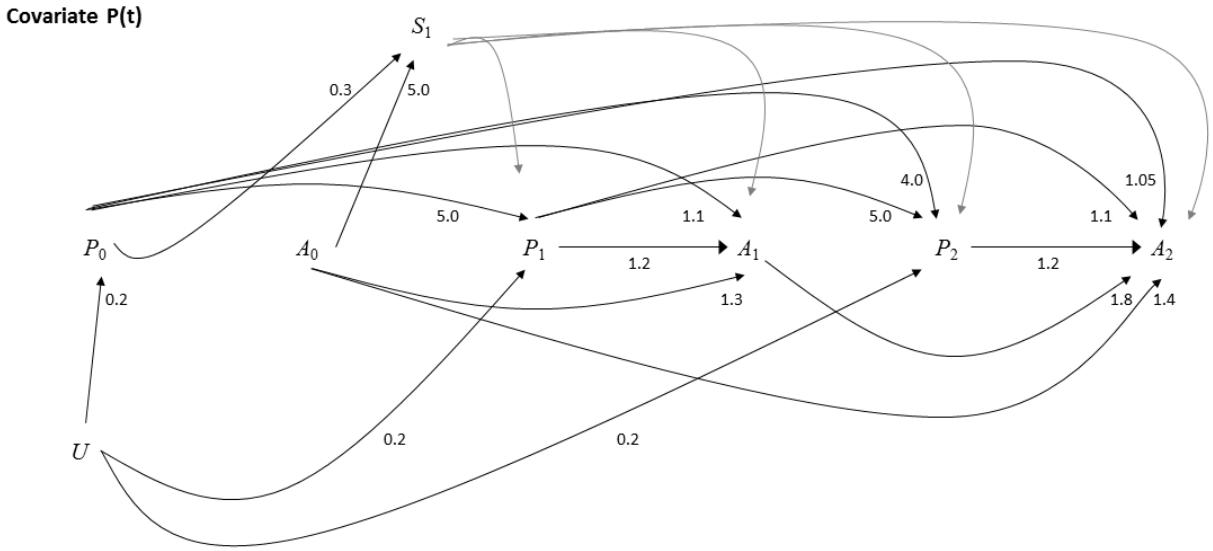
Appendix Figure 5c.



Appendix Figure 5d.



Appendix Figure 5e.



Appendix Figure 5f.