HHS Public Access

Author manuscript

Stat Med. Author manuscript; available in PMC 2023 October 15.

Published in final edited form as:

Stat Med. 2022 October 15; 41(23): 4554–4577. doi:10.1002/sim.9525.

Targeted maximum likelihood estimation of causal effects with interference: a simulation study

Paul N Zivich[1,2], **Michael G Hudgens**[3], **M Alan Brookhart**[4,5], **James Moody**[6], **David J Weber**[7], **Allison E Aiello**[1,2]

[1]Department of Epidemiology, Gillings School of Global Public Health, UNC Chapel Hill, NC, USA

[2]Carolina Population Center, UNC Chapel Hill, NC, USA

[3]Department of Biostatistics, Gillings School of Global Public Health, UNC Chapel Hill, NC, USA

[4]NoviSci, Durham, NC, USA

[5]Department of Population Health Sciences, Duke University, Durham, NC, USA

[6]Department of Sociology, Duke University, Durham, NC, USA

[7]Division of Infectious Diseases, Department of Medicine, UNC Chapel Hill, NC, USA

Abstract

Interference, the dependency of an individual's potential outcome on the exposure of other individuals, is a common occurrence in medicine and public health. Recently, targeted maximum likelihood estimation (TMLE) has been extended to settings of interference, including in the context of estimation of the mean of an outcome under a specified distribution of exposure, referred to as a policy. This paper summarizes how TMLE for independent data is extended to general interference (network-TMLE). An extensive simulation study is presented of network-TMLE, consisting of four data generating mechanisms (unit-treatment effect only, spillover effects only, unit-treatment and spillover effects, infection transmission) in networks of varying structures. Simulations show that network-TMLE performs well across scenarios with interference, but issues manifest when policies are not well-supported by the observed data, potentially leading to poor confidence interval coverage. Guidance for practical application, freely available software, and areas of future work are provided.

Keywords

interference; networks; peer effects; spillover effects; targeted maximum likelihood estimation

1. Introduction

Causal effect estimation often relies on the assumption of no interference, such that an individual's potential outcomes are independent of all other individuals' exposure.[1–3] However, interference is common across many areas of medicine and public health, most

All authors declare no conflicts of interest.

notably in infectious disease and medical social sciences. Examples include interference between injection drug users at risk of HIV,[4,5] students within the same school,[6] and individuals connected within social networks.[7] The ongoing global SARS-CoV-2 pandemic has brought further attention to interference; evaluation of physical distancing and shelter-in-place policies have highlighted how such policies, or lack thereof, can impact other nearby geographic regions.[8,9] In addition to infectious disease, interference occurs across substantive areas, with examples including transmission of opioid use within households in pharmacoepidemiology,[10] passive tobacco smoke exposure in cancer epidemiology,[11,12] and behaviors among children within classrooms in developmental psychology.[13]

When interference is present, multiple estimands may be considered.[1,14] One estimand of public health importance is the mean of an outcome under a specific policy. For example, what would the three-month risk of influenza have been if 60% of the population had been randomly selected to receive an influenza vaccine? To estimate this quantity or other related estimands, methods allowing for interference have been developed for two broad settings: partial interference and general interference. The partial interference assumption stipulates interference occurs within but not between groups of individuals,[1,15] which allows for the application of standard statistical theory.[16–19] While the partial interference assumption is sometimes reasonable, interference patterns do not always allow the separation of individuals into independent groups. General interference allows, in principle, for any two units in a population to affect each other. Methods for general interference may further be delineated by whether the exposure is randomized. In randomized experiments, methods can leverage the random assignment as the basis of inference.[20,21] In the observational setting, inference is more challenging because of the potential for confounding and lack of independent replicates. Extensions of targeted maximum likelihood estimation for independent data (IID-TMLE) have recently been developed to allow for general interference in observational studies.[22–24]

In this paper, we present simulation studies of TMLE for general interference (network-TMLE) in observational settings. While simulations have been conducted to evaluate the finite sample performance of network-TMLE,[23,24] previous empirical studies have been limited to relatively simple random networks. In practice, networks often exhibit more complex properties,[25–27] limiting the utility of previous simulation studies to guide application. Additionally, previous simulations have explored only a narrow set of data generating mechanisms and model specifications. To address these gaps, we conducted simulations for the estimation of the mean potential outcome under varied data generating mechanisms with a wider variety of networks, including an observed network of face-to-face contacts among university students, and various model specifications. Two policies were assessed: setting all individuals to some constant probability of exposure and shifting the probability of exposure by a constant.

The outline of the remainder of this paper is as follows. Section 2 reviews IID-TMLE in the context of stochastic policies. Network-TMLE for stochastic policies is described in Section 3. The simulation study design and results are presented in Sections 4 and 5, respectively. Section 6 concludes with a discussion and areas of future work. A freely available implementation of network-TMLE is provided in Python.

2. Targeted Maximum Likelihood Estimation

The TMLE of average causal effects is a doubly-robust substitution estimator that incorporates an outcome model and a propensity score (or exposure) model through a targeting step.[28,29] These models are often referred to as nuisance models since they are not of direct interest. The double-robustness property means that if one nuisance model is correctly specified, then the estimator will be statistically consistent. Under the assumption that both nuisance models are correctly specified, TMLE has the advantage of retaining root-n convergence rates when paired with data-adaptive (machine learning) estimators for the nuisance models that have at least quarter-root-n convergence.[28,30,31] In the absence of interference, TMLE methods have been developed for average causal effects,[28] causal effects under different longitudinal treatment plans,[32] and stochastic policies.[33] The following is a brief review of IID-TMLE for stochastic policies.

2.1 Estimands and Assumptions

Consider drawing inference about the effect of a binary exposure on either a A binary or continuous outcome Y in an observational study. For individual $i = 1,\ldots,n$, let W_i indicate a vector of observed baseline covariate(s), A_i the observed exposure, and Y_i the observed outcome. Assume (W_i, A_i, Y_i) for $i = 1,\ldots,n$ are independent and identically distributed (IID) and there is no interference. Let $Y_i(a)$ indicate the potential outcome for individual i had, possibly counter to fact, their exposure been $a \in \mathscr{A} = \{0, 1\}$. The goal is to draw inference about the mean outcome under a policy, denoted in general by ω, which alters or shifts the distribution of A. For example, the target estimand might be a deterministic policy where everyone is exposed. Denote the conditional distribution of A given W under policy ω by $\Pr^*(A = a \mid W)$. For the aforementioned deterministic policy, $\Pr^*(A = 1 \mid W) = 1$. Stochastic policies may also be of interest where $0 < \Pr^*(A = 1 \mid W) < 1$. The target estimand is the average outcome under policy ω, which may be defined as the population mean:

$$\psi^p = E\left[\sum_{a \in \mathscr{A}} Y(a)\Pr^*(A = a \mid W)\right] \tag{1}$$

the sample mean:

$$\psi^s = \frac{1}{n}\sum_{i=1}^{n}\sum_{a \in \mathscr{A}} Y_i(a)\Pr^*(A_i = a \mid W_i) \tag{2}$$

or the W-conditional mean:

$$\psi^c = \frac{1}{n}\sum_{i=1}^{n} E\left[\sum_{a \in \mathscr{A}} Y_i(a)\Pr^*(A_i = a \mid W_i) \,\middle|\, W_i\right] \tag{3}$$

For the deterministic policy of everyone exposed, equations 1–3 reduce to $E[Y(a = 1)]$, $\frac{1}{n}\sum_{i=1}^{n} Y_i(a = 1)$, and $\frac{1}{n}\sum_{i=1}^{n} E[Y_i(a = 1) \mid W_i]$, respectively. The population mean is of greatest relevance when units are randomly sampled from an infinite target population, in

which case ψ^p is the average outcome in the target population if policy ω were adopted. However, units may not always be drawn from a clearly defined target population, in which case the sample mean may be of interest instead. The sample mean estimand is defined only in terms of the finite population of n units observed, with the covariates W_i and potential outcomes $Y_i(a)$ viewed as fixed features of the finite population. Finally, a third possible estimand is the W-conditional mean, ψ^c, which corresponds to the mean under the policy ω treating only W_i as fixed for the population of n units. For large n, the empirical distribution of W should closely approximate the distribution of W, such that the population mean and the W-conditional mean should be similar. In general, the target estimand is denoted by ψ.

To express the estimand as a function of the observed data, identification assumptions are necessary. The following sufficient identification assumptions are used:

1. If $A_i = a$ then $Y_i = Y_i(a)$
2. $Y(a) \perp A | W$ for all $a \in \mathscr{A}$
3. If $\Pr^*(A = a \mid W) > 0$ then $\Pr(A = a | W) > 0$, for all $a \in \mathscr{A}$

where assumption 1 is causal consistency,[34] 2 is conditional exchangeability,[35] and 3 is the positivity assumption for stochastic policies.[33]

2.2 Estimation

IID-TMLE can be divided into five steps: outcome model estimation, weight estimation, targeting, estimation of ψ, and inference. The steps for point estimation of the mean are the same for ψ^p, ψ^s, and ψ^c, with differences occurring in variance estimation.[36] Therefore, we only distinguish between the estimands during estimation of the variance. In the case of a continuous Y, the observed values of Y are first rescaled to lie in (0,1), which is necessary for the targeting step.

IID-TMLE begins with the estimation of an outcome model for $E[Y_i \mid A_i, W_i]$, using either a parametric model or machine learning. Predicted values from the outcome model $(\hat{Y}_i)$ are then generated using the estimated model. Next an exposure model for $\Pr(A_i \mid W_i)$ is estimated with either a parametric model or machine learning, and the following weight is computed

$$\frac{\Pr^*(A_i \mid W_i)}{\widehat{\Pr}(A_i \mid W_i)}$$

where the denominator is computed based on the estimated exposure model and the numerator (the policy of interest) is assumed to be known. Next the logistic regression model

$$\text{logit}(\Pr(Y_i = 1)) = \eta_0 + \text{logit}(\hat{Y}_i)$$

is fit using weighted maximum likelihood, with the previously estimated weights. In the case of continuous Y_i, $\Pr(Y_i = 1)$ is replaced with the rescaled Y_i. This targeting step solves

the efficient score equation in a single step,[37] without introducing additional parametric modeling assumptions. The estimated intercept, $\hat{\eta}_0$, can be thought of as a correction term for the outcome predictions. When $\hat{Y}_i$ is close to the observed outcomes, then $\hat{\eta}_0$ will be near zero. When the outcome model is incorrectly specified and the exposure model is correct, $\hat{\eta}_0$ shifts the values of $\text{logit}(\hat{Y}_i)$. Alternatively, regression using the so-called 'clever covariate' could be considered.[38] While the clever covariate and weighted targeting approaches are asymptotically equivalent, weighted targeting may have better finite sample performance due to reduced sensitivity to stochastic positivity violations.[39]

To estimate ψ for a deterministic policy, the following procedure is used. The exposure under ω, denoted by A_i^*, and the estimated outcome model are used to impute the outcomes under ω, denoted by $\hat{Y}_i^*$. Next, the targeted predictions are computed via

$$\tilde{Y}_i^* = \text{expit}\left(\hat{\eta}_0 + \text{logit}\left(\hat{Y}_i^*\right)\right)$$

and ψ is estimated as the mean of the targeted predictions,

$$\hat{\psi} = \frac{1}{n}\sum_{i=1}^{n} \tilde{Y}_i^*$$

For stochastic policies, the estimation procedure above requires modification. Because the distribution of A^* is no longer degenerate under a stochastic policy, the following Monte Carlo approach is used. For $k = 1,\ldots,m$ sample A_{ik}^* from Bernoulli $(\Pr^*(A_i = 1 \mid W_i))$. For each A_{ik}^*, compute the imputed outcome and targeted prediction for individual i, say $\hat{Y}_{ik}^*$ and $\tilde{Y}_{ik}^*$, as in the deterministic policy estimation procedure above. Then $\tilde{Y}_i^*$ is calculated by $\tilde{Y}_i^* = \sum_{k=1}^{m} \tilde{Y}_{ik}^*/m$ and the estimator for ψ is the mean of the targeted predictions as before.

Lastly, $(1-\alpha)$ confidence intervals (CI) can be constructed by $\hat{\psi} \pm z_{1-\alpha/2}\left(\hat{\sigma}^2/n\right)^{0.5}$, where $z_{1-\alpha/2}$ denotes the $1-\alpha/2$ quantile of a standard normal distribution and $\hat{\sigma}^2$ is an estimand-specific variance estimator. For the population mean ψ^p the variance is estimated by[33]

$$\hat{\sigma}_p^{\,2} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\Pr^*(A_i \mid W_i)}{\widehat{\Pr}(A_i \mid W_i)}\left(Y_i - \hat{Y}_i\right) + \tilde{Y}_i^* - \hat{\psi}\right)^2$$

and for the W-conditional mean ψ^c the variance is estimated by[36]

$$\hat{\sigma}_c^{\,2} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\Pr^*(A_i \mid W_i)}{\widehat{\Pr}(A_i \mid W_i)}\left(Y_i - \hat{Y}_i\right)\right)^2$$

While the sample mean variance is not identified, the asymptotic variance of the conditional mean is always greater than or equal to the sample mean asymptotic variance.[36] Therefore,

the sample mean variance may be conservatively estimated by the conditional sample mean variance estimator.

3. Targeted Maximum Likelihood Estimation with Dependent Data

In the presence of interference, the potential outcomes depend on both an individual's exposure and the exposure of others in the population. Consider the setting where individuals are connected via a network of edges (e.g., an edge may indicate two individuals are friends within a social network, live within a certain distance of each other, or had a face-to-face conversation in the past week). Suppose the network structure is static (i.e., fixed over time) and can be summarized by an $n \times n$ adjacency matrix $\mathcal{G}$. Let $\mathcal{G}_{ij}$ denote the (i,j) entry of $\mathcal{G}$, where $\mathcal{G}_{ij} = 1$ if an edge exists between i and j. Assume no interference between individuals i and j if $\mathcal{G}_{ij} = 0$. Since interference is a relation between individuals, by definition $\mathcal{G}_{ii} = 0 \; \forall \; i \in n$. Throughout, individual i's 'immediate contacts' refers to individuals that have an edge with i.

From $\mathcal{G}$ and the covariates, various summary measures can be calculated. The total number of immediate contacts for individual i (also referred to as degree) is defined as $F_i = \sum_{j=1}^{n} \mathcal{G}_{ij}$. The exposure status for i's immediate contacts can be expressed by different summary measures, which are functions of $(A_1, A_2, \ldots, A_n)$ and $\mathcal{G}$, and in general is denoted by A_i^s with possible realizations $a_i^s \in \mathcal{A}^s$. For example, $A_i^s = \sum_{j=1}^{n} I\left(A_j = 1\right)\mathcal{G}_{ij}$ is the number of individual i's immediate contacts with $A = 1$.

3.1 Estimands and Assumptions

In the presence of interference, the potential outcomes for individual i can be denoted by $Y_i(a_i, a_{-i})$, where a_{-i} indicates the exposure for all individuals excluding i. Assume an exposure mapping such that only the summary measure of an individual's immediate contacts' exposures is necessary to define all of an individual's potential outcomes (referred to as weak dependence hereafter),[20] such that the potential outcomes may be denoted by $Y_i(a_i, a_i^s)$, with $a \in \mathcal{A} = \{0, 1\}$ and $a^s \in \mathcal{A}^s$. The target estimand is the average outcome under policy ω for the sample, which may be either the network mean:

$$\psi^p = E\left[\frac{1}{n}\sum_{i=1}^{n} \sum_{a \in \mathcal{A}, a^s \in \mathcal{A}^s} Y_i(a, a^s)\Pr^*\left(A_i = a, A_i^s = a^s \mid W_i, W_i^s\right)\right] \tag{4}$$

or the $\boldsymbol{W}$-conditional mean:

$$\psi^c = \frac{1}{n}\sum_{i=1}^{n} E\left[\sum_{a \in \mathcal{A}, a^s \in \mathcal{A}^s} Y_i(a, a^s)\Pr^*\left(A_i = a, A_i^s = a^s \mid W_i, W_i^s\right) \mid \boldsymbol{W}\right] \tag{5}$$

where W_i^s denotes a vector of summary measures for unit i's immediate contacts' baseline covariates and $\boldsymbol{W} = (W_1, W_2, \ldots, W_n)$.[24] Note that where there is no interference and the units are IID, equation 4 reduces to equation 1. Thus, equation 4 can be viewed as a generalization

of the population mean allowing for network dependence. Here, a superpopulation of networks each consisting of n units with the same structure can be envisioned, of which only a single observation, the observed network, is available. Therefore, ψ^p can be interpreted as the expected mean outcome for a network of n units under the policy ω. The $\boldsymbol{W}$-conditional mean similarly envisions a superpopulation of networks all under ω but now the superpopulation is composed of both the network with the same structure and same values of $\boldsymbol{W}$ among the n units. In the IID setting, equation 5 reduces to equation 3. In general, the target estimand is denoted by ψ.

The following assumptions allow for ψ to be identified:

1. If $A_i = a$, $A_i^s = a^s$ then $Y_i = Y_i(a, a^s)$
2. $Y(a, a^s) \perp A, A^s | W, W^s$, for all $a \in \mathscr{A}$, $a^s \in \mathscr{A}^s$
3. If Pr * $(A = a, A^s = a^s \mid W, W^s) > 0$ then Pr $(A = a, A^s = a^s \mid W, W^s) > 0$, for all $a \in \mathscr{A}$, $a^s \in \mathscr{A}^s$

where assumption 1 is causal consistency, 2 is conditional exchangeability, and 3 is positivity assumption for stochastic policies. In practice, choosing the set of covariates W and W^s such that the conditional exchangeability assumption is plausible may be informed by subject matter knowledge. In many settings, the degree of units, F, might be included among the covariates composing the vector W, since F may affect A^s and may also affect the outcome of interest, as individuals who have many contacts may be different from those who have fewer contacts. Accounting for degree of units is analogous to methods for clustered data which allow for informative cluster sizes,[40] i.e., associations between the outcome and the cluster size which exist even after conditioning on other baseline covariates.

3.2 Network-TMLE

Network-TMLE extends the TMLE framework to dependent data by allowing A_i to depend on W_i and W_i^s, and Y_i to depend on A_i, A_i^s, W_i, and W_i^s. Similar to IID-TMLE, network-TMLE is doubly robust and is divided into five steps: estimate the outcome model, estimate the weights, targeting, estimation of ψ, and inference. As before, point estimation for ψ^p and ψ^c remains the same, with differences occurring in estimation of the variance.

Step 1) Estimate the outcome model—A model for $E\left[Y_i \mid A_i, A_i^s, W_i, W_i^s\right]$ can be estimated by handling each observation as if it were IID (see Section 4 of van der Laan[22]). For example, ordinary least squares could be used to estimate the parameters of the model

$$E\left[Y_i \mid A_i, A_i^s, W_i, W_i^s; \beta\right] = \beta_0 + \beta_1 A_i + \beta_2 A_i^s + \beta_3 W_i + \beta_4 W_i^s$$

After estimating the model parameters, predicted outcomes under the observed A_i and A_i^s for each unit are calculated, indicated by $\hat{Y}_i$.

Step 2) Estimate the weights—The weights can be expressed as

$$\frac{\pi^*(W_i, W_i^S; \gamma^*, \delta^*)}{\pi(W_i, W_i^S; \gamma, \delta)} = \frac{\Pr^*\left(A_i \mid W_i,\ W_i^S; \gamma^*\right)\Pr^*\left(A_i^S \mid A_i,\ W_i,\ W_i^S; \delta^*\right)}{\Pr\left(A_i \mid W_i,\ W_i^S; \gamma\right)\Pr\left(A_i^S \mid A_i,\ W_i,\ W_i^S; \delta\right)}$$

where γ, δ, γ^*, and δ^* denote the parameters for the models for $\Pr(A_i \mid W_i,\ W_i^S)$, $\Pr(A_i^S \mid A_i,\ W_i,\ W_i^S)$, $\Pr^*(A_i \mid W_i, W_i^S)$, and $\Pr^*(A_i^S \mid A_i, W_i, W_i^S)$, respectively. The model for $\Pr(A_i \mid W_i,\ W_i^S)$ can be estimated with a logistic regression model treating observations as IID. Different models may be assumed for estimation of $\Pr(A_i^S \mid A_i,\ W_i,\ W_i^S)$. For example, if A_i^S is a binary variable indicating whether at least one of individual i's immediate contacts is exposed, then logistic regression might be used. If A_i^S is instead a count variable (e.g., the number of immediate contacts exposed), then Poisson or negative binomial regression models might be assumed. For more general A and A^S, estimation of the weights can be done using generalized propensity functions.[41]

Estimation of the numerator of the weights can be accomplished using a simulation approach as follows. Here, only policies where A_i and A_j are independent conditional on $\boldsymbol{W}$ for all $i \neq j$ are considered, since it is often easier to specify policies that assign exposures to each individual, $\Pr^*(A = a \mid W, W^S)$, as opposed to specifying a policy in terms of both an individual's exposure and their contacts summary measure, $\Pr^*(A = a, A^S = a^S \mid W, W^S)$.[24] The first step to the simulation approach entails generating m copies of the data set, indexed by $k = 1,\ldots,m$. Next, the assigned exposure under the i policy for individual in copy k, A_{ik}^*, is sampled from $\text{Bernoulli}(\Pr^*(A_i = 1 \mid W_i, W_i^S))$. Then A_{ik}^* and $\mathcal{G}$ are used to calculate the new exposure summary measure, A_{ik}^{*S}, under the policy. The parameters, γ^* and δ^*, are then estimated using all m copies of the data simultaneously. Finally, the numerator of the weights, $\pi^*(W_i, W_i^S; \gamma^*, \delta^*)$, is estimated by using the estimated parameters $\hat{\gamma}^*$, $\hat{\delta}^*$ and the observed values of A_i, A_i^S, W_i and W_i^S.

Step 3) Targeting—To target, the following logistic regression model is fit using weighted maximum likelihood,

$$\text{logit}(\Pr(Y_i = 1)) = \eta_0 + \text{logit}(\hat{Y}_i)$$

where the weights are those calculated in step 2. As before, $\Pr(Y_i = 1)$ is replaced with the rescaled Y_i in the case of a continuous outcome. As with IID-TMLE, the targeting step solves the efficient score equation. Similarly, regression with a clever covariate could be used instead, with these approaches being asymptotically equivalent.[22] However, the weighted targeting approach may have better finite sample performance and be less computationally intensive in the network dependence setting compared to the clever covariate approach.[23]

Step 4) Estimation of ψ—As in the IID setting, stochastic policies are evaluated using a Monte Carlo approach. First, m copies of the data set, indexed by $k = 1,\ldots,m$, are generated. Then A_{ik}^{*} is sampled from $\text{Bernoulli}\left(\Pr^{*}\left(A_i = 1 \mid W_i, W_i^{s}\right)\right)$ and used to calculate the summary measures for A_{ik}^{*s}. To reduce computational burden, A_{ik}^{*} and $A_{ik}^{s\,*}$ generated in step 2 can be reused here. For each A_{ik}^{*} and A_{ik}^{*s}, the imputed outcome, $\hat{Y}_{ik}^{*}$, is computed using the previously estimated outcome model. The targeted prediction, $\tilde{Y}_{ik}^{*}$, is then computed by

$$\tilde{Y}_{ik}^{*} = \text{expit}\left(\hat{\eta}_0 + \text{logit}\left(\hat{Y}_{ik}^{*}\right)\right)$$

The mean of the targeted predictions, $\hat{\psi}_k = \sum_{i=1}^{n} \tilde{Y}_{ik}^{*}/n$ is calculated for each k. The estimator for ψ is the mean of the m estimates, i.e.,$\hat{\psi} = \sum_{k=1}^{m} \hat{\psi}_k/m$.

Step 5) Inference for ψ—Estimating the variance of the network mean estimator is challenging. Closed-form and bootstrap variance estimators have been proposed, [24] but these variance estimators require either that $W_1,\ldots,W_n$ are IID or that the distribution of $\boldsymbol{W}$ can be correctly modeled.[24] In settings where network interference may be present, the IID assumption will be unrealistic and correctly modeling the distribution of $\boldsymbol{W}$ may be difficult due to possible complicated network dependencies between individuals. Alternatively, an estimator of an approximate upper bound on the variance of the network mean has been proposed.[23] However, this estimator can be uninformative under some skewed degree distributions, which commonly occur in many social networks.[42]

Due to the overly restrictive conditions on the variance estimators for the network mean, the $\boldsymbol{W}$-conditional mean may be the preferred estimand. Specifically, the variance estimator for the $\boldsymbol{W}$-conditional mean does not require assumptions regarding the distribution of $\boldsymbol{W}$. For the $\boldsymbol{W}$-conditional mean, the variance can be estimated by[23]

$$\hat{\sigma}_{c_1}{}^{2} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\pi^{*}\left(W_i, W_i^{s}; \hat{\gamma}^{*}, \hat{\delta}^{*}\right)}{\pi\left(W_i, W_i^{s}; \hat{\gamma}, \hat{\delta}\right)}\left(Y_i - \hat{Y}_i\right)\right)^{2}$$

with $(1 - \alpha)$ CI constructed by $\hat{\psi}^{c} \pm z_{1-\alpha/2}\left(\hat{\sigma}_{c_1}{}^{2}/n\right)^{0.5}$. This variance estimator assumes that dependence between observations is solely due to measured covariates of immediate contacts, sometimes referred to as direct transmission.[24] However, latent dependence, where individuals close in the network are more likely to share unmeasured covariates compared to individuals far away, is likely in practice. Therefore, we also considered the following variance estimator,[24] which allows for latent dependence limited up to second-order contacts (immediate contacts of i's immediate contacts):

$$\hat{\sigma}_{c_2}{}^{2} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{G}_{ij}\left(\frac{\pi^{*}\left(W_i, W_i^{s}; \hat{\gamma}^{*}, \hat{\delta}^{*}\right)}{\pi\left(W_i, W_i^{s}; \hat{\gamma}, \hat{\delta}\right)}\left(Y_i - \hat{Y}_i\right) \times \frac{\pi^{*}\left(W_j, W_j^{s}; \hat{\gamma}^{*}, \hat{\delta}^{*}\right)}{\pi\left(W_j, W_j^{s}; \hat{\gamma}, \hat{\delta}\right)}\left(Y_j - \hat{Y}_j\right)\right)$$

where $\mathbb{G}_{ij} = \mathcal{G}_{ij}$ for $i \neq j$ and $\mathbb{G}_{ii} = 1$.

4. Simulation Study Design

Simulations were conducted to assess the performance of the network-TMLE methods described in the previous section across varying networks and data generating mechanisms. All simulations were repeated 3000 times. We considered two different policy types. For the first policy, all individuals in the network were assigned the same probability of exposure, $\Pr^*(A_i = 1 \mid W_i, W_i^s) = \Pr^*(A_i = 1) = p$. For the second policy, each individual's log-odds of exposure was shifted by a constant value, $\Pr^*(A_i = 1 \mid W_i, W_i^s) = \text{expit}(\text{logit}(\Pr(A_i = 1 \mid W_i, W_i^s)) + q)$. Shifts in the log-odds of exposure were used to ensure all probabilities under the policy remained between (0, 1). The true values of the $\boldsymbol{W}$-conditional mean for each policy were obtained empirically by taking the mean of ψ from 10,000 different data sets under the corresponding policy, with the network and $\boldsymbol{W}$ held fixed across data sets.

4.1 Networks

Seven different networks were used: uniform random graphs ($n = 500$, $n = 1000$, $n = 2000$), modified clustered power-law random graphs ($n = 500$, $n = 1000$, $n = 2000$), and the eX-FLU network of self-reported contacts among undergraduate students (Appendix Figure B.1).[43] The uniform random graphs followed uniform degree distributions with minimum degree 1 and maximum degree 6. The modified clustered power-law random graphs consisted of separately generated clustered power-law random subgraphs, with edges randomly generated between the random subgraphs. Each of the clustered power-law subgraphs was separately generated from a Barabasi-Albert random graph model with a fixed probability for closing triads between nodes.[44] For each node, three connections were generated and the probability of triad closure was set to 0.75. The advantage of this approach is that the random graph takes on common characteristics of empirical networks, including a power-law degree distribution, a high clustering coefficient, and an underlying community structure. Lastly, the eX-FLU network was based on data from the eX-FLU cluster-randomized trial, a study to assess the efficacy of three-day self-isolation when symptomatic of respiratory illness on subsequent respiratory infection among university students.[43] Over the ten-week study period, enrolled students reported face-to-face contacts each week. From the ten weeks of self-reported contacts, we generated a single static network and selected the largest connected component. Summary statistics for the networks are provided in Appendix Table B.1.

Different nuisance model specifications of network-TMLE were evaluated for each scenario: both nuisance models were correctly specified, misspecification of the outcome model, and misspecification of the exposure model (model specifications are provided in Appendix B). Model misspecification was induced by specifying the wrong form for W^s, the summary measure(s) of covariates, along with excluding a covariate from W^s. As the correct specification of nuisance models will often be unknown in practice, we further evaluated a more flexible specification of nuisance models. This approach consisted of the following procedure. For a categorical covariate, V^s, included in W^s, the corresponding summary

measure was defined as $V_i^s(v) = \sum_{j=1}^{n} I(V_j = v)\mathcal{G}_{ij}$ for all $v \in V$ besides a chosen referent. For example, suppose V^s is the only measure included in W^s and the possible values for V_i were {0,1,2}. Therefore, the summary measures for covariates would be $W^s = (V^s(v = 1), V^s(v = 2))$. Since (v) tended to be right-skewed for networks with skewed degree distributions and to reduce the dimensionality of the nuisance models, $V^s(v)$ was further grouped into categories via a generalization of histogram binning.[45] This approach kept narrower bins at lower values of $V^s(v)$ with wider bins at the right tail. Each bin was included in the regression model as an indicator term (e.g., if $V^s(v)$ was binned into 5 categories, 4 indicator variables were included in the nuisance models). While continuous V could also be discretized and modeled using the previous approach, the following summary measure was used instead $V_i^s = \sum_{j=1}^{n}(V_i - V_j)\mathcal{G}_{ij} / \sum_{j=1}^{n}\mathcal{G}_{ij}$. Outcome nuisance models for the flexible specification were estimated using L2-penalized regression.

Performance of network-TMLE may be improved when restricting inference to nodes below a pre-specified degree for skewed degree distributions.[24] Therefore, restricting inference by degree was further compared for the clustered power-law random graphs and eX-FLU network. Nodes with degrees above a pre-defined maximum had their value for held as fixed and were considered as background A features. For the clustered power-law random graph ($n = 500$), nodes with a degree above 18 were considered as features of the background (2%). Similarly, nodes with a degree greater than 22 were considered as background features for the eX-FLU network (5%) and the remaining clustered power-law random graphs ($n = 1000$: 2%, $n = 2000$: 2%).

4.2 Data generating mechanisms

Four data generating mechanisms inspired by real-world scenarios were considered. Each data generating mechanism was selected to feature different possible exposure effects, including individual-specific (i.e., unit-treatment) effects and spillover effects from contacts. Below is a brief description of each of the data generating mechanisms with further details provided in Appendix B.

4.2.1 Statin and Cardiovascular Disease—

To simulate a no interference setting, a data generating mechanism based on a hypothetical study on statin initiation and subsequent atherosclerotic cardiovascular disease (ASCVD) was created. Statins are cholesterol-lowering drugs that have been shown to reduce cardiovascular disease risk by reducing cholesterol synthesis.[46,47] The mechanism of action may reasonably allow researchers to believe that whether individual i's friends take a statin has no influence on individual i's risk of ASCVD. Therefore, ASCVD risk was independent of immediate contacts in our simulation. Confounders were based on the 2018 primary prevention guidelines for the management of blood cholesterol,[48] and included age, low-density lipoprotein levels, and ASCVD risk score.

Let X_i indicate age, L_i indicate log-transformed low-density lipoprotein, R_i indicate risk score, and $W_i = (X_i, L_i, R_i)$. The conditional probability of taking a statin was specified by:

$$\text{logit}(\Pr(A_i = 1 \mid W_i)) = -5.3 + 0.2\, L_i + 0.15(X_i - 30) + 0.4\, I(0.05 \le R_i < 0.075) + 0.9\, I(0.075 \le R_i < 0.2) + 1.5\, I(R_i \ge 0.2)$$

The conditional probability of ASCVD was specified by:

$$\text{logit}(\Pr(Y_i = 1 \mid A_i, W_i)) = -5.05 - 0.8A_i + 0.37\sqrt{X_i - 39.9} + 0.75R_i + 0.75L_i$$

4.2.2 Naloxone and Opioid Overdose—For spillover effect only, a data generating mechanism based on the effect of naloxone on subsequent opioid overdose deaths was created. Opioid overdose deaths have dramatically increased in recent years.[49,50] Naloxone has been used as an emergency intervention to rapidly reverse opioid overdoses by blocking opioid receptors,[51] and has been made increasingly available to the general population to prevent overdose deaths.[52,53] Nasal spray formulations rely on another person for administration, with self-administration having occurred only in rare cases.[54] Therefore, the prevention of opioid overdose deaths with naloxone is an example where the protective effect may operate solely via spillover effects. Confounders included gender, recent overdose, and recent release from prison, which have been observed as predictors of opioid overdose in previous studies.[55,56] In the context of this mechanism, the interference pattern could be thought of as a co-injection network.

Let S_i indicate gender, O_i indicate recent overdose, P_i indicate recent release from prison, and $W_i = (S_i, O_i, P_i, F_i)$. The conditional probability of naloxone was generated according to:

$$\text{logit}\left(\Pr\left(A_i = 1 \mid W_i, W_i^s\right)\right) = -0.5 - 1.5\, P_i + 1.5\, P_i S_i + 0.3 \sum_{j=1}^{n} I\left(O_j = 1\right)\mathcal{G}_{ij} + 0.5\frac{\sum_{j=1}^{n} I\left(S_j = 1\right)\mathcal{G}_{ij}}{\sum_{j=1}^{n}\mathcal{G}_{ij}} + 0.05F_i$$

The conditional probability of death from opioid overdose was specified by:

$$\text{logit}\left(\Pr\left(Y_i = 1 \mid A_i^s, W_i, W_i^s\right)\right) = -0.4 - 0.2 \sum_{j=1}^{n} I\left(A_j = 1\right)\mathcal{G}_{ij} + 1.7\, P_i - 1.1\, S_i + 0.6 \sum_{j=1}^{n} I\left(O_j = 1\right)\mathcal{G}_{ij} - 1.5\frac{\sum_{j=1}^{n} I\left(S_j = 1\right)\mathcal{G}_{ij}}{\sum_{j=1}^{n}\mathcal{G}_{ij}} - 0.4F_i$$

4.2.3 Comprehensive Dietary Intervention and Body Mass Index—For simultaneous unit-treatment and spillover effects, a data generating mechanism based on a comprehensive dietary intervention on body mass index (BMI) was created. Research has found BMI to be socially clustered,[57,58] with the transmission of obesity theorized to result from social pressures or the shared environments of social contacts.[58] Comprehensive dietary interventions that limit caloric intake and increase the quality of food may reduce BMI.[59] Our simulation focused on a theoretical dietary intervention that impacts an

individual's BMI as well as their immediate friends' BMI. Confounders included baseline BMI, gender, and baseline exercise. In this context, the interference pattern can be viewed as a network of friendships.

Let S_i indicate gender, B_i indicate baseline BMI, E_i indicate exercise at baseline, U_i indicate the unobserved variable (proximity to work), and $W_i = (S_i, B_i, E_i, F_i)$. The conditional probability of starting the proposed diet at baseline was specified by:

$$\text{logit}\left(\Pr\left(A_i = 1 \mid W_i, W_i^S\right)\right) = -1.5 + 0.05(B_i - 30) + 2.0 S_i E_i + 1.0 \frac{\sum_{j=1}^{n} I\left(E_j = 1\right) \mathcal{G}_{ij}}{\sum_{j=1}^{n} \mathcal{G}_{ij}} + 1.0 \frac{\sum_{j=1}^{n} I\left(S_j = 1\right) \mathcal{G}_{ij}}{\sum_{j=1}^{n} \mathcal{G}_{ij}} + 0.05 F_i$$

BMI at follow-up was generated by:

$$Y_i = 3.9 + B_i - 3A_i - 2\, I\left(3 < \sum_{j=1}^{n} I\left(A_j = 1\right) \mathcal{G}_{ij}\right) - 2U_i + 2S_i - 2E_i - 1.0 \sum_{j=1}^{n} I\left(E_j = 1\right) \mathcal{G}_{ij} - 0.75 \sum_{j=1}^{n} I\left(S_j = 1\right) \mathcal{G}_{ij} + \frac{\sum_{j=1}^{n} \mathcal{G}_{ij}\left(B_j - B_i\right)}{\sum_{j=1}^{n} \mathcal{G}_{ij}} + 0.2 F_i + 3\, I\left(0.4 < \frac{\sum_{j=1}^{n} \mathcal{G}_{ij} U_i}{\sum_{j=1}^{n} \mathcal{G}_{ij}}\right) + \epsilon_i$$

where $\epsilon_i \sim$ Normal(0, 1). Here, U_i is related to Y_i and not A_i, and is unobserved (i.e., not included in the network-TMLE outcome nuisance model). Additionally, B_i was made to be assortative in the underlying network. Therefore, this data generating mechanism was expected to exhibit latent variable dependence.

4.2.4 Infectious Disease Transmission—The fourth simulation mechanism entailed a Susceptible-Infected-Recovered (SIR) model of human-to-human transmission of an infectious agent. The hypothetical vaccine followed a 'leaky' model, such that the vaccine reduced the probability of infection given a single exposure to an infectious agent.[60] The spillover effect of the vaccine was composed of contagion (vaccinated individuals were less likely to become infected and thus less likely to transmit) and infectiousness effects (vaccinated-but-infected individuals had reduced probability of transmitting the disease).[61]

The stochastic SIR model was implemented as follows. Of n units in a network, 1% were selected to be initially infected. Actively infectious individuals were able to infect their immediate contacts based on a probability of transmission conditional on characteristics of the infected and uninfected individuals. Infected individuals remained infectious for a period of five discrete time-steps after becoming infected and recovered after the infectious period (no longer infectious nor capable of being infected by contacts). All transmission events occurred over a period of ten time-steps. Unlike the previous data generating mechanisms, the infection transmission mechanism does not necessarily adhere to the weak dependence assumption. By chance, an infection could spread up to a maximum of 10 contacts away from the source.

Let V_i indicate asthma, H_i indicate hand hygiene, and $W_i = (V_i, H_i, F_i)$. The probability of being vaccinated was specified by:

$$\text{logit}\left(\Pr\left(A_i = 1 \mid W_i, W_i^S\right)\right) = -1.9 + 1.75V_i + 1.0H_i + 1.0\sum_{j=1}^{n} I\left(H_j = 1\right)\mathcal{G}_{ij} + 1.3\sum_{j=1}^{n} I\left(V_j = 1\right)\mathcal{G}_{ij} - 0.65F_i$$

and the probability of individual i becoming infected at discrete time-point t ($D_{i,t}$) by individual j was generated by:

$$\text{logit}\left(\Pr\left(D_{i,t} = 1 \mid Z_{j,t} = 1,\ \mathcal{G}_{ij} = 1, A_i, A_j, W_i\right)\right) = -2.5 - 1.0A_i - 0.2A_j + 1.0V_i - 0.2H_i$$

where $Z_{j,t} = 1$ indicates whether j was in the infectious category at time t. For individuals with multiple infectious contacts, whether transmission of the infection occurred was evaluated for each contact independently. Note this is not the outcome model used in network-TMLE. Instead, network-TMLE used the following outcome model:

$$\text{logit}\left(\Pr\left(Y_i = 1 \mid A_i, A_i^S, W_i, W_i^S\right)\right) = \beta_0 + \beta_1 A_i + \beta_2\sum_{j=1}^{n} I\left(A_j = 1\right)\mathcal{G}_{ij} + \beta_3 V_i + \beta_4\sum_{j=1}^{n} I\left(V_j = 1\right)\mathcal{G}_{ij} + \beta_5 H_i + \beta_6\sum_{j=1}^{n} I\left(H_j = 1\right)\mathcal{G}_{ij} + \beta_7 F_i$$

where Y_i is the indicator variable of ever infected by the end of follow-up.

4.3 Performance Metrics

To assess the performance of network-TMLE, the following metrics were used: bias, empirical standard error (ESE), and 95% CI coverage. Bias was defined as the mean of $\hat{\psi}^{\omega} - \psi^{\omega}$ for each policy ω. ESE was estimated by the standard deviation of the simulation estimates for each policy scenario. CI coverage was calculated as the proportion of 95% CIs containing the true mean of the outcome. Tables comparing the correctly specified and flexibly specified nuisance models are available in Appendix D.

4.4 Software

All simulations were conducted using Python 3.6.6 with the following libraries: NumPy,[62] SciPy,[63] statsmodels,[64] patsy,[65] and NetworkX.[66] Since no current implementation of network-TMLE was available in Python, we designed one. MossSpider is freely available on the Python Package Index (PyPI) and GitHub (github.com/pzivich/MossSpider). Our implementation was supported by replicating the simulations from Sofrygin and van der Laan 2017 (Appendix A).[23] All simulation code is available at github.com/pzivich/publications-code.

5. Simulation Study Results

5.1 Statin and ASCVD

For the hypothetical study of statins the assumption regarding no interference is valid, so network-TMLE is not necessary for estimation in this context but is expected to correctly estimate the proportion under each policy. When both nuisance models were correctly specified, there was little bias for all networks (Figures 1–2, Appendix Figures C.1–C9). For the uniform network ($n = 500$), CI coverage was less than the nominal level for policies where substantially more individuals would be exposed relative to the observed data for the uniform random graph (Figure 1). On the other hand, CI coverage exceeded the nominal level for policies where the probability of exposure was dissimilar to the observed data for the restricted-by-degree eX-FLU graph (Figure 2). This over-coverage was in part due to confidence intervals that spanned the entire parameter space (0 to 1) (Appendix Table D.4). For the unrestricted eX-FLU graph, the conservative coverage remained (Appendix Figure C.1), but the ESE was increased. The unrestricted power-law graph ($n = 500$) results were also similar to the unrestricted eX-FLU results (Appendix Figure C.3) but restricting by degree for the power-law random graph reduced the extent of the conservative coverage (Appendix Figure C.2). CI coverage and the patterns exhibited were similar for both variance estimators.

Misspecification of either the exposure or outcome models did not substantially alter the performance of network-TMLE. As expected, misspecification of the outcome model resulted in an increased ESE compared to both model beings correctly specified or only the outcome model being correctly specified. The proposed flexible approach for modeling the W^S terms performed adequately, but policies where the probability of exposure was greater than 0.5 had the lowest CI coverage of the differing model specifications for the uniform and the power-law random graphs (Figure 1, Appendix Figure C.2–C.3). However, CI coverage was improved in the eX-FLU network for most policies (Figure 2).

Under increased n for the uniform and power-law random graphs, bias remained negligible, and the ESE decreased as n increased (Appendix Figure C.4–C.9, Appendix Tables D.2–D.3, Tables D.8–D.12). For the uniform random graph, CI coverage was closer to expected levels with the larger n but coverage was still below the nominal level for policies where individuals were assigned a high (>0.7) of receiving statins. As n increased for the restricted and unrestricted power-law random graph, CI coverage followed a similar pattern to $n = 500$, with approximately nominal coverage for most policies when restricting by degree and conservative coverage when unrestricted by degree. As before, the flexible modeling approach had worse coverage for policies where the probability of exposure was greater than 0.5.

5.2 Naloxone and Opioid Overdose

For simulations of naloxone and opioid overdose (where there was a spillover effect only), network-TMLE had variable performance across the scenarios. For the uniform random graph ($n = 500$), network-TMLE exhibited negligible bias and CI coverage approximating the nominal level for policies where the likelihood of exposure was not substantially

different from the observed distribution of exposure (Figure 3). However, coverage dropped to 75% for policies where nearly everyone was exposed. Performance in terms of bias and CI coverage was similar across the varying model specifications. For $n = 1000$, the bias remained negligible, and CI coverage slightly increased for policies where nearly everyone was exposed but remained below nominal levels (Appendix Figure C.13, Appendix Table D.13). Similar patterns occurred for $n = 2000$ (Appendix Figure C.16, Appendix Table D.14). Again, the variations of the nuisance models had similar patterns in terms of performance.

For the eX-FLU network restricted by degree, network-TMLE had minimal bias across policies (Figure 4). Similar to the uniform random graph, CI coverage dropped for policies where most individuals were exposed. Furthermore, performance was poor when only the exposure model was correctly specified, with some bias and CI coverage below nominal levels for all policies. The flexible specification of W^S was improved over scenario where only the exposure model was correctly specified. Performance in terms of CI coverage was worse when the eX-FLU network was not restricted by degree, mainly for policies were nearly everyone would have been exposed (Appendix Figure C.10, Appendix Tables D.15–D.16).

Performance of network-TMLE for the clustered power-law random graph ($n = 500$) was comparable to the eX-FLU simulations (Appendix Figure C.11–C.12). As with the eX-FLU network, performance was improved when restricting by degree. However, CI coverage was nominal for some ω when only the exposure model was correctly specified. For $n = 1000$ and $n = 2000$, patterns of performance were largely the same, but CI coverage did not appreciably change for the more extreme policies (Appendix Figure C.14–C.15 Figure C.17–C.18, Tables D.19–D.22).

5.3 Comprehensive Dietary Intervention and BMI

For simulations of a comprehensive dietary intervention on BMI with both unit-treatment and spillover effects, network-TMLE point estimates exhibited minimal bias and the corresponding CI coverage was approximately 95% in most scenarios. For the uniform random graph ($n = 500$), there was little bias, and coverage was near 95% when both models were correctly specified (Figure 5). Performance was similar when the exposure model was misspecified. When the outcome model was misspecified, the ESE substantially increased, and coverage was above 95%. The flexible specification of W^S had similar performance to other models but coverage was slightly decreased near the extremes of ω. Results were similar for $n = 1000$ and $n = 2000$ (Appendix C Figure C.22, Figure C.25, Tables D.23–D.25).

The overall patterns with the varying model specifications were similar for the eX-FLU network when restricted by degree (Figure 6). However, the flexible specification of W^S resulted in decreased coverage, particularly for policies where the probability of exposure was set to be low. Restricting by degree reduced the ESE, particularly when the outcome model was misspecified, but did not substantially alter the CI coverage (Appendix Figure C.19, Tables D.26–D.27). However, the unrestricted analyses with the flexible modeling approach had increased CI coverage for policies with low probabilities of exposure

compared to the restricted flexible models, likely due to the additional observations available for estimation.

Network-TMLE's performance across the power-law random graphs ($n = 500$, $n = 1000$, $n = 2000$) generally followed the same patterns as the uniform random graphs. There was little bias across the varying model specifications, misspecification of the outcome model led to increased ESE and conservative CI coverage, and the flexible specification of W^S performed adequately for most policies (Appendix Figure C.23–C.24, Figure C.20–C.21, Figure C26–C.27). As with the eX-FLU network, the ESE decreased when restricting by degree (Appendix Tables D.28–D.33).

5.4 Vaccine and Infectious Disease Transmission

For the uniform network ($n = 500$), there was relatively little bias but 95% CI coverage based on the direct transmission variance estimator was below 0.80 across all policies and model specifications (Figure 7). The 95% CI coverage based on the latent dependence variance estimator was improved but was still below 0.95 (Appendix Table D.34). Similar patterns occurred for the uniform network with $n = 1000$ and $n = 2000$ (Appendix Figure C.31, Figure C.34, Table D.35–D.36).

For the eX-FLU network restricted by degree, bias was present for some policies across all model specifications (Figure 8). Again, the 95% CI coverage was noticeably different between the variance estimators, with the latent dependence variance estimator resulting in improved coverage. However, CI coverage remained below 0.95 for all ω. Comparing different policy specifications, shifts in the propensity scores had less bias and better coverage compared to assigning the same probability of vaccination to all units. When the eX-FLU network was not restricted by degree, bias was largely similar, but the distribution of estimates was more heavily skewed (Appendix Figure C.28). For the clustered power-law random graph, similar patterns of bias and CI coverage were observed for $n = 500$ (Appendix Figures C.29–30), $n = 1000$ (Appendix Figures C.32–33), and $n = 2000$ (Appendix Figures C.35–36).

6. Discussion

Here, TMLE for IID data with stochastic policies and an extension of TMLE to network-dependent data were reviewed. The performance of network-TMLE was assessed in a variety of different network structures, combinations of unit-treatment and spillover effects, and nuisance model specifications. Finally, software implementing network-TMLE in Python is made freely available, which may help facilitate wider application.

Network-TMLE inference about the $\boldsymbol{W}$-conditional mean under different stochastic policies performed well in a variety of simulations based on real-world examples. These results demonstrate the potential utility of network-TMLE over a range of settings. The simulations suggest when employing network-TMLE in practice that networks with a skewed distribution should be restricted by degree and nuisance models should be flexibly specified. Network-TMLE did not perform well when the weak dependence assumption was violated. Performance was also degraded for policies where the probability of exposure was

substantially different from the observed proportion exposed due to a lack of support in the data. For example, when exposure was uncommon, CI coverage was poor for policies where nearly everyone was exposed. In these cases, few individuals with high degree have all their immediate contacts exposed. However, the policy of interest has most individuals having all their immediate contacts being exposed. Therefore, the parametric nuisance models extrapolate from those few individuals to estimate the policy. While these sparsity problems are overcome asymptotically, care should be exercised when employing network-TMLE for inference about policies 'far' from the observed data.

To help identify how well a proposed policy is supported by the observed data, we propose the following diagnostic. For the chosen summary measure A^s, a bar chart or histogram of the observed values are plotted stratified by A. For the proposed policy, the new values A^{s*} are plotted as done for A^s. Therefore, the observed distribution of A^s can be visually compared to the distribution of exposures under the proposed policy A^{s*}. An illustrative example of this diagnostic plot for both well supported and poorly supported policies is provided in Figure 9. Similarly, restricting inference to individuals below a specified degree may reduce bias and improve confidence interval coverage for networks with skewed degree distributions. While individuals above the maximum degree are then considered to be fixed features of the network and the target parameter has a modified interpretation, the improvements may nonetheless be preferred. However, restricting by degree and the resulting decrease in observations available to estimate the nuisance models may exacerbate finite data issues in some cases.

Limiting inference to policies 'close' to the observed data or focusing on policies which more modestly perturb the exposure distribution may also be more relevant from a practical perspective. In the absence of interference, commonly targeted estimands, like the average causal effect, contrast two extreme exposure distributions: everyone exposed versus no one exposed.[33,67–69] Such extreme counterfactual exposure settings may be unrealistic or irrelevant in practice. For instance, when assessing the effect of smoking on some health outcome, the counterfactual scenario where all individuals smoke is likely unrealistic. Rather, there may be more interest in the effect of policies or interventions which modestly decrease the likelihood of smoking. As another example, consider policies to encourage influenza vaccination uptake. Previous approaches have resulted in only minor to moderate increases in the vaccine receipt.[70–73] Therefore, the counterfactual scenario of everyone in the population being vaccinated may be of less relevance, in addition to being difficult to draw valid inferences about.

In practical application, nuisance models should be flexibly specified. Here, we demonstrated one option that binned W^s. To help reduce unnecessary increases in the dimensionality of the model, the procedure was further paired with a L2-penalized regression model. Overall, the flexible modeling approach performed well over a wide range of settings and was comparable in terms of bias and coverage to network-TMLE when both nuisance models were correctly specified. As correct model specification is unlikely in practice, flexible specification of the nuisance models is recommended. Nonparametric or data-adaptive approaches could be utilized; for instance $\Pr\left(A_i^s \mid A_i, W_i, W_i^s\right)$ could be

estimated with a conditional density super learner.[74] Provided the nonparametric estimators of the nuisance models converge at a sufficiently fast rate and other conditions hold, the network-TMLE estimator will still be consistent and asymptotically normal.[24] Additional empirical research is needed to study the finite-sample performance of network-TMLE when such data-adaptive approaches are used for nuisance model estimation.

In the vaccination and infection mechanism, the infection status of a unit depends on units outside of its immediate contacts. This more widespread dependence, a violation of weak dependence, likely resulted in the occurrence of some bias, reduced CI coverage, and explains the difference observed between the variance estimators. To address weak dependence, the extension of network-TMLE for longitudinal data is needed.[22] A longitudinal extension could instead require that weak dependence holds only within each measured time interval,[24] as opposed to weak dependence holding over the entire duration of follow-up. Furthermore, the longitudinal extension would allow for a summary measure of infectious immediate contacts within each time interval to be included in the nuisance models.

Future work could consider the following. Our simulation study assessed a flexible approach for including summary measure of covariates, W^s, in nuisance models. However, the summary measure A^s was assumed to be known. While reliable background information on A^s may be known in some settings, this will not always be the case. One method of avoiding specification of a particular summary measure for A^s in the exposure model is to factor A^s into b different binary conditional distributions (where b is the maximum degree) and estimate these conditional distributions with a series of b logistic models.[23] However, this approach is limited to scenarios where the degree distribution is near uniform. Instead, a similar categorization and binning approach could also be applied to A^s,[24] which would apply to both nuisance models and allow for non-uniform degree distributions. Here, two closed-form variance estimators for the $\boldsymbol{W}$-conditional mean were compared. Other research has proposed bootstrap variance estimators,[75,76] with a parametric bootstrap estimator for network-TMLE in the context of the network mean outperforming the closed-form variance.[24] Further comparison of bootstrap variance estimators in terms of their assumptions and performance remains of interest. Additional empirical evaluation could be conducted of network-TMLE for other estimands such as marginal unit-treatment effects (i.e. direct effects).[23] Generalization of network-TMLE to other related estimands (i.e., spillover effects, total effects) is also of interest. Finally, direct comparisons between network-TMLE and auto-g-computation, a recent extension of the parametric g-formula for general interference,[77] could be undertaken.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

PNZ was supported by T32-HD091058 and T32-AI007001. MGH was supported by NIH grant R01-AI085073. AEA received funding from NIBIB R01-EB025021 and acknowledges NICHD T32-HD091058 and P2C HD050924. The eX-FLU study was funded by United States Centers for Disease Control and Prevention U01-

CK000185 grant. 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) or the United States Centers for Disease Control and Prevention.

We thank Elizabeth Ogburn for her feedback and discussion. We would further like to thank the University of North Carolina at Chapel Hill and the Research Computing group for providing computational resources that have contributed to these results. The Python implementation of network-TMLE is available as MossSpider on the Python Package Index and on GitHub (github.com/pzivich/MossSpider). Code to replicate the simulations is available at https://github.com/pzivich/publications-code.

References:

1. Hudgens MG, Halloran ME. Toward Causal Inference With Interference. Journal of the American Statistical Association. 2008;103(482):832–842. doi:10.1198/016214508000000292 [PubMed: 19081744]
2. Cox DR. Planning of Experiments. Vol 20. Wiley New York; 1958.
3. Rubin DB. Randomization analysis of experimental data: The Fisher randomization test comment. Journal of the American Statistical Association. 1980;75(371):591–593.
4. Miller M, Neaigus A. Sex partner support, drug use and sex risk among HIV-negative non-injecting heroin users. AIDS Care. 2002;14(6):801–813. doi:10.1080/0954012021000031877 [PubMed: 12511213]
5. Miller M The dynamics of substance use and sex networks in HIV transmission. Journal of Urban Health. 2003;80(3):iii88–iii96. doi:10.1093/jurban/jtg086 [PubMed: 14713675]
6. Paluck EL, Shepherd H, Aronow PM. Changing climates of conflict: A social network experiment in 56 schools. PNAS. 2016;113(3):566–571. doi:10.1073/pnas.1514483113 [PubMed: 26729884]
7. Zhang S, de la Haye K, Ji M, An R. Applications of social network analysis to obesity: a systematic review. Obesity Reviews. 2018;19(7):976–988. [PubMed: 29676508]
8. Holtz D, Zhao M, Benzell SG, et al. Interdependence and the cost of uncoordinated responses to COVID-19. PNAS. 2020;117(33):19837–19843. doi:10.1073/pnas.2009522117 [PubMed: 32732433]
9. Li Y, Undurraga EA, Zubizarreta JR. Effectiveness of Localized Lockdowns in the COVID-19 Pandemic. American Journal of Epidemiology. 2022;191(5):812–824. doi:10.1093/aje/kwac008 [PubMed: 35029649]
10. Seamans MJ, Carey TS, Westreich DJ, et al. Association of Household Opioid Availability and Prescription Opioid Initiation Among Household Members. JAMA Intern Med. 2018;178(1):102–109. doi:10.1001/jamainternmed.2017.7280 [PubMed: 29228098]
11. Jarvis MJ, Feyerabend C, Bryant A, Hedges B, Primatesta P. Passive smoking in the home: plasma cotinine concentrations in non-smokers with smoking partners. Tob Control. 2001;10(4):368–374. doi:10.1136/tc.10.4.368 [PubMed: 11740030]
12. Du Y, Cui X, Sidorenkov G, et al. Lung cancer occurrence attributable to passive smoking among never smokers in China: a systematic review and meta-analysis. Transl Lung Cancer Res. 2020;9(2):204–217. doi:10.21037/tlcr.2020.02.11 [PubMed: 32420060]
13. Foster EM. Causal inference and developmental psychology. Dev Psychol. 2010;46(6):1454–1480. doi:10.1037/a0020204 [PubMed: 20677855]
14. Halloran ME, Haber M, Longini IM Jr, Struchiner CJ. Direct and indirect effects in vaccine efficacy and effectiveness. Am J Epidemiol. 1991;133(4):323–331. [PubMed: 1899778]
15. Sobel ME. What Do Randomized Studies of Housing Mobility Demonstrate? Journal of the American Statistical Association. 2006;101(476):1398–1407. doi:10.1198/016214506000000636
16. Halloran ME, Hudgens MG. Dependent Happenings: A Recent Methodological Review. Curr Epidemiol Rep. 2016;3(4):297–305. doi:10.1007/s40471-016-0086-4 [PubMed: 28133589]
17. Tchetgen EJT, VanderWeele TJ. On causal inference in the presence of interference. Statistical methods in medical research. October 11;21(1):55–75. doi:10.1177/0962280210386779
18. Perez-Heydrich C, Hudgens MG, Halloran ME, Clemens JD, Ali M, Emch ME. Assessing effects of cholera vaccination in the presence of interference. Biometrics. 2014;70(3):731–744. doi:10.1111/biom.12184 [PubMed: 24845800]

19. Liu L, Hudgens MG, Saul B, Clemens JD, Ali M, Emch ME. Doubly robust estimation in observational studies with partial interference. Stat. 2019;8(1):e214. [PubMed: 31440374]
20. Aronow PM, Samii C. Estimating average causal effects under general interference, with application to a social network experiment. The Annals of Applied Statistics. 2017;11(4):1912–1947.
21. Bowers J, Fredrickson MM, Panagopoulos C. Reasoning about interference between units: A general framework. Political Analysis. 2013;21(1):97–124.
22. van der Laan MJ. Causal Inference for a Population of Causally Connected Units. J Causal Inference. 2014;2(1):13–74. doi:10.1515/jci-2013-0002 [PubMed: 26180755]
23. Sofrygin O, van der Laan MJ. Semi-Parametric Estimation and Inference for the Mean Outcome of the Single Time-Point Intervention in a Causally Connected Population. Journal of causal inference. 2017;5(1):20160003. doi:10.1515/jci-2016-0003 [PubMed: 29057197]
24. Ogburn EL, Sofrygin O, Diaz I, van der Laan MJ. Causal inference for social network data. Published online June 1, 2022. doi:10.48550/arXiv.1705.08527
25. Smith JA, Moody J. Structural Effects of Network Sampling Coverage I: Nodes Missing at Random(1). Soc Networks. 2013;35(4):10.1016/j.socnet.2013.09.003. doi:10.1016/j.socnet.2013.09.003
26. Newman MEJ, Park J. Why social networks are different from other types of networks. Physical Review E. 9;68(3):036122. doi:10.1103/PhysRevE.68.036122
27. Badham J, Kee F, Hunter RF. Network structure influence on simulated network interventions for behaviour change. Social Networks. 2021;64:55–62. doi:10.1016/j.socnet.2020.08.003
28. Schuler MS, Rose S. Targeted maximum likelihood estimation for causal inference in observational studies. Am J Epidemiol. 2017;185(1):65–73. [PubMed: 27941068]
29. Van der Laan MJ, Rose S. Targeted Learning: Causal Inference for Observational and Experimental Data. Springer Science & Business Media; 2011.
30. Kennedy EH. Semiparametric theory and empirical processes in causal inference. In: Statistical Causal Inferences and Their Applications in Public Health Research. Springer; 2016:141–167.
31. Zivich PN, Breskin A. Machine Learning for Causal Inference: On the Use of Cross-fit Estimators. Epidemiology. 2021;32(3):393–401. doi:10.1097/EDE.0000000000001332 [PubMed: 33591058]
32. Schomaker M, Luque-Fernandez MA, Leroy V, Davies MA. Using longitudinal targeted maximum likelihood estimation in complex settings with dynamic interventions. Stat Med. 2019;38(24):4888–4911. [PubMed: 31436859]
33. Muñoz ID, van der Laan M. Population Intervention Causal Effects Based on Stochastic Interventions. Biometrics. 2012;68(2):541–549. doi:10.1111/j.1541-0420.2011.01685.x [PubMed: 21977966]
34. Cole SR, Frangakis CE. The consistency statement in causal inference: a definition or an assumption? Epidemiology. 2009;20(1):3–5. [PubMed: 19234395]
35. Hernán MA, Robins JM. Estimating causal effects from epidemiological data. Journal of Epidemiology and Community Health. 2006;60(7):578–586. doi:10.1136/jech.2004.029496 [PubMed: 16790829]
36. Balzer LB, Petersen ML, van der Laan MJ, SEARCH Collaboration. Targeted estimation and inference for the sample average treatment effect in trials with and without pair-matching. Stat Med. 2016;35(21):3717–3732. doi:10.1002/sim.6965 [PubMed: 27087478]
37. Rose S, van der Laan MJ. Understanding TMLE. In: van der Laan MJ, Rose S, eds. Targeted Learning: Causal Inference for Observational and Experimental Data. Springer Series in Statistics. Springer; 2011:83–100. doi:10.1007/978-1-4419-9782-1_5
38. Díaz I, van der Laan MJ. Stochastic treatment regimes. In: Targeted Learning in Data Science. Springer; 2018:219–232.
39. Sekhon JS, Gruber S, Porter KE, van der Laan MJ. Propensity-Score-Based Estimators and C-TMLE. In: van der Laan MJ, Rose S, eds. Targeted Learning: Causal Inference for Observational and Experimental Data. Springer Series in Statistics. Springer; 2011:343–364. doi:10.1007/978-1-4419-9782-1_21

40. Seaman S, Pavlou M, Copas A. Review of methods for handling confounding by cluster and informative cluster size in clustered data. Stat Med. 2014;33(30):5371–5387. doi:10.1002/sim.6277 [PubMed: 25087978]
41. Imai K, van Dyk DA. Causal Inference With General Treatment Regimes. Journal of the American Statistical Association. 2004;99(467):854–866. doi:10.1198/016214504000001187
42. Caldarelli G Scale-Free Networks: Complex Webs in Nature and Technology. OUP Oxford; 2007.
43. Aiello AE, Simanek AM, Eisenberg MC, et al. Design and methods of a social network isolation study for reducing respiratory infection transmission: The eX-FLU cluster randomized trial. Epidemics. 2016;15:38–55. doi:10.1016/j.epidem.2016.01.001 [PubMed: 27266848]
44. Holme P, Kim BJ. Growing scale-free networks with tunable clustering. Physical Review E. 2002;65(2):026107.
45. Denby L, Mallows C. Variations on the Histogram. Journal of Computational and Graphical Statistics. 2009;18(1):21–31. doi:10.1198/jcgs.2009.0002
46. Ference BA, Ginsberg HN, Graham I, et al. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. European Heart Journal. 2017;38(32):2459–2472. doi:10.1093/eurheartj/ehx144 [PubMed: 28444290]
47. Buhaescu I, Izzedine H. Mevalonate pathway: A review of clinical and therapeutical implications. Clinical Biochemistry. 2007;40(9):575–584. doi:10.1016/j.clinbiochem.2007.03.016 [PubMed: 17467679]
48. Grundy SM, Stone NJ, Bailey AL, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA guideline on the management of blood cholesterol: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Journal of the American College of Cardiology. 2019;73(24):e285–e350. [PubMed: 30423393]
49. Rudd RA, Aleshire N, Zibbell JE, Matthew Gladden R. Increases in Drug and Opioid Overdose Deaths—United States, 2000–2014. American Journal of Transplantation. 2016;16(4):1323–1327. doi:10.1111/ajt.13776
50. Wilson N. Drug and opioid-involved overdose deaths—United States, 2017–2018. MMWR Morbidity and Mortality Weekly Report. 2020;69.
51. Handal KA, Schauben JL, Salamone FR. Naloxone. Annals of Emergency Medicine. 1983;12(7):438–445. [PubMed: 6309038]
52. Dwyer K, Walley AY, Langlois BK, et al. Opioid education and nasal naloxone rescue kits in the emergency department. West J Emerg Med. 2015;16(3):381–384. doi:10.5811/westjem.2015.2.24909 [PubMed: 25987910]
53. Clark AK, Wilder CM, Winstanley EL. A Systematic Review of Community Opioid Overdose Prevention and Naloxone Distribution Programs. Journal of Addiction Medicine. 2014;8(3):153–163. doi:10.1097/adm.0000000000000034 [PubMed: 24874759]
54. Green TC, Ray M, Bowman SE, McKenzie M, Rich JD. Two cases of intranasal naloxone self-administration in opioid overdose. Substance Abuse. 2014;35(2):129–132. [PubMed: 24821348]
55. Binswanger IA, Blatchford PJ, Mueller SR, Stern MF. Mortality after prison release: opioid overdose and other causes of death, risk factors, and time trends from 1999 to 2009. Annals of Internal Medicine. 2013;159(9):592–600. [PubMed: 24189594]
56. Zedler B, Xie L, Wang L, et al. Risk Factors for Serious Prescription Opioid-Related Toxicity or Overdose among Veterans Health Administration Patients. Pain Medicine. 2014;15(11):1911–1929. doi:10.1111/pme.12480 [PubMed: 24931395]
57. Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. New England Journal of Medicine. 2007;357(4):370–379. doi:10.1056/NEJMsa066082 [PubMed: 17652652]
58. Smith NR, Zivich PN, Frerichs L. Social Influences on Obesity: Current Knowledge, Emerging Methods, and Directions for Future Research and Practice. Current Nutrition Reports. 2020;9(1):31–41. doi:10.1007/s13668-020-00302-8 [PubMed: 31960341]

59. Lemmens VEPP, Oenema A, Klepp KI, Henriksen HB, Brug J. A systematic review of the evidence regarding efficacy of obesity prevention interventions among adults. Obes Rev. 2008;9(5):446–455. doi:10.1111/j.1467-789X.2008.00468.x [PubMed: 18298429]
60. Halloran ME, Haber M, Longini IM Jr. Interpretation and estimation of vaccine efficacy under heterogeneity. American Journal of Epidemiology. 1992;136(3):328–343. [PubMed: 1415152]
61. Ogburn EL, VanderWeele TJ. Vaccines, contagion, and social networks. Ann Appl Stat. 2017;11(2):919–948. doi:10.1214/17-AOAS1023
62. Harris CR, Millman KJ, van der Walt SJ, et al. Array programming with NumPy. Nature. 2020;585(7825):357–362. doi:10.1038/s41586-020-2649-2 [PubMed: 32939066]
63. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods. 2020;17(3):261–272. [PubMed: 32015543]
64. Seabold S, Perktold J. Statsmodels: Econometric and Statistical Modeling with Python. In: Walt S van der, Millman J, eds. Proceedings of the 9th Python in Science Conference.; 2010:92–96. doi:10.25080/Majora-92bf1922-011
65. Patsy. Python for Data; 2022. Accessed June 6, 2022. https://github.com/pydata/patsy
66. Hagberg AA, Schult DA, Swart PJ. Exploring network structure, dynamics, and function using NetworkX. In: Varoquaux G, Vaught T, Millman J, eds.; 2008:11–15.
67. Edwards JK, Cole SR, Lesko CR, et al. An Illustration of Inverse Probability Weighting to Estimate Policy-Relevant Causal Effects. Am J Epidemiol. 2016;184(4):336–344. doi:10.1093/aje/kwv339 [PubMed: 27469514]
68. Westreich D From exposures to population interventions: pregnancy and response to HIV therapy. Am J Epidemiol. 2014;179(7):797–806. doi:10.1093/aje/kwt328 [PubMed: 24573538]
69. Kennedy EH. Nonparametric Causal Effects Based on Incremental Propensity Score Interventions. Journal of the American Statistical Association. 2019;114(526):645–656. doi:10.1080/01621459.2017.1422737
70. Shropshire AM, Brent-Hotchkiss R, Andrews UK. Mass media campaign impacts influenza vaccine obtainment of university students. J Am Coll Health. 2013;61(8):435–443. doi:10.1080/07448481.2013.830619 [PubMed: 24152021]
71. Monn JL. An Evidence-based Project to Improve Influenza Immunization Uptake. The Journal for Nurse Practitioners. 2016;12(4):e159–e162. doi:10.1016/j.nurpra.2015.11.030
72. Zimmerman RK, Nowalk MP, Raymund M, et al. Tailored interventions to increase influenza vaccination in neighborhood health centers serving the disadvantaged. Am J Public Health. 2003;93(10):1699–1705. doi:10.2105/ajph.93.10.1699 [PubMed: 14534225]
73. Stinchfield PK. Practice-proven interventions to increase vaccination rates and broaden the immunization season. The American Journal of Medicine. 2008;121(7):S11–S21.
74. Díaz Muñoz I, van der Laan MJ. Super learner based conditional density estimation with application to marginal structural models. Int J Biostat. 2011;7:Article 38. doi:10.2202/1557-4679.1356
75. Del Prete D, Forastiere L, Sciabolazza VL. Causal Inference on Networks under Continuous Treatment Interference. arXiv:200413459 [econ, stat]. Published online April 28, 2020. Accessed January 6, 2022. http://arxiv.org/abs/2004.13459
76. Forastiere L, Airoldi EM, Mealli F. Identification and Estimation of Treatment and Interference Effects in Observational Studies on Networks. Journal of the American Statistical Association. 2021;116(534):901–918. doi:10.1080/01621459.2020.1768100
77. Tchetgen Tchetgen EJ, Fulcher IR, Shpitser I. Auto-G-Computation of Causal Effects on a Network. Journal of the American Statistical Association. 2021;116(534):833–844. doi:10.1080/01621459.2020.1811098 [PubMed: 34366505]

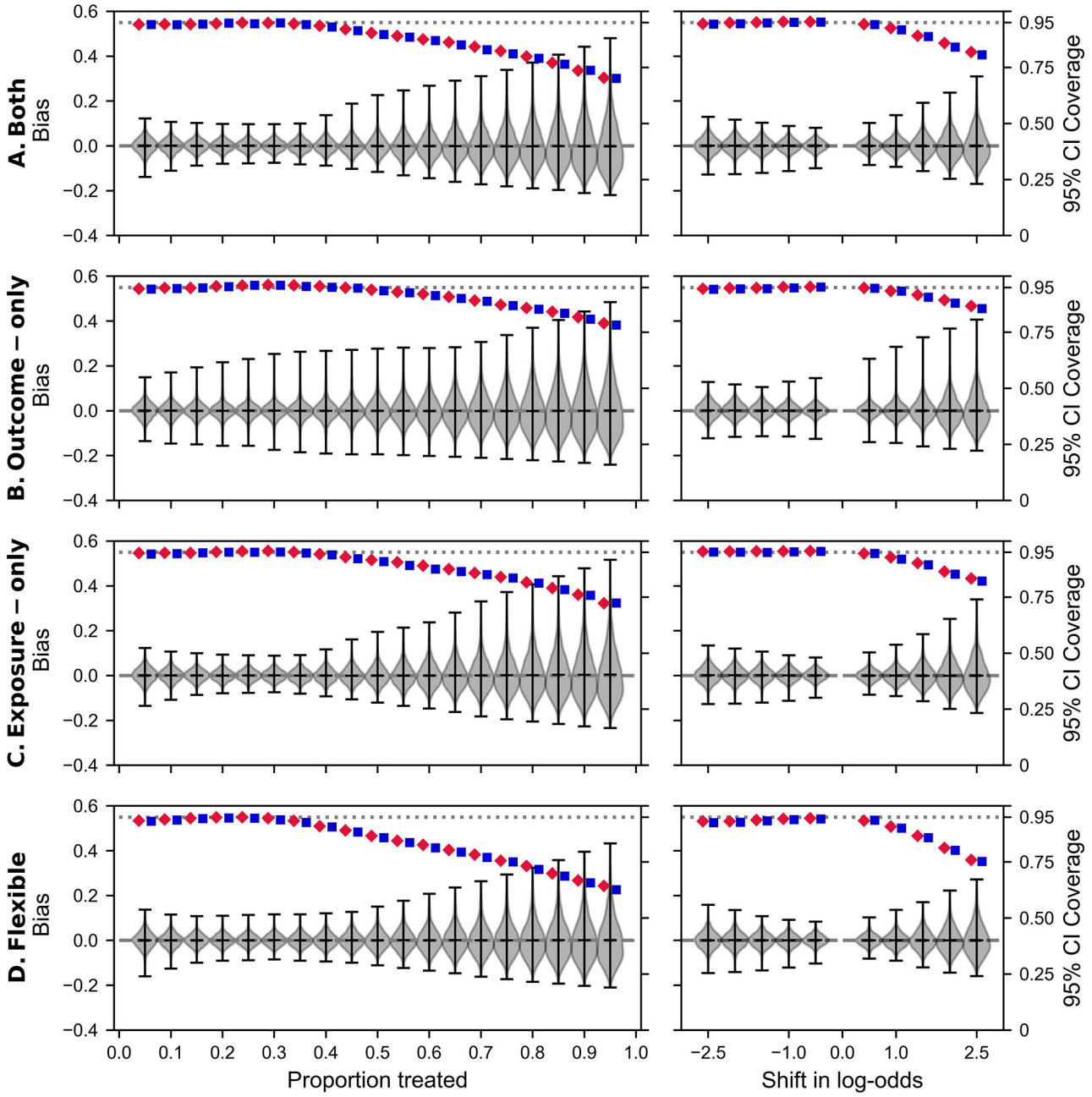

Figure 1:

Target maximum likelihood estimation for statins and atherosclerotic heart disease, and the uniform random graph. Left y-axes and violin plots correspond to bias, defined as the estimated conditional sample mean minus the true conditional sample mean. The right y-axes and diamonds correspond to 95% confidence interval (CI) coverage The red diamond corresponds to the direct-transmission-only variance estimator and the blue square corresponds to the latent-variable-dependence variance estimators. The first column corresponds to all individuals in the population having the same set probability of statins. The second column corresponds to the shift in log-odds of the predicted probability of statins for each individual. The proportion of statins in the observed data was 25%. A:

Network-TMLE with both nuisance models correctly specified. B: Network-TMLE with the exposure model misspecified. C: Network-TMLE with the outcome model misspecified. D: Network-TMLE with a flexible specification of W^s.

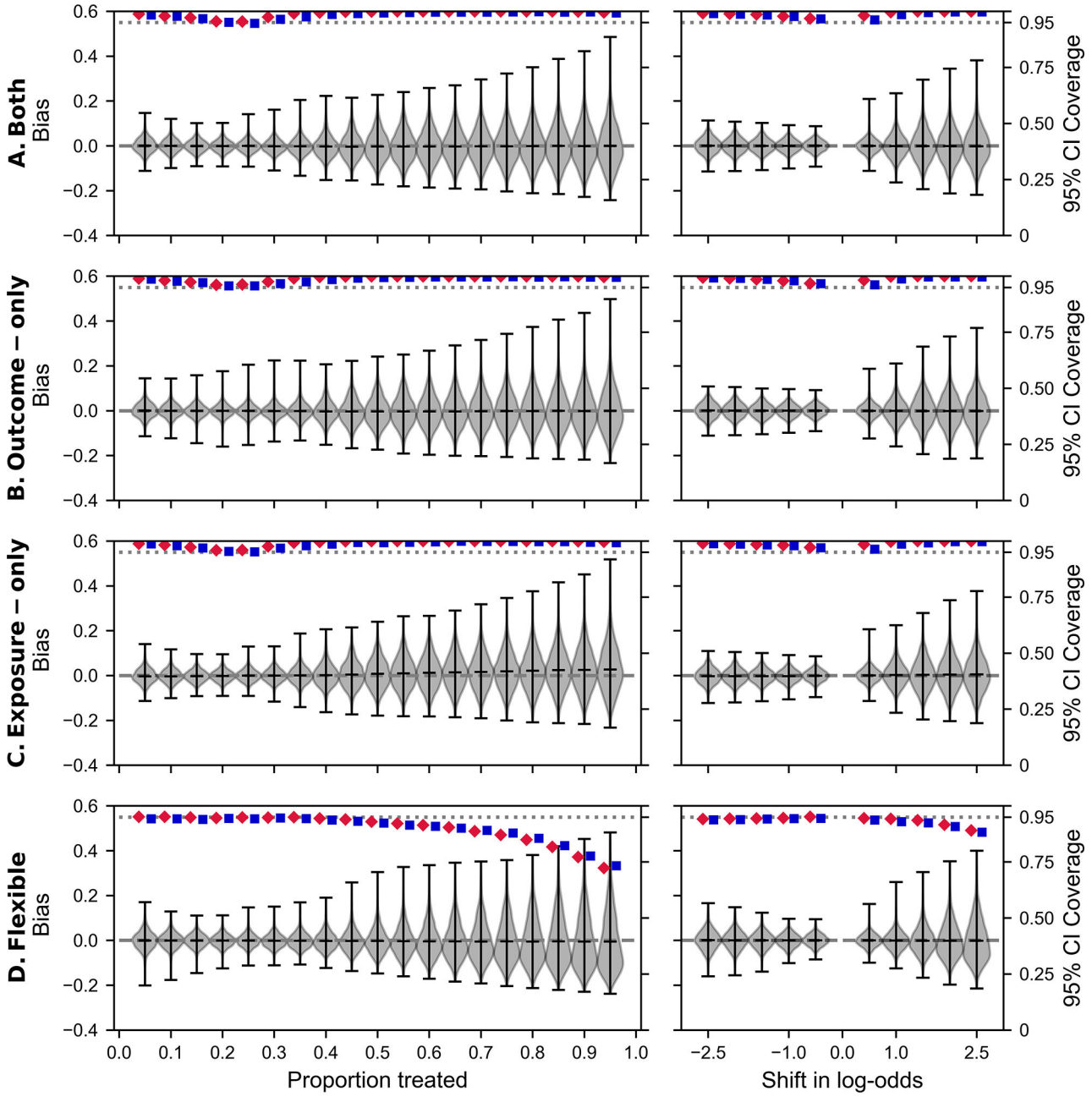

Figure 2:

Target maximum likelihood estimation for statins and atherosclerotic heart disease, and the eX-FLU network restricted by degree. The maximum degree for participants was restricted to be 22 or less. Left y-axes and violin plots correspond to bias, defined as the estimated conditional sample mean minus the true conditional sample mean. The right y-axes and diamonds correspond to 95% confidence interval (CI) coverage. The red diamond corresponds to the direct-transmission-only variance estimator and the blue square corresponds to the latent-variable-dependence variance estimators. The first column corresponds to all individuals in the population having the same set probability of statins. The second column corresponds to the shift in log-odds of the predicted probability of

statins for each individual. The proportion of statins in the observed data was 24%. A: Network-TMLE with both nuisance models correctly specified. B: Network-TMLE with the exposure model misspecified. C: Network-TMLE with the outcome model misspecified. D: Network-TMLE with a flexible specification of W^s.

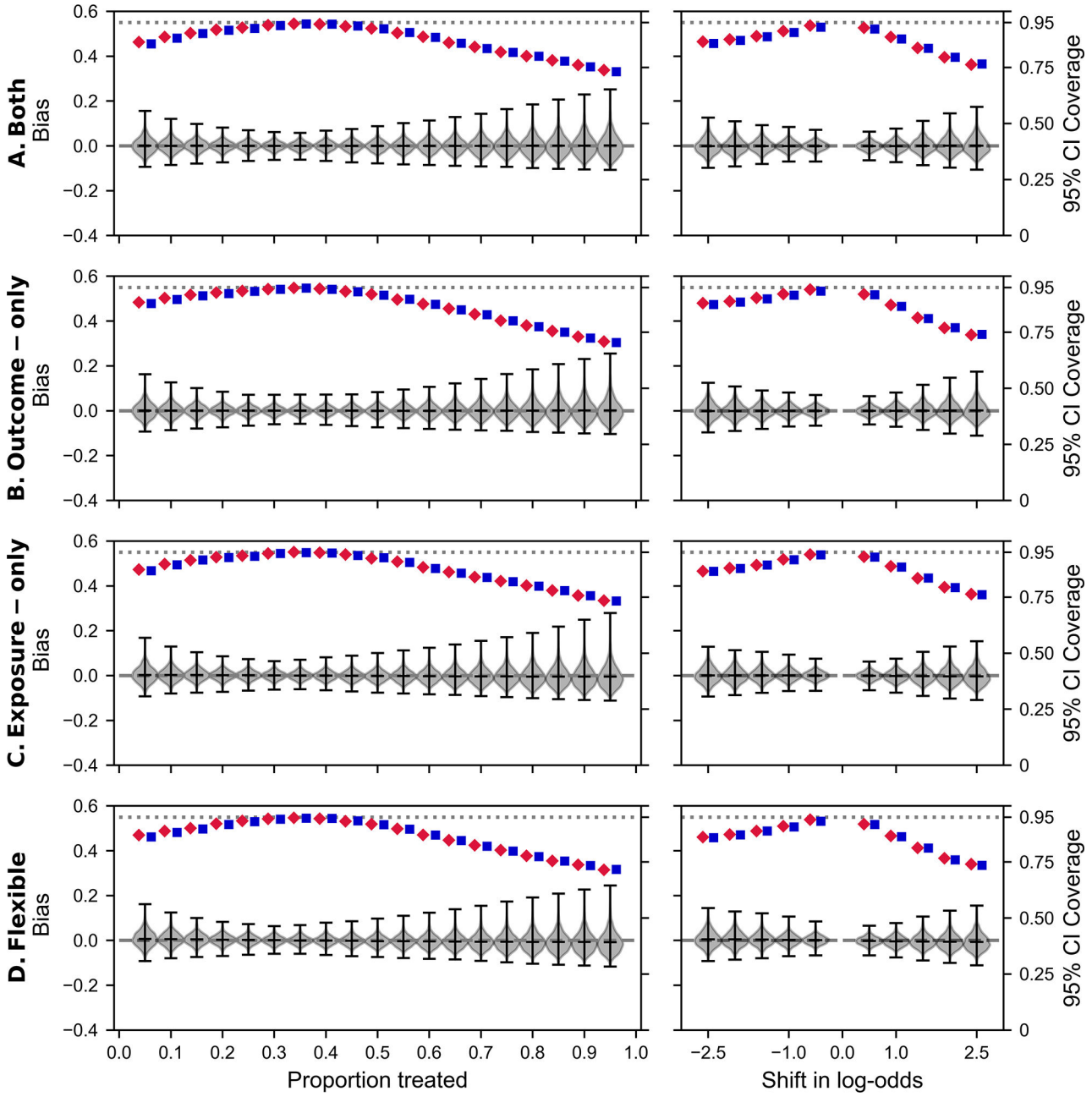

Figure 3:

Target maximum likelihood estimation for naloxone and opioid overdose, and the uniform random graph. Left y-axes and violin plots correspond to bias, defined as the estimated conditional sample mean minus the true conditional sample mean. The right y-axes and diamonds correspond to 95% confidence interval (CI) coverage. The red diamond corresponds to the direct-transmission-only variance estimator and the blue square corresponds to the latent-variable-dependence variance estimators. The first column corresponds to all individuals in the population having the same set probability of naloxone. The second column corresponds to the shift in log-odds of the predicted probability of naloxone for each individual. The proportion of naloxone in the observed data was 35%. A:

Network-TMLE with both nuisance models correctly specified. B: Network-TMLE with the exposure model misspecified. C: Network-TMLE with the outcome model misspecified. D: Network-TMLE with a flexible specification of W^s.

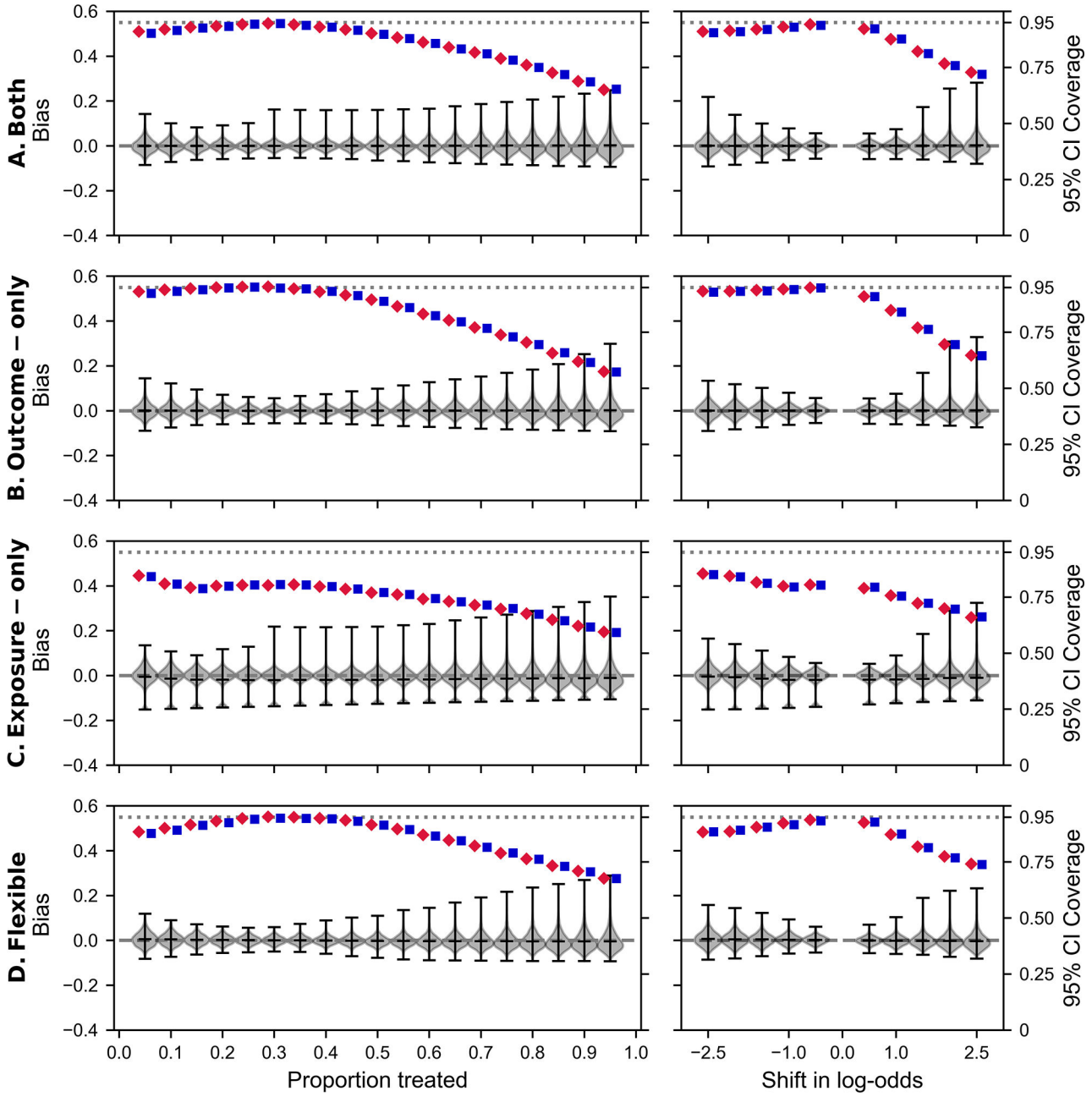

Figure 4:

Target maximum likelihood estimation for naloxone and opioid overdose, and the eX-FLU network restricted by degree. The maximum degree for participants was restricted to be 22 or less. Left y-axes and violin plots correspond to bias, defined as the estimated conditional sample mean minus the true conditional sample mean. The right y-axes and diamonds correspond to 95% confidence interval (CI) coverage. The red diamond corresponds to the direct-transmission-only variance estimator and the blue square corresponds to the latent-variable-dependence variance estimators. The first column corresponds to all individuals in the population having the same set probability of naloxone. The second column corresponds to the shift in log-odds of the predicted probability of naloxone for each individual.

The proportion of naloxone in the observed data was 34%. A: Network-TMLE with both nuisance models correctly specified. B: Network-TMLE with the exposure model misspecified. C: Network-TMLE with the outcome model misspecified. D: Network-TMLE with a flexible specification of W^s.

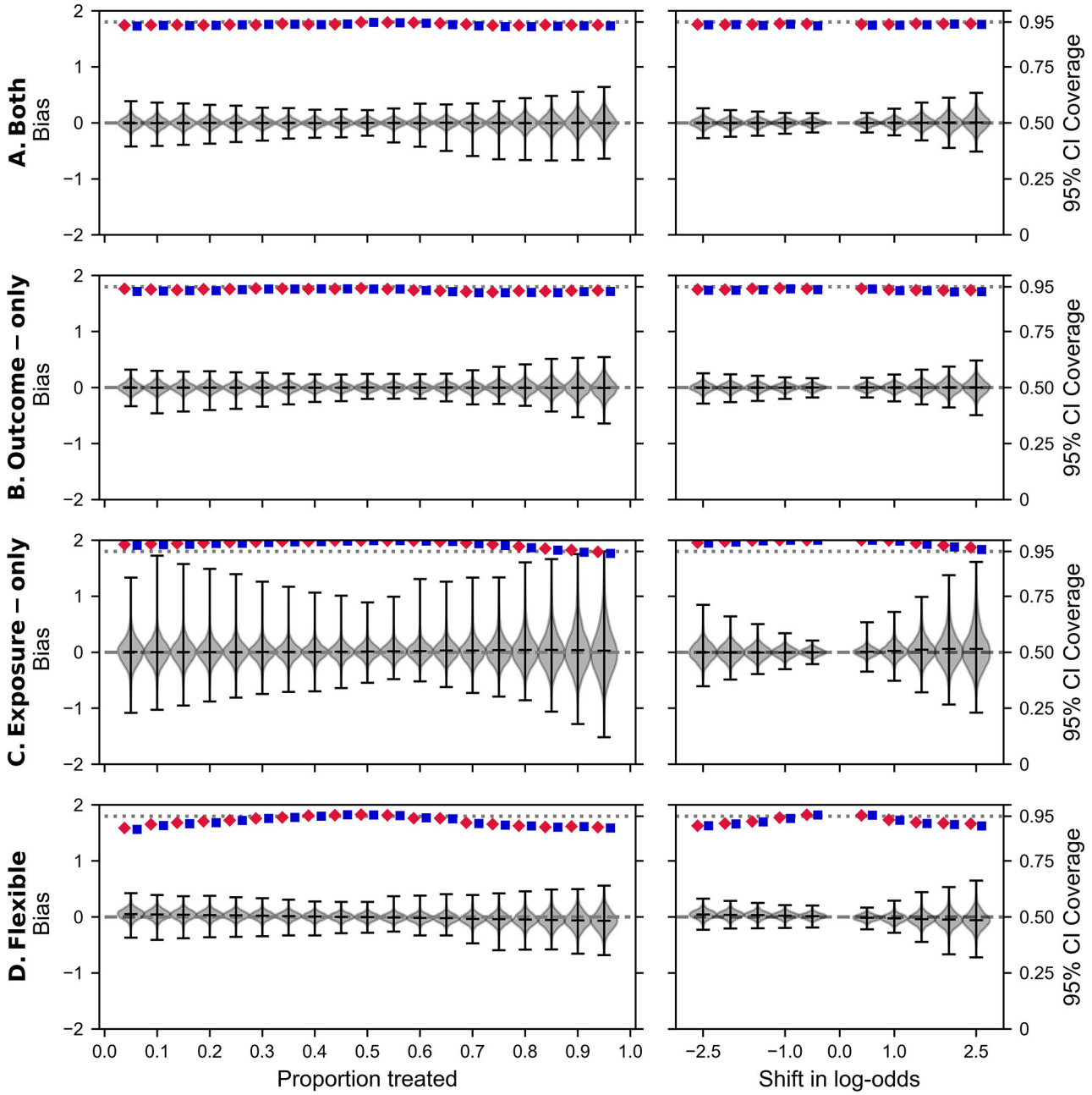

Figure 5:

Target maximum likelihood estimation for diet and body mass index, and the uniform random graph. Left y-axes and violin plots correspond to bias, defined as the estimated conditional sample mean minus the true conditional sample mean. The right y-axes and diamonds correspond to 95% confidence interval (CI) coverage. The red diamond corresponds to the direct-transmission-only variance estimator and the blue square corresponds to the latent-variable-dependence variance estimators. The first column corresponds to all individuals in the population having the same set probability of diet. The second column corresponds to the shift in log-odds of the predicted probability of diet for each individual. The proportion on a diet in the observed data was 48%. A: Network-TMLE

with both nuisance models correctly specified. B: Network-TMLE with the exposure model misspecified. C: Network-TMLE with the outcome model misspecified. D: Network-TMLE with a flexible specification of W^s.

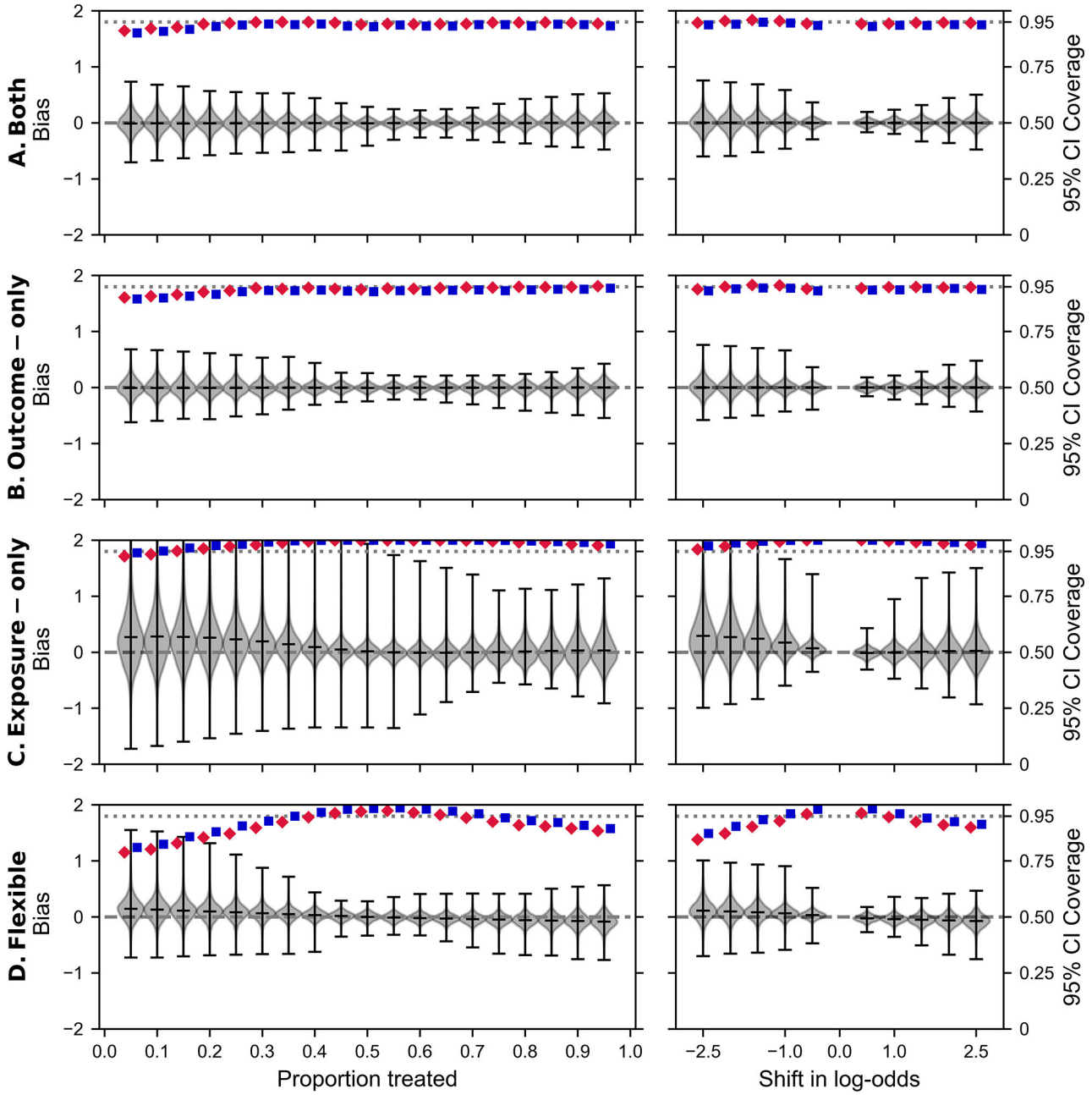

Figure 6:

Target maximum likelihood estimation for diet and body mass index and the eX-FLU network restricted by degree. The maximum degree for participants was restricted to be 22 or less. Left y-axes and violin plots correspond to bias, defined as the estimated conditional sample mean minus the true conditional sample mean. The right y-axes and diamonds correspond to 95% confidence interval (CI) coverage. The red diamond corresponds to the direct-transmission-only variance estimator and the blue square corresponds to the latent-variable-dependence variance estimators. The first column corresponds to all individuals in the population having the same set probability of diet. The second column corresponds to the shift in log-odds of the predicted probability of diet for each individual. The

proportion on a diet in the observed data was 52%. A: Network-TMLE with both nuisance models correctly specified. B: Network-TMLE with the exposure model misspecified. C: Network-TMLE with the outcome model misspecified. D: Network-TMLE with a flexible specification of W^s.

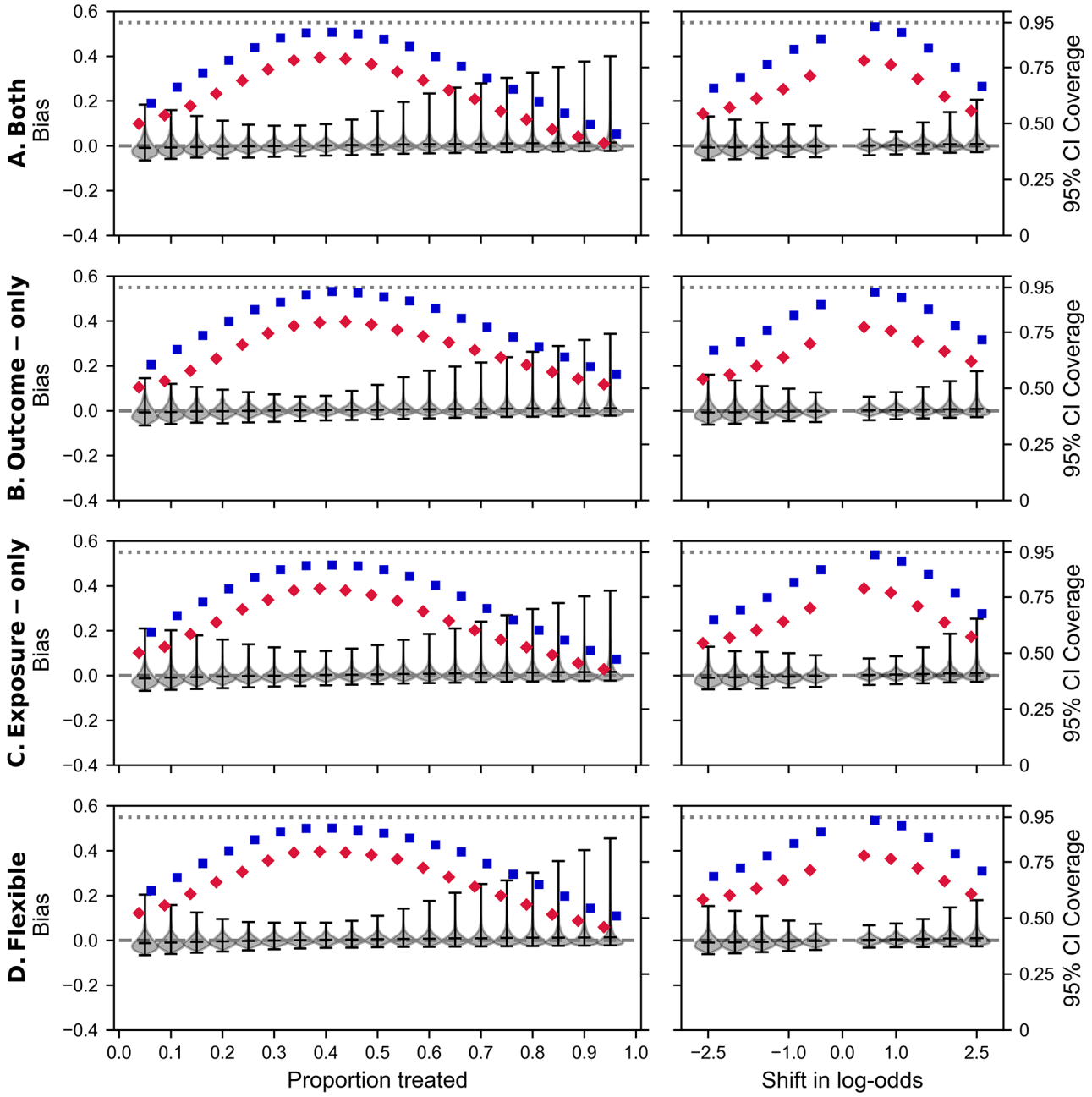

Figure 7:

Target maximum likelihood estimation for vaccination and infection and the uniform random graph. Left y-axes and violin plots correspond to bias, defined as the estimated conditional sample mean minus the true conditional sample mean. The right y-axes and diamonds correspond to 95% confidence interval (CI) coverage. The red diamond corresponds to the direct-transmission-only variance estimator and the blue square corresponds to the latent-variable-dependence variance estimators. The first column corresponds to all individuals in the population having the same set probability of vaccination. The second column corresponds to the shift in log-odds of the predicted probability of vaccination for each individual. The proportion vaccinated in the observed

data was 30%. A: Network-TMLE with both nuisance models correctly specified. B: Network-TMLE with the exposure model misspecified. C: Network-TMLE with the outcome model misspecified. D: Network-TMLE with a flexible specification of W^s.

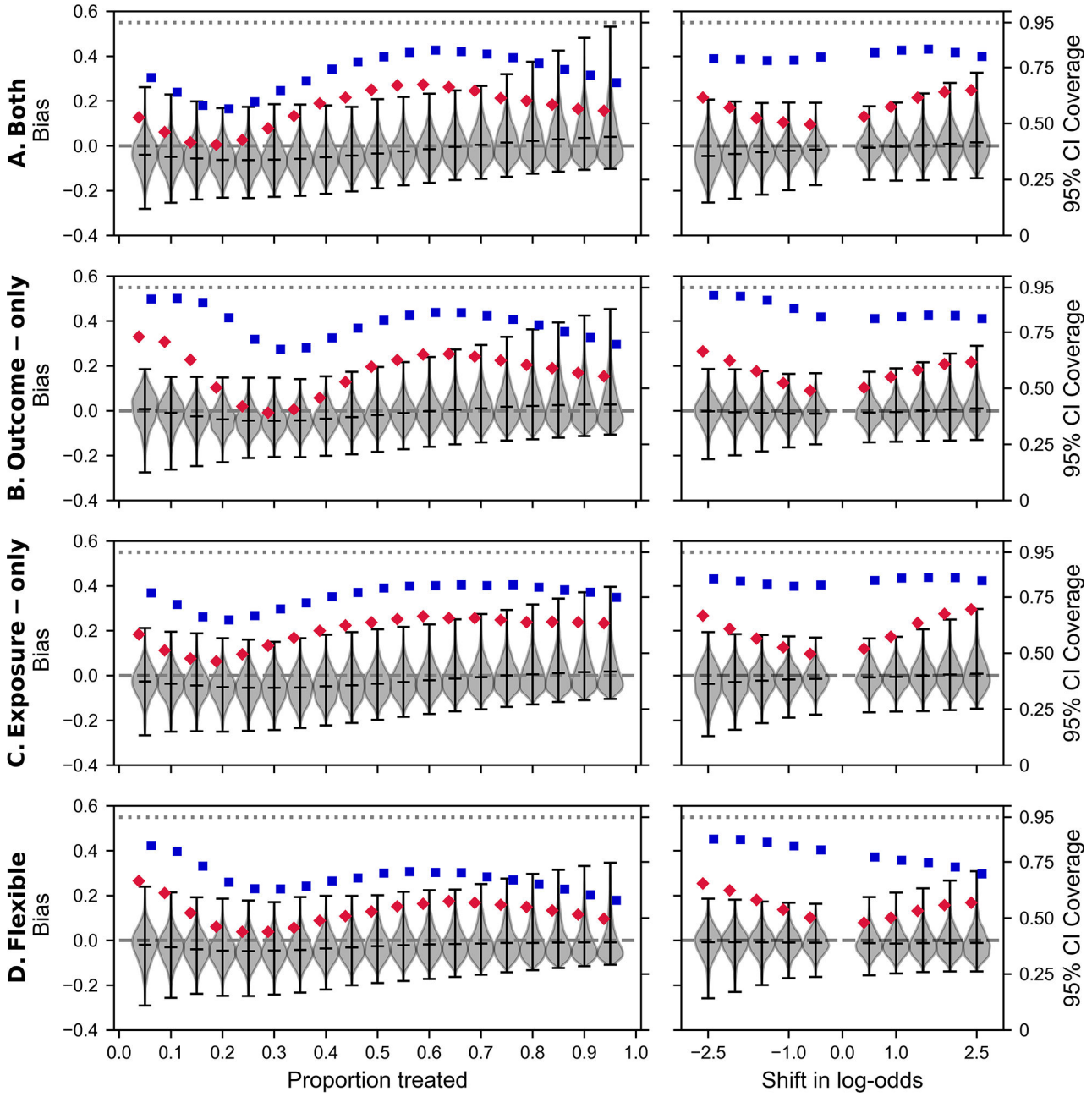

Figure 8:

Target maximum likelihood estimation for vaccination and infection, and the eX-FLU network restricted by degree. The maximum degree for participants was restricted to be 22 or less. Left y-axes and violin plots correspond to bias, defined as the estimated conditional sample mean minus the true conditional sample mean. The right y-axes and diamonds correspond to 95% confidence interval (CI) coverage. The red diamond corresponds to the direct-transmission-only variance estimator and the blue square corresponds to the latent-variable-dependence variance estimators. The first column corresponds to all individuals in the population having the same set probability of vaccination. The second column corresponds to the shift in log-odds of the predicted probability of vaccination for each

individual. The proportion vaccinated in the observed data was 35%. A: Network-TMLE with both nuisance models correctly specified. B: Network-TMLE with the exposure model misspecified. C: Network-TMLE with the outcome model misspecified. D: Network-TMLE with a flexible specification of W^s.

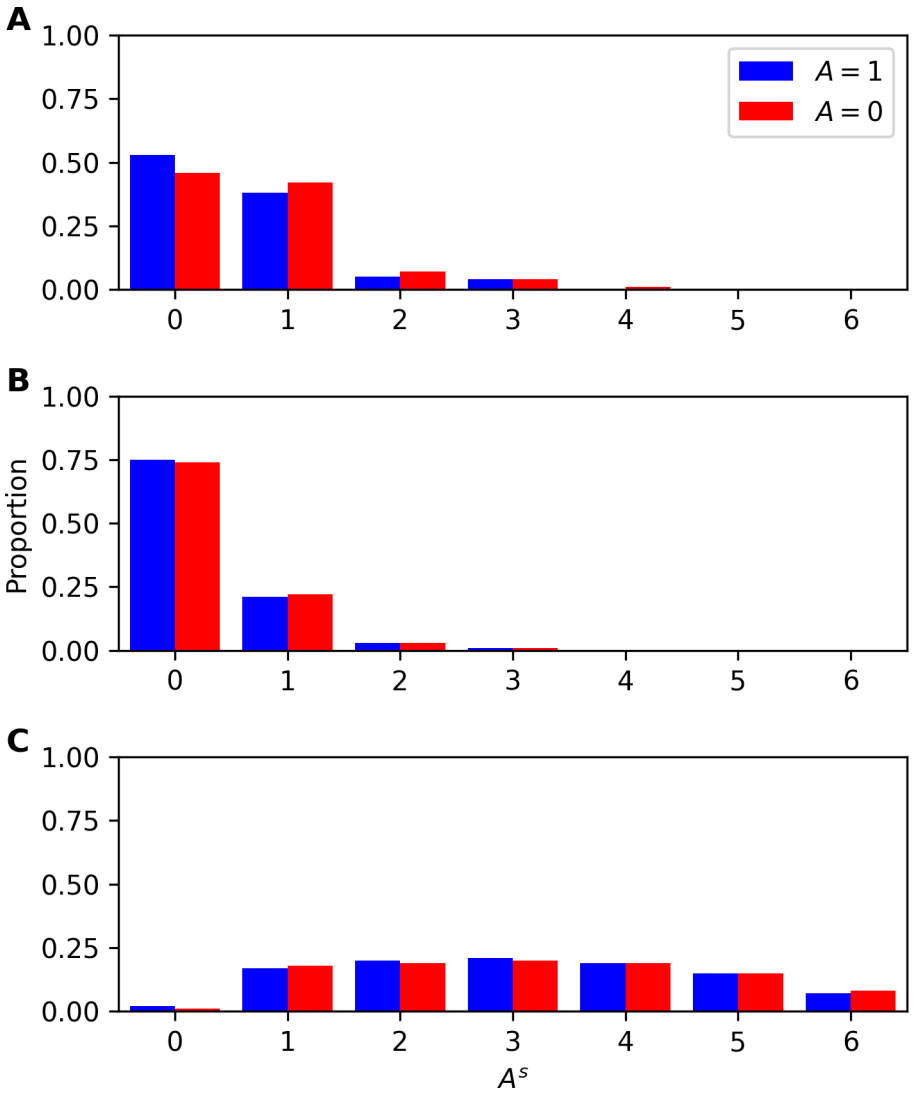

Figure 9:
Proposed diagnostic plots for policies for a network. A: observed distribution of Âs by individual's A for 500 individuals. B: distribution of Âs under a well-supported policy. C: distribution of Âs under a poorly supported policy. Here, the network has a maximum degree of six. Since the policy in C has little-to-no support, estimation of the latter policy should be avoided, or recognize that results are highly dependent on extrapolations from the nuisance models.