



## Research Article

David Stein\*, Lauren D'Arinzo, Fraser Gaspar, Max Oliver, Kristin Fitzgerald, Di Lu, Steven Piantadosi, Alpesh Amin, and Brandon Webb

# A Compound Model of Multiple Treatment Selection with Applications to Marginal Structural Modeling

<https://doi.org/10.1515/sample-YYYY-XXXX>

Received Month DD, YYYY; revised Month DD, YYYY; accepted Month DD, YYYY

**Abstract:** Methods of causal inference are used to estimate treatment effectiveness for non-randomized study designs. The propensity score (i.e., the probability that a subject receives the study treatment conditioned on a set of variables related to treatment and/or outcome) is often used with matching or sample weighting techniques to, ideally, eliminate bias in the estimates of treatment effect due to treatment decisions. If multiple treatments are available, the propensity score is a function of the adjustment set and the set of possible treatments. This paper develops a compound model that separates the treatment decision into a binary decision: treat or don't treat; and a potential treatment decision: choose the treatment that would be given if the subject is treated. It is applicable if the treatment set is finite, treatments are given at one time point, and the outcome is observed at a fixed time point. This representation can reduce bias when not all treatments are available to all patients. Multiple treatment stabilized marginal structural weights were calculated with this approach, and the method was applied to an observational study to evaluate the effectiveness of different neutralizing monoclonal antibodies to treat infection with various severe acute respiratory syndrome coronavirus 2 variants.

**Keywords:** Observational studies, causal inference, marginal structural models, multiple treatment types, monoclonal antibodies, COVID-19

## 1 Introduction

Large health data sets may include structured and unstructured clinical data, indicators of social determinants of health, genomics, and data from wearable sensors. Analysis of these data will contribute to enhanced understanding of health and disease [1]. The US Food and Drug Administration (FDA) continues to expand the use of real-world evidence (RWE), obtained from applying valid inference methods to real-world data, in making regulatory decisions [7]. Estimating treatment effectiveness using data from a non-interventional study, such as an observational study (OBS), requires that variables influencing both treatment decisions and outcomes be controlled to avoid biasing the study conclusions.

Propensity score (PS) methods can be used to control confounding [4, 10, 23]. The PS is often presented as the conditional probability of the outcomes of a binary decision: treat or don't treat [22]. Robins [20] defines a PS for general treatment and observation processes; Imbens [13] defines the generalized propensity

---

\*Corresponding author: David Stein, Lauren D'Arinzo, Fraser Gaspar, Max Oliver, Kristin Fitzgerald, The MITRE Corporation, Bedford, MA USA, e-mail: [dstein@mitre.org](mailto:dstein@mitre.org), [ldarinzo@mitre.org](mailto:ldarinzo@mitre.org), [fgaspar@mitre.org](mailto:fgaspar@mitre.org), [moliver@mitre.org](mailto:moliver@mitre.org), [kafitzgerald@mitre.org](mailto:kafitzgerald@mitre.org)

Di Lu, Aveshka in support of Biomedical Advanced Research and Development Authority, Administration for Strategic Preparedness and Response, US Department of Health and Human Services, Washington D.C., USA, email: [Di.Lu@hhs.gov](mailto:Di.Lu@hhs.gov)

Steven Piantadosi, Harvard Medical School, Boston, MA, USA, email: [spiantadosi@bwh.harvard.edu](mailto:spiantadosi@bwh.harvard.edu)

Alpesh Amin, University of California Irvine Medical Center, Irvine, CA, USA, email: [anamin@uci.edu](mailto:anamin@uci.edu)

Brandon Webb, Intermountain Healthcare, Salt Lake City UT, USA, email: [brandon.webb@imail.org](mailto:brandon.webb@imail.org)

score (GPS), which allows multiple treatments, and Imai and van Dyk [12] define the propensity function that allows multivariate treatments that can be continuous, categorical, or ordinal. Estimation of causal effects with multiple treatments are surveyed in Lopez and Gutman [16]. Methods to estimate the binary propensity score (BPS) are reviewed in Austin and Stuart [23] and Austin [4].

Marginal structural models (MSMs) use sample-weighted logistic regression and other approaches [20] to estimate treatment effects where the independent variables of the regression model are treatment and effect modifiers of interest, and the sample weights are derived from the BPS or GPS [14, 17, 20, 21, 25]. Sample weighting equalizes the distribution of covariates across different treatment groups. This approach avoids the interpolation and numerical issues often encountered when using regression approaches with a large number of covariates [19]. Using stabilized weights in the MSM reduces the dynamic range in comparison with unstabilized weights [10], and augmented weights provide robustness to model mismatch [17]. Hernan et al. found good agreement between treatment effect estimated using MSMs and treatment effect estimated from randomized controlled trials (RCTs) [9]. MSMs have been used to estimate the effects of multiple time varying exposures [5, 8, 11].

This paper shows that, for a discrete treatment set in which one option is don't treat, the GPS can be computed with a BPS, defined as the probability of receiving any treatment, and a potential treatment selection model (PTSM), defined as the probability of receiving each non-null treatment conditioned on the subject being treated, if treatment is given at one time point and outcome is assessed at another single time point. This approach simplifies the computation of multiple treatment MSM weights. Standard methods, [23], may be used to compute the BPS, and the PTSM may be a function of fewer variables than the BPS. All patients may be eligible for non-null treatment; however, all treatments may not be available to all patients. Confounding can occur if the GPS positivity constraint, [10], is violated. In this case, unconfounded estimates can be obtained by partitioning the population into subsets that have a positive probability of receiving every treatment in a subset of potential treatments. This approach was used to estimate the effects of various neutralizing monoclonal antibodies (nMAbs) to treat COVID-19, and sample results are presented.

## 2 A Representation of a Subclass of Multiple treatment Models

Assume a discrete set of treatments that includes the option to not treat. The representation describes treatment selection as a two-stage process: determine whether to treat or not treat; and select the treatment, excluding no treatment, to be provided if the subject is treated. This allows for the computation of the GPS from a BPS and a PTSM.

Assume a discrete set of treatments,  $\mathcal{T} = \{t_0, \dots, t_m\}$ , where  $t_0$  signifies no treatment, and let  $\mathcal{X}$  be the space of subject covariates. The generalized treatment model (GTM) is a random variable

$$A : \mathcal{X} \rightarrow \mathcal{T},$$

and the GPS is the probability density,

$$\begin{aligned} \rho(t, x) &= \text{pr}(A(x) = t) \\ &= \text{pr}(t|x). \end{aligned}$$

Define a binary treatment model (BTM) by

$$A_1(x) = \begin{cases} 0 & : \text{if } A(x) = t_0 \\ 1 & : \text{otherwise.} \end{cases} \quad (1)$$

The GPS,  $\rho$ , induces a BPS

$$\begin{aligned} \rho_1(x) &= 1 - \rho(t_0, x) \\ &= \text{pr}(A_1(x) = 1). \end{aligned} \quad (2)$$

The positivity constraint on  $\rho(x, t)$  ( $0 < \rho(x, t) < 1$  for all  $x, t$ ) [12] implies that  $0 < \rho_1(x) < 1$ . Let  $A_1(A)$  and  $\rho_1(A)$  denote the BTM and BPS, respectively, derived from the GTM  $A$ .

Let  $\mathcal{T}_1 = \{t_1, \dots, t_m\}$ . Given the GTM,  $A$ , and the GPS,  $\rho(x, t)$ , define the PTSM,  $T_1(A) : \mathcal{X} \rightarrow \mathcal{T}_1$  by

$$\begin{aligned} pr(T_1(x) = t_j) &= \frac{\rho(t_j, x)}{pr(A_1(x) = 1)} \\ &= pr(A(x) = t_j | A_1(x) = 1). \end{aligned}$$

Given a BTM  $\alpha : \mathcal{X} \rightarrow \{0, 1\}$  and a PTSM  $\tau : \mathcal{X} \rightarrow \{t_1, \dots, t_m\}$  define the GPS by

$$\rho(x, t_j) = \begin{cases} pr(\alpha(x) = 1)pr(\tau(x) = t_j) & : \text{if } j > 0 \\ pr(\alpha(x) = 0) & : \text{if } j = 0. \end{cases} \quad (3)$$

Let  $\rho(\alpha, \tau)$  denote the GPS derived from  $(\alpha, \tau)$ , and let  $A(\alpha, \tau)$  be the corresponding treatment model.

**Theorem 1.** A GTM  $A : \mathcal{X} \rightarrow \mathcal{T}$  such that  $\mathcal{T} = \{t_0, \dots, t_m\}$ , where  $t_0$  is the option to not treat, is equivalent to a BTM,  $A_1 : \mathcal{X} \rightarrow \{0, 1\}$ , and a PTSM,  $\mathcal{T}_1 : \mathcal{X} \rightarrow \{t_1, \dots, t_m\}$ .

*Proof.* Let  $\alpha$  and  $\tau$  be a BTM and PTSM, respectively. Then,  $A_1(\rho(\alpha, \tau)) = \alpha$ , and  $T_1(\rho(\alpha, \tau)) = \tau$ . Let  $A$  be a GTM. Then  $A(A_1(A), T_1(A)) = A$ . □

### 3 Marginal Structural Model Weights for Multiple Treatments

Assume an independent set of  $N$  samples  $\{(y_j, t_j, x_j, v_j) \mid 1 \leq j \leq N\}$ , where  $y_j$  is the outcome,  $t_j = A(x_j) \in \mathcal{T}$  and  $v_j$  is a vector of effect modifiers of interest. The samples can be expressed as  $\{(y_j, a_j, t_j, x_j, v_j) \mid 1 \leq j \leq N\}$  where  $a_j \in \{0, 1\}$  and  $t_j \in \mathcal{T}_1$ . Note that, if  $a_j = 0$ ,  $t_j$  is interpreted as the potential treatment if subject  $j$  were to receive a treatment;  $t_j$  may be hidden or may be obtained as a sample from  $\mathcal{T}_1$  drawn from the distribution  $p(T_1 | A_1 = 1, x)$ .

**Theorem 2.** Stabilized marginal structural model weights can be calculated from the BPM,  $A_1$ , and the PTSM,  $\mathcal{T}_1$ , as

$$SW_j^1 = \begin{cases} \frac{pr_1(a_j, t_j | v_j)}{pr_1(a_j | x_j, v_j) pr_1(t_j | x_j, v_j)} & \text{if } a_j = 1, \\ \frac{pr_1(a_j | v_j)}{pr_1(a_j | x_j, v_j)} & \text{if } a_j = 0. \end{cases} \quad (4)$$

*Proof.* Let  $A = A(A_1, T_1)$  be the corresponding GTM from Theorem 1. The stabilized MSM weights are, [8, 20],

$$SW_j = \frac{pr(t_j | v_j)}{pr(t_j | x_j, v_j)}. \quad (5)$$

It suffices to show that  $SW_j = SW_j^1$ . Assume  $a_j = 0$ .  $a_j = 0 \iff t_j = t_0$ , and therefore, in this case,  $SW_j = SW_j^1$ . Assume  $a_j = 1$ . We show that the numerators and denominators of  $SW^1$  and  $SW$  are equal. From equation (3),

$$\begin{aligned} pr_1(a_j | x_j, v_j) pr_1(t_j | x_j, v_j) &= r(x_j, v_j, t_j) \\ &= pr(t_j | x_j, v_j). \end{aligned}$$

$$\begin{aligned} pr(t_j | v_j) &= \int pr(t_j, x_j | v_j) dx_j \\ &= \int pr(t_j | x_j, v_j) pr(x_j) dx_j \\ &= \int pr_1(a_j, t_j | x_j, v_j) pr(x_j) dx_j \\ &= pr_1(a_j, t_j | v_j). \end{aligned}$$

□

A proof of Theorem 2 using importance sampling is presented in Appendix B.

Using equation (4) requires that  $pr_1(a_j|x_j, v_j)pr_1(t_j|x_j, v_j) > 0$ . Assume that  $\mathcal{X}$  is the covariate space of all treatment-eligible patients. Because not all treatment-eligible patients may be eligible for all treatments, to avoid confounding,  $\mathcal{X}$  is partitioned into maximal subsets so that for each subset,  $S_k \subset \mathcal{X}$  in the partition, there is an associated subset of treatments,  $T(S_k) \subset \mathcal{T}_1$ , such that for each  $x \in X_k$  and  $t \in T(S_k)$ ,  $p(T = t|X = x) > 0$ . Estimated effects of treatments in  $T(S_k)$  are valid only on  $S_k$ , and confounding can occur if the positivity constraint is violated (see Appendix A). The effect of a treatment available to multiple subsets in the partition may differ on these subsets.

## 4 Application to the Evaluation of nMAbs for the Treatment of COVID-19

The MITRE Corporation and four health systems, sponsored by the US Department of Health and Human Services Administration for Strategic Preparedness and Response, completed an OBS of the effectiveness of nMAbs for treating COVID-19 in accordance with FDA emergency use authorizations (EUAs). The data consisted of over 160,000 deidentified patient records of more than 70 covariates from patients with a positive COVID-19 laboratory test, of whom over 25,000 received nMAbs. The study covered the 15-month time period November 2020–January 2022. A detailed description of the study may be found in [2], and additional results on the effect of social determinants of health on nMAbs utilization and efficacy are described in [3].

nMAbs effectiveness was evaluated using MSMs. Let  $\Theta = \{X_1, \dots, X_K\}$  be the partition of the sample covariate space,  $\mathcal{X}$ , defined as follows. For each  $x \in \mathcal{X}$ , define  $S_x = \{t \in \mathcal{T}_1 \mid p(t|x) > 0\}$ . Index the set of subsets  $\{S_x\}$  by  $\{S_k \mid 1 \leq k \leq K\}$ , and define  $X_k = \{x \in \mathcal{X} \mid S_x = S_k\}$ . Let  $m_k = \|S_k\|$ .  $A \in \{0, 1\}$  is the binary variable indicating no-treatment/treatment,  $t_{jk}$  is the binary variable indicating use of treatment  $t_j$ , and  $g$  is the logit link function. The MSMs used to evaluate different treatments effectiveness were the weighted logistic regression models,

$$g(E(Y^A|\mathcal{T}_1, k)) = \alpha_{0k} + \alpha_{1k}At_{1k} + \dots + \alpha_{m_k k}At_{m_k k}, \quad (6)$$

where the weights were computed from equation (4).

The statistical analysis pipeline included multiple imputation using the R package mice, [24], BPS estimation, modeling of treatment selection probabilities, calculation of MSM weights, and the fitting of the weighted logistic regression models using the Sandwich package [26, 27]. PS modeling was done using random forest, gradient boosted trees, and logistic regression over a set of hyperparameters. A logistic regression model produced the best covariate balance and was used for further analysis. Effects estimated using different imputed data sets were combined to obtain overall effect estimates and standard deviations using Rubin's Rules [15].

The treatment selection model was based on treatment type utilization frequency, which varied over the study period. Patient index date (PID), defined as the date of positive diagnosis, was one of the subject covariates used in the PS model. To protect subject privacy, PID was quantized to one month, and in certain instances, randomly perturbed. Figure 1 shows the percentage of each nMAB type among treated patients across the health systems as a function of study month (SM). The treatment selection model was  $p(t_k|x) = p(t_k|SM)$ , where  $p(t_k|SM)$  was approximated by the sample fractions shown in Figure 1.

The study period was partitioned into several epochs matching the dominant variants—pre-Delta: November 2020–June 2021; Delta: July 2021–November 2021; Delta/Omicron: December 2021; and Omicron: January 2022. Treatments available during part of a study phase were assumed to be available at any time during the phase. Thus, the partitioning of the covariate space required by the positivity constraint was

assumed to be consistent with the study phases. Figures 2 and 3 show which treatments were evaluated during each phase of the study.

The study evaluated the outcomes, emergency department (ED) visits, hospitalizations, deaths, and deaths or hospitalizations within 14 and 30 days after PID. The estimated probabilities of these outcomes and 95% confidence intervals for non-treated patients and patients treated with each of the nMAbs available during each phase are shown in Figures 2 and 3. These probabilities were calculated from the MSMs given in equation 6.

Bamlanivimab alone was used only from November 2020 through March 2021, and therefore bamlanivimab efficacy is comparable only with treatments given during this time period. The results suggest that patients treated with casirivimab-imdevimab, bamlanivimab-etesevimab, and sotrovimab had similar 14- and 30- day hospitalization rates during the Delta and Delta/Omicron phases, and that patients treated with casirivimab-imdevimab had lower 14- and 30- day hospitalization rates than did patients treated with bamlanivimab-etesevimab or sotrovimab during the Omicron phase.

## 5 Discussion

This paper shows that if the treatment set is discrete, the treatment is given at one time point, and the outcome is observed at a fixed time point, then the multiple treatment PS can be expressed as a BPS and a potential treatment model. The BPS is a function of the study covariates and can be computed using standard methods. Marginal structural model weights were derived from this representation. Two proofs were given: one was derived from MSM weights for the GPS, equation (5); and the other, in Appendix B, was derived from importance sampling. Appendix C provides confirmation of these results by comparing effect estimates of a simulated RCT and an OBS. The estimates obtained from the OBS and RCT were equal up to estimation error, and the results from the RCT generally had lower standard error than the results from the OBS. The extension of the binary propensity potential treatment model to more general treatment and outcome scenarios could be investigated.

Confounding can occur in multitreatment studies if not all patients are eligible to receive all treatments. To avoid this, the study population was partitioned so that, for each subset in the partition, there is a subset of treatments such that each subject in the partition subset has a non-zero probability of receiving any treatment in the associated treatment subset. Treatment effectiveness is estimated for each subset in the partition. A treatment associated with multiple subsets of the covariate partition may differ in effectiveness on the subsets, and relative effectiveness of two or more treatments may also differ on these subsets. Appendix A provides an example of this confounding.

Multivariate logistic regression is a common way to estimate a categorical treatment PS [12]. This approach will produce a non-zero probability of treating every patient with any treatment, which, when not true, can lead to biased estimates. Separate modeling of the propensity to treat from the potential treatment selection provides a more flexible and, potentially, more accurate approach.

The method was used for a large observational study. The BPS was fit to data over the entire study period, whereas the potential treatment model was separately estimated for each phase of the pandemic during the study period, as defined by the dominant variant. The study phases did not coincide exactly with the availability of treatments, which could contribute to bias in the effect estimates. Study month was the only independent variable used in the potential treatment model. Health system was also considered and found not to be significant. Other explanatory variables were not considered due to time constraints.

## 6 Acknowledgements

This work was performed by the mAb Real-World Evidence Collaborative. The views expressed are solely those of the authors and do not necessarily represent those of the U.S. Department of Health and Human

Services. This study was supported wholly or in part with federal funds from the Administration for Strategic Preparedness and Response, Biomedical Advanced Research and Development Authority, under Contract Number 75FCMC18D0047, Task Order 75A50121F80012, awarded to The MITRE Corporation.

## References

- [1] Abernethy, A., L. Adams, M. Barrett, C. Bechtel, P. Brennan, A. Butte, J. Faulkner, E. Fontaine, S. Friedhoff, J. Halamka, M. Howell, K. Johnson, P. Lee, P. Long, D. McGraw, R. Miller, J. Perlin, D. Rucker, L. Sandy, L. Savage, L. Stump, P. Tang, E. Topol, R. Tuckson, and K. Valdes (2022). The promise of digital health: Then, now, and the future. *NAM Perspectives*.
- [2] Ambrose, N., A. Amin, B. Anderson, J. Barrera-Oro, M. Bertagnolli, F. Campion, D. Chow, R. Danan, L. D'Arinzo, A. Drews, K. Erlandson, K. Fitzgerald, M. Garcia, F. Gaspar, C. Gong, G. Hanna, S. Jones, B. Lopansri, M. McClellan, J. Musser, J. O'Horo, S. Piantadosi, B. Pritt, R. Razonable, S. Roberts, S. Sandmeyer, D. Stein, F. Vahidy, B. Webb, and J. Yttri. Real world effectiveness of neutralizing monoclonal antibodies for covid-19 from a multicenter consortium study. To appear in *JAMA Open*.
- [3] Ambrose, N., A. Amin, B. Anderson, M. Bertagnolli, F. Campion, D. Chow, R. Danan, L. D'Arinzo, A. Drews, K. Erlandson, K. Fitzgerald, F. Gaspar, C. Gong, G. Hanna, H. Hawley, S. Jones, B. Lopansri, M. McClellan, J. Musser, J. O'Horo, S. Piantadosi, B. Pritt, R. Razonable, S. Rele, S. Roberts, S. Sandmeyer, D. Stein, F. Vahidy, B. Webb, and J. Yttri. The influence of social determinants on receiving outpatient treatment with monoclonal antibodies, disease risk, and effectiveness for covid-19. Submitted.
- [4] Austin, P. C. and E. A. Stuart (2015, December). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine* 34(28), 3661–3679.
- [5] Banack, H. R. and J. S. Kaufman (2015). Estimating the time-varying joint effects of obesity and smoking on all-cause mortality using marginal models. *American Journal of Epidemiology* 183(2), 122–129.
- [6] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- [7] Concato, J. and J. Corrigan-Curay (2022). Real-world evidence — where are we now? *New England Journal of Medicine* 386(18), 1680–1682.
- [8] Hernan, M., B. A. Brumback, and J. M. Robins (2001, June). Marginal structural models to estimate the joint causal effect of nonrandomized treatments. *Journal of the American Statistical Association* 96(454), 440–448.
- [9] Hernan, M., B. A. Brumback, and J. M. Robins (2002, June). Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures. *Statistics in Medicine* 21(12), 1689–1709.
- [10] Hernan, M. and R. James (2020). *Causal Inference: What If*. Chapman & Hall/CRC Press.
- [11] Howe, C. J., S. R. Cole, S. H. Mehta, and G. D. Kirk (2012). Estimating the effects of multiple time-varying exposures using joint marginal structural models: alcohol consumption, injection drug use, and hiv acquisition. *Epidemiology* 23(4), 574–582.
- [12] Imai, K. and D. A. van Dyk (2004). Causal inference with general treatment regimes: Generalizing the propensity score. *Journal of the American Statistical Association* 99(467), 854–866.
- [13] Imbens, G. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika* 87(3), 706–710.
- [14] Joffe, M. M., T. R. Ten Have, H. I. Feldman, and S. E. Kimmel (2004, November). Model selection, confounder control, and marginal structural models: Review and new applications. *The American Statistician* 58(4), 272–279.
- [15] Leyrat, C., S. R. Seaman, I. R. White, I. Douglas, L. Smeeth, J. Kim, M. Resche-Rigon, J. R. Carpenter, and E. J. Williamson (2017). Propensity score analysis with partially observed covariates: How should multiple imputation be used? *Statistical Methods in Medical Research* 28(1), 3–19.
- [16] Lopez, M. J. and R. Gutman (2017). Estimation of causal effects with multiple treatments: A review and new ideas. *Statistical Science* 32(3), 432–454.
- [17] Lunceford, J. and M. Davidian (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: A comparative study. *Statistics in Medicine* 23, 2937–2960.
- [18] MATLAB (2021). *Version 9.11.0 (R2021b)*. Natick, Massachusetts: The MathWorks Inc.
- [19] Ranganathan, P., C. S. Pramesh, and R. Aggarwai (2017). Common pitfalls in statistical analysis: Logistic regression. *Perspectives in Clinical Research* 8, 148–151.
- [20] Robins, J. M. (1998). Marginal structural models. *1997 Proceedings of the American Statistical Association, Section on Bayesian Statistical Science*, 1–10.
- [21] Robins, J. M., M. A. Hernán, and B. A. Brumback (2000, September). Marginal structural models and causal inference in epidemiology. *Epidemiology* 11(5), 550–560.
- [22] Rosenbaum, P. R. (2010). *Design of Observational Studies*. Springer Series in Statistics. New York: Springer. OCLC: ocn444428720.

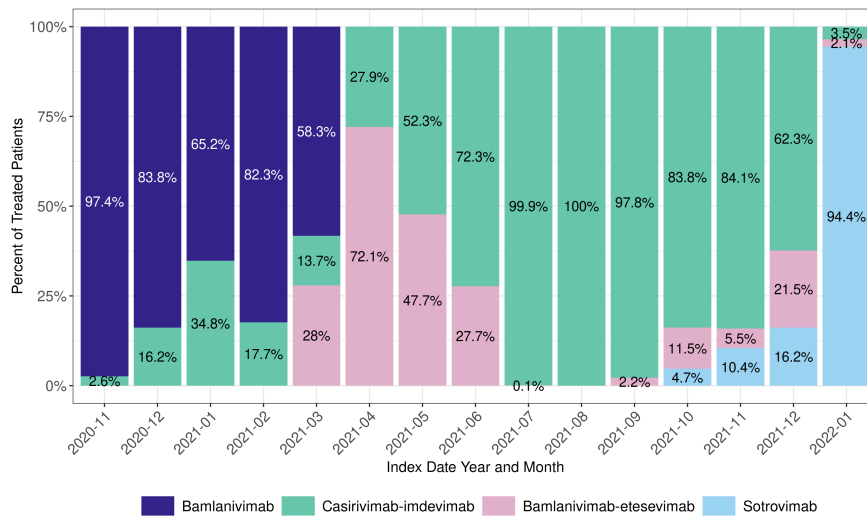
- [23] Stuart, E. A. (2010, February). Matching methods for causal inference: A review and a look forward. *Statistical Science* 25(1), 1–21.
- [24] van Buuren, S. and K. Groothuis-Oudshoorn (2011). mice: Multivariate imputation by chained equations in r. *Journal of Statistical Software* 45(3), 1–67.
- [25] VanderWeele, T. J. (2012, August). Confounding and effect modification: Distribution and measure. *Epidemiologic Methods* 1(1), 55–82.
- [26] Zeileis, A. (2006). Object-oriented computation of sandwich estimators. *Journal of Statistical Software* 16(9), 1–16.
- [27] Zeileis, A., S. Köll, and N. Graham (2020). Various versatile variances: An object-oriented implementation of clustered covariances in R. *Journal of Statistical Software* 95(1), 1–36.

## NOTICE

This (software/technical data) was produced for the US Government under Contract Number 75FCMC18D0047, and is subject to Federal Acquisition Regulation Clause 52.227-14, Rights in Data-General. No other use other than that granted to the US Government, or to those acting on behalf of the US Government under that Clause is authorized without the express written permission of The MITRE Corporation. For further information, please contact The MITRE Corporation, Contracts Management Office, 7515 Colshire Drive, McLean, VA 22102-7539, (703) 983-6000.

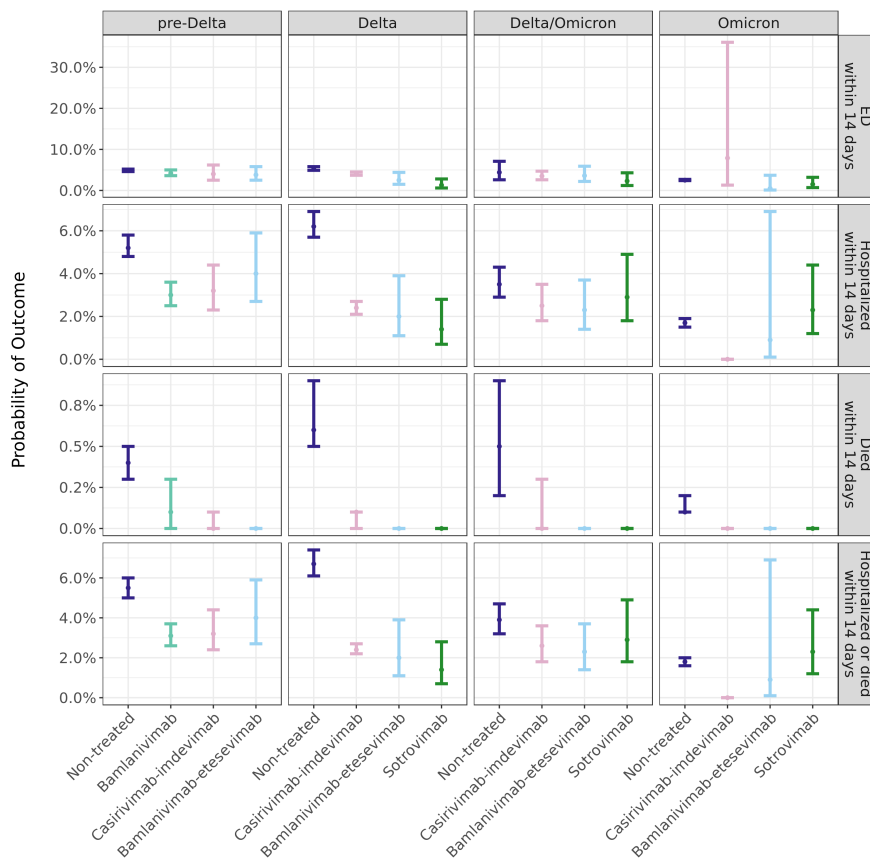
**©2023 The MITRE Corporation**



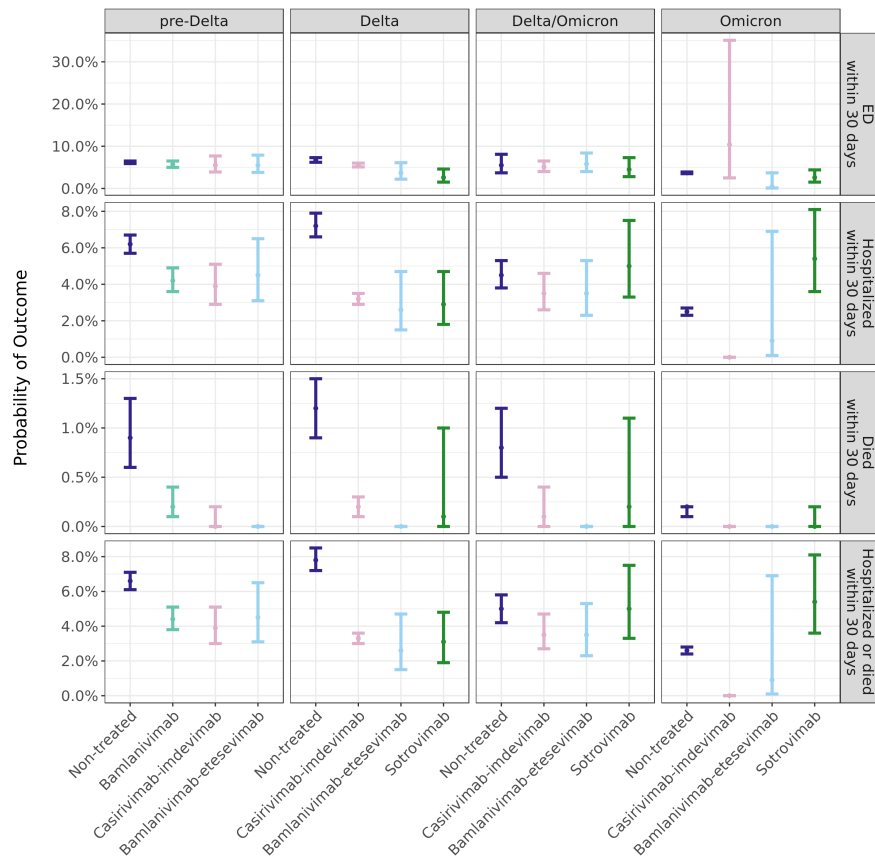


**Fig. 1: The changing frequency of clinical use of different nMAbs from November 2020 (study month 1) through January 2022 (study month 15) in the data set.**





**Fig. 2: Probability of study 14-day outcomes for patients not treated with nMAb and treated with different types of nMAbs during the four pandemic phases.**



**Fig. 3: Probability of study 30-day outcomes for patients not treated with nMAB and treated with different types of nMAbs during the four pandemic phases.**

# Appendices

## A Demonstration of Confounding if $p(t|x) = 0$

The following example shows how estimates of treatment effect for different treatment types can be biased if the positivity constraint is violated. Assume two time periods and two treatments,  $t_1$  and  $t_2$ . Untreated is denoted by  $t_0$ . Assume that the probability that a subject appears in interval 2 is twice the probability that a subject appears in interval 1. The probabilities of a subject receiving no treatment or the treatments and the probabilities of the adverse outcome,  $p_j(Y = 1)$ , for the two time intervals,  $j = 1, 2$ , and for the combined interval,  $j = C$ , are given in Table 1. Treatment 1 is given in interval 1 but not in interval 2, whereas treatment 2 is given in both intervals. The disease becomes more contagious and more virulent in interval 2 as compared with interval 1. This table also presents the odds ratios of the adverse outcomes, comparing treated versus non-treated patients, for each of the time periods and for the combined time period. One sees that, in time interval 1, treatments 1 and 2 are of equal effectiveness, with odds ratios of 0.11. The odds ratio for treatment 2 in interval 2 is 0.08—note that the probability of the adverse outcome increases for both untreated and those given treatment 2 in interval 2 as compared with interval 1. The combined odds ratio is also shown in the table. From the combined odds ratios, treatment 1 is more effective than treatment 2, whereas they are of equal effectiveness in interval 1, and treatment 1 is not used against the more virulent strain in interval 2. The combined odds ratio for treatment 1 is confounded by comparing treatment effectiveness of the treated population in interval 1 with untreated patients from interval 2, a period of time during which treatment 1 was not available.

**Tab. 1:** Illustration of confounding when  $p(t|x) = 0$

<b>Treat</b>	$p_1(t)$	$p_2(t)$	$p_1(Y = 1)$	$p_2(Y = 1)$	$p_C(Y = 1)$	$OR_1$	$OR_2$	$OR_C$
$t_0$	<b>0.34</b>	<b>0.34</b>	<b>0.5</b>	<b>0.75</b>	<b>0.67</b>			
$t_1$	<b>0.33</b>	<b>0</b>	<b>0.1</b>	<b>NA</b>	<b>0.1</b>	<b>0.11</b>	<b>NA</b>	<b>0.06</b>
$t_2$	<b>0.33</b>	<b>0.66</b>	<b>0.1</b>	<b>0.2</b>	<b>0.17</b>	<b>0.11</b>	<b>0.08</b>	<b>0.1</b>

## B An Alternate Derivation of $SW^1$

Robins [20] defines marginal structural models for general treatment and response time dependent processes, where the response process includes the outcome process of interest and the process of other recorded variables. Examples include treatment processes such that the treatment is given at multiple discrete time points and outcome process that are measured at a fixed time or are failure time process. He shows, using influence functions, that weighting observation by the inverse of a subject's probability of having had his observed treatment history allows for the estimation of causal effects from non-randomized observations. A simpler proof using importance sampling that is applicable to the case of one of several possible treatments, including the null treatment, given at a single time point and a fixed time to outcome is presented.

Assume an independent set of  $N$  samples  $\{(y_j, t_j, x_j) \mid 1 \leq j \leq N\}$ , where  $y_j$  is the outcome, and  $t_j = A(x_j) \in \mathcal{T}$ . The samples can be expressed as  $\{(y_j, a_j, t_j, x_j) \mid 1 \leq j \leq N\}$  where  $a_j \in \{0, 1\}$  and  $t_j \in \mathcal{T}_1$ . If  $a_j = 0$ ,  $t_j$  is interpreted as the potential treatment. Assume that there is a discrete set  $L$  and a mapping  $\phi : X \rightarrow L$  such that  $p(t_j \mid x_j) = p(t_j \mid \phi(x_j))$ —that is,  $t_j \sim \psi_{\phi(x_j)}$ . Let  $Y^i$  denote the outcome random variable if all patients have exposure  $A = i$ , for  $i = 0, 1$ .

The observations can be used to estimate causal effects if the following assumptions hold:

A.1  $Y^i \perp\!\!\!\perp A \mid (X, T)$ .

A.2  $p(A \mid X)p(T \mid X) > 0$ .

A.3 The outcome of one individual is independent of the treatment assignment of any other.

From assumption A.1, within strata of  $(X, T)$  treated and untreated patients are exchangeable [25], and from assumption A.3., outcomes of different patients are independent. Thus, the expected value of the causal variables can be computed from observations, according to

$$E(Y^1 \mid T) = \sum_x E(Y \mid A = 1, T, x) f(x), \quad (7)$$

and

$$\begin{aligned} E(Y^0 \mid T) &= E(Y^0) \\ &= \sum_x E(Y \mid A = 0, x) f(x). \end{aligned} \quad (8)$$

Importance sampling [6] states that if  $\{z_j\}$  is a sample drawn from probability density function  $f$ , and  $g$  is a probability density function such that  $g(x) = 0$  if  $f(x) = 0$ , then  $\{\frac{g(z_j)}{f(z_j)} z_j\}$  is a sample drawn from  $g$ . This result is used to transform samples drawn from  $f(a = 1, t_j, x)$  and  $f(a = 0, x)$  to samples drawn from  $f(x)$  so that A.1 holds for the weighted samples.

Let  $SW_j^1 \odot (a_j, t_j, x_j)$  denote the sample  $(a_j, t_j, x_j)$  counted with multiplicity  $SW_j^1$ . Let  $a_o \in \{0, 1\}$  and  $t_o \in \{t_1, \dots, t_M\}$ . If  $a_o = 1$ , define  $S_o = \{(a_j, t_j, x_j) \mid a_j = a_o, t_j = t_o\}$ , and if  $a_o = 0$ , define  $S_o = \{(a_j, t_j, x_j) \mid a_j = a_o\}$ . Let  $N_o = |S_o|$ . Let  $I_o$  be the indicator function for  $S_o$  defined by  $I_o(j) = 1$  if  $(a_j, t_j, x_j) \in S_o$ , and  $I_o(j) = 0$ , otherwise.

Define

$$w \odot S_o = \{SW_j^1 \odot (a_j, t_j, x_j) \mid I_o(j) = 1\}. \quad (9)$$

**Lemma 1.**  $w \odot S_o$  is a sample from  $p(x)$ .

**Proof.** Case 1:  $a_o = 1$ .  $\{(a_o, t_o, x_j)\}$  is a sample from

$$\begin{aligned} p(x \mid a_o, t_o) &= \frac{p(a_o, t_o, x)}{p(a_o, t_o)} \\ &= \frac{p(a_o \mid t_o, x) p(t_o, x)}{p(a_o, t_o)} \\ &= \frac{p(a_o \mid x) p(t_o \mid x) p(x)}{p(a_o, t_o)}. \end{aligned}$$

Importance sampling implies that each sample from the pseudo population,  $(a_o, t_o, x_j)$ , is a sample from

$$\begin{aligned}w_j p(x|a_o, t_o) &= \frac{p(a_o, t_o)}{p(a_o|x)p(t_o|x)} \times \frac{p(a_o|x)p(t_o|x)p(x)}{p(a_o, t_o)} \\ &= p(x).\end{aligned}$$

Case 2:  $a_o = 0$ .  $\{(a_o, x_j)\}$  is a sample from

$$\begin{aligned}p(x|a_o) &= \frac{p(a_o, x)}{p(a_o)} \\ &= \frac{p(a_o|x)p(x)}{p(a_o)}\end{aligned}$$

Importance sampling implies that each sample from the pseudo population,  $(a_o, x_j)$ , is a sample from

$$\begin{aligned}w_j p(x|a_o) &= \frac{p(a_o)}{p(a_o|x)} \times \frac{p(a_o|x)p(x)}{p(a_o)} \\ &= p(x).\end{aligned}$$

**Theorem 3.** *Under the assumptions A.1–A.3*

$$\lim_{N \rightarrow \infty} \frac{1}{N_o} \sum_{j=0}^N I_o(j) w_j Y_j = E(Y^{a_o} | t_o). \quad (10)$$

**Proof.** This follows from equations 7 and 8 and Lemma 1.

## C Simulation Comparing MSM Weighting and Random Treatment Assignment

A simulation was conducted to experimentally verify that effect estimates obtained using the MSM weighting derived above under non-random treatment assignment, referred to as the OBS, agree, within estimation error, with effect estimates obtained without weighting under random treatment assignment, referred to as the RCT. The simulation was carried out using MATLAB [18]. The simulation assumes two binary covariates,  $(X_1, X_2)$ , two binary treatment types,  $T_1$  and  $T_2$ , where  $T_j = 1$  indicates treatment with drug  $j$ , and a binary treatment variable,  $A$ , such that  $A = 0, 1$  indicates no-treatment and treatment, respectively. The total sample size was 100,000.

### C.1 Covariate Model

Samples of the correlated binary covariates  $(X_1, X_2)$  were obtained from samples of a normal random vector  $(Z_1, Z_2)$  having mean  $(0, 0)$  and covariance matrix  $\Gamma = \begin{pmatrix} 1 & 0.2 \\ 0.2 & 1 \end{pmatrix}$ . Thresholds  $\tau_1$  and  $\tau_2$  were determined to satisfy the equations:  $p(Z_1 \leq \tau_1) = 0.67$ , and  $p(Z_2 \leq \tau_2) = 0.25$ . Samples of  $(X_1, X_2)$  were obtained by thresholding samples of  $(Z_1, Z_2)$  according to  $X_i = 1$  if  $Z_i > \tau_i$  and 0, otherwise, for  $i = 1, 2$ .

### C.2 Outcome Model

Let  $g$  be the logit function. The outcome,  $Y$ , was binary, and modeled as

$$g(E(Y)) = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_1 X_2 + \alpha_4 A T_1 + \alpha_5 A T_2 + \alpha_6 A X_1 T_1 + \alpha_7 A X_1 T_2 + \alpha_8 A X_2 T_1 + \alpha_9 A X_2 T_2 + \alpha_{10} A T_1 X_1 X_2 + \alpha_{11} A T_2 X_1 X_2. \quad (11)$$

Table 2 lists the values of the coefficients, and Table 3 lists the log odds, probabilities, and odds ratios of  $Y = 1$  for all combinations of the input variables. The outcome probabilities were selected and the coefficients were obtained by solving a system of equations.

**Tab. 2:** Coefficients of the output model

Index	0	1	2	3	4	5	6	7	8	9	10	11
Value	-2.944	0.747	1.210	-0.399	-1.651	-1.651	0.903	-0.044	-0.507	-0.799	1.124	0.936

### C.3 Treatment Assignment

The RCT was simulated by randomly assigning each subject to one of three study arms: no treatment ( $A=0$ ), treatment with drug 1 ( $T_1 = 1$ ), or treatment with drug 2 ( $T_2 = 1$ ). The simulation of the OBS was done by assigning a subject to the treatment group ( $A = 1$ ) or the non-treatment group ( $A = 0$ ) using the following propensity model:

$$g(E(A|X_1, X_2)) = -1.1 + X_1 + 2X_2. \quad (12)$$

**Tab. 3:** Log odds, probabilities, and odds ratios of the outcome model conditioned on the input values.

A	$X_1$	$X_2$	$T_1$	$T_2$	Log odds	Probability	Odds ratio
0	0	0	0	0	-2.9444	0.05	NA
0	1	0	0	0	-2.1972	0.1	NA
0	0	1	0	0	-1.7346	0.15	NA
0	1	1	0	0	-1.3863	0.2	NA
1	0	0	1	0	-4.5951	0.01	0.1919
1	1	0	1	0	-2.9444	0.05	0.4737
1	0	1	1	0	-3.8918	0.02	0.1156
1	1	1	1	0	-1.5163	0.18	0.8780
1	0	0	0	1	-4.5951	0.01	0.1919
1	1	0	0	1	-3.8918	0.02	0.1837
1	0	1	0	1	-4.1846	0.015	0.0863
1	1	1	0	1	-2.9444	0.05	0.2105

### C.4 Selection of Treatment Type

For the simulation of the OBS, an additional choice of treatment type was made according to the following protocol. Let  $P_{i,j}$ ,  $i, j = 0, 1$ , be the probability of selecting  $T_1$  when  $X_1 = i$  and  $X_2 = j$ . The OBS treatment selection used  $P_{00} = 0.5$ ,  $P_{01} = 0.4$ , and  $P_{10} = 0 = P_{11}$ .

### C.5 Equalization of the Distribution of the Population Covariates

Lemma 1 of Appendix B asserts that the distribution of  $(X_1, X_2)$  of the weighted samples is independent of study arm: untreated, treated with  $T_1$ , or treated with  $T_2$ . This was demonstrated by applying the  $\chi^2$  test for independence to the initial count data and to the weighted count data shown in Tables 4 and 5, respectively. The p-values of the test applied to these tables were  $< 2.2 \times 10^{-16}$  and 0.271, respectively.

**Tab. 4:** Unweighted population counts of OBS data with  $X_1 = 0$ . The p-value of the  $\chi^2$  test of independence is less than  $2.2 \times 10^{-16}$

	Unweighted counts: $A = 0$	Unweighted counts: $T_1 = 1$	Unweighted counts: $T_2 = 1$
$X_2 = 0$	61,375	10,230	10,160
$X_2 = 1$	5,174	5,173	18,235

**Tab. 5:** Weighted population counts of the observational data with  $X_1 = 0$ . The p-value of the  $\chi^2$  test of independence is 0.271.

	Weighted counts: $A = 0$	Weighted counts: $T_1 = 1$	Weighted counts: $T_2 = 1$
$X_2 = 0$	54,440	12,576	14,735
$X_2 = 1$	11,912	2,830	3,316

### C.6 Comparison of MSM and RCT Effect Estimates

Treatment effects were estimated by fitting the simulated RCT data to unweighted logistic regression models and the simulated observational data to weighted logistic regression models using weights calculated from



equation (4). Treatment effects are reported as log odds ratios and 95% confidence intervals obtained from the coefficients of these models. If  $X_1 = 0$ , the average treatment effects (ATE), over the values of  $X_2$ ,  $T_1$ , and  $T_2$  and the effects of  $T_1$  and  $T_2$  for  $X_2 = 0$  and 1 are given in table 6. For  $X_1 = 1$  the ATE for  $T_2$  over the values of  $X_2$  and the effects of  $T_2$  for each value of  $X_2$  are given in table 7. The 95% confidence intervals are seen to overlap in all cases. The standard errors of the OBS estimates are generally larger than the standard errors of the RCT estimates.

**Tab. 6:** Treatment effects reported as log odds ratios and 95% confidence intervals:  $X_1 = 0$ . The two study designs used independent data sets each consisting of 100,000 samples drawn from the same population.

	RCT: Log odds ratio (95% confidence intervals)		OBS: Log odds ratio (95% confidence intervals)	
	$T_1$	$T_2$	$T_1$	$T_2$
<b>ATE</b>	<b>-1.78 (-1.89,-1.67)</b>	<b>-1.93 (-2.04,-1.81)</b>	<b>-1.88 (-2.03,-1.73)</b>	<b>-1.82 (-1.96,-1.68)</b>
$X_2 = 0$	<b>-1.59 (-1.72,-1.46)</b>	<b>-1.73 (-1.87,-1.60)</b>	<b>-1.71 (-1.90,-1.53)</b>	<b>-1.51 (-1.66,-1.35)</b>
$X_2 = 1$	<b>-2.18 (-2.38,-1.99)</b>	<b>-2.35 (-2.56,-2.14)</b>	<b>-2.26 (-2.53,-1.98)</b>	<b>-2.59 (-2.88,-2.29)</b>

**Tab. 7:** Treatment effects reported as log odds ratios and 95% confidence intervals:  $X_1 = 1$ . Note that only treatment 2 is available on the OBS. The two study designs used independent data sets each consisting of 100,000 samples drawn from the same population.

	RCT: Log odds ratio (95% confidence intervals)	OBS: Log odds ratio (95% confidence intervals)
<b>ATE</b>	<b>-1.61 (-1.67,-1.55)</b>	<b>-1.63 (-1.68,-1.57)</b>
$X_2 = 0$	<b>-1.67 (-1.76,-1.59)</b>	<b>-1.72 (-1.80,-1.64)</b>
$X_2 = 1$	<b>-1.57 (-1.65,-1.48)</b>	<b>-1.58 (-1.66,-1.50)</b>