



Published in final edited form as:

Epidemiology. 2021 March 01; 32(2): 202–208. doi:10.1097/EDE.0000000000001315.

Incremental Propensity Score Effects for Time-Fixed Exposures

Ashley I. Naimi, PhD^{1,*}, Jacqueline E. Rudolph, PhD¹, Edward H. Kennedy, PhD², Abigail Cartus, MPH³, Sharon I. Kirkpatrick, PhD⁴, David M. Haas, MD⁵, Hyagriv Simhan, MD⁶, Lisa M. Bodnar, PhD³

¹Department of Epidemiology, Emory University, Atlanta, GA.

²Department of Statistics & Data Science, Carnegie Mellon University, Pittsburgh, PA.

³Department of Epidemiology, University of Pittsburgh, Pittsburgh, PA.

⁴School of Public Health and Health Systems, University of Waterloo, Waterloo, Ontario, Canada.

⁵Department of Obstetrics & Gynecology, School of Medicine, Indiana University, Indianapolis, IN.

⁶Magee Womens Research Institute, University of Pittsburgh, Pittsburgh, PA.

Abstract

When causal inference is of primary interest, a range of target parameters can be chosen to define the causal effect of interest, such as average treatment effects (ATEs). However, ATEs may not always align with the research question at hand. Furthermore, the assumptions needed to interpret estimates as ATEs, such as exchangeability, consistency, and positivity, are often not met. Here, we present the incremental propensity score (incremental PS) approach to quantify the effect of shifting each person's exposure propensity by some pre-determined amount. Compared to the ATE, incremental PS may better reflect the impact of certain policy interventions, and do not require that positivity hold. Using the Nulliparous Pregnancy Outcomes Study: monitoring mothers-to-be (nuMoM2b), we quantified the relation between total vegetable intake and the risk of preeclampsia, and compared it to average treatment effect estimates. The ATE estimates suggested a reduction of between two and three preeclampsia cases per 100 pregnancies for consuming at least 1/2 a cup of vegetables per 1,000 kcal. However, positivity violations obfuscate the interpretation of these results. In contrast, shifting each woman's exposure propensity by odds ratios ranging from 0.20 to 5.0 yielded no difference in the risk of preeclampsia. Our analyses show the utility of the incremental propensity score effects in addressing public health questions with fewer assumptions.

Keywords

machine learning; nonparametric methods; double-robust estimation; causal inference; epidemiologic methods

*Correspondence: Department of Epidemiology, Emory University, 1518 Clifton Road, Atlanta, GA 30322, ashley.naimi@emory.edu.
Conflicts of Interest: None

Introduction

Applied researchers are increasingly expressing interest in the average treatment effects (ATE) for defining an exposure effect of interest. This effect quantifies an outcome summary of interest (e.g., risk, rate, odds) that would be observed if all individuals in the population were exposed versus unexposed.¹ Methods to quantify ATEs often include inverse-probability weighting,² g computation (a.k.a., the g formula),³ or a range of doubly robust approaches.⁴ The validity of these methods relies on a set of basic assumptions, for example, counterfactual consistency, exchangeability, no interference, correct parametric model specification, and positivity.^{5,6} However, in many instances, there are indications that one or more of these assumptions is violated, complicating the interpretation and use of these increasingly common methods.

Epidemiologic studies tend to be complicated by a number of challenges related to the definition, measurement, and the analysis of exposures. At the crossroads of nutritional and reproductive epidemiology, questions of primary importance are often asked with regards to the relation between diet and pregnancy outcomes. However, dietary factors are often strongly correlated with one another.⁷ Such strong relations can lead to violations in the positivity assumption. Positivity is met when there are exposed and unexposed individuals in all strata defined by confounding values. Foods have complex, reciprocal interrelations because of dietary behavioral patterns, meaning that people whose dietary patterns are high in vegetables also tend to consume higher amounts of other healthy foods.⁷ This can lead to problems with positivity.

Additionally, average treatment effects quantify the effect that would be observed under a hypothetical intervention in which all individuals are fully exposed versus fully unexposed to the treatment under study. This property of the average treatment effect has been the source of some recent criticism, largely because public health interventions will rarely result in fully exposed or unexposed populations.^{8–11} Incremental propensity score effects are a natural alternative to the ATE. In line with a growing body of literature on the effects of shifts in treatments or exposures based on baseline characteristics of individuals in the sample,^{8,27} incremental PS estimands define the effect of shifts in the conditional probability of being exposed. Such shifts represent changes in an individual's underlying probability of being exposed or unexposed, as a function of their covariate values, rather than their actual exposure status.

Here, we demonstrate how novel incremental propensity score (incremental PS) effects can be estimated via doubly-robust methods,¹² and how they can provide valuable information about the relation between an exposure and outcome of interest, even when positivity is violated. Using an example based on the relation between total vegetable intake and the risk of preeclampsia, we illustrate the causal contrast defined by an incremental PS effect, the reasoning behind its use, and how to interpret results obtained from estimators targeting incremental PS effects.

Methods

Data were obtained from the nuMoM2b Study, a large prospective U.S. pregnancy cohort, previously described elsewhere.¹³ Briefly, nuMoM2b enrolled 10,038 women from eight medical centers across the U.S. between 2010–2013. Women were eligible if they had a viable singleton pregnancy, were between 6 to 13 completed weeks of gestation, and had no previous pregnancy that lasted ≥ 20 weeks' gestation. Each site's local institutional review board approved the study and all women gave written, informed consent. Our analytic sample included 8,259 women who delivered at ≥ 20 weeks gestation and had complete dietary and birth outcome data.

Usual dietary intake in the 3 months before conception was assessed at 6 to 13 weeks' gestation using a self-administered modified Block 2005 food frequency questionnaire (FFQ),¹⁴ which was available in English and Spanish. The FFQ methodology was provided in detail previously.¹⁵

The FFQ's food list was developed from the National Health and Nutrition Examination Survey 1999–2002 dietary recall data, and the nutrient database was developed from the USDA Food and Nutrient Database for Dietary Studies.¹⁶ Food groups were derived from the MyPyramid Equivalents Database, version 2.0.¹⁷ For our primary exposure, usual daily vegetable intake was expressed in cup equivalents per 1,000 kcal, reflecting total vegetable density as per the construction of the Healthy Eating Index (HEI)-2010.¹⁸ The usual daily intake of vegetables was expressed relative to energy as cups per 1,000 kcal. Our primary interest was in evaluating whether a hypothetical intervention encouraging women to consume at least $\frac{1}{2}$ a cup of vegetables per 1,000 kcal would be effective at reducing the risk of preeclampsia.

Preeclampsia was defined as the presence of the following symptoms at ≥ 20 weeks gestation through 14 days: gestational hypertension (≥ 140 mm hg systolic or ≥ 90 mm Hg diastolic on two occasions ≥ 6 hours apart or one occasion with subsequent antihypertensive therapy, excluding blood pressures recorded during the second stage of labor) and proteinuria (≥ 300 mg/24 hour collection or protein/creatinine ratio ≥ 0.3 or incremental PStick 2+), thrombocytopenia (platelet count $<100,000/\text{mm}^3$), or pulmonary edema. Preeclampsia included superimposed preeclampsia or eclampsia, regardless of the timing of onset. We did not distinguish between early and late onset preeclampsia, which may have distinct etiologies.³¹ Cases that presented atypically and were difficult to classify according to study criteria were adjudicated by review of clinical data by the principal investigators and final classification was reached by their consensus judgment.

Confounders of the relationship between vegetable density and preeclampsia included baseline anthropometric and demographic measures, as well as other measures derived from the HEI. HEI confounders included whole grains, dairy products, total protein foods, seafood and plant proteins, fatty acids, refined grains, sodium, and “empty” calories.¹⁸ Other baseline confounders included self-reported highest level of education, self-reported race/ethnicity (indicator of non-Hispanic Black status). Other self-reported confounder information included marital status, smoking before pregnancy, and medical insurance. At

the initial visit, women had their weight measured using an electronic or balance scale while wearing only light clothes and no shoes, and height measured using a stadiometer or measuring tape, used to compute body mass index (BMI) in kg/m², which we included as a confounder.

Statistical Analysis

Doubly Robust Estimation of the Average Treatment Effect—We used three different estimators to quantify the relation between total vegetable density and preeclampsia in the nuMoM2b data. We first used two relatively well known doubly robust estimators to quantify the average treatment effect, defined as:

$$\psi \equiv E(Y^1 - Y^0)$$

which, in our setting can be interpreted as the risk difference over all of gestation that would be observed if all women consumed at least 1/2 a cup of vegetables per 1,000 kcal, versus if all women consumed less than or equal to 1/2 a cup per 1,000 kcal.

The average treatment effect can also be quantified with standard approaches (e.g., inverse probability weighting, or g computation). However, as we discuss further below, the incremental propensity score is implemented with machine learning algorithms. In general, inverse probability weighting and g computation methods should not be implemented with machine learning algorithms because of the problems that result from the curse of dimensionality, which include finite-sample bias, high mean squared error, and no general way to obtain valid confidence intervals.¹⁹ Thus, to more fairly compare the results from the doubly-robust estimator of the incremental PS effect and estimators of the ATE, we require estimators that are also doubly robust, such as augmented inverse probability weighting (AIPW) and targeted minimum loss based estimation (TMLE).

The augmented inverse probability weighted estimator we used was defined as:

$$\hat{\psi}^{aipw} = \frac{1}{N} \sum_{i=1}^N \left\{ \frac{(X_i - \hat{\pi}_i)}{\hat{\pi}_i(1 - \hat{\pi}_i)} \times (Y_i - \hat{\mu}_i(X, \mathbf{C})) \right\} + \hat{\mu}_i(X = 1, \mathbf{C}) - \hat{\mu}_i(X = 0, \mathbf{C}). \quad (1)$$

where X is an indicator of whether total vegetable density is greater than half a cup per 1,000 kcals, Y is an indicator of preeclampsia, \mathbf{C} is the vector of confounders defined above, and i indexes each woman in the sample. In the above equation, $\hat{\pi}_i$ represents individual predictions from the propensity score model, which is a model for the exposure regressed against all confounders. Furthermore, $\hat{\mu}_i(X = x, \mathbf{C})$ represents individual predictions from the outcome model if the exposure were set to some value x .

We also used targeted minimum loss-based estimation,^{20(p72-3)} defined as:

$$\hat{\psi}^{tmle} = \frac{1}{N} \sum_{i=1}^N \{ \hat{\mu}_i^u(X = 1, \mathbf{C}) - \hat{\mu}_i^u(X = 0, \mathbf{C}) \}, \quad (2)$$

where $\hat{\mu}_i^u(X, \mathbf{C})$ are predictions from an outcome model that is updated using an inverse probability weight, defined as:

$$H(X, \mathbf{C}) = \begin{cases} \frac{1}{\hat{\pi}_i} & \text{if } X = 1 \\ -\frac{1}{1 - \hat{\pi}_i} & \text{otherwise} \end{cases}$$

Which is included in a no-intercept logistic regression model for the observed outcome that includes the outcome predictions $\hat{\mu}_i(X, \mathbf{C})$ from the initial model as an offset. The $\hat{\mu}_i^u(X = 1, \mathbf{C})$ and $\hat{\mu}_i^u(X = 0, \mathbf{C})$ predictions are then generated from this model by first setting X to 1 and then to 0, respectively, for all women in the sample.

Under exchangeability, counterfactual consistency, no interference, and positivity, the average treatment effect is identified, and can be quantified via AIPW and TMLE. Positivity holds when the conditional probability of being exposed or unexposed is bounded away from [0, 1]. Formally, we can define positivity as a function of the propensity score, as:

$$0 < \pi(\mathbf{C}) \equiv P(X = 1 | \mathbf{C}) < 1.$$

The positivity assumption can be violated if $P(X = 1 | \mathbf{C}) \approx 0$ or $P(X = 1 | \mathbf{C}) \approx 1$ (or equivalently, if $P(X = 0 | \mathbf{C}) \approx 1$ or $P(X = 0 | \mathbf{C}) \approx 0$). If positivity is violated, then the average treatment effect is not identified, making the interpretation of ATE estimates more challenging.

Incremental Propensity Score Effects—Instead of targeting the average treatment effect, we might be interested in quantifying the risk of preeclampsia that would be observed if we could shift each women’s propensity for eating at least 1/2 a cup of vegetables per 1,000 kcal by a certain amount.

For example, suppose that the probability that a given woman consumes at least 1/2 a cup of vegetables per 1,000 kcal is 35% (in other words, her propensity score for being exposed is 0.35). We may ask what her preeclampsia risk would be if we were able to shift her propensity for being exposed to 0.45? Such a shift intervention would reflect an odds ratio change in the propensity score of:

$$\frac{0.45/(1 - 0.45)}{0.35/(1 - 0.35)} \approx 1.5$$

We can thus define a “shift intervention” for each woman in the sample, such that their propensity score might increase by an odds ratio of, say, 1.5.

Incremental propensity score effects quantify risks under such “shift interventions.”¹² In our setting, we quantify the risk of preeclampsia that would be observed if each women’s propensity for being exposed were shifted by an odds ratio, denoted δ . If we let $\pi_\delta =$

$(\delta\pi)/(\delta\pi+1-\pi)$ be the propensity score under the shift intervention in which the observed propensity score is increased by the odds ratio δ , then:

$$E(Y^{\pi\delta}) = \sum_{\mathbf{c}, \pi_\delta} E(Y | X, \mathbf{C}) P_\delta(X = 1 | \mathbf{C}) p(\mathbf{C}). \quad (3)$$

One can derive an estimator of $E(Y^{\pi\delta})$ as:

$$\xi_\delta^{ips} = \sum_{i=1}^N \frac{\delta \hat{\pi}_i \hat{\mu}_i(X = 1 | \mathbf{C}) + (1 - \hat{\pi}_i) \hat{\mu}_i(X = 0 | \mathbf{C})}{\delta \hat{\pi}_i + 1 - \hat{\pi}_i} \quad (4)$$

where $\hat{\pi}_i$ and $\hat{\mu}_i(X = 1 | \mathbf{C})$ are as defined for the AIPW and TMLE estimators above.

Using the estimator defined in equation 4, one can show why the estimators for incremental PS effects do not require positivity. Consider the scenario where an individual's propensity score is 0. Under such a positivity violation, equation 4 for this woman becomes:

$$\begin{aligned} \xi_{i,\delta}^{ips} &= \frac{\delta \times 0 \times \hat{\mu}_i(X = 1 | \mathbf{C}) + (1 - 0) \times \hat{\mu}_i(X = 0 | \mathbf{C})}{\delta \times 0 + 1 - 0} \\ &= \hat{\mu}_i(X = 0 | \mathbf{C}), \end{aligned}$$

which is the risk of preeclampsia that would be observed if they were not exposed. For individuals whose propensity score is 1, the above equation yields the risk of preeclampsia that would be observed if they were exposed. In this way, incremental propensity score effects do not require positivity to hold, as the potential outcomes defined above for individuals with extreme propensity scores are identified under exchangeability, no interference, and consistency.¹²

One important feature of incremental PS effects is that shift interventions need not be defined via variables in the propensity score model. For example, confounders in our propensity score model included whole grain consumption, percent empty calories, self-reported race/ethnicity, and BMI. Hypothetical interventions that shift the propensity score need not necessarily occur by changing these variables.

Note that because the AIPW, TMLE, and our estimator for the incremental PS effect are “doubly robust” we can use machine learning methods to quantify the effects of interest.¹⁹ In our study, we used a stacked generalization²¹ that included random forests, generalized additive models, and generalized linear models.

Finally, even though the estimator we used for incremental PS effects is “doubly robust,” the incremental PS estimand is defined as a function of the propensity score. It is therefore important to try to model the propensity score as carefully as possible. In the event that the propensity score model is misspecified, one could still interpret the resulting estimates as quantifying a mathematical projection (or approximation) of the true model, but it may be difficult to interpret what this projection represents.

We used R version 4.0.1 (2020–06–06) for all analyses. The AIPW and estimator for incremental PS effects were implemented via the npcausal package, available on GitHub. TMLE was implemented via the tmle package.²² All the code used to conduct the analyses presented here is also available at https://github.com/ainaimi/numom_IPSI.

Results

Table 1 shows summary outcome and covariate values within each vegetable density level and overall. There were important confounder differences between exposed and unexposed pregnancies. In particular, large confounder differences were observed for self reported race, marital status, smoking status, public insurance status, as well as several dietary factors (Table 1). These large differences suggest potentially strong relationships between confounders and the exposure. Such strong relations can be suggestive of positivity violations. Indeed, the propensity score overlap plot depicted in Figure 1 confirms that positivity is violated with our exposure. This figure indicates that the average treatment effect is not identified in these data.

For illustrative purposes, however, we present the average treatment effect estimates from AIPW and TMLE in Table 2. This table shows that both ATE estimators quantified a relatively strong association between total vegetable density and preeclampsia, with between two and three fewer preeclampsia cases per 100 pregnancies among those who consume less than 1/2 a cup of vegetables per 1,000 kcal relative who consume at least 1/2 a cup. Though these results suggest vegetable consumption as a fairly important modality to prevent preeclampsia, the positivity violations observed in Figure 1 raise important questions about the validity of this interpretation.

Using incremental propensity scores, we quantified the risk of preeclampsia that would be observed if each woman's propensity for being exposed were altered by odds ratios ranging between 0.20 and 5.0. The left panel of Figure 2 presents all of the unique propensity score values in the nuMoM2b data as a function of the selected δ values. The vertical line indexes $\delta = 1$, which returns the observed propensity score. This figure shows the extent to which each unique propensity score changes under different δ values. For example, for women whose observed propensity score (rounded to the nearest two digits) was exactly 0.50 ($N = 26$), their shifted propensity scores range between 0.17 for $\delta = 0.20$ and 0.83 for $\delta = 5$.

The right panel of Figure 2 demonstrates the overall propensity score distributions under three different shift ORs: 0.2 (minimum δ), 1 (corresponding to the actual propensity scores), and 5 (maximum δ). This left panel shows that, under all three scenarios, a majority of the propensity score density resides in the far right tail of the distribution. This is because the overall probability of consuming at least 1/2 cup of vegetables per 1,000 kcal in our sample was roughly 80%. However, the figure also indicates the extent to which the propensity score distribution changes under the extremes of our selected δ values. For example, under a δ value of 0.2, there is a noticeable shift in the propensity score distribution to the left. Finally, as depicted in Figure 3, changing the propensity for being exposed yielded no meaningful difference in the risk of preeclampsia.

Discussion

We sought to quantify the causal relation between total dietary vegetable density and the risk of preeclampsia. As is typical, we defined the average treatment effect to capture this relation. Using empirical data to quantify the average treatment effect requires that certain assumptions hold, such as exchangeability, counterfactual consistency, and positivity. We found strong positivity violations for consuming at least 1/2 a cup of vegetables per 1,000 kcal.

Under such conditions, the average treatment effect is not identified. Formally, an effect (defined via potential outcomes) is identified when it can be written as a function of observed data. To do this, we require that certain assumptions (including positivity) hold.⁶ Thus, even though one may obtain a numerical estimate of the average treatment effect, it may not be interpretable as such. Typical strategies to deal with positivity violations include subsetting the sample to those with propensity score overlap,²³ propensity score truncation,²⁴ excluding certain problematic confounders.²⁵ Here, we demonstrated an alternative to these approaches, namely, doubly robust estimation of incremental PS effects.

While we have shown why incremental propensity scores do not require positivity for identifiability, it is important to note that other assumptions are essential. These include exchangeability, no interference, and counterfactual consistency. Exchangeability would be met if there were no information, selection, collider, or confounding bias present in the study. In the context of nutritional epidemiology, collider bias may be of particular concern because we assume reported intake of other dietary components (e.g., total protein consumption, plant protein consumption, percent empty calories) *confound* the relation between total vegetable consumption and preeclampsia. It is entirely possible that our measures of these other dietary components mediate rather than confound the relation, which could lead to collider bias and blocking of the effect of interest. However, except in certain circumstances (e.g., instrumental variable estimation), the validity of these assumptions do not depend on the which estimator is used. These assumptions are as important when one is targeting average treatment effects via double-robust estimators, or whether incremental PS methods are used.

Similarly, the no interference and counterfactual consistency assumptions do not depend on estimator selection. If one can justify that there is no causal effect between one person's exposure and another's outcome, and that the observed outcome is equivalent to the potential outcome under the observed exposure, these assumptions can be justified. Otherwise, identifiability may be jeopardized.

One important feature that differentiates incremental propensity score and other shift intervention effects from, say, the average treatment effect, is the inability to explicitly define a randomized trial, even hypothetically, to evaluate the effect of shift interventions. Because of the inability to conceptualize such shift interventions in terms of a randomized trial, questions may arise about the validity of counterfactual consistency²⁹ or treatment variation irrelevance.³⁰ First, while propensity score shifts may not be easily conceived in terms of an intervention, the same is not necessarily for the exposure used to compute the

propensity score. Interventions could be conceptualized for getting women to consume more than 1/2 a cup of vegetables per 1,000 kcal. Second, while shift intervention estimands may not be directly computable in a randomized trial, they can still provide information useful to narrowing down the set of interventions to be studied more in depth via randomization.²⁷

Finally, we did not address the complications that can arise in the presence of loss-to-follow-up or competing risk events.²⁸ Such complications are, in fact, present in the NuMoM study, where, out of the 8,259 women, 45 women experienced a spontaneous pregnancy loss, and 45 women experience a stillbirth over the course of follow-up. Informal ancillary analyses, provided in eAppendix 1, suggested no impact of these competing events on our findings (for both the ATE and incremental PS effects). However, as of the present article, neither the AIPW function, the TMLE function, nor the incremental PS effects function we used include options to compare sub-distribution and cause-specific measures when competing events are present.

Here, we estimated incremental propensity score effects for a time-fixed nutritional exposure and a binary outcome in the absence of censoring. However, incremental propensity score effects can also be used in settings with a time-varying exposure subject to time-dependent confounding,¹² as well as to settings with censored survival outcomes.²⁶

Incremental propensity scores can be useful even when positivity holds. The target estimand for incremental propensity scores effects is closely related to shift interventions that have been studied in various settings.^{8–11} Incremental propensity score effects provide a relatively easy approach to quantifying such population relevant effects, as they define interventions that are arguably more likely to align with what might occur in a population health setting: i.e., shifts in the conditional probability of being exposed, instead of setting each person's exposure to a fixed value. As such, one need not rely on incremental propensity score effects only in situations where positivity is violated. Rather, the approach can prove valuable as one of many tools available for modeling population relevant effects in a public health framework.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

Sources of Funding: AIN was supported by NIH grant number R01HD093602; EHK was supported by NSF DMS Grant #1810979; The study is supported by grant funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD): U10 HD063036, RTI International; U10 HD063072, Case Western Reserve University; U10 HD063047, Columbia University; U10 HD063037, Indiana University; U10 HD063041, University of Pittsburgh; U10 HD063020, Northwestern University; U10 HD063046, University of California Irvine; U10 HD063048, University of Pennsylvania; and U10 HD063053, University of Utah.

References

1. Hernán MA and Robins J. Causal Inference. Forthcoming. Chapman/Hall, <http://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/>, accessed 12 Mar 2014.

2. Hernan MA, Brumback B, and Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of hiv-positive men. *Epidemiology*. 2000; 11:561–570. [PubMed: 10955409]
3. Robins JM. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical Modelling*. 1986; 7:1393–1512.
4. Daniel RM. Double robustness. In: *Wiley StatsRef: Statistics Reference Online*. John Wiley & Sons, Ltd. 2018;.
5. Robins JM and Hernán MA. Estimation of the causal effects of time-varying exposures. In: Fitzmaurice G, Davidian M, Verbeke G, and Molenberghs G (Eds.) *Advances in Longitudinal Data Analysis*. Boca Raton, FL: Chapman & Hall. 2009; 553–599.
6. Naimi AI, Cole SR, and Kennedy EH. An Introduction to G Methods. *Int J Epidemiol*. 2017; 46:756–762. [PubMed: 28039382]
7. Committee DGA. Report of the dietary guidelines advisory committee on the dietary guidelines for americans, 2010, to the secretary of agriculture and the secretary of health and human services. Tech. Rep., Washington DC. 2010.
8. Munoz ID and van der Laan M. Population intervention causal effects based on stochastic interventions. *Biometrics*. 2012; 68:541–549. [PubMed: 21977966]
9. Naimi AI, Moodie EE, Auger N, and Kaufman JS. Stochastic mediation contrasts in epidemiologic research: Inter-pregnancy interval and the educational disparity in preterm birth. *Am J Epidemiol*. 2014; 180:436–45. [PubMed: 25038216]
10. Rogawski ET, Gray CL, and Poole C. An argument for renewed focus on epidemiology for public health. *Annals of epidemiology*. 2016; 26:729–733. [PubMed: 27659585]
11. Westreich D. From patients to policy: Population intervention effects in epidemiology. *Epidemiology*. 2017; 28:525–528. [PubMed: 28282339]
12. Kennedy EH. Nonparametric causal effects based on incremental propensity score interventions. *Journal of the American Statistical Association*. 2018; :1–12. [PubMed: 30034060]
13. Haas DM, Parker CB, Wing DA et al. A description of the methods of the nulliparous pregnancy outcomes study: monitoring mothers-to-be (numom2b). *Am J Obstet Gynecol*. 2015; 212:539.e1–539.e24.
14. Block G, Hartman AM, Dresser CM, Carroll MD, Gannon J, and Gardner L. A data-based approach to diet questionnaire design and testing. *Am J Epidemiol*. 1986; 124:453–469. [PubMed: 3740045]
15. Bodnar LM, Cartus AR, Kirkpatrick SI et al. Machine learning as a strategy to account for dietary synergy: an illustration based on dietary intake and adverse pregnancy outcomes. *Am J Clin Nutr*. 2020; Forthcoming.
16. USDA Food and Nutrient Database for Dietary Studies version 1.0. Beltsville, MD: Agricultural Research Service, Food Surveys Research Group. U.S. Department of Agriculture. 2004.
17. U.s. department of agriculture agriculture research service food surveys research group.
18. Guenther P, Casavale K, Reedy J et al. Update of the healthy eating index: Hei-2010. *J Acad Nutr Diet*. 2013; 113:569–80. [PubMed: 23415502]
19. Naimi AI, Kennedy EH, and Mishler A. Challenges in obtaining valid causal effect estimates with machine learning algorithms. *Am J Epidemiol*. Under Review;.
20. Rose S and van der Laan MJ. Targeted learning: causal inference for observational and experimental data. New York, NY: Springer. 2011.
21. Naimi AI and Balzer LB. Stacked generalization: an introduction to super learning. *Eur J Epidemiol*. 2018; 33:459–464. [PubMed: 29637384]
22. Gruber S and van der Laan MJ. tmle: An r package for targeted maximum likelihood estimation. *Journal of Statistical Software*. 2012; 51:1–35. [PubMed: 23504300]
23. Sturmer T, Rothman KJ, Avorn J, and Glynn RJ. Treatment effects in the presence of unmeasured confounding: dealing with observations in the tails of the propensity score distribution—a simulation study. *Am J Epidemiol*. 2010; 172:843–854. [PubMed: 20716704]

24. Cole SR and Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol.* 2008; 168:656–64. [PubMed: 18682488]
25. Petersen ML, Porter KE, Gruber S, Wang Y, and van der Laan MJ. Diagnosing and responding to violations in the positivity assumption. *Stat Methods in Med Res.* 2012; 21:31–54. [PubMed: 21030422]
26. Kim K, Kennedy EH, and Naimi AI. Incremental intervention effects in studies with many timepoints, repeated outcomes, and dropout. <https://arxiv.org/abs/190704004>. 2019.
27. Haneuse S and Rotnitzky A. Estimation of the effect of interventions that modify the received treatment. *Statistics in Medicine.* 2013; 32:5260–5277. [PubMed: 23913589]
28. Rudolph JE, Lesko CR, and Naimi AI. Causal Inference in the Face of Competing Events. *Current Epidemiology Reports.* 2020; In Press.
29. Cole SR and Frangakis C. The consistency statement in causal inference: a definition or an assumption?. *Epidemiology.* 2009; 20:3–5. [PubMed: 19234395]
30. VanderWeele TJ. Concerning the Consistency Assumption in Causal Inference. *Epidemiology.* 2009; 20:880–83. [PubMed: 19829187]
31. Lisonkova S and Joseph KS. Incidence of preeclampsia: risk factors and outcomes associated with early- versus late-onset disease. *Am J Obstet Gynecol.* 2013; 209:544.e1–544.e12.

Observed Exposure 0 1

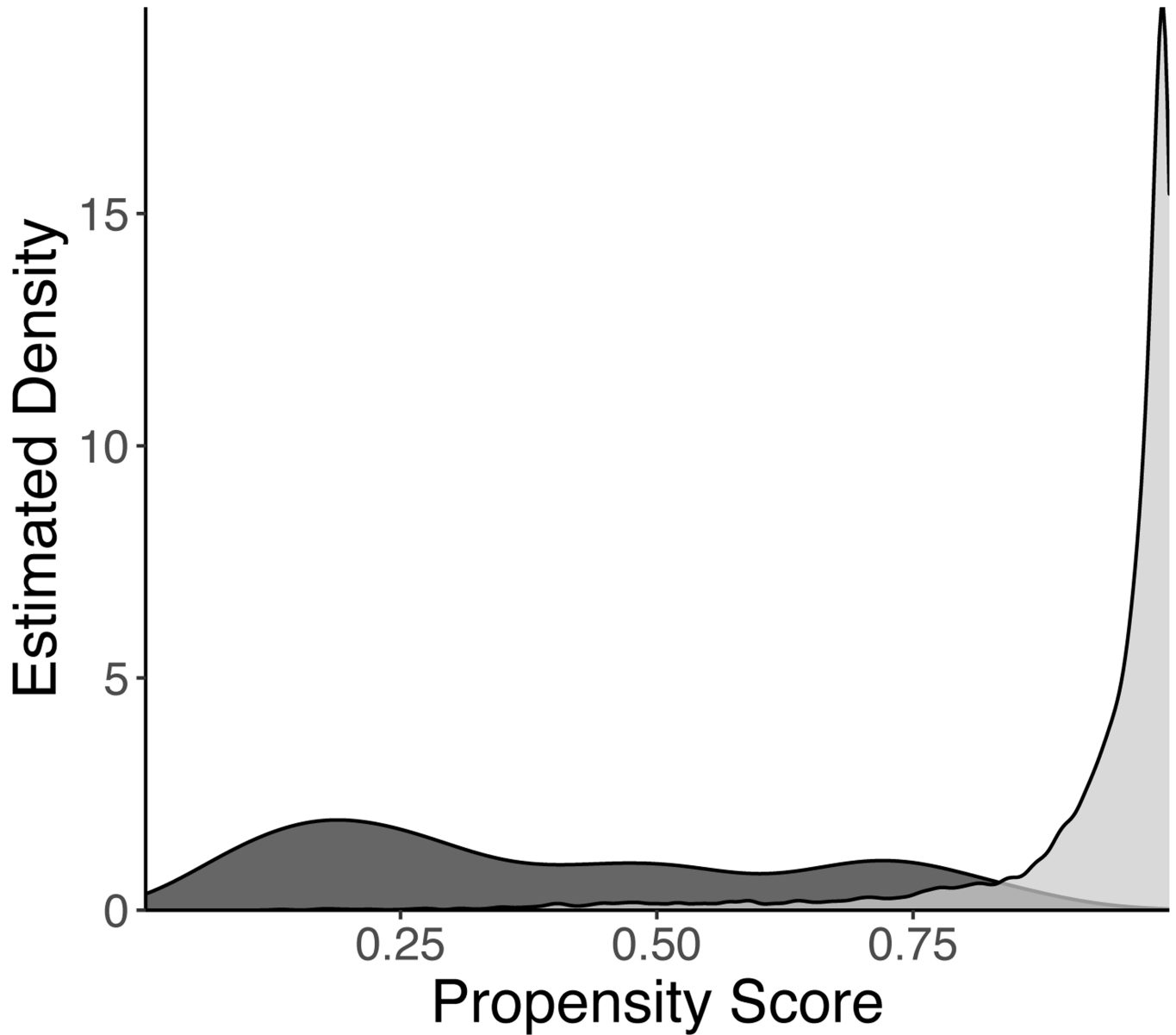


Figure 1: Propensity score overlap plot for consuming at least 1/2 a cup of vegetables per 1,000 kcal in the Nulliparous Pregnancy Outcomes Study: monitoring mothers-to-be (nuMoM2b), 2010–2013.

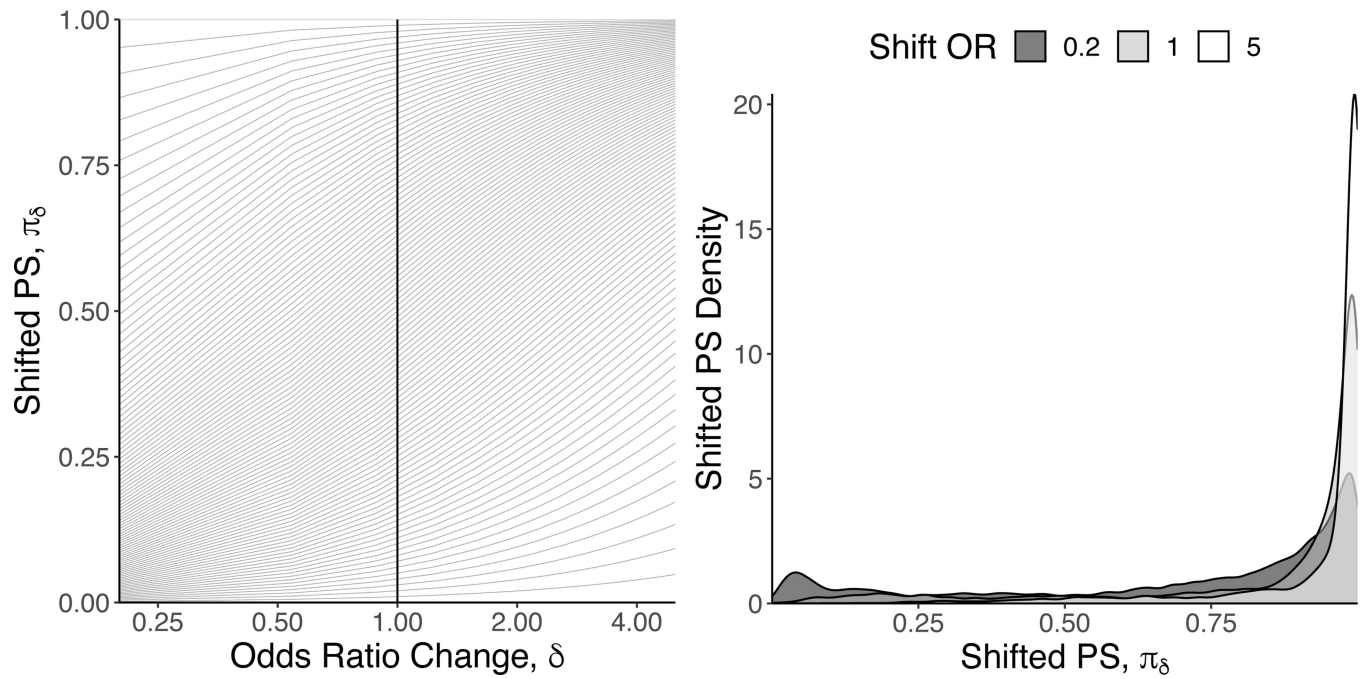


Figure 2:

Left Panel: Shifted propensity score values as a function of the δ odds ratio change for all unique propensity score values in 8,259 from the Nulliparous Pregnancy Outcomes Study: monitoring mothers-to-be(nuMoM2b), 2010–2013. Solid vertical line represents point on x -axis with all of the unique observed propensity score values (i.e., no shift intervention).

Right Panel: Shifted propensity score distributions for δ values of {0,2, 1, and 5.0} estimated with a kernel density smoother.

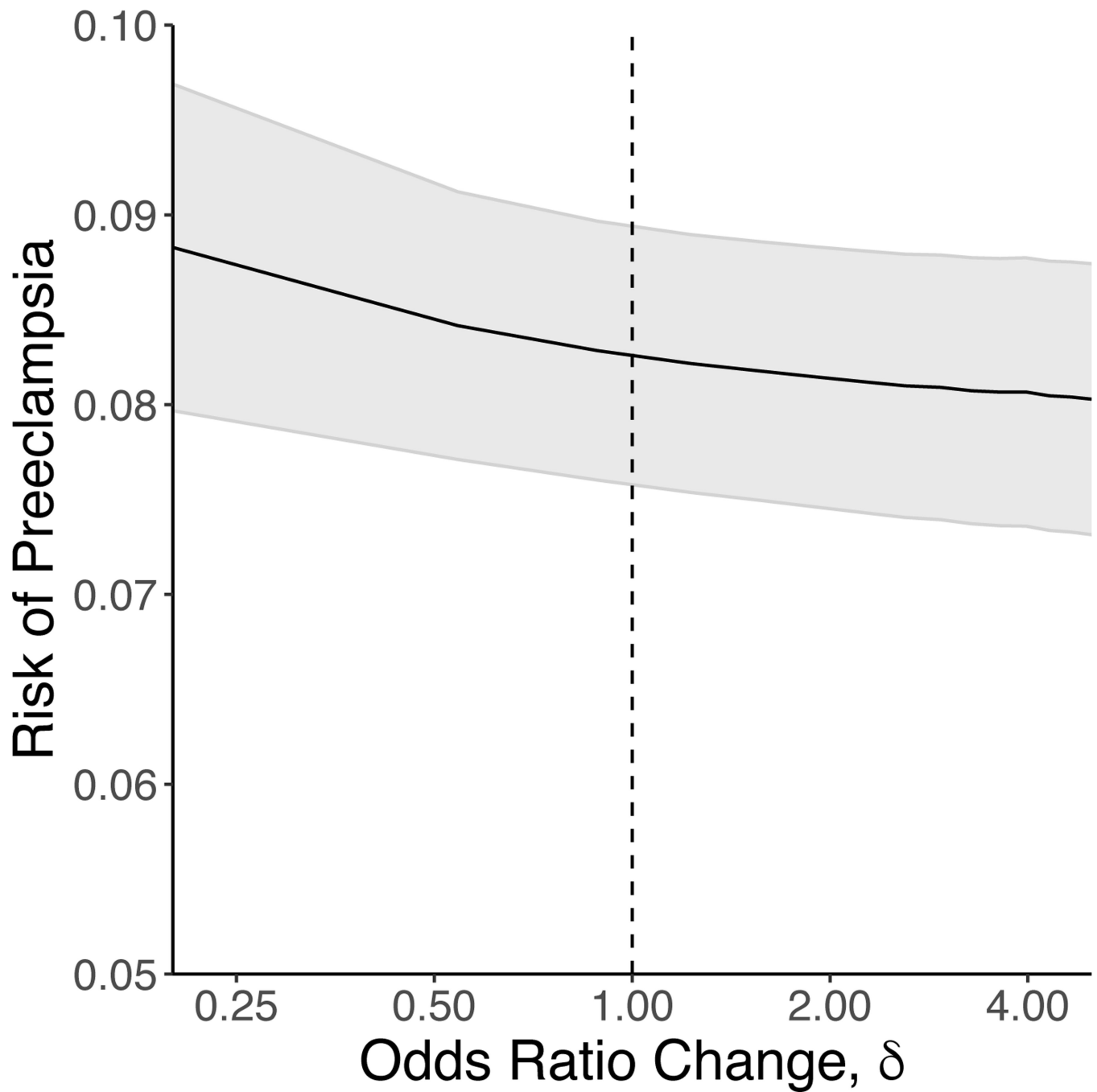


Figure 3: Risk of preeclampsia as a function of changes in the propensity score (PS) for consuming at least 1/2 a cup of vegetables per 1,000 kcal in 8,259 women from the Nulliparous Pregnancy Outcomes Study: monitoring mothers-to-be (nuMoM2b), 2010–2013.

Table 1:

Mean confounder values among exposed and unexposed pregnancies and in the overall sample of 8,259 women from the Nulliparous Pregnancy Outcomes Study: monitoring mothers-to-be (nuMoM2b), 2010–2013.

	Vegetable Density ¹ < 1/2 Cup		
	Yes N = 1,370	No N = 6,889	Overall N = 8,259
Preeclampsia, N (%)	151 (11)	531 (7.7)	682 (8.3)
Age in years, mean (SD)	24.8 (5.0)	28.1 (5.2)	27.5 (5.3)
Self Reported Black, N (%)	321 (23)	623 (9)	944 (11)
Married, N (%)	481 (35)	4721 (69)	5202 (63)
Prepregnancy Smoking, N (%)	375 (27)	1019 (15)	1394 (17)
Prepregnancy BMI in kg/m ² , mean (SD)	26.2 (7.1)	25.4 (6.0)	25.5 (6.2)
Public Insurance, N (%)	632 (46)	1419 (21)	2051 (25)
Empty Calories as % of energy, mean (SD)	39.3 (9.4)	30.4 (7.6)	31.8 (8.6)
Dairy in cups/1,000 kcal, mean (SD)	0.9 (0.6)	0.9 (0.5)	0.9 (0.5)
Total protein foods in oz/1,000 kcal, mean (SD)	2.1 (0.7)	2.5 (0.8)	2.4 (0.8)
Total fruit in cups/1,000 kcal, mean (SD)	0.7 (0.5)	0.9 (0.5)	0.8 (0.5)
Seafood and plant proteins in oz/1,000 kcal, mean (SD)	0.6 (0.4)	0.8 (0.5)	0.8 (0.5)
Sodium in g/1,000 kcal, mean (SD)	1.4 (0.3)	1.7 (0.2)	1.6 (0.3)
Fatty Acids ² , mean (SD)	1.3 (2.0)	1.5 (0.9)	1.4 (1.1)
Refined Grains in oz/1,000 kcal, mean (SD)	2.3 (0.8)	2.2 (0.7)	2.2 (0.7)
Whole Grains in oz/1,000 kcal, mean (SD)	0.5 (0.5)	0.6 (0.4)	0.6 (0.4)

¹Usual daily vegetable intake, expressed in cups per 1,000 kcal

²Ratio of poly- and monounsaturated fatty acids to saturated fatty acids

Table 2:

Average treatment effect estimates on the risk difference scale for the relation between consuming at least 1/2 a cup of vegetables per 1,000 kcal and preeclampsia risk in 8,259 women from the Nulliparous Pregnancy Outcomes Study: monitoring mothers-to-be (nuMoM2b), 2010–2013.

	Risk Difference*	95% CI
Augmented Inverse Probability Weighting	–3.3	–5.1, –1.6
Targeted Minimum Loss-Based Estimation	–2.2	–3.7, –0.7

* Expressed per 100 pregnancies in the sample.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript