Practice of Epidemiology

Comparison of Parametric and Nonparametric Estimators for the Association Between Incident Prepregnancy Obesity and Stillbirth in a Population-Based Cohort Study

Ya-Hui Yu*, Lisa M. Bodnar, Maria M. Brooks, Katherine P. Himes, and Ashley I. Naimi*

* Correspondence to Dr. Ashley I. Naimi, Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, 5131 Parran Hall, 130 DeSoto Street, Pittsburgh, PA 15261 (e-mail: ashley.naimi@pitt.edu).

Initially submitted May 1, 2018; accepted for publication January 10, 2019.

While prepregnancy obesity increases risk of stillbirth, few studies have evaluated the role of newly developed obesity independent of long-standing obesity. Additionally, researchers have relied almost exclusively on parametric models, which require correct specification of an unknown function for consistent estimation. We estimated the association between incident obesity and stillbirth in a cohort constructed from linked birth and death records in Pennsylvania (2003–2013). Incident obesity was defined as body mass index (weight (kg)/height (m)²) greater than or equal to 30. We used parametric G-computation, semiparametric inverse-probability weighting, and parametric/nonparametric targeted minimum loss-based estimation (TMLE) to estimate the association between incident prepregnancy obesity and stillbirth. Compared with pregnancies from women who stayed nonobese, women who became obese prior to their next pregnancy were estimated to have 2.0 (95% confidence interval (CI): 0.5, 3.5) more stillbirths per 1,000 pregnancies using parametric G-computation. However, despite well-behaved stabilized inverse probability weights, risk differences estimated from inverse-probability weighting, nonparametric TMLE, and parametric TMLE represented 6.9 (95% CI: 3.7, 10.0), 0.4 (95% CI: 0.1, 0.7), and 2.9 (95% CI: 1.5, 4.2) excess stillbirths per 1,000 pregnancies, respectively. These results, particularly those derived from nonparametric TMLE, were highly sensitive to covariates included in the propensity score models. Our results suggest that caution is warranted when using nonparametric estimators to quantify exposure effects.

nonparametric estimation; obesity; pregnancy; stillbirth; targeted minimum loss-based estimation

Abbreviations: BMI, body mass index; CI, confidence interval; GWG, gestational weight gain; IPW, inverse-probability weighting; TMLE, targeted minimum loss-based estimation.

Editor's note: An invited commentary on this article appears on page 1337, and the authors' response appears on page 1343.

Women of childbearing age in the United States have a higher risk of stillbirth than women in other developed countries (1). Stillbirth can affect mental and physical health, as well as a family's psychological and social well-being (2–4). Therefore, a better understanding of fetal death etiology is paramount. Evidence suggests a strong association of prepregnancy obesity with higher risk of stillbirth (5). In light of the epidemics of overweight (56%) and obesity (30%) among women of childbearing age in the United States (6), combined with increased stillbirth during

similar time periods, maternal obesity may be an important driver of excess stillbirths.

Despite the importance of this relationship, 2 major challenges limit the ability to identify and estimate the causal effect of prepregnancy obesity on stillbirth. First, the time of onset of obesity is unknown in most studies. As a result, it is unclear whether estimated associations are driven largely by persons with long-standing obesity, persons with recently developed obesity, or some combination of both. Studies have shown that cardiometabolic risk increases with the duration of obesity (7–9)—suggesting, in principle, that associations of obesity with stillbirth may be different for women who became obese in the past year as compared with women who have been consistently obese for a much longer period. Second, because

obesity cannot be assigned via randomization, cause-effect interpretations rely more heavily on parametric assumptions about the nature of the exposure-outcome relationship (10). Such assumptions are commonplace and can involve distributional, functional form, linearity, and other components. Recently developed machine learning methods enable researchers to avoid many of these assumptions (11), yet they are potentially more susceptible to instability because of their highly adaptive nature (12).

We used data from a population-based cohort study to estimate the association between obesity and stillbirth in a way that addresses these obstacles in the literature. First, we focused on incident obesity by restricting all analyses to a subset of women whose obesity onset was bounded. Second, we compared the associations between prepregnancy obesity and stillbirth derived using parametric and nonparametric methods.

METHODS

Conceptual framework

We used the potential outcomes framework (13), informed by a reproductive-life-based approach (14), to articulate possible causal effects of interest. The perfect (albeit impossible) study for quantifying the *effect* of prepregnancy obesity on stillbirth would be an ideal randomized trial among reproductive-aged women who were not obese and not pregnant at baseline (15). Women would be randomized into exposure groups (e.g., would remain nonobese or become obese) and would be followed until their terminal event (stillbirth or a competing event) or the end of study follow-up. The effect of prepregnancy obesity could then be nonparametrically estimated by simply contrasting the proportions of events in each group. While parametric models could be employed to estimate the effect of interest, causal effects are fully nonparametrically identified under randomization.

Were such a study possible, several challenges in defining and estimating the effect of obesity on stillbirth could be handled with relative ease. These include: 1) the ability to define cause-effect relationships using the counterfactual modeling framework and, importantly, to articulate what compliance with the study protocol would precisely entail (16); 2) the ability to properly account for competing risks, including infertility (i.e., no pregnancy) and miscarriage (17); 3) the absence of confounding (in expectation), which would enable effect estimation without resorting to statistical (namely parametric) adjustment for a (possibly) high-dimensional set of covariates; and 4) the ability to clearly define the onset of the risk period in each group, which corresponds to the onset of obesity status in the exposed group. As with all observational studies examining the relationship between obesity and health outcomes, we could not address the first two challenges in our study. We address here the challenges imposed by the third and fourth.

Study population

We used data from a population-based cohort study linking birth and infant/fetal death records in Pennsylvania from the period 2003–2013 ($n = 1,551,919$ singleton pregnancies). We created a unique identifier per woman to link consecutive pregnancies by using a sequential, deterministic linkage

strategy (18–20) (see Web Appendix 1, available at <https://academic.oup.com/aje>). The state of Pennsylvania started using revised birth certificates in 2003 and a fetal death certificate in 2006 which included information on maternal body mass index (BMI). Therefore, we excluded women who had stillbirths before 2006 (4,006 women, 5,168 pregnancies) because of unavailable BMI measurements.

To assess the relationship between incident prepregnancy obesity and stillbirth, we restricted our analysis to women who had 2 or more pregnancies during the period 2003–2013 and a prepregnancy BMI less than 30 in their first identified pregnancy during this period (332,357 women, 769,758 pregnancies). We excluded women with questionable data (a nonlogical age (1,167 women, 3,349 pregnancies), a nonlogical interpregnancy interval (92 women, 316 pregnancies), or prior twin gestation (5,359 women, 8,636 pregnancies)).

Our final analytical sample consisted of the second and/or third pregnancies occurring during 2006–2013 among women who were not obese at the time of their first pregnancy. For the women who became obese in their second pregnancies, their third pregnancies were not included in our analysis. Exclusion of pregnancies occurring after the onset of obesity was consistent with our study objective, which was to examine this association among pregnant women without an obesity history. Overall, 363,610 pregnancies were available for the stillbirth analysis (2006–2013) (Web Figure 1).

The protocol for this study was approved by the institutional review board of the University of Pittsburgh (Pittsburgh, Pennsylvania).

Outcome, exposure, and confounders

Stillbirth was defined as fetal death at ≥ 20 weeks' gestation. Prepregnancy BMI was calculated as self-reported weight (kg) divided by height (m) squared and was categorized as nonobese (< 30) or obese (≥ 30) (21).

Potential confounders were identified as the minimally sufficient adjustment set from the causal diagram depicted in Web Figure 2 (22). We adopted a reproductive-life-based approach considering the potential impact of prior pregnancies (14). Identified confounders were maternal age, maternal race/ethnicity, maternal height, maternal education, pregnancy order, interval between the current pregnancy and the last pregnancy (interpregnancy interval), urban residence, percentage of black residents in the neighborhood, prepregnancy diabetes, prepregnancy hypertension, smoking status, marital status, and type of health insurance during the current and prior pregnancies.

We also adjusted for characteristics of the *prior* pregnancy, including prepregnancy BMI, gestational weight gain (GWG), gestational diabetes, gestational hypertension, smoking status during pregnancy, gestational age, birth weight, birth facility's level of neonatal care, admission to the neonatal intensive care unit, participation in the Women, Infants, and Children (WIC) program, breastfeeding, mode of delivery, Apgar score, stillbirth, and infant death. Note that prior prepregnancy BMI and prior GWG are likely strong confounders. These factors are highly correlated with current prepregnancy BMI (23) and have impacts on prior adverse pregnancy outcomes, which increases the risk of recurrent adverse pregnancy outcomes (24). However, because they were highly predictive of our exposure, we

decided that prior prepregnancy BMI and prior GWG had to be included in exposure models cautiously, since they might introduce positivity violations (25, 26).

For all parametric models, categorical variables were grouped on the basis of the levels outlined in Table 1 and Web Table 1, while continuous variables (interpregnancy interval, maternal height, prior prepregnancy BMI, gestational age, GWG, and birth weight) were fitted with 4-knot restricted quadratic splines (27).

Details on data collection methods have previously been published (28). Briefly, medical records were used for information on maternal characteristics, parity, delivery payment method, prepregnancy diabetes or hypertension, address of the primary residence, gestational age at delivery, and level of neonatal care available at the birth facility. We used the US Census tract of the primary residential address to determine level of urbanicity and proportion of black residents in the neighborhood, by applying Rural-Urban Continuum Codes (29). Interpregnancy interval was calculated as the interval between the delivery date of the last pregnancy and the conception date of the current pregnancy. GWG was calculated as a *z* score by applying a gestational-age- and BMI-specified *z*-score chart (30, 31).

Missing data

Our analytical sample consisted of 394,072 pregnancies from 320,677 women. Of these, 24% of the pregnancies had records with missing values on variables of interest. We imputed these missing data 10 times using multiple imputation via chained equations (32). Our imputation models included variables of analytical interest without missing values (year of birth, infant death/stillbirth, breastfeeding intentions, etc.), as well as auxiliary variables (mother's identifier, county-level Federal Information Processing Standards codes, census tract, facility code, etc.), to impute missing values on prepregnancy weight and height, weight at delivery, age, race/ethnicity, parity, smoking status prior to pregnancy, education, health insurance, marital status, urban residence, enrollment in the Women, Infants, and Children program, prepregnancy diabetes/hypertension, Apgar score, birth weight, gestational age, and infant sex.

Statistical analysis

Estimands. Our interest lies in estimating the marginal association between incident obesity and stillbirth on both the risk difference and risk ratio scales:

$$\sum_c P(Y = 1|X = 1, C) - \sum_c P(Y = 1|X = 0, C)$$

and

$$\frac{\sum_c P(Y = 1|X = 1, C)}{\sum_c P(Y = 1|X = 0, C)},$$

respectively—where incident obesity status (*X*) is coded as 1 if the mother is obese and 0 otherwise, stillbirth status (*Y*) is coded as 1 if stillbirth occurred and 0 otherwise, and *C* denotes all confounders. These estimands can be interpreted as the associations between obesity and stillbirth that are not attributable to the confounding variables listed above.

Estimators. To quantify the estimands defined above, we used parametric G-computation (33), semiparametric inverse-probability-weighted estimation (34), and parametric and nonparametric targeted minimum loss-based estimation (TMLE) (35). Each approach relies on the specification of an exposure and/or outcome model, defined nonparametrically as

$$P(X = 1|C) = f(C)$$

and

$$P(Y = 1|X, C) = g(X, C),$$

respectively, where *f*() and *g*() are arbitrary functions of the confounders (*C*) and the exposure (*X*) defining the relationship between the right- and left-hand sides of each equation. These models are typically referred to as the propensity score model and the outcome model, respectively. However, in an observational study, we do not know the true form of the relationships encoded in *f*() and *g*() (36).

We first used parametric G-computation to estimate the association of prepregnancy obesity with stillbirth. G-computation relies on correct parametric specification of the outcome model function *g*(*X*, *C*). Though the correct outcome model is often never known, it is often implemented by 1) regressing the observed outcome against the exposure and confounders (37), 2) predicting outcomes for all pregnancies under exposed and unexposed scenarios, 3) computing the average of each of these predicted outcomes, and 4) contrasting these averages to quantify the estimand of interest. With a time-fixed exposure such as the one we present, the G-computation estimator is equivalent to model-based standardization (38). All listed confounders were included in the outcome model. Standard errors were obtained using the normal interval bootstrap with 100 resamples (39).

Semiparametric inverse-probability weighting (IPW) estimation was implemented by constructing inverse probability weights, which specifies the propensity score function *f*(*C*) as a standard (i.e., parametric) logistic model. These models are semiparametric because they are obtained from the combination of a parametric (finite-dimensional) propensity score model with a nonparametric (infinite-dimensional) outcome model (40, 41). All listed confounders were included in the propensity score model. Parameters corresponding to the risk difference and risk ratio from linear and log-linear regression models were quantified via nonparametric maximum likelihood estimation, with each pregnancy's contribution to the likelihood weighted by the inverse of the probability of being in its observed obesity category. Standard errors were obtained using the robust variance estimator (42).

TMLE was performed by updating the G-computation estimator using information from the exposure model (via the "clever covariate") to better target the estimand of interest (35). We implemented TMLE both parametrically and nonparametrically.

In the parametric approach, both the propensity score model and the outcome model were fitted via standard logistic regression. When TMLE was implemented nonparametrically, both the propensity score and outcome models were specified using an ensemble machine learning method (Super Learner) (43) to create an optimally weighted combination of predictions from a library of algorithms (Web Appendix 2). "Optimal" was defined

Table 1. Characteristics of Mothers in a Population-Based Cohort Study of Stillbirths and Live Births, by Prepregnancy Obesity Status, Pennsylvania, 2006–2013

Characteristic	Prepregnancy Obesity Status ^a			
	Nonobese (n = 326,178 ^b)		Obese (n = 37,432)	
	No. of Pregnancies	%	No. of Pregnancies	%
Maternal race/ethnicity				
Non-Hispanic white	244,918	75.1	24,837	66.4
Non-Hispanic black	39,912	12.2	7,643	20.4
Hispanic	26,171	8.0	4,012	10.7
Other	15,177	4.7	940	2.5
Maternal age, years				
<20	10,531	3.2	1,270	3.4
20–29	160,717	49.3	21,615	57.7
>29	154,930	47.5	14,547	38.9
Maternal education				
Less than high school	47,574	14.6	6,179	16.5
High school or equivalent	80,504	24.7	12,456	33.3
Some college	80,448	24.7	11,273	30.1
College graduate	117,652	36.1	7,524	20.1
Size of metropolitan area, no. of residents				
>1 million	169,994	52.1	18,729	50.0
250,000–1 million	95,344	29.2	10,899	29.1
<250,000	60,840	18.7	7,804	20.9
Neighborhood % black				
Lowest tertile (<1%)	111,964	34.3	11,470	30.6
Middle tertile (1%–7%)	115,208	35.3	11,076	29.6
Highest tertile (>7%)	99,006	30.4	14,886	39.8
Prepregnancy diabetes				
No	324,933	99.6	37,078	99.1
Yes	1,245	0.4	354	0.9
Prepregnancy hypertension				
No	323,724	99.3	36,548	97.6
Yes	2,454	0.7	884	2.4
Prepregnancy smoking, cigarettes/day				
0 (nonsmoker)	259,755	79.6	28,674	76.6
<10	18,023	5.5	2,565	6.9
10–20	19,591	6.0	2,510	6.7
>20	28,809	8.8	3,683	9.8
Type of health insurance				
Nonprivate	127,233	39.0	18,653	49.8
Private	198,945	61.0	18,779	50.2
Marital status				
Unmarried	106,304	32.6	16,625	44.4
Married	219,874	67.4	20,807	55.6
Interpregnancy interval, months				
<18	217,048	66.5	27,178	72.6
≥18	109,130	33.5	10,254	27.4

^a Obesity was defined as body mass index (weight (kg)/height (m)²) ≥30.^b Total number of pregnancies.

using a 5-fold cross-validated L^2 loss function (44). Our Super Learner library included the arithmetic mean, neural networks, multivariate adaptive regression splines (MARS), least absolute shrinkage and selection operator (LASSO), generalized additive models, random forests, and gradient boosted machines. We also included generalized linear models with second-order polynomials and 2-way interactions between the main terms. A cross-validated search over the (hyper-) parameter space was also conducted for random forests, gradient boosting, least absolute shrinkage and selection operator, and generalized additive models. All listed confounders were included in both the propensity score and outcome models. By default, propensity scores are truncated at the 2.5th and 97.5th percentiles when implemented in TMLE.

RESULTS

In our cohort, approximately 11% of women became obese prior to their second or third pregnancy. Compared with their nonobese counterparts, women with incident prepregnancy obesity were more likely to be non-Hispanic black, to be Hispanic, to be younger, to have a high school education or less, and to live in neighborhoods with a higher percentage of black residents. They were also subject to more prepregnancy diabetes and hypertension and more often had nonprivate health insurance, were unmarried, and had shorter intervals between pregnancies (Table 1). In their last pregnancy, the obese women experienced more adverse pregnancy outcomes, including infant death, fetal death, and large size for gestational age, and they had higher prepregnancy BMI (Web Figure 3), greater GWG, and more gestational diabetes and gestational hypertension (Web Table 1).

As shown in Figure 1, among pregnancies from nonobese women, the crude stillbirth rate increased with pregnancy order, from 3.2 per 1,000 pregnancies to 3.7 per 1,000 pregnancies for second and third pregnancies, respectively. Pregnancies in women who became obese had a higher crude stillbirth rate, which also increased with pregnancy order: from 5.7 per 1,000 pregnancies to 7.8 per 1,000 pregnancies for second and third pregnancies, respectively.

To avoid problems induced by violations of the positivity assumption for estimators that rely on the propensity score (IPW, TMLE), we examined the mean and maximum values for the stabilized inverse probability weights. As shown in Table 2, the mean

value was close to 1.00 for all estimation methods (parametric, nonparametric) and the maximum value was no greater than 48.5, suggesting well-behaved inverse probability weights (26).

Despite these well-behaved weights, we found important differences in the estimated associations of incident obesity with stillbirth between different methods when adjusting for all identified confounders. The unadjusted risk difference for stillbirth when comparing obese pregnancies with nonobese pregnancies was 2.7 (95% confidence interval (CI): 1.9, 3.5) per 1,000 pregnancies. When using parametric G-computation to adjust for all of the identified confounders, we found that women with incident prepregnancy obesity had 2.0 (95% CI: 0.5, 3.5) excess stillbirths per 1,000 pregnancies relative to nonobese women (Table 3). Parametric TMLE yielded a risk difference of 2.9 (95% CI: 1.5, 4.2) excess stillbirths per 1,000 pregnancies. In contrast, the corresponding risk differences from semiparametric IPW or nonparametric TMLE deviated much further from the G-computation results, with 6.9 (95% CI: 3.7, 10.0) and 0.4 (95% CI: 0.1, 0.7) excess stillbirths per 1,000 pregnancies, respectively.

To investigate this deviation, we examined the propensity score overlap plots comparing obese and nonobese groups estimated from IPW (Web Figure 4, panel A), parametric TMLE (Web Figure 5, panel A), and nonparametric TMLE (Web Figure 6, panel A). In contrast to the results shown in Table 2, the plots indicated problematic overlap, especially for nonparametric TMLE.

Next, we reexamined this association by excluding prior prepregnancy BMI, prior GWG, or both of these factors from the set of confounding factors. These two confounders were known to be highly correlated with current prepregnancy obesity status and the suspected cause of both the poor propensity score overlap and the high variability in the estimated associations across methods. After removing these potential confounders, we found that results from G-computation were most consistent, with point estimates ranging from 1.9 (95% CI: 0.7, 3.1) for the estimate without prior prepregnancy BMI to 2.1 (95% CI: 0.5, 3.5) for the estimate without prior GWG (Table 3). In contrast, nonparametric TMLE yielded the most varying set of results, with estimates ranging from 3.2 (95% CI: 2.3, 4.0) for the estimate without prior prepregnancy BMI to 2.1 (95% CI: 1.7, 2.5) for the estimate without prior GWG (Table 3).

A similar pattern was observed for risk ratios (Table 4). When all of the identified confounders were included, risk ratios for stillbirth comparing women who became obese with those who

Figure 1. Crude rate of stillbirth in a population-based cohort study of stillbirths and live births, by pregnancy order and maternal obesity status, Pennsylvania, 2006–2013. The crude stillbirth rate (with 95% bootstrapped confidence interval (CI)) is presented as the number of stillbirths per 1,000 pregnancies. Obesity was defined as body mass index (weight (kg)/height (m)²) ≥30. *n* = number of pregnancies.

Table 2. Distribution of Stabilized Weights in a Population-Based Cohort Study of Stillbirths and Live Births, by Estimation Method and Confounder Set, Pennsylvania, 2006–2013

Confounder Set	Stabilized Weight					
	IPW		Parametric TMLE ^a		Nonparametric TMLE ^b	
	Mean	Range	Mean	Range	Mean	Range
All identified confounders ^c	1.01	0.11–48.5	1.01	0.11–48.5	0.91	0.11–4.21
Excluding prior prepregnancy BMI	1.00	0.13–8.09	1.00	0.13–8.09	0.97	0.11–22.64
Excluding prior GWG	1.00	0.11–31.12	1.00	0.11–31.14	0.93	0.11–3.52
Excluding prior prepregnancy BMI and GWG	1.00	0.13–8.35	1.00	0.13–8.37	0.96	0.13–3.93

Abbreviations: BMI, body mass index; GWG, gestational weight gain; IPW, inverse-probability weighting; TMLE, targeted minimum loss-based estimation; WIC, Women, Infants, and Children.

^a TMLE with parametric regression.

^b TMLE with Super Learner, including the algorithms of SL.mean, SL.nnet, SL.earth, SL.ranger, SL.xgboost, SL.glmnet, SL.gam, and SL.glm.interaction.

^c Maternal race/ethnicity, maternal height, pregnancy order, interval between the current and last pregnancies (interpregnancy interval), variables for both the prior pregnancy and the current pregnancy (maternal age, maternal education, urban residence, percentage of black residents in the neighborhood, prepregnancy diabetes, prepregnancy hypertension, smoking status, marital status, and health insurance status), and variables for the prior pregnancy (pregnancy BMI, GWG, gestational diabetes, gestational hypertension, smoking status during pregnancy, gestational age, birth weight, birth facility's level of neonatal care, admission to the neonatal intensive care unit, participation in the WIC program, breastfeeding, mode of delivery, Apgar score, stillbirth, and infant death).

stayed nonobese were 2.99 (95% CI: 2.19, 4.08), 1.60 (95% CI: 0.93, 2.75), 1.80 (95% CI: 1.45, 2.22), and 1.12 (95% CI: 1.02, 1.23) as estimated from IPW, parametric G-computation, parametric TMLE, and nonparametric TMLE, respectively. After removal of both influential confounders, risk ratios derived from IPW

became much smaller (risk ratio = 1.71, 95% CI: 1.45, 2.02), while parametric and nonparametric TMLE generated risk ratios of 1.44 (95% CI: 1.20, 1.74) and 1.43 (95% CI: 1.27, 1.62), respectively. Parametric G-computation yielded a stable result (risk ratio = 1.56, 95% CI: 0.98, 2.47).

Table 3. Risk Difference for Stillbirth Among Women With Incident Obesity^a Versus Women Without Incident Obesity, According to Estimation Method and Confounder Set, in a Population-Based Cohort Study of Stillbirths and Live Births, Pennsylvania, 2006–2013

Confounder Set	Estimation Method							
	IPW		G-Computation		Parametric TMLE ^b		Nonparametric TMLE ^c	
	RD ^d	95% CI	RD	95% CI	RD	95% CI	RD	95% CI
All identified confounders ^e	6.9	3.7, 10.0	2.0	0.5, 3.5	2.9	1.5, 4.2	0.4	0.1, 0.7
Excluding prior prepregnancy BMI	2.4	1.4, 3.4	1.9	0.7, 3.1	1.5	0.6, 2.4	3.2	2.3, 4.0
Excluding prior GWG	6.0	3.4, 8.6	2.1	0.5, 3.5	3.0	1.7, 4.3	2.1	1.7, 2.5
Excluding prior prepregnancy BMI and GWG	2.4	1.5, 3.4	1.9	0.7, 3.1	1.5	0.6, 2.5	1.5	0.9, 2.0

Abbreviations: BMI, body mass index; CI, confidence interval; GWG, gestational weight gain; IPW, inverse-probability weighting; RD, risk difference; TMLE, targeted minimum loss-based estimation; WIC, Women, Infants, and Children.

^a Obesity was defined as body mass index (weight (kg)/height (m)²) ≥ 30.

^b TMLE with parametric regression.

^c TMLE with Super Learner, including the algorithms of SL.mean, SL.nnet, SL.earth, SL.ranger, SL.xgboost, SL.glmnet, SL.gam, and SL.glm.interaction.

^d Risk difference per 1,000 pregnancies.

^e Maternal race/ethnicity, maternal height, pregnancy order, interval between the current and last pregnancies (interpregnancy interval), variables for both the prior pregnancy and the current pregnancy (maternal age, maternal education, urban residence, percentage of black residents in the neighborhood, prepregnancy diabetes, prepregnancy hypertension, smoking status, marital status, and health insurance status), and variables for the prior pregnancy (pregnancy BMI, GWG, gestational diabetes, gestational hypertension, smoking status during pregnancy, gestational age, birth weight, birth facility's level of neonatal care, admission to the neonatal intensive care unit, participation in the WIC program, breastfeeding, mode of delivery, Apgar score, stillbirth, and infant death).

Table 4. Risk Ratio for Stillbirth Among Women With Incident Obesity^a Versus Women Without Incident Obesity, According to Estimation Method and Confounder Set, in a Population-Based Cohort Study of Stillbirths and Live Births, Pennsylvania, 2006–2013

Confounder Set	Estimation Method							
	IPW		G-Computation		Parametric TMLE ^b		Nonparametric TMLE ^c	
	RR	95% CI	RR	95% CI	RR	95% CI	RR	95% CI
All identified confounders ^d	2.99	2.19, 4.08	1.60	0.93, 2.75	1.80	1.45, 2.22	1.12	1.02, 1.23
Excluding prior prepregnancy BMI	1.70	1.43, 2.02	1.55	1.01, 2.37	1.43	1.18, 1.72	1.94	1.68, 2.22
Excluding prior GWG	2.75	2.07, 3.66	1.61	0.93, 2.75	1.80	1.49, 2.27	1.62	1.48, 1.77
Excluding prior prepregnancy BMI and GWG	1.71	1.45, 2.02	1.56	0.98, 2.47	1.44	1.20, 1.74	1.43	1.27, 1.62

Abbreviations: BMI, body mass index; CI, confidence interval; GWG, gestational weight gain; IPW, inverse-probability weighting; RR, risk ratio; TMLE, targeted minimum loss-based estimation; WIC, Women, Infants, and Children.

^a Obesity was defined as body mass index (weight (kg)/height (m)²) ≥30.

^b TMLE with parametric regression.

^c TMLE with Super Learner, including the algorithms of SL.mean, SL.nnet, SL.earth, SL.ranger, SL.xgboost, SL.glmnet, SL.gam, and SL.glm.interaction.

^d Maternal race/ethnicity, maternal height, pregnancy order, interval between the current and last pregnancies (interpregnancy interval), variables for both the prior pregnancy and the current pregnancy (maternal age, maternal education, urban residence, percentage of black residents in the neighborhood, prepregnancy diabetes, prepregnancy hypertension, smoking status, marital status, and health insurance status), and variables for the prior pregnancy (prepregnancy BMI, GWG, gestational diabetes, gestational hypertension, smoking status during pregnancy, gestational age, birth weight, birth facility's level of neonatal care, admission to the neonatal intensive care unit, participation in the WIC program, breastfeeding, mode of delivery, Apgar score, stillbirth, and infant death).

DISCUSSION

We used different analytical methods to quantify the association between prepregnancy obesity and stillbirth among women who were not obese in their prior pregnancies. While we found positive associations of incident prepregnancy obesity with stillbirth using all methods, meaningful differences were observed between methods that relied, at least partially, on propensity scores (semiparametric IPW, nonparametric/parametric TMLE). Furthermore, the variability in estimated associations correlated with the degree of nonpositivity and was exacerbated when nonparametric methods were used. Nonparametric TMLE demonstrated the highest volatility in estimated associations under different sets of confounders.

Several challenges must be considered when estimating the effect of exposures (such as obesity) that cannot be simply assigned. Among these are the need to 1) clearly articulate the onset of the exposure period and 2) deal with fact that some degree of reliance on parametric models is necessary to quantify the effect of interest. If the true time of exposure onset is available, dealing with the first challenge is relatively straightforward. However, avoiding reliance on parametric modeling assumptions presents additional technical complexities.

Estimation of effects can be accomplished in several ways. One can use outcome modeling approaches (e.g., parametric G-computation) that require researchers to correctly model the entire data-generating distribution (37, 45). One can also use IPW, which relies on correct specification of a propensity score model, relaxing the requirement to include all exposure-confounder interactions, and still obtain a consistent estimator of the marginal exposure-outcome association (34).

However, for parametric models, researchers must correctly specify the functional form of the relationship between the confounders and the exposure (via propensity score adjustment) or the confounders and the outcome (via outcome modeling adjustment). Correct model specification is an unverifiable assumption that requires choosing the correct link function, including the correct set of exposure-confounder and/or confounder-confounder interactions, and appropriately modeling potential nonlinearities. Misspecification of any of these components of the correct model specification assumption can lead to bias in estimated effects.

Use of double-robust methods offers an opportunity to rely on nonparametric (e.g., machine learning) techniques (35, 43). In doing so, researchers may avoid relying on unwarranted assumptions about how the exposure and confounders relate to the exposure or the outcome (46). Numerous theoretical and simulation studies have shown that nonparametric TMLE has greater efficiency and less bias when compared with (misspecified) parametric singly robust methods (11) and nonparametric singly robust methods (46, 47).

However, the benefits of using such nonparametric techniques come at a cost. In this study, we showed that when strong relationships between confounders and exposures exist, nonparametric TMLE can yield highly volatile estimates. In our study, body weight characteristics from prior pregnancies were strongly associated with obesity in the pregnancy of interest. These strong associations resulted in positivity violations that were not apparent when examining the distribution of stabilized inverse probability weights, a commonly used diagnostic strategy (45). Thus, while nonparametric TMLE has attractive properties, researchers

should be aware of the costs associated with its flexible nature. We found that nonparametric TMLE estimates were highly variable, and this variability was related to the degree of propensity score overlap. These results suggest that when using nonparametric TMLE, researchers should employ bias analyses (12) or attempt to triangulate effect estimates with a variety of approaches (e.g., G-computation).

Our results should be interpreted with consideration of some limitations. First, to improve internal validity for quantifying the association between obesity and stillbirth, we sacrificed some generalizability. The boundaries for the onset of obesity were set by focusing on multiparous women who were nonobese at the time of their first pregnancy; therefore, we cannot infer that our results apply to nulliparous women. It is plausible that the risk factors for becoming obese may differ between these two groups and that the discrepancies in risk factors may have different associations with stillbirth. Furthermore, as with previous observational studies (5), we could not identify the true time of obesity onset. Thus, our study was not ideal for evaluation of the association of long-standing obesity with stillbirth. Dichotomization of BMI values is a well-recognized issue when studying the effect of obesity. It is especially important to recognize this in our study, since we defined incident obesity as a change in obesity status between 2 pregnancies. Because of the relatively short time window between pregnancies, change in obesity status in our study may represent a smaller BMI change than that which would be present in studies examining change in obesity status over a lifetime. However, examining the effect of BMI changes of different magnitudes on the risk of stillbirth was outside the scope of our study, which emphasized comparing the estimates derived from parametric and nonparametric methods. In addition, even though we further refined the prepregnancy obesity construct as incident prepregnancy obesity, obesity is not a perfectly defined exposure. Different reasons for women's transitioning from nonobese to obese may have different associations with stillbirth (16). Future studies with sufficient information on BMI changes and obesity onset are warranted in order to tackle these issues.

Our study has important practical implications for researchers using nonparametric estimation techniques. Notably, estimates obtained from nonparametric TMLE should be interpreted in light of careful examination of the propensity score distribution among the exposed and unexposed, and they should be compared with results from alternative parametric and nonparametric approaches.

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania (Ya-Hui Yu, Lisa M. Bodnar, Maria M. Brooks, Ashley I. Naimi); Department of Obstetrics, Gynecology, and Reproductive Sciences, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania (Lisa M. Bodnar, Katherine P. Himes); and Magee-Womens Research Institute, Pittsburgh, Pennsylvania (Lisa M. Bodnar, Katherine P. Himes).

All of the authors contributed equally to this work.

This study was supported by grant HD065807 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

We thank Dr. Laura Balzer for expert advice and Sara Parisi and Melissa Mangini for data management. Computing resources were provided by the University of Pittsburgh Center for Research Computing.

Conflict of interest: none declared.

REFERENCES

- MacDorman MF, Gregory ECW. Fetal and perinatal mortality: United States, 2013. *Natl Vital Stat Rep.* 2015;64(8):1–24.
- Badenhorst W, Riches S, Turtton P, et al. The psychological effects of stillbirth and neonatal death on fathers: systematic review. *J Psychosom Obstet Gynaecol.* 2006;27(4):245–256.
- Daugirdaitė V, van den Akker O, Purewal S. Posttraumatic stress and posttraumatic stress disorder after termination of pregnancy and reproductive loss: a systematic review. *J Pregnancy.* 2015;2015:646345.
- Blackmore ER, Côté-Arsenault D, Tang W, et al. Previous prenatal loss as a predictor of perinatal depression and anxiety. *Br J Psychiatry.* 2011;198(5):373–378.
- Aune D, Saugstad OD, Henriksen T, et al. Maternal body mass index and the risk of fetal death, stillbirth, and infant death: a systematic review and meta-analysis. *JAMA.* 2014;311(15):1536–1546.
- Flegal KM, Carroll MD, Kit BK, et al. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999–2010. *JAMA.* 2012;307(5):491–497.
- Mongraw-Chaffin M, Foster MC, Kalyani RR, et al. Obesity severity and duration are associated with incident metabolic syndrome: evidence against metabolically healthy obesity from the Multi-Ethnic Study of Atherosclerosis. *J Clin Endocrinol Metab.* 2016;101(11):4117–4124.
- Hu Y, Bhupathiraju SN, de Koning L, et al. Duration of obesity and overweight and risk of type 2 diabetes among US women. *Obesity (Silver Spring).* 2014;22(10):2267–2273.
- Tanamas SK, Wong E, Backholer K, et al. Duration of obesity and incident hypertension in adults from the Framingham Heart Study. *J Hypertens.* 2015;33(3):542–545.
- Hernán MA, Robins JM. Estimating causal effects from epidemiological data. *J Epidemiol Community Health.* 2006;60(7):578–586.
- Schuler MS, Rose S. Targeted maximum likelihood estimation for causal inference in observational studies. *Am J Epidemiol.* 2017;185(1):65–73.
- Bahamyriou A, Blais L, Forget A, et al. Understanding and diagnosing the potential for bias when using machine learning methods with doubly robust causal estimators [published online ahead of print May 2, 2018]. *Stat Methods Med Res.* (doi:10.1177/0962280218772065).
- Rubin DB. Causal inference using potential outcomes: design, modeling, decisions. *J Am Stat Assoc.* 2005;100(469):322–331.
- Hutcheon JA, Platt RW. The impact of past pregnancy experience on subsequent perinatal outcomes. *Paediatr Perinat Epidemiol.* 2008;22(4):400–408.
- Daniel RM, De Stavola BL, Vansteelandt S. Commentary: the formal approach to quantitative causal inference in epidemiology: misguided or misrepresented? *Int J Epidemiol.* 2016;45(6):1817–1829.

16. Hernán MA, Taubman SL. Does obesity shorten life? The importance of well-defined interventions to answer causal questions. *Int J Obes (Lond)*. 2008;32(suppl 3):S8–S14.
17. Naimi A, Auger N. Cumulative risk of stillbirth in the presence of competing events. *Br J Obstet Gynaecol*. 2016;123(7):1071–1074.
18. Blakely T, Salmond C. Probabilistic record linkage and a method to calculate the positive predictive value. *Int J Epidemiol*. 2002;31(6):1246–1252.
19. Herman AA, McCarthy BJ, Bakewell JM, et al. Data linkage methods used in maternally-linked birth and infant death surveillance data sets from the United States (Georgia, Missouri, Utah and Washington State), Israel, Norway, Scotland and Western Australia. *Paediatr Perinat Epidemiol*. 1997;11(suppl 1):5–22.
20. Bodnar LM, Siminerio LL, Himes KP, et al. Maternal obesity and gestational weight gain are risk factors for infant death. *Obesity (Silver Spring)*. 2016;24(2):490–498.
21. World Health Organization. *Obesity: Preventing and Managing the Global Epidemic*. (WHO Technical Report Series 894). Geneva, Switzerland: World Health Organization; 2000.
22. Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Med Res Methodol*. 2008;8:Article 70.
23. Rong K, Yu K, Han X, et al. Pre-pregnancy BMI, gestational weight gain and postpartum weight retention: a meta-analysis of observational studies. *Public Health Nutr*. 2015;18(12):2172–2182.
24. Malacova E, Regan A, Nassar N, et al. Risk of stillbirth, preterm delivery, and fetal growth restriction following exposure in a previous birth: systematic review and meta-analysis. *Br J Obstet Gynaecol*. 2018;125(2):183–192.
25. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008;168(6):656–664.
26. Westreich D, Cole SR. Invited commentary: positivity in practice. *Am J Epidemiol*. 2010;171(6):674–677.
27. Howe CJ, Cole SR, Westreich DJ, et al. Splines for trend analysis and continuous confounder control. *Epidemiology*. 2011;22(6):874–875.
28. Lemon LS, Naimi AI, Abrams B, et al. Prepregnancy obesity and the racial disparity in infant mortality. *Obesity (Silver Spring)*. 2016;24(12):2578–2584.
29. Economic Research Service, US Department of Agriculture. Rural-Urban Continuum Codes. 2013. <https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx#U7WgCvldUg0>. Updated May 10, 2013. Accessed April 1, 2018.
30. Hutcheon JA, Platt RW, Abrams B, et al. A weight-gain-for-gestational-age z score chart for the assessment of maternal weight gain in pregnancy. *Am J Clin Nutr*. 2013;97(5):1062–1067.
31. Hutcheon JA, Platt RW, Abrams B, et al. Pregnancy weight gain charts for obese and overweight women. *Obesity (Silver Spring)*. 2015;23(3):532–535.
32. Royston P. Multiple imputation of missing values: further update of ice, with an emphasis on categorical variables. *Stata J*. 2009;9(3):466–477.
33. Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math Model*. 1986;7(9–12):1393–1512.
34. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11(5):550–560.
35. van der Laan MJ, Rose S. *Targeted Learning: Causal Inference for Observational and Experimental Data*. New York, NY: Springer-Verlag New York; 2011.
36. Robins JM. Data, design, and background knowledge in etiologic inference. *Epidemiology*. 2001;12(3):313–320.
37. Snowden JM, Rose S, Mortimer KM. Implementation of G-computation on a simulated data set: demonstration of a causal inference technique. *Am J Epidemiol*. 2011;173(7):731–738.
38. Muller CJ, MacLehose RF. Estimating predicted probabilities from logistic regression: different methods correspond to different target populations. *Int J Epidemiol*. 2014;43(3):962–970.
39. Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. Boca Raton, FL: CRC Press; 1994.
40. Tsiatis AA. *Semiparametric Theory and Missing Data*. New York, NY: Springer Publishing Company; 2006.
41. Breskin A, Cole SR, Westreich D. Exploring the subtleties of inverse probability weighting and marginal structural models. *Epidemiology*. 2018;29(3):352–355.
42. White H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econom J Econom Soc*. 1980;48(4):817–838.
43. van der Laan MJ, Polley EC, Hubbard AE. Super Learner. *Stat Appl Genet Mol Biol*. 2007;6(1):Article 25.
44. Naimi AI, Balzer LB. Stacked generalization: an introduction to super learning. *Eur J Epidemiol*. 2018;33(5):459–464.
45. Hernán M, Robins J. *Causal Inference*. Boca Raton, FL: Chapman & Hall/CRC Press. In press.
46. Naimi AI, Kennedy EH. Nonparametric double robustness. *arXiv*. 2017. (doi:arXiv:1711.07137v1). Accessed April 1, 2018.
47. Luque-Fernandez MA, Belot A, Valeri L, et al. Data-adaptive estimation for double-robust methods in population-based cancer epidemiology: risk differences for lung cancer mortality by emergency presentation. *Am J Epidemiol*. 2018;187(4):871–878.