Methodology article

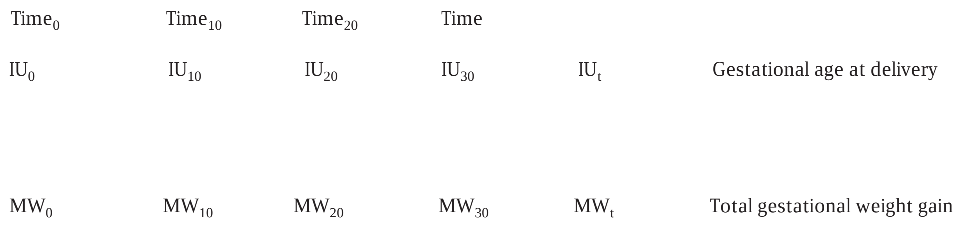
which only total weight gain is available to obtain a gestational age standardised exposure for epidemiological analyses. While z-scores can help mitigate the correlation between two variables when they adequately represent the underlying population they valid estimation of the gestational-age-speciÞc weight gain means and variances for a z-score chart requires additional assumptions. These assumptions include a validation sample representative of the study population and correct speciÞcation of the model underlying the relationship between gestational age and weight gain, including the mean, variance, link function, and functional form of the covariates and any potential confounders. The implications of violating these assumptions, such as applying z-scores derived from one population to a fundamentally different one, have not been fully evaluated.

An alternative method for addressing correlation between total weight gain and gestational age is regression-based adjustment for gestational age. However, this straightforward approach has also not been evaluated in the context of gestational weight gain and in the past its validity has created controversy.[5–7] There was concern that adjusting for gestational age may induce collider stratiÞcation bias by adjusting for a potential intermediate between weight gain and mortality.[7] The adjustment approach has the advantage that researchers can specify the model to Þt the structure of the gestational weight gain and gestational age relationship in their observed data.

In this paper, we use directed acyclic graphs (DAGs) and simulation approaches to evaluate the confounding due to gestational age in studies of assessing the relation between gestational weight gain and pregnancy outcomes associated with gestational age. First, we use DAGs to describe the correlation between weight gain and gestational age longitudinally across gestation. This can help clarify assumptions about the directionality of the relations between variables at the cross-sectional time point of delivery, the most common form of perinatal variables. Second, we utilise analytical and simulation approaches to assess the implications and potential biases from using simpliÞed composite measures of total weight gain and gestational age at delivery as related to neonatal mortality, an outcome where critical data gaps remain in the literature.[1] We compare the approach of adjusting for gestational age in a model of total gestational weight to applying previously published z-score reference chart values for total gestational weight gain.[3] By using a large, diverse cohort, comprising several sites from across the United States, we evaluate the impact of mis-specifying the distributions of weight gain and gestational age, a key assumption of the z-score approach.

Directed acyclic graphs

Gestational weight gain is a time-dependent variable that is often treated as a Þxed variable of total weight gain at delivery, as data on total weight gain are more often collected than repeated measures of maternal weight gain. In reality, total gestational weight gain represents the summation of maternal weight gain at each week of gestation culminating in a measure of cumulative weight gain (or loss). Similarly, gestational age at delivery is the summation of indicator variables denoting whether or not the baby is still *in utero* at each gestational week (t). In addition, birthweight is the summation of foetal weight accumulated longitudinally across gestation. The DAG in Figure 1 displays a simpliÞcation of these longitudinal time-dependent variables at intermittent gestational ages. We assume that there are no unmeasured confounders and that the DAG is complete. Whether or not a woman is still pregnant (i.weigoted tostudyis in isat



gestational age at delivery, but rather a ÒUÓ representing the unobserved longitudinal processes and feedbacks between weight gain and gestational age. This is in contrast to previous assumptions that the summary measure of total gestational weight gain is directly impacted by the summary measure of gestational age at delivery or vice versa, which fails to account for the past time-dependent confounding affected by prior exposures. In this DAG, adjusting for gestational age at delivery blocks the backdoor path through U, removing confounding.

Figure 2b displays a DAG representing the assumptions made by the weight-gain-for-gestational-age z-score. As gestational age at delivery was used in the transformation to create the z-score there is an arrow between both gestational age and total weight gain and the z-score. Taking the approach of modelling the z-score as the exposure of interest implies that the effects of gestational weight gain and gestational age on neonatal mortality are assumed to be entirely contained in the z-score. The z-score is shown as the exposure of interest to represent its suggested use in epidemiologic studies.[3]

In Figures 2c–d, we build on the DAGs described above to show an additional scenario of interest which includes a confounder, C, of the gestational age at delivery and neonatal mortality relation. Under the scenario that C is measured (e.g. maternal age), an unbiased estimate of the total weight gain and neonatal mortality relation can be achieved by adjusting for C. However, if C is unmeasured (e.g. genetics), under the scenario in Figure 2c, collider bias may be induced when conditioning on gestational age at delivery,[8] while under the z-score model in Figure 2d, C remains a confounder of the z-score neonatal mortality relation. Both scenarios displayed in Figures 2c–d result in open paths that can lead to biased estimates when C is unmeasured, where the magnitude of the bias depends on the strength of the relationship between C, gestational age at delivery and neonatal mortality, as shown algebraically in Appendix A. It is important to note that collider bias generally tends to be smaller in magnitude than confounding bias,[9,10] but this tendency has yet to be evaluated under these causal structures, which is beyond the scope of this paper.

Methods

We utilised data from the Consortium on Safe Labor (CSL) to compare effect estimates from models that use total gestational weight gain with adjustment for gestational age, to models that apply the weight-gain-for-gestational-age z-score to assess the total effect of total weight gain on neonatal mortality. The CSL was comprised of 12 US hospital-based sites (2002–2008).[11] Data were extracted from maternal and neonatal electronic medical records. All study procedures were reviewed and approved by each participating siteÕs Institutional Review Board.

For this analysis, we utilised data from the Þrst singleton birth with information available on prepregnancy weight, height, gestational age, birthweight, and neonatal mortality (n = 121 922). We limited our analysis to normal weight mothers (prepregnancy body mass index (kg/m^2) of 18.5–24.9; (53.9%), to avoid the potential interaction with prepregnancy weight status, and to mothers who delivered between 24 and 40 weeks (n = 65 669). We limited the data to deliveries ≤40 weeks as the weight-gain-for-gestational-age z-score chart stops at 40 weeks.[3]

Total gestational weight gain was calculated as the difference between a motherÕs prepregnancy weight as recorded on her medical chart and her weight the linear models by study site because the underlying distribution of weight gain and gestational age may vary across site and this allowed us to assess the z-score related assumption of correct speciÞcation of the underlying the relationship between gestational

Table 1. Sample characteristics by study site (n = 65 643)

Characteristic	Site 1 (n = 2369) %	2 (n = 9719) %	3 (n = 2755) %	4 (n = 2849) %	5 (n = 24 129) %	6 (n = 2711) %	7 (n = 3486) %	8 (n = 1649) %	9 (n = 1667) %	10 (n = 2735) %	11 (n = 7434) %	12 (n = 4140) %
Age[a]	27.4 (6.6)	32.2 (5.7)	28.2 (6.2)	24.7 (5.8)	27.0 (5.1)	27.4 (5.6)	27.0 (6.8)	24.6 (6.2)	26.1 (6.4)	25.2 (6.1)	27.2 (6.7)	25.8 (6.3)
Race-ethnicity												
White	57.1	59.3	59.9	62.1	56.1	81.5	68.9	46.0	43.4	68.7	11.7	5.1
Black	15.2	8.0	11.2	19.4	29.4	0.5	4.2	35.2	38.3	24.6	47.1	24.9
Hispanic	17.7	19.2	15.7	9.8	10.8	10.0	7.4	2.5	13.3	2.2	31.6	58.0
Asian/ PaciÞc Islander	4.5	2.5	10.8	6.9	2.0	1.8	15.3	3.4	3.0	2.8	4.3	1.7
Other/ Unknown	5.6	11.1	2.4	1.8	1.7	6.2	4.3	12.9	2.1	1.8	5.2	10.3
Nulliparous	46.7	57.7	55.3	41.1	43.5	40.2	43.9	53.6	45.9	44.5	52.9	47.8
Married	66.8	43.0	86.4	64.3	34.7	84.8	84.2	42.4	29.4	47.2	28.2	45.0
Insurance												
Private	55.5	80.8	37.6	69.6	28.8	79.0	93.7	65.0	24.7	43.3	31.9	5.9
Public/Self pay	27.7	18.5	0.3	30.4	64.7	21.0	5.8	31.9	71.7	56.4	68.2	33.3
Other/Unknown	16.8	0.6	62.1	0.0	6.5	0.0	0.5	3.2	3.6	0.3	0.0	60.9
Smoked prior	5.8	10.7	1.7	11.0	20.0	3.7	2.4	0.2	19.0	19.3	12.1	4.2
Gestational weight gain												
Total, kg[a]	15.8 (5.7)	15.6 (5.4)	15.6 (5.7)	15.8 (6.5)	15.0 (5.3)	14.3 (5.0)	15.8 (6.0)	15.2 (6.2)	16.2 (6.3)	15.2 (6.5)	14.6 (6.1)	15.2 (6.3)
z-score[a,b]	−0.01 (1.00)	−0.11 (0.96)	−0.06 (1.07)	−0.13 (1.20)	−0.23 (1.02)	−0.32 (0.94)	−0.09 (1.16)	−0.22 (1.16)	−0.02 (1.16)	−0.22 (1.21)	−0.23 (1.15)	−0.18 (1.19)
Gestational age at delivery, wks [a]	38.3 (1.9)	39.0 (1.6)	38.5 (1.8)	38. 8 (2.0)	38.7 (1.6)	38.8 (1.5)	38.6 (2.1)	38.6 (2.0)	38.5 (2.2)	38.6 (2.2)	37.9 (2.5)	38.1 (2.5)

[a]Values represent mean (standard deviation).

[b]z-score estimated from published charts. Hutcheon et al.[3]

Table 2. Overall simulation results: total gestational weight gain and the risk for neonatal mortality estimated using a continuous measure and quintiles of total weight gain, unadjusted and adjusted for gestational age at delivery and using the weight-gain-for-gestational-age z-score

	GWG			GWG adjusted for GA			z-score		
Model	Mean RR[a]	Standard error	95% CI coverage	Mean RR[a]	Standard error	95% CI coverage	Mean RR[a]	Standard error	95% CI coverage
Linear[b]	0.87	0.02	0%	1.00	0.02	94%	0.97	0.08	84%
Quintiles[c]									
1	Reference			Reference			Reference		
2	0.30	0.26	0%	1.03	0.28	94%	0.78	0.26	84%
3	0.24	0.29	0%	1.02	0.30	95%	0.64	0.28	63%
4	0.21	0.32	0%	0.97	0.33	95%	0.76	0.26	83%
5	0.20	0.33	0%	1.01	0.34	95%	0.90	0.24	94%

CI, conÞdence interval; GA, gestational age at delivery; GWG, gestational weight gain; RR, mean relative risk across 5000 replicates.
[a]Expected RR based on simulations was 1.0. Simulations were repeated 5000 times.
[b]Linear estimates for GWG are per kg increase in weight gain, while z-score estimates are per z-score unit increase.
[c]Gestational weight gain quintile cut points (kg): 1, −9.1–10.9; 2, 10.9–13.6; 3, 13.6–15.9; 4, 15.9–19.5; 5, 19.5–49.0.

Table 3. Site speciÞc simulation results: total gestational weight gain (GWG) and the risk for neonatal mortality estimated using a continuous measure of total weight gain, unadjusted and adjusted for gestational age at delivery and using the weight-gain-for-gestational-age z-score

		GWG			GWG adjusted for GA			z-score		
Site	n	Mean RR[a]	Standard error	95% CI coverage	Mean RR[a]	Standard error	95% CI coverage	Mean RR[a]	Standard error	95% CI coverage
1	2369	0.78	0.27	47%	0.98	0.18	95%	0.66	0.59	70%
2	9719	0.89	0.08	44%	1.00	0.06	96%	1.02	0.31	92%
3	2755	0.87	0.12	55%	0.98	0.22	94%	0.93	0.50	73%
4	2849	0.87	0.09	47%	0.99	0.11	95%	0.87	0.32	84%
5	24 129	0.88	0.05	12%	1.00	0.04	94%	0.96	0.13	86%
6	2711	0.91	0.17	89%	0.98	0.18	95%	1.43	0.72	80%
7	3486	0.87	0.07	35%	0.99	0.07	94%	1.03	0.23	93%
8	1649	0.89	0.11	75%	0.97	0.16	95%	1.03	0.44	92%
9	1667	0.90	0.08	70%	0.98	0.15	94%	1.05	0.42	90%
10	2735	0.90	0.09	61%	0.99	0.08	93%	1.01	0.39	87%
11	7434	0.88	0.04	8%	1.00	0.04	94%	1.10	0.18	88%
12	4140	0.88	0.05	16%	1.00	0.05	95%	1.01	0.21	94%

CI, conÞdence interval; GA, gestational age at delivery; GWG, gestational weight gain; RR, mean relative risk across 5000 replicates.
[a]Estimates for GWG are per kg increase in weight gain, while z-score estimates are per z-score unit increase. Expected RR based on simulations is 1.0. Simulations were repeated 5000 times.

for how we implement this variable in epidemiological models. From the DAG presented, it is clear that total gestational weight gain is a time-varying exposure, although it is often treated as a Þxed variable. We show that although there is feedback between gestational age and weight gain across time, when limited to studies of total gestational weight gain and assessing the association with neonatal mortality, adjusting for gestational age blocks the open backdoor path between these prior longitudinal relationships. This Þnding is supported by our simulated model of no true association between weight gain and neonatal mortality, where we demonstrated that efÞcient unbiased estimates can be achieved by adjusting for gestational age.

Adjusting for gestational age at delivery in studies of total gestational weight gain has previously been of debate.[5–7] Using simulated data, we found that when adjusted for gestational age, total weight gain can independent of gestational age. The regression-based

HHSN267200603425C). The authors would like to thank Drs. Nansi S. Boghossian, Jennifer A. Hutcheon, Robert W. Platt and Lindsey A. Sjaarda for their helpful suggestions to earlier versions of the manuscript. The authors have no conßicts of interest.

References

1 Institute of Medicine and National Research Council. Weight Gain During Pregnancy: Reexamining the Guidelines. Washington, DC: National Academies Press, 2009.

2 Hutcheon JA, Bodnar LM, Joseph KS, Abrams B, Simhan HN, Platt RW. The bias in current measures of gestational weight gain. Paediatric and Perinatal Epidemiology 2012; 26:109–16.

3 Hutcheon JA, Platt RW, Abrams B, Himes KP, Simhan HN, Bodnar LM. A weight-gain-for-gestational-age z score chart for the assessment of maternal weight gain in pregnancy. The American Journal of Clinical Nutrition 2013; 97:1062–7.

4 Hutcheon JA, Platt RW, Abrams B, Himes KP, Simhan HN, Bodnar LM. Pregnancy weight gain charts for obese and overweight women. Obesity (Silver Spring) 2015; 23:532–5.

5 Davis RR, Hofferth SL, Shenassa ED. Gestational weight gain and risk of infant death in the United States. American Journal of Public Health 2014; 104:S90–5.

6 Davis RR, Hofferth SL, Shenassa ED. Davis et al. respond. American Journal of Public Health 2014;104:e2.

7 Hutcheon JA, Bodnar LM, Abrams B. Untangling gestational weight gain from gestational age in infant mortality studies. American Journal of Public Health 2014; 104:e1–2.

8 VanderWeele TJ, Mumford SL, Schisterman EF. Conditioning on intermediates in perinatal epidemiology. Epidemiology 2012; 23:1–9.

9 Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratiÞcation bias. Epidemiology 2003; 14:300–6.

10 Whitcomb BW, Schisterman EF, Perkins NJ, Platt RW. QuantiÞcation of collider-stratiÞcation bias and the birthweight paradox. Paediatric and Perinatal Epidemiology 2009; 23:394–402.

11 Zhang J, Troendle J, Reddy UM, Laughon SK, Branch DW, Burkman R, et al. Contemporary cesarean delivery practice in the United States. American Journal of Obstetrics and Gynecology 2010;203:326.e1-.e10.

12 Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins, 2008.

13 Allison DB, Paultre F, Goran MI, Poehlman ET, HeymsÞeld SB. Statistical considerations regarding the use of ratios to adjust data. International Journal of Obesity and Related Metabolic Disorders 1995; 19:644–52.

14 Delbaere I, Vansteelandt S, De Bacquer D, Verstraelen H, Gerris J, De Sutter P, et al. Should we adjust for gestational age when analysing birth weights? The use of z-scores revisited. Human Reproduction 2007; 22:2080–3.

15 Johnson JL, Farr SL, Dietz PM, Sharma AJ, BarÞeld WD, Robbins CL. Trends in gestational weight gain: the Pregnancy Risk Assessment Monitoring System, 2000-2009. American Journal of Obstetrics and Gynecology 2015; 28:00086–1.

16 Dibley MJ, Staehling N, Nieburg P, Trowbridge FL. Interpretation of Z-score anthropometric indicators derived from the international growth reference. American Journal of Clinical Nutrition 1987; 46:749–62.

17 Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000; 11:550–60.

18 Mitchell EM, Hinkle SN, Schisterm EF. ItÕs About Time: a survival approach to gestational weight gain and preterm delivery. Epidemiology 2016; 27:182–7.

19 Hill B, Skouteris H, Fuller-Tyszkiewicz M. Interventions designed to limit gestational weight gain: a systematic review of theory and meta-analysis of intervention components. Obesity Reviews 2013; 14:435–50.

Appendix A

Consider a z-score chart where z-scores are calculated by relating gestational weight gain (GWG) during pregnancy to gestational age (t) through a linear model such that:

$$\mathrm{GWG}(t) = \psi_0 + \psi_1 t + \psi_2 t^2 + \varepsilon_1, \qquad 1$$

where the ε_1′s are identically distributed with mean 0 and variance σ^2. The variable t is chosen to represent gestational age to emphasise the equivalence of gestational age with time, and GWG(t) denotes that GWG is a function of time. Then the z-score calculated from this model is equivalent to the standardised residuals:

$$Z = \frac{\mathrm{GWG}(t) - \hat{E}(\mathrm{GWG}|t)}{\sqrt{\widehat{\mathrm{Var}}(\mathrm{GWG}|t)}} = \frac{\varepsilon}{\sigma}.$$

The goal of using a z-score is to provide a marker of relative gestational weight gain that is independent of gestational age. This method will be successful so long as the assumptions in (1) are met, namely, that ε_1 is independent of gestational age with mean 0.

Suppose, however, that model (1) is misspeciÞed, and in fact, the true model for GWG includes an interaction between gestational age (t) and pre-pregnancy body mass index (BMI), such that:

$$\mathrm{GWG}(t) = \psi_0 + \psi_1 t + \psi_2 t^2 + \psi_3(\mathrm{BMI} * t) + \varepsilon_2.$$

Then Þtting model (1) will violate the assumption of independence between the errors and t, since the error terms are now inclusive of the interaction term:

$$\varepsilon_1 = \psi_3(\mathrm{BMI} * t) + \varepsilon_2.$$

which is clearly correlated with t. Thus, z-scores calculated based on fitting model (1) will not succeed in removing the association between the z-score and gestational age. Subsequently, z-scores based on total gestational weight gain and gestational age at delivery will fail to adequately remove these correlations. This performance is not only hindered by the omission of interaction terms of gestational age, but could also be enforced by failing to include a variable that acts as a confounder between gestational age and GWG, or between gestational age and the outcomes of interests (e.g. neonatal mortality). For instance, if gestational age is caused by some unmeasured variable C, failure to incorporate C into Model (1) will result in a z-score that remains correlated with C. Subsequently, if C is a confounder of the relationship between gestational age and neonatal mortality, C will remain a confounder of the z-score and neonatal mortality. Essentially, the model used to calculate the z-score chart must correctly specify the relationship between gestational age and GWG in order for the z-scores to perform as expected.

Now consider the logistic regression model of interest relating the z-score to the binary outcome of interest (Y), where total GWG (GWG_{tot}) is measured at delivery ($t = GA_{del}$), and:

$$\mathrm{logit}\left[\Pr(Y = 1|Z)\right] = \alpha + Z\beta = \alpha + \left(\frac{GWG_{tot} - \hat{E}(GWG_{tot}|GA_{del})}{\sqrt{\widehat{\mathrm{Var}}\left(GWG_{tot}|GA_{del}\right)}} \right)\beta.$$

For simplicity, let us suppose that $\mathrm{Var}(GWG_{tot}|GA_{del}) =$