Study Protocol

Title: Optimizing Gestational Weight Gain in Obese Pregnant Women: A Nationwide Real-World Study Using the U.S. Natality Cohort

Protocol Version: v1.0

Date: 26 August 2026

Study Type: Real-World Observational Cohort Study

Data Source: U.S. CDC NCHS National Vital Statistics System (NVSS) Birth Data (2023–2025)

NCT Number:

Reporting Guidelines: STROBE and RECORD-PE [1, 2]

1. Background and Rationale

Pre-pregnancy obesity (Body Mass Index [BMI] ≥ 30.0 kg/m²) is a formidable challenge in modern obstetrics, significantly increasing the risk of adverse maternal and neonatal outcomes. To mitigate these morbidities, the 2009 Institute of Medicine (IOM) guidelines recommend a uniform gestational weight gain (GWG) target of 5 to 9 kg for all obese pregnant individuals.

However, this monolithic recommendation fails to account for the profound metabolic heterogeneity across different obesity classes (Class I through III). Observational data yield conflicting results on whether highly restrictive weight gain—or intentional weight loss (< 0 kg)—confers a net clinical benefit for the mother or paradoxically increases the risk of small-for-gestational-age (SGA) neonates. This study aims to utilize a massive, contemporary U.S. cohort to mathematically derive and validate class-specific optimal GWG ranges that safely minimize competing maternal and neonatal morbidities.

2. Study Objectives

Primary Objective: To delineate the continuous, non-linear dose-response relationship between GWG and composite adverse outcomes across distinct obesity classes, and to mathematically derive the optimal GWG ranges that balance competing maternal and neonatal risks. **Secondary Objectives:**

1. To evaluate the independent associations between GWG and specific maternal/neonatal outcomes categorized by severity (3-tier grading system).
2. To assess whether extreme GWG intervals, specifically weight loss (< 0 kg), confer any net clinical benefit.

3. Study Design and Data Source

This is a population-based retrospective cohort study using data from the U.S. National Vital Statistics System (NVSS) natality files (2022–2024). As the dataset is fully de-identified and publicly available, this study is exempt from institutional review board (IRB) approval.

4. Eligibility Criteria

Inclusion Criteria:

- Nulliparous and multiparous individuals aged 18 to 45 years.

- Singleton, live-born, term infants (37 to 41 completed weeks of gestation).
- Pre-pregnancy obesity (BMI ≥ 30.0 kg/m²).
- **Initiation of prenatal care strictly during the first trimester (months 1–3 of pregnancy)**.

Exclusion Criteria:

- Missing data on pre-pregnancy or delivery weights, or unknown birth weights.
- Implausible gestational weight gain records (GWG < -10 kg or > 50 kg).
- Unknown mode of delivery or individuals with a documented history of prior cesarean sections (to restrict the focus to primary cesarean outcomes).

The final cohort is stratified into WHO obesity classes: Class I (30.0–34.9 kg/m²), Class II (35.0–39.9 kg/m²), and Class III (≥ 40.0 kg/m²).

5. Exposure Definition and Prespecified Categorization

The primary exposure is total gestational weight gain (GWG), calculated as the exact difference between maternal weight at delivery and pre-pregnancy weight. Continuous GWG is categorically partitioned into **eight distinct clinical intervals**. Importantly, the interval representing the IOM guideline is designated as a fully closed mathematical range:

1. < 0 kg (Weight loss)
2. 0 to 2.9 kg
3. 3.0 to 4.9 kg
4. **[5.0, 9.0] kg (Prespecified Reference Category):** This closed interval inclusive of both endpoints (5.0 to 9.0 kg) accurately encompasses the 2009 IOM recommended target for obese individuals.
5. 9.1 to 10.9 kg
6. 11.0 to 12.9 kg
7. 13.0 to 15.9 kg
8. ≥ 16.0 kg (Extreme excessive gain)

(Note: GWG rate is not calculated or analyzed in this protocol).

6. Outcome Definitions and the 3-Tier Severity Grading System

The primary endpoints are two comprehensive composite indicators (M1_equal and M2_equal) reflecting adverse maternal and neonatal events, defined as the occurrence of *any* of their respective subcomponents. These outcomes are further strictly stratified into a **3-tier severity grading system**.

6.1 Maternal Composite Outcome (M1_equal):

- **Tier 1 (Severe/Life-threatening):** Intensive care unit (ICU) admission, maternal blood transfusion, or eclampsia. These criteria universally align with the CDC Severe Maternal Morbidity (SMM) indices, the ACOG/SMFM joint consensus on obstetric morbidity, and the WHO "Maternal Near Miss" framework [3-5].

- **Tier 2 (Moderate):** Gestational hypertension, gestational diabetes mellitus (GDM), primary cesarean delivery, or severe perineal laceration without Tier 1 events. These constitute significant adverse outcomes requiring intervention per ACOG guidelines but do not inherently denote terminal systemic failures [6, 7].

- **Tier 0:** Uncomplicated.

6.2 Neonatal Composite Outcome (M2_equal):

- **Tier 1 (Severe/Life-threatening):** Neonatal ICU (NICU) admission, prolonged assisted ventilation (> 6 hours), or neonatal seizures. These align with the WHO "Neonatal Near-Miss" criteria for severe multi-organ dysfunction and the AAP levels of severe neonatal care [8, 9].

- **Tier 2 (Moderate):** Small for gestational age (SGA, < 10th percentile), large for gestational age (LGA, > 90th percentile) based on the **2017 US Reference for Singleton Birth Weight Percentiles Using Obstetric Estimates of Gestation** [10], immediate assisted ventilation, or 5-minute Apgar score < 7 [11].

- **Tier 0:** Uncomplicated.

7. Statistical Analysis

All statistical analyses are conducted independently for each obesity class.

7.1 Multivariable Regression and Missing Data Strategy We employ multivariable logistic regression models to evaluate the independent associations between the 8-category GWG intervals and the composite outcomes (M1_equal and M2_equal). Models are adjusted for maternal age, race/ethnicity, education, marital status, primary payment source, WIC participation, parity, pre-pregnancy diabetes, chronic

hypertension, history of infertility treatment, and smoking patterns. **Handling of Missing Data:** In the primary analysis, missing covariate data are rigorously addressed by retaining "Unknown" as an independent analytical category. This approach prevents arbitrary data deletion, preserves maximum statistical power, and avoids introducing selection bias inherent to complete-case exclusions in real-world data. Conversely, in sensitivity analyses, Multiple Imputation by Chained Equations (MICE, m=20) is exclusively utilized for covariates with missingness rates exceeding 10% (e.g., marital status).

7.2 Non-linear Dose-Response and Optimal Range Derivation To precisely characterize the continuous dose-response trajectory, restricted cubic spline (RCS) functions are incorporated with five knots placed exactly at the 5th, 27.5th, 50th, 72.5th, and 95th percentiles of the GWG distribution. Based on the RCS curves for M1_equal and M2_equal, we estimate the nadirs, defining a "safe threshold" as the GWG interval where the relative risk increases by ≤ 10% compared to the respective nadir (aOR ≤ 1.10). The final recommended GWG range is determined by the mathematical intersection of the safe thresholds for both M1_equal and M2_equal. We perform 1,000 bootstrap resamples to calculate the 95% CIs for the nadirs and corresponding bounds.

7.3 Hierarchical Gatekeeping Strategy for Multiplicity Control To rigorously control the family-wise error rate across multiple hypotheses, a predefined hierarchical gatekeeping strategy leveraging sequential Likelihood Ratio Tests (LRT) is implemented [12]:

- **Layer 1 (Global Interaction Test):** We assess the GWG × BMI interaction using an omnibus LRT. An interaction $P < 0.05$ justifies the subsequent BMI-stratified analysis.
- **Layer 2 (Within-Group Fixed Sequence):** Within each BMI subgroup, formal hypothesis testing proceeds in a strict sequence using omnibus LRTs to verify overall model improvement ($\alpha = 0.05$):
    - *Step 1:* Test binary M1_equal via Logistic LRT. If $P < 0.05$, proceed to
    - *Step 2:* Test binary M2_equal via Logistic LRT. If $P < 0.05$, proceed to
    - *Step 3:* Test graded M1_grade via Ordinal Logistic LRT. If $P < 0.05$, proceed to Step 4.
    - *Step 4:* Test graded M2_grade via Ordinal Logistic LRT.
- **Layer 3 (Inter-Group Correction):** A Bonferroni correction ($\alpha \approx 0.0167$) is strictly applied across the 3 BMI subgroups for the sequence testing.
- **Exploratory Outcomes:** For secondary individual outcome components, P-values are adjusted using the Benjamini-Hochberg procedure, with a False

Discovery Rate (FDR) q-value threshold of < 0.05 [13]. Note: For graded outcomes, if the proportional odds assumption is violated (Brant test $P < 0.05$), parameter estimation correctly downgrades to multinomial logistic regression to calculate tier-specific aORs.

7.4 Model Diagnostics To confirm structural integrity, the following diagnostics are performed:

1. Hosmer-Lemeshow Test (based on 10 deciles) for goodness-of-fit.
2. AUC and Brier Score for discrimination and calibration.
3. Variance Inflation Factor (VIF) for multi-collinearity.
4. Cook's Distance and DFBETAs for influential data points.

7.5 Sensitivity Analyses

1. **Pure Obesity Cohort:** Restricting to women without pre-pregnancy diabetes and chronic hypertension.
2. **Missing Data Check:** As described in 7.1, applying Complete Case Analysis and MICE (for missing >10%).
3. **E-value Assessment:** Calculating E-values for unmeasured confounding [14].
4. **Full Population Verification:** Expanding the cohort to include individuals with prior cesarean sections, adjusting for this history, to confirm that primary exclusion criteria did not introduce critical selection bias.

7.6 Subgroup Analyses Effect modifications are explored across: Race/Ethnicity, Pre-pregnancy Diabetes, Chronic Hypertension, Maternal Age (< 35 vs. ≥ 35 years), and Parity.

References

[1] von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement. *Ann Intern Med*. 2007;147(8):573-577.

[2] Langan SM, Schmidt SA, Wing K, et al. The reporting of studies conducted using observational routinely-collected health data (RECORD-PE). *BMJ*. 2018;363:k3532.

[3] American College of Obstetricians and Gynecologists and the Society for Maternal-Fetal Medicine (ACOG/SMFM), Kilpatrick SK, Ecker JL. Severe maternal morbidity: screening and review. *Am J Obstet Gynecol*. 2016;215(3):B17-B22.

[4] Kern-Goldberger AR, Howell EA, Srinivas SK, Levine LD. What we talk about when we talk about severe maternal morbidity. *Am J Obstet Gynecol MFM*. 2023;5(5):100882.

[5] Say L, Souza JP, Pattinson RC. Maternal near miss--towards a standard tool for monitoring quality of maternal health care. *Best Pract Res Clin Obstet Gynaecol*. 2009;23(3):287-296.

[6] ACOG. Practice Bulletin No. 202: Gestational Hypertension and Preeclampsia. *Obstet Gynecol*. 2019;133(1):e1-e25.

[7] Lailler G, et al. Adverse Maternal and Infant Outcomes in Women With Chronic Hypertension in France. *J Am Heart Assoc*. 2023;12(5):e027266.

[8] Pileggi-Castro C, Camelo Jr JS, Perdoná GC, et al. Development of criteria for identifying neonatal near-miss cases. *Bull World Health Organ*. 2014;92(8):582-590.

[9] American Academy of Pediatrics Committee on Fetus and Newborn. Levels of neonatal care. *Pediatrics*. 2012;130(3):587-597.

[10] Talge NM, Mudd LM, Sikorskii A, Basso O. United States Birth Weight Reference Corrected for Exact Gestational Age. Pediatrics. 2014;133(5):844-853. (Updated/Validated by: Katz J, et al. A 2017 US Reference for Singleton Birth Weight Percentiles Using Obstetric Estimates of Gestation. Pediatrics. 2019;144(2):e20190184).

[11] ACOG, AAP. Committee Opinion No. 816: The Apgar Score. *Obstet Gynecol*. 2021;137(1):e1-e12.

[12] Dmitrienko A, Tamhane AC, Wiens BL. General multistage gatekeeping procedures. *Biom J*. 2008;50(5):667-677.

[13] Benjamini Y, Hochberg Y. Controlling the false discovery rate. *J R Stat Soc B*. 1995;57(1):289-300.

[14] VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann Intern Med*. 2017;167(4):268-274.