Discover Health Systems

Discover

Research

Validation of gestational weight gain records on South Carolina birth certificate data

Oluwatosin A. Momodu[1] · **Ronnie D. Horner**[2] · **Jihong Liu**[3] · **Elizabeth Crouch**[4,5] · **Brian Chen**[1]

Received: 6 August 2023 / Accepted: 27 August 2023
Published online: 19 September 2023

Abstract

Objective To examine the accuracy of gestational weight gain (GWG) reported on birth certificates (BCs) in comparison with electronic medical records (EMRs), the gold standard.

Methods BC data and EMR data were from a random sample of pregnant women who enrolled in CenteringPregnancy program, a type of group-based prenatal care, at three obstetric clinics in South Carolina between 2015 and 2019 (n = 206). Retrospective review of EMR data on patients' prepregnancy BMI and GWG according to 2009 Institute of Medicine guidelines was conducted. Analyses involved summarizing the variables by their mean and mean differences per data source, and then calculating percent-weighted agreement and kappa statistics.

Results The mean values of BMI, delivery weight and total weight gain were similar between BC and EMRs. Data correlation for variables was high for both data sources (height: r = 0.94, prepregnancy weight: r = 0.93, prepregnancy BMI: r = 0.92, delivery weight: r = 0.96, total weight gain: r = 0.60). The BCs slightly underestimated the proportion of women in the normal-weight BMI category but overestimated the proportion in the overweight BMI category. Additionally, BCs slightly overestimated women with inadequate GWG and underestimated those with excessive GWG. Overall, the BC and EMR data were in agreement regarding prepregnancy BMI (weighted-agreement = 90%, Kappa = 0.78) and GWG categories (weighted-agreement = 84%, Kappa = 0.63).

Conclusion BC estimates of prepregnancy BMI and GWG categories were similar to those recorded in the EMRs. The South Carolina BC database is a valid database for gestational weight and can provide reasonable estimates for the state in the evaluation of the CenteringPregnancy program.

The research was conducted at University of South Carolina, Arnold School of Public Health, Department of Health Services Policy and Management, Columbia, South Carolina, 29208, USA.

✉ Oluwatosin A. Momodu, drtosintaiwo@yahoo.com; Ronnie D. Horner, rhorner@unmc.edu; Jihong Liu, jliu@mailbox.sc.edu; Elizabeth Crouch, crouchel@mailbox.sc.edu; Brian Chen, bchen@mailbox.sc.edu | [1]Department of Health Services Policy and Management, Arnold School of Public Health, University of South Carolina, 915 Greene Street, Discovery Building, Columbia, SC 29208, USA. [2]Department of Health Services Research and Administration | College of Public Health, University of Nebraska Medical Center, 984350 Nebraska Medical Center, Omaha, NE 68198-4350, USA. [3]Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, 915 Greene Street, Discovery Building 459, Columbia, SC 29208, USA. [4]Department of Health Services Policy and Management, University of South Carolina, Discovery Building, Suite 345, Columbia, SC 29208, USA. [5]University of South Carolina, Rural and Minority Health Research Center, 220 Stoneridge Drive, Suite 204, Columbia, SC, 29201, USA.

1 Introduction

The validation of gestational weight gain (GWG) is of particular interest to maternal and child health researchers because weight gain within the recommended levels or range according to the Institute of Medicine (IOM) guidelines has favorable pregnancy outcomes, particularly regarding infant birth weight [1]. The IOM recommends that assessment of adequate GWG should be based on the mother's prepregnancy body mass index (BMI). Excessive and inadequate weight gain in pregnancy is often associated with clinical conditions in mothers (and infants as well). For example, excessive GWG is associated with an increased risk of pregnancy-induced hypertension, gestational diabetes, complications during labor and delivery, nonelective cesarean section delivery, postdelivery weight retention and subsequent maternal obesity, while inadequate GWG has a greater effect on the unborn infant and may likely result in poor fetal development, intrauterine growth restriction (IUGR), prematurity, low birth weight and increased risk for small for gestational age (SGA), as highlighted by the American College of Obstetrics and Gynecology (ACOG) [2, 3].

Birth certificate data are often used in studies of reproductive outcomes because they are a readily accessible data source. Investigators [4, 5] of prenatal interventions have been particularly concerned about the quality of GWG data reported on birth certificates because comparisons of new programs such as the CenteringPregnancy program versus standard prenatal care may be affected. Indeed, inaccurate GWG data on birth certificates has been identified as a likely cause of the mixed or inconsistent findings regarding the merit of the CenteringPregnancy program (i.e., group-based prenatal visits) versus standard prenatal care. These investigators have recommended that a validation study be conducted to determine the quality of GWG reports on birth certificates [4–6].

Previous studies validating maternal data from birth certificates, particularly maternal pregnancy weight, height and GWG using EMRs as the gold standard, are quite few in number [7–9] and were carried out in different populations and states (Florida, Pennsylvania, New York and Vermont); findings from these studies are mixed or inconsistent. The most comprehensive evaluation of 2003 birth certificate data was conducted by the National Center for Health Statistics (NCHS) in a total of eight hospitals in two states, and wide variation in the quality of data was found by item and hospital [10]. This comprehensive study did not assess GWG or BMI. Clearly, the use of birth certificate data for maternal and child health studies requires an assessment of the quality of the data for each state because of inconsistent data results from previous studies in different states.

As a prelude to conducting an evaluation of a diet-enhanced version of the CenteringPregnancy (CP) program—a type of group-based prenatal care—based in the Midlands of South Carolina, we examined the validity of pre-pregnancy BMI and GWG on birth certificate data among women in the state who participated in the program to ensure an accurate assessment or evaluation of the prenatal program. The study focused on CenteringPregnancy program participants since the women in the program and those receiving standard prenatal care were sufficiently similar regarding sociodemographic characteristics, and the result from one source in this case can be applied to the other. The South Carolina standard certificate of live birth serves as a legal document and a national and state data source for monitoring maternal and infant health. The Division of Biostatistics at the Department of Health and Environmental Control registers births and completes the items on birth certificates. Even though the division ensures quality control of the statistical processing and dissemination of vital statistics, there may still be errors in the information pertaining to gestational weight gain in electronic medical records (EMRs), which are completed by medically trained professionals or paramedics who abstract the data from medical charts. These factors include the volume of information in medical records, the abstractor's knowledge of the topic, which is related to abstractor credentials/training, inadequate time for abstraction tasks, unavailability of abstraction tools, and incorrect prepregnancy weight data because the information can be found in multiple places in the medical records, leading to inaccurate prepregnancy body mass index (BMI) data [11]. In this report, we present the findings on the validity of data (prepregnancy body mass index and gestational weight gain) in birth certificates vis-à-vis an EMR chart review.

2 Methods

This study uses data of women who registered early in their pregnancy (i.e., by 20 weeks of gestation) into the CenteringPregnancy program for prenatal care in three obstetric clinics in the midlands of South Carolina between 2015 and 2019.

2.1 Sampling strategy, sample size and data collection

From the raw list of 804 participating mothers from the three sites, 53 did not have a social security number (SSN) or date of birth, while an additional 83 did not have an SSN; because these data were needed for proper linkage to the birth certificate database, these women were excluded from the study. The study used stratified random sampling to draw a subsample from the remaining 668 women who participated in the CenteringPregnancy program. Thirty percent (206) of the 668 women across the three sites were selected, with a sampling ratio of 0.50 (50%) for each of the two sites that had started the CenteringPregnancy program early (2015) and 1.0 (100%) for the site that started the program 2 years later (2018). Different sampling ratios were used to adjust for differences among the sites in the characteristics of the patients associated with birth outcomes ("Appendix"). For example, the sites with a 50% sampling ratio had more obese women compared to the small site that had a higher sampling ratio, which might influence the outcome.

A Health Insurance Portability and Accountability Act (HIPAA)-compliant medical abstraction form was used to guide and facilitate the abstraction of patients' information related to the study items—prepregnancy BMI and GWG. A retrospective chart review of EMR is considered the gold standard for the validation of the quality of GWG data recorded on birth certificates [6, 8, 12, 13]. The abstraction form was designed to follow the format of the EMR so that abstraction would be accurate and efficient. The abstraction protocol involved manually searching through a patient's EMR with their identifiers (first name, last name, date of birth and social security number) to abstract data on the maternal measures, such as prepregnancy BMI, gestational age, height, and weight at last prenatal care or delivery, to determine GWG. Other maternal- and pregnancy-related characteristics or variables (shown in Table 1) that were not included in the variables for validation or accuracy checks were taken from birth certificate data.

The Revenue and Fiscal Affairs Office (RFA) provided the patients' hospital discharge/birth certificate data after receiving the list of CenteringPregnancy participants from the clinic, which included personal identifiers such as first name, last name, social security number (SSN) and date of birth (DOB) for the women (hereafter referred to as participating women) for the period under consideration. These hospital discharge record/birth certificate data were linked with the data abstracted from the EMRs for the analysis process. Approximately 92% of the participating women were successfully matched or linked across data sources. The most likely reasons for failure to link include delivery occurring outside South Carolina, at a birthing center (not hospital) or outside of our study timeframe (2015 and 2019).

2.2 Inclusion and exclusion criteria

In the analysis, we excluded observations where the data had values that were out of the normal range as described by the Centers for Disease Control and Prevention on nutrition, physical activity and obesity and the Behavioral Risk Factor Surveillance System (BRFSS). As shown in the flowchart in Fig. 1, women aged less than 18 years and greater than 49 years were excluded. Additionally, women who had a prepregnancy weight less than 50 pounds or greater than 650 pounds, a height less than 3 feet or greater than or equal to 8 feet or a prepregnancy body mass index (BMI) less than 12 or greater than 100 at 2 months of pregnancy were excluded from the study [14]. Additionally, women whose infants had a birth weight less than 500 g or greater than 5000 g were excluded. We further excluded women who had prepregnancy diabetes because their physicians would have recommended modifying their lifestyle-related factors, which would have impacted their weight gain, unlike for patients with prepregnancy hypertension or chronic hypertension who were included. We also excluded women with missing data for variable information in the birth certificate dataset, as shown in Fig. 1. All these exclusion criteria were formally used in evaluating GWG. Very few participants were excluded based on the criteria.

2.3 Data analysis

The outcome variables included prepregnancy BMI (prepregnancy weight and height) and gestational weight gain (prepregnancy weight and weight at delivery).

Prepregnancy BMI. Abstracted data for height, prepregnancy weight or measured first-trimester weight were used to compute prepregnancy body mass indices (BMIs) for our study sample (173 participating women) and to categorize these women into four groups: underweight (< 18.5 kg/m^2), normal weight (18.5–24.9 kg/m^2), overweight (25.0–29.9 kg/m^2), and obese (≥ 30.0 kg/m^2).

Gestational Weight Gain (GWG). Crude total (gestational) weight gain was calculated for both data sources as the difference in weight prior to pregnancy and at delivery. The women were categorized into groups with inadequate,

Table 1 Sociodemographic and other maternal characteristics of the participating women between 2015 and 2019, N = 173

Maternal characteristics	Subsample data	*P* values
Maternal age (mean; SD)	26.14 ± 5.52	0.171
Maternal age		0.525
< 20	16 (9.25%)	
20–24	60 (34.68%)	
25–29	54 (31.21%)	
30–34	26 (15.03%)	
≥ 35	17 (9.83%)	
Race		0.018[a]
White	51 (29.48%)	
African American	102 (58.96%)	
Asian (Indian, Chinese, Japanese, Hawaiian)	20 (11.56%)	
Level of education		0.005[a]
< High school	13 (5.51%)	
High school graduate	99 (57.23%)	
College graduate	41 (23.70%)	
Postgraduate	20 (11.56%)	
Parity		0.001[a]
No children	136 (78.61%)	
One child	28 (16.18%)	
Two children	6 (3.47%)	
≥ Three children	3 (1.73%)	
Health insurance type		0.001[a]
Private insurance	95 (54.91%)	
Government/medicaid	68 (39.31%)	
Self-pay	4 (2.31%)	
Other type	6 (3.47%)	
Smoking during pregnancy		0.294
No	171 (98.84%)	
Yes	2 (1.16%)	
Prepregnancy BMI (mean; SD)	28.50 ± 7.18	0.866
Prepregnancy BMI		0.619
Underweight	5 (2.89%)	
Normal	53 (30.64%)	
Overweight	54 (31.21%)	
Obese	61 (35.26%)	
Month that PNC began		0.004[a]
0–3 months	138 (79.77%)	
4–6 months	29 (16.76%)	
7–9 months	6 (3.47%)	
Prepregnancy hypertension		0.500
No	168 (97.11%)	
Yes	5 (2.89%)	
Gestational age at delivery (wks) (mean; SD)	38.31 ± 1.70	0.707
Gestational age at delivery		0.642
Preterm (≤ 37 wks)	17 (9.83%)	
Term (> 37 wks)	156 (90.17%)	
Total weight gain at delivery (kg) (mean; SD)	11.61 ± 8.64	0.063
Total weight gain categories at delivery according to the IOM recommendation		0.123
Inadequate	58 (33.53%)	
Adequate	42 (24.28%)	
Excess	73 (42.20%)	

Table 1 (continued)

P values indicate a ≤ 0.05 threshold or significance level, indicating evidence of a difference between birth certificate data and subsample data

[a]Indicates significant values

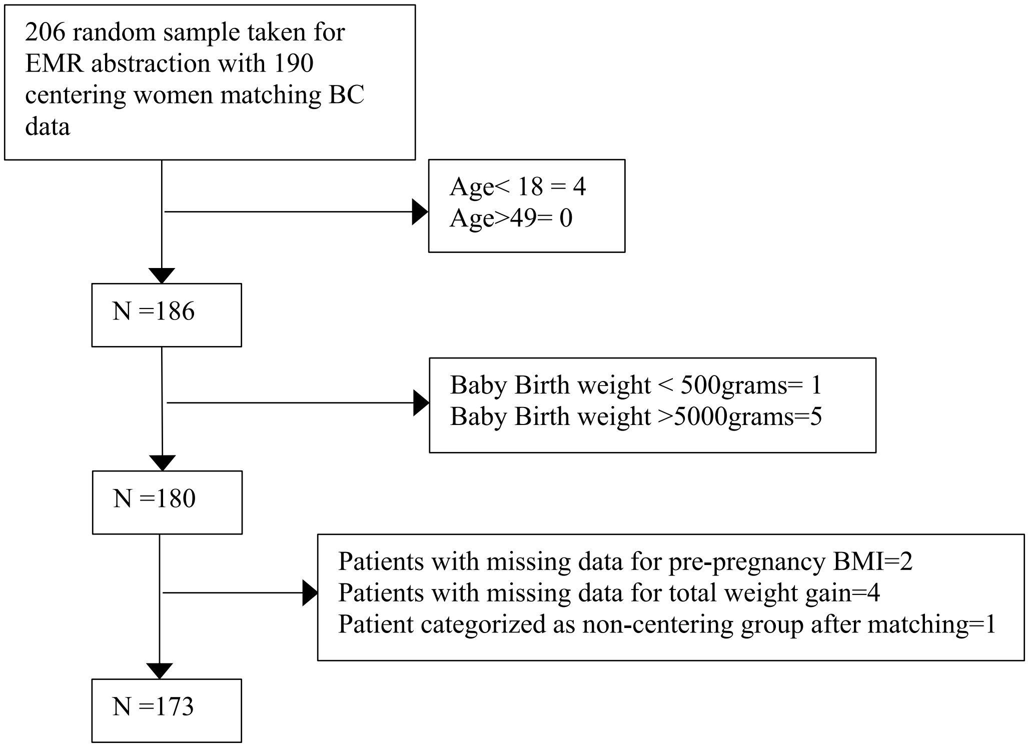

Fig. 1 Flow chart of inclusion/exclusion criteria for subset of centering women for validation study

adequate, and excessive GWG according to the IOM guideline that specify the amount of acceptable weight gain based on the maternal prepregnancy BMI. The weekly rate of gestational weight gain, as an alternative measure of weight gain, was used in previous publications considering that total weight gain varies by pregnancy duration and was calculated as follows: [(total weight gain – expected first trimester gestational weight gain)/(gestational age at birth in weeks – 13 weeks)] [4, 5].

The sociodemographic characteristics and medical history variables of the CenteringPregnancy participants were assessed using descriptive statistics such as t tests for continuous variables and the chi square (χ^2) test for categorical variables. A simple correlation or comparison between birth certificate data and medical record data for GWG, including height, pregnancy weight, prepregnancy BMI, and weight at delivery, was performed. Means, standard deviations, mean differences (birth certificate minus medical record) and 95% confidence intervals were calculated for all continuous variables.

Overall distributions of prepregnancy BMI (height and prepregnancy weight) and GWG categories derived from each source were assessed and matched. For the prepregnancy BMI categories (underweight, normal weight, overweight, and obese) and GWG categories (inadequate, adequate, and excessive) determined using the medical records and birth certificates, we calculated the percent agreement and kappa statistics to account for chance agreement between the two data sources. The sensitivity, specificity, negative predictive value (NPV) and positive predictive value (PPV) were also calculated for these two variables using the birth certificate data as the test data and EMR data as the truth or reference data. For these metrics of validation, we considered all measures above 70% as acceptable, which was suitable for the question of interest [15–17]. STATA/SE 14.1 software was used for all analyses. The protocol was reviewed and approved by the Prisma Health Medical Group and the University of South Carolina Institutional Review Board (IRB) for Human Research (Study ID #-Pro00096005). All guidelines, including treating data as confidential and not making attempt to identify individual participants were observed.

Table 2 Mean, standard deviation mean difference, and distribution of reporting errors in abstracted variables reported in electronic medical records compared to birth certificates, N = 173

	N	Birth certificate		Medical records		Mean difference between birth certificates and medical records (95% CI)	Simple correlation of both data sources
		Mean	SD	Mean	SD		
Height (cm)	173	162.3	7.7	162.6	7.6	− 0.26 (− 0.66, 0.14)	0.94
Prepregnancy weight (kg)	173	75.2	19.8	74.7	19.7	0.51 (− 0.63, 1.64)	0.93
Prepregnancy body mass index (BMI) (kg/m^2)	173	28.5	7.2	28.2	6.9	0.32 (− 0.11, 0.75)	0.92
Delivery weight (kg)	173	86.8	18.6	87.2	19.6	− 0.39 (− 1.16, 0.39)	0.96
Total weight gain (kg)	173	11.6	8.6	12.5	7.9	− 0.91 (− 2.02, 0.23)	0.60

SD means standard deviation

CI means confidence interval

Table 3 Agreement of prepregnancy BMI categories by birth certificates compared to electronic medical records N = 173

Birth certificate classification	Medical record classification								
	Under weight	Normal weight	Overweight	Obese	Row total	Sen (%)	Spec (%)	PPV (%)	NPV (%)
Underweight	4	1	0	0	5	67	99	80	99
Normal weight	2	47	3	1	53	77	95	89	88
Overweight	0	11	37	6	54	80	87	69	92
Obese	0	2	6	53	61	88	93	87	94
Total	6	61	46	60	173				

Crude and weight agreement proportions between birth certificates and medical records for underweight, normal weight, overweight and obese body mass index are 81.50% and 89.98% respectively

Statistical values for crude and weighted kappa for underweight, normal weight, overweight and obese body mass index are 0.73 and 0.78 respectively

Sen = sensitivity, Spec = specificity, PPV = positive predictive value, NPV = negative predictive value

3 Results

After applying the above inclusion and exclusion criteria, our study population or sample size included 173 (84%) participating women/births who were successfully matched across data sources.

The sociodemographic and maternal characteristics of the participating women for the period under consideration are shown in Table 1.

Table 2 shows the comparison of the data from each source. The mean values for EMR abstracted data for height, prepregnancy weight, prepregnancy BMI, delivery weight and total weight gain correlated with the birth certificate data. The height, prepregnancy weight, prepregnancy BMI, and delivery weight values all had a simple correlation > 0.9, as shown in the table. The mean differences in height, pregnancy weight, delivery weight and BMI values from the birth certificates compared to the medical records were quite small.

Table 3 displays the agreement between the birth certificate and EMR data on key characteristics. The birth certificates were used to classify the women into prepregnancy BMI categories, which was similar to the classifications using the medical records (weighted agreement proportion equals 89.98%) but with slightly fewer women classified as normal weight according to birth certificates versus EMR data and slightly more women classified as overweight. Overall, the sensitivity was highest for the obese BMI category (88%) and lowest for the underweight BMI category (67%). The specificity was high for all four BMI categories, with the overweight BMI category having the lowest value (87%). The positive predictive value (PPV) was highest for the normal-weight BMI category (89%) and lowest for the overweight BMI category (69%). The negative predictive value (NPV) was high for all four BMI categories, with the normal-weight BMI category having the lowest value (88%).

Table 4 shows the agreement for gestational weight gain (GWG) between birth certificate and medical record abstracted data among participating women who had singleton infants (weighted agreement proportion of 84.10%). Most of these women delivered at term, i.e., greater than 37 weeks' gestation (90.17%). Birth certificate data classified more participating women as having inadequate weight gain, i.e., below the IOM recommendation, compared to medical record abstracted data, and fewer as having excess weight gain according to the IOM recommendation. Sensitivity was highest for women with excessive weight gain in pregnancy (76%) and lowest for those with adequate GWG (71%). Specificity was high for all three GWG categories, with the inadequate GWG category having the lowest value (83%). The positive predictive value (PPV) was highest for excessive GWG (84%) and lowest for inadequate GWG (64%). The negative predictive value (NPV) was high for all three GWG categories, with the excessive GWG category having the lowest value (81%).

4 Discussion

As recommended by other investigators, this study examined the accuracy of prepregnancy BMI (height and prepregnancy weight) and GWG records using South Carolina birth certificate data compared to EMR abstracted data, which is the gold standard [4–6]. The study used data from a subsample of pregnant women who participated in the CenteringPregnancy group-based prenatal care program from 2015 to 2019 in three out of five obstetric sites in the Midlands of South Carolina. Women participating in the program and those receiving standard prenatal care were sufficiently similar regarding their characteristics, and the results from one source applied to the other. Overall, birth certificate mean estimates for height ($r = 0.94$), prepregnancy weight ($r = 0.93$), prepregnancy BMI ($r = 0.92$), and delivery weight ($r = 0.96$) largely correlated with the EMR data. Total weight gain was also correlated ($r = 0.60$) but not as strongly as the other variables. The mean differences in the variables between both data sources were quite small. A considerable number of women had height and weight at delivery values on birth certificates that were within a good reporting range of the EMR. Underreporting was common for prepregnancy weight, prepregnancy BMI and total weight gain. Prepregnancy body mass index (BMI) categories (underweight, normal weight, overweight, obese) for birth certificates agreed with those of EMRs, although birth certificates classified slightly fewer women as having normal weight and slightly more as being overweight compared to EMR abstracted data. For BMI categories, the BC data were both reasonably precise and accurate (PPV range between 69 and 89%) and somewhat all-inclusive (sensitivity range between 67 and 88%).

Prepregnancy weight values that were underreported were most likely to have contributed to the misclassification of prepregnancy BMI categories. This variable can be improved upon by measuring the weight at the first prenatal visit or just prior to conception for quality assurance and avoiding the use of self-reported figures. Additionally, birth certificate gestational weight gain categories (inadequate, adequate, excessive) were similar to those in the EMR data, although birth certificates classified slightly more women as having inadequate weight gain, i.e., below the IOM recommendation, and slightly fewer women as having excess weight gain in comparison to medical records. Regarding GWG categories, birth certificate data are reasonably accurate (PPV between 64 and 84%) and moderately

Table 4 Agreement of gestational weight gain (GWG) categories by birth certificates compared to electronic medical records, N = 173

Birth certificate classification	Medical record classification							
	Inadequate GWG	Adequate GWG	Excessive GWG	Row Total	Sen (%)	Spec (%)	PPV (%)	NPV (%)
Inadequate GWG	37	8	13	58	73	83	64	88
Adequate GWG	6	30	6	42	71	91	71	91
Excessive GWG	8	4	61	73	76	87	84	81
Column total	51	42	80	173				

Institute of Medicine (IOM) Gestational Weight Gain (GWG) Recommendations

Crude and weighted agreement proportions between birth certificates and medical records for inadequate, adequate, and excessive gestational weight gain are 73.99% and 84.10%, respectively

Statistical values for crude and weighted kappa for inadequate, adequate, and excessive gestational weight gain are 0.61 and 0.63 respectively

Sen = sensitivity, Spec = specificity, PPV = positive predictive value, NPV = negative predictive value inclusive (sensitivity ranges between 71 and 76%). As mentioned before, improvement in prepregnancy weight documentation can improve data on prepregnancy BMI categories, therefore enhancing GWG measurements and its categorization and avoiding misclassification. Our findings show that birth certificate data can provide reasonable estimates of these variables, at least in South Carolina.

Previous studies suggested the need for the validation of the quality of these variables (prepregnancy BMI and GWG and their categories) in birth certificate data because of mixed results or findings from prior studies on group-based prenatal care programs, with some showing a positive association, some showing a negative association and some showing no significant difference [18]. The findings from these studies (summarized in Appendices Tables 6 and 7) are contrary or inconsistent with our results with respect to BMI and GWG categories, although the mean estimates of related variables were close to that of the gold standard. For example, Park's study in Florida in 2005 investigated the reliability and validity of height, weight and prepregnancy BMI records in the Women, Infants and Children (WIC) Nutrition Program dataset compared to birth certificates (gold standard) and found that WIC data minimally over-estimated the prevalence of underweight and normal weight and slightly underestimated the prevalence of over-weight and obesity according to BMI. The study did not evaluate GWG [9]. The difference in findings was also noted in Bodnar's study in Pennsylvania in 2014 and the Deputy study in New York and Vermont in 2018, which compared prepregnancy BMI and GWG data from birth certificates and PRAMS with data from EMRs (gold standard). Some of the variables were slightly overestimated or underestimated compared to the gold standard (EMRs). The reasons for the variation in results may be because the studies were carried out in different states with different populations of women and also because some studies used different gold standards, such as birth certificates. South Carolina data add to current knowledge, as the state has a population that may differ from other states in which similar work was done. It is important that researchers continue to monitor the accuracy of data for these variables on birth certificates. Researchers should continue to put effort into screening or evaluating the quality of maternal prepregnancy BMI and GWG categories in birth certificate data in different settings, as high-quality data give accurate, consistent, and reliable results in quantitative research that can better inform decision-making for health services policies. Overall, the South Carolina birth certificate form still provides a reasonable estimate of the prevalence of these variables for research purposes, for example, in examining the effect or impact of different prenatal care programs.

The strength of this validation study is that it validated the use of South Carolina birth certificate data for studies of prenatal care programs. We recognized the various limitations of our study. In addition to being limited to a single state, our study population was largely reflective of the experience of African American women who are between the ages of 20 and 29 years with at least a high school diploma, so future studies should consider a different age group in the same or different populations and different settings. Additionally, the sample size is another limitation, so we recommend that future studies expand on this limitation. Nevertheless, the results from our validation study show that in South Carolina, birth certificate estimates for height, prepregnancy BMI (and categories), prepregnancy weight, delivery weight and gestational weight gain categories were similar to those of electronic medical records; thus, the South Carolina birth certificate database is a valid database that can provide reasonable estimates for these variables for public health practice, future research purposes and particularly for the state's evaluation of the CenteringPregnancy program.

Acknowledgements The authors would like to express sincere gratitude to CenteringPregnancy Manager, Ms. Beverly Huegel, who provided support to the conception and logistics to data acquisition. We also extend our appreciation to Ivory Harding, Allyson Malbouf and other members of the OBGYN department and information technology department at Prisma Health-University of South Carolina, Columbia, South Carolina, United States of America (USA) for their assistance with research logistics and participants' data gathering. We could not have successfully assembled patients without their help and support.

Author contributions All authors made substantial contributions to virtually all aspects of the research. Conceptualization, design, acquisition of data, analysis, interpretation of data, editing and revision by [OAM], [RDH], [JL], [ELC] and [BC]. Formal analysis and investigation and the first draft of the manuscript was written by [OAM]. All authors read and approved the final version of the manuscript.

Funding There's no funding to report on this research.

Data availability The datasets used and/or analyzed during this study are available from South Carolina Revenue and Fiscal Affairs (SC RFA) and with the corresponding author.

Code availability STATA software was used to carry out statistical analyses for the research. All codes used in analyses are available with corresponding author.

Declarations

Competing interests The authors have no competing or conflict of interests to disclose.

Ethics approval and consent to participate The protocol was reviewed and approved by the Prisma Health Medical Group and the University of South Carolina Institutional Review Board (IRB) for Human Research (Study ID #-Pro00096005). All guidelines, including treating data as confidential and not making attempt to identify individual participants were observed.

Consent to participate Not applicable.

Consent for publication Not applicable.

Appendix

See Tables 5, 6, and 7.

Table 5 Study sampling across sites

Sites	Total patients	Sampling ratio	Sample size	Total sample size by site
Clinic 1				64
2015–2018	168	0.25	42	
Jan/Feb to Dec 2018	88	0.25	22	
Clinic 2				90
2015–2018	101	0.25	25	
Jan/Feb to Dec 2018	259	0.25	65	
Clinic 3				52
Jan/Feb to Dec 2018	52	1.00	52	
Grand total	668		206	206

Table 6 Comparison of previous studies that validated pre-pregnancy body mass index (BMI)

Author and year of study	State	Gold standard data	Comparison data	Underweight	Normal	Overweight	Obese
Park et al. [9]	Florida	BC	1st trimester WIC	↑	↑	↓	↓
Bodnar et al. [8]	Pennsylvania	EMR	BC	↔	↔	↔	↓
Deputy et al. [7]	New York and Vermont	EMR	BC	↔	↔	↔	↔
		EMR	PRAMS	↑	↑	↓	↓
Current study 2022	South Carolina	EMR	BC	↓	↓	↑	↑

**BC—Birth Certificate

**WIC—Women, Infants and Children (WIC) Nutrition Program

**EMR—Electronic Medical Records

**PRAMS—Pregnancy Risk Assessment Monitoring System

**↑ = Overestimate ↓ = Underestimate ↔ = Similar

Table 7 Comparison of studies that validated gestational weight gain (GWG)

Author and year of study	State	Gold Standard Data	Comparison Data	Inadequate	Adequate	Excessive
Park et al. [9]	Florida	BC	1st trimester WIC	–	–	–
Bodnar et al. [8]	Pennsylvania	EMR	BC	–	–	–
Deputy et al., [7]	New York and Vermont	EMR	BC	↑	↓	↑
Current study 2022	South Carolina	EMR	BC	↑	↔	↓

**BC—Birth Certificate

**WIC—Women, Infants and Children (WIC) Nutrition Program

**EMR—Electronic Medical Records

**PRAMS—Pregnancy Risk Assessment Monitoring System

**↑ = Overestimate ↓ = Underestimate ↔ = Similar – = information not available

References

1. Gilmore LA, Redman LM. Weight gain in pregnancy and application of the 2009 IOM guidelines: toward a uniform approach. Obesity (Silver Spring, Md). 2015;23(3):507–11. https://doi.org/10.1002/oby.20951.
2. Marshall NE, Abrams B, Barbour LA, Catalano P, Christian P, Friedman JE, Hay WW, Hernandez TL, Krebs NF, Oken E, Purnell JQ, Roberts JM, Soltani H, Wallace J, Thornburg KL. The importance of nutrition in pregnancy and lactation: lifelong consequences. Am J Obstet Gynecol. 2021;S0002–9378(21):02728–9. https://doi.org/10.1016/j.ajog.2021.12.035.
3. Martínez García RM, Jiménez Ortega AI, Peral Suárez Á, Bermejo López LM, Rodríguez-Rodríguez E. [Importance of nutrition during pregnancy. Impact on the composition of breast milk]. Nutr Hospital. 2021;37(Spec No 2):38–42.
4. Kominiarek MA, Crockett A, Covington-Kolb S, Simon M, Grobman WA. Association of group prenatal care with gestational weight gain. Obstet Gynecol. 2017;129(4):663–70. https://doi.org/10.1097/AOG.0000000000001940.
5. Kominiarek MA, Lewkowitz AK, Carter E, Fowler SA, Simon M. Gestational weight gain and group prenatal care: a systematic review and meta-analysis. BMC Pregnancy Childbirth. 2019. https://doi.org/10.1186/s12884-018-2148-8.
6. Dietz P, Bombard J, Mulready-Ward C, Gauthier J, Sackoff J, Brozicevic P, Gambatese M, Nyland-Funke M, England L, Harrison L, Farr S. Validation of selected items on the 2003 U.S. standard certificate of live birth: New York City and Vermont. Public Health Rep (Washington, DC: 1974). 2015;130(1):60–70. https://doi.org/10.1177/003335491513000108.
7. Deputy NP, Sharma AJ, Bombard JM, Lash TL, Schieve LA, Ramakrishnan U, Stein AD, Nyland-Funke M, Mullachery P, Lee E. Quality of maternal height and weight data from the revised birth certificate and pregnancy risk assessment monitoring system. Epidemiology. 2019;30(1):154–9. https://doi.org/10.1097/EDE.0000000000000936.
8. Bodnar LM, Abrams B, Bertolet M, Gernand AD, Parisi SM, Himes KP, Lash TL. Validity of birth certificate-derived maternal weight data: validity of maternal weight data. Paediatr Perinat Epidemiol. 2014;28(3):203–12. https://doi.org/10.1111/ppe.12120.
9. Park S, Sappenfield WM, Bish C, Bensyl DM, Goodman D, Menges J. Reliability and validity of birth certificate prepregnancy weight and height among women enrolled in prenatal WIC program: Florida, 2005. Matern Child Health J. 2011;15(7):851–9. https://doi.org/10.1007/s10995-009-0544-4.
10. Martin JA, Wilson EC, Osterman MJK, Saadi EW, Sutton SR, Hamilton BE. Assessing the quality of medical and health data from the 2003 birth certificate revision: results from two states. Natl Vital Stat Rep Centers Disease Control Prevent Natl Center Health Stat Natl Vital Stat Syst. 2013;62(2):1–19.
11. Zozus MN, Pieper C, Johnson CM, Johnson TR, Franklin A, Smith J, Zhang J. Factors affecting accuracy of data abstracted from medical records. PLoS ONE. 2015;10(10):e0138649. https://doi.org/10.1371/journal.pone.0138649.
12. Frosst G, Hutcheon J, Joseph K, Kinniburgh B, Johnson C, Lee L. Validating the British Columbia perinatal data registry: a chart re-abstraction study. BMC Pregnancy Childbirth. 2015. https://doi.org/10.1186/s12884-015-0563-7.
13. Kjerulff KH, Attanasio LB. Validity of birth certificate and hospital discharge data reporting of labor induction. Women's Health Issues. 2018;28(1):82–8. https://doi.org/10.1016/j.whi.2017.10.005.
14. Lange SJ, Moore LV, Galuska DA. Data for decision-making: exploring the division of nutrition, physical activity, and obesity's data, trends, and maps. Prev Chronic Dis. 2019;16:E131. https://doi.org/10.5888/pcd16.190043.
15. Williamson T, Miyagishima RC, Derochie JD, Drummond N. Manual review of electronic medical records as a reference standard for case definition development: a validation study. CMAJ Open. 2017;5(4):E830–3. https://doi.org/10.9778/cmajo.20170077.
16. Coleman N, Halas G, Peeler W, Casaclang N, Williamson T, Katz A. From patient care to research: a validation study examining the factors contributing to data quality in a primary care electronic medical record database. BMC Fam Pract. 2015. https://doi.org/10.1186/s12875-015-0223-z.
17. Tirschwell DL, Longstreth WT. Validating administrative data in stroke research. Stroke. 2002;33(10):2465–70. https://doi.org/10.1161/01.STR.0000032240.28636.BD.
18. Carter EB, Temming LA, Akin J, Fowler S, Macones GA, Colditz GA, Tuuli MG. Group prenatal care compared with traditional prenatal care: a systematic review and meta-analysis. Obstet Gynecol. 2016;128(3):551–61. https://doi.org/10.1097/AOG.0000000000001560.